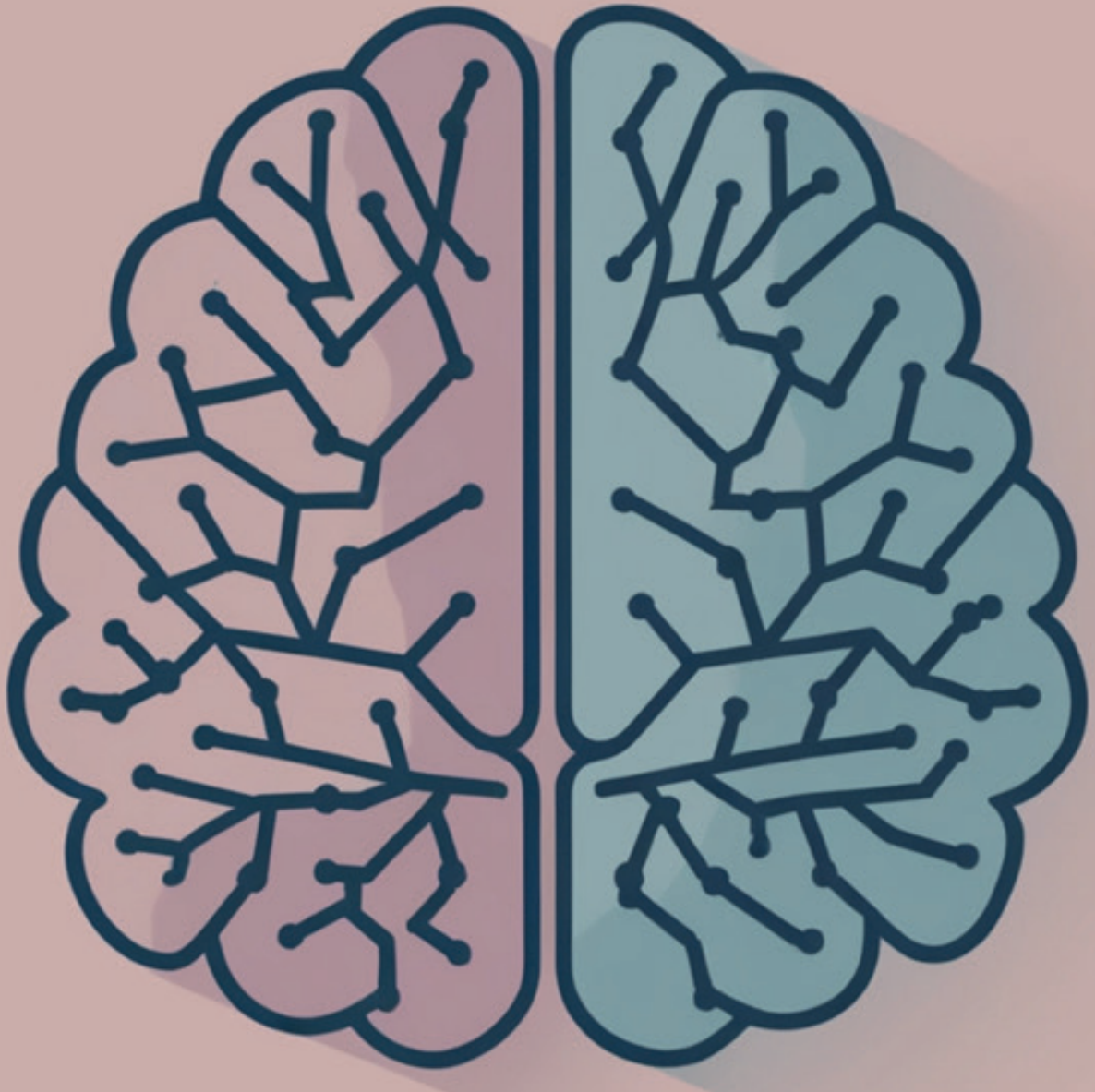
Discrete Mathematics Days 2024



Delia Garijo Royo
David Orden Martín
Francisco Santos Leal
(Eds.)

OBRAS COLECTIVAS
CIENCIAS 20

UAH

Discrete Mathematics Days 2024

Delia Garijo Royo
David Orden Martín
Francisco Santos Leal
(Eds.)

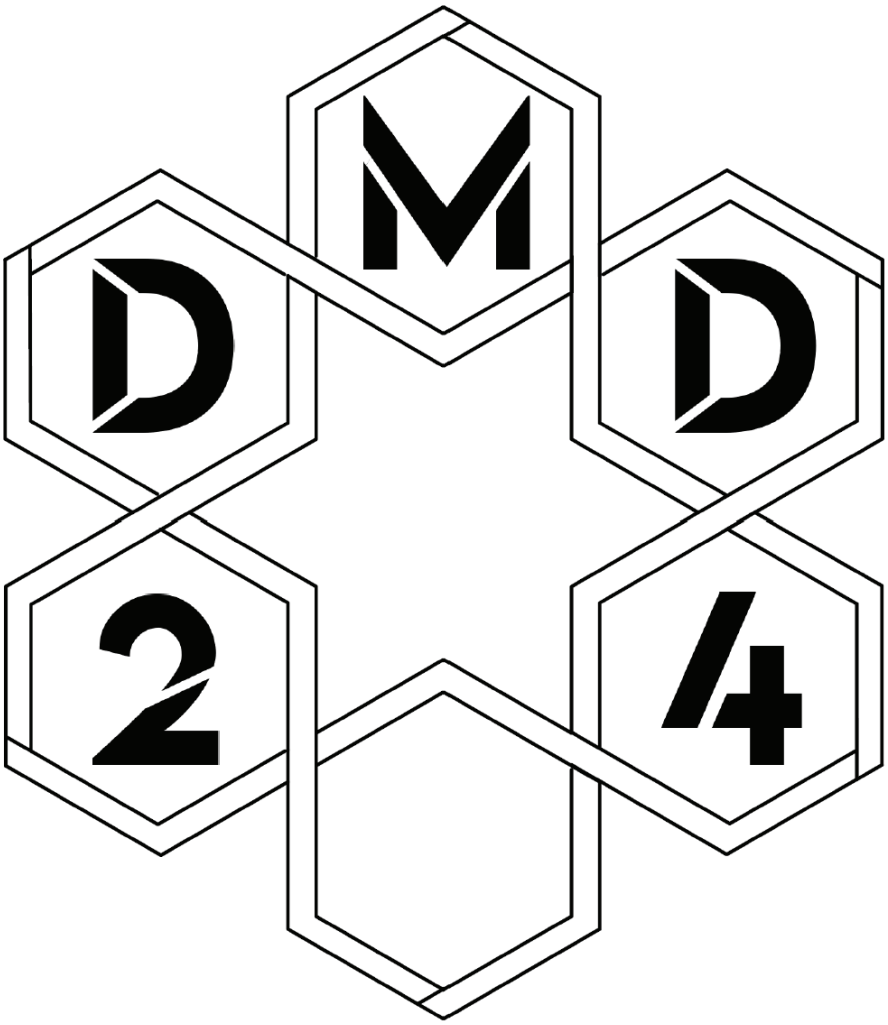
Universidad de Alcalá
EDITORIAL UNIVERSIDAD DE ALCALÁ

El contenido de este libro no podrá ser reproducido,
ni total ni parcialmente, sin el previo permiso escrito del editor.
Todos los derechos reservados.

© De los textos: sus autores
© De las imágenes: sus autores
© De la ilustración de portada: Leonardo AI
© Editorial Universidad de Alcalá, 2024
Plaza de San Diego, s/n
28801 Alcalá de Henares
www.uah.es

I.S.B.N.: 978-84-18979-38-5
https://doi.org/10.37536/TYSP5643

Discrete Mathematics Days 2024



Universidad de Alcalá

Universidad de Alcalá
DEPARTAMENTO DE FÍSICA Y MATEMÁTICAS

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
Departament de Matemàtiques

DAM network

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

AGENCIA ESTATAL DE INVESTIGACIÓN

Table of Contents

Invited talks

Graph universality 1
Julia Böttcher

Coboundary expansion, codes, and agreement tests 2
Irit Dinur

Counting polytopes 3
Arnau Padrol

Recent work on the Erdős-Hajnal Conjecture 4
Alex Scott

Ramon Llull prize talk

Hamilton cycles in random graphs 5
Alberto Espuny Díaz

Contributed talks

Switching methods for the construction of cospectral graphs 7
Aida Abiad, Nils van de Berg and Robin Simoens

The Four-Color Ramsey Multiplicity of Triangles 13
Aldo Kiem, Sebastian Pokutta and Christoph Spiegel

Speed and size of dominating sets in domination games 19
Ali Deniz Bagdas, Dennis Clemens, Fabian Hamann and Yannick Mogge

On the solutions of linear systems over additively idempotent semirings 25
Alvaro Otero Sanchez, Daniel Camazón Portela and Juan Antonio López Ramos

Rainbow loose Hamilton cycles in Dirac hypergraphs 30
Amarja Kathapurkar, Patrick Morris and Guillem Perarnau

d-regular graph on n vertices with the most k-cycles 36
Arturo Ortiz San Miguel and Gabor Lippner

The weight spectrum of the Reed-Muller codes RM(m5,m) 42
Claude Carlet

Separating Cycle Systems 48
Fábio Botler and Tássio Naia

The algorithmic fried potato problem in two dimensions 53
Francisco Criado Gallart and Francisco Santos Leal

Three-term arithmetic progressions in two-colorings of the plane. 59
Gabriel Currier, Kenneth Moore and Chi Hoi Yip

Bounding the balanced upper chromatic number 65
Gabriela Araujo-Pardo, Silvia Fernández-Merchant, Adriana Hansberg, Dolores Lara, Amanda Montejano and Déborah Oliveros

Creating trees with high maximum degree 71
Grzegorz Adamski, Małgorzata Bednarska-Bzdega, Sylwia Antoniuk, Dennis Clemens, Fabian Hamann and Yannick Mogge

Random lifts of very high girth and their applications to frozen colourings 77
Guillem Perarnau and Giovanne Santos

A covering problem for zonotopes and Coxeter permutahedra 83
Gyula Károlyi

Classification of Edge-to-edge Monohedral Tilings of the Sphere 89
Hoi Ping Luk, Ho Man Cheung and Min Yan

Betti numbers of monomial curves 95
Ignacio García Marco, Philippe Gimenez and Mario González-Sánchez

The flexibility among 3-decompositions 101
Irene Heinrich and Lena Volk

Computing edge-colored ultrahomogeneous graphs 107
Irene Heinrich, Eda Kaja and Pascal Schweitzer

Regular polytopes, sphere packings and Apollonian sections 113
Iván Rasskin

Disconnected common graphs via supersaturation 119
Jae-Baek Lee and Jonathan Noel

Bicolored point sets admitting non-crossing alternating Hamiltonian paths 124
Jan Soukup

On homogeneous matroid ports 130
Jaume Martí-Farré and Anna de Mier

Enumeration of unlabelled chordal graphs with bounded tree-width 136
Jordi Castellví and Clément Requilé

The Borsuk number of a graph 142
José Cáceres, Delia Garijo, Alberto Marquez and Rodrigo Silveira

A canonical van der Waerden theorem in random sets 148
José D. Alvarado, Yoshiharu Kohayakawa, Patrick Morris, Guilherme O. Mota and Miquel Ortega

Multi-objective Linear Integer Programming based in Test Sets 154
José María Ucha, María Isabel Hartillo and Haydee Jiménez

Rainbow connectivity of multilayered random geometric graphs 160
Josep Diaz, Oznur Yasar Diner, Maria Serna and Oriol Serra

Polytope Neural Networks 166
Juan L. Valerdi

Expressing the coefficients of the chromatic polynomial in terms of induced subgraphs: a systematic approach 172
Kerri Morgan and Lluis Vena

Extending the Continuum of Six-Colorings 178
Konrad Mundiger, Sebastian Pokutta, Christoph Spiegel and Max Zimmer

On Ewald's and Nill's Conjectures about smooth polytopes 184
Luis Crespo Ruiz, Álvaro Pelayo and Francisco Santos

Integer programs with nearly totally unimodular matrices: the cographic case 190
Manuel Aprile, Samuel Fiorini, Gwenael Joret, Stefan Kober, Michał T. Seweryn, Stefan Weltge and Yelena Yuditsky

Limit theorems for the Erdős–Rényi random graph conditioned on being a cluster graph .. 196
Marc Noy, Martijn Gösgens, Lukas Lüchtrath, Elena Magnanini and Élie de Panafieu

On the sum of several finite subsets in $\R^2$ 202
Mario Huicochea, René González-Martínez, Amanda Montejano and David Suárez

On a conjecture concerning the roots of Ehrhart polynomials of symmetric edge polytopes from complete multipartite graphs 208
Max Kölbl

An Approximate Counting Version of the Multidimensional Szemer\'edi Theorem 214
Natalie Behague, Joseph Hyde, Natasha Morrison, Jonathan Noel and Ashna Wright

A short proof of an inverse theorem in bounded torsion groups 218
Pablo Candela, Diego Gonzalez-Sanchez and Balázs Szegedy

Complexity measures of trilean functions 224
Sara Asensio, Ignacio García-Marco and Kolja Knauer

Geometric quasi-cyclic low density parity check codes 228
Simeon Ball and Tomàs Ortega

On additive codes over finite fields 231
Simeon Ball, Michel Lavrauw and Tabriz Popatia

Increasing paths in the temporal stochastic block model 236
Sofiya Burova, Gabor Lugosi and Guillem Perarnau

Ranges of polynomials control degree ranks of Green and Tao over finite prime fields 241
Thomas Karam

Product representation of perfect cubes 247
Zsigmond György Fleiner, Márk Hunor Juhász, Blanka Kövér, Péter Pál Pach and Csaba Sándor

Poster presentations

Computing 2-homogeneous equitable partitions of graphs with a unique tree representation 253
Aida Abiad and Sjanne Zeijlemaker

Categorification of Flag Algebras 259
Aldo Kiem, Christoph Spiegel and Sebastian Pokutta

The rectilinear convex hull of disks 265
Carlos Alegría, Justin Dallant, Jean-Paul Doignon, Pablo Pérez-Lantero and Carlos Seara

Characterization of the equality in some discrete isoperimetric and Brunn-Minkowski type inequalities 270
Eduardo Lucas Marín and David Iglesias López

Totally Greedy Sequences Generated by a Class of Second-Order Linear Recurrences With Constant Coefficients 276
Hebert Pérez-Rosés

An algebraic approach to the Weighted Sum Method in Multi-objective Integer Programming 282
José Manuel Jiménez, José María Ucha and Haydee Jiménez

Sidorenko-type inequalities for Trees 288
Lina Simbaqueba, Natalie Behague, Gabriel Crudele and Jon Noel

A note on generalized crowns in linear r-graphs 293
Linpeng Zhang, Hajo Broersma and Ligong Wang

A Kneser-type theorem for restricted sumsets 298
Mario Huicochea

Preface

This book contains the extended abstracts of the invited talks, contributed talks, and contributed posters accepted for presentation at the *Discrete Mathematics Days* 2024. This international conference was held in Alcalá de Henares, Spain, on July 3-5, 2024, focusing on current topics in Discrete Mathematics and being a satellite event of the 9th European Congress of Mathematics.

Organized every two years, this conference inherited in 2016 the long tradition of the *Jornadas de Matemática Discreta*, organized biennially in Spain since 1998. It combines a strong scientific program with a friendly atmosphere, gathering audience from reputed senior researchers to master and doctoral students. This book includes 52 contributions together with the abstracts of the five invited talks, which include the first edition of the Ramon Llull Prize talk.

As for each and all of them, this edition has been the result of the efforts of a number of people. First, the members of the organizing committee, who put their best for the success of this event:

- Guillermo Esteban (co-chair), Universidad de Alcalá.
- Andrea de las Heras, Universitat Politècnica de Catalunya.
- David Orden (co-chair), Universidad de Alcalá.
- Marino Tejedor-Romero, Universidad de Alcalá.
- Lluís Vena, Universitat Politècnica de Catalunya.

We also want to thank the members of the scientific committee, for contributing their expertise in a careful and constructive way:

- Aida Abiad, Eindhoven University of Technology.
- Marie Albenque, Université Paris Cité.
- Sergio Cabello, Univerza v Ljubljani.
- Pablo Candela, Universidad Autónoma de Madrid.
- Vida Dujmović, University of Ottawa.
- Alberto Espuny Díaz, Universität Heidelberg.
- Stefan Felsner, Technische Universität Berlin.
- Delia Garijo (co-chair), Universidad de Sevilla.
- Gyula Károlyi, Eötvös University and Renyi Institute Budapest.
- Dan Král', Masaryk University Brno.
- Marc Noy, Universitat Politècnica de Catalunya.
- Diego Ruano, Universidad de Valladolid.
- Francisco Santos (co-chair), Universidad de Cantabria.

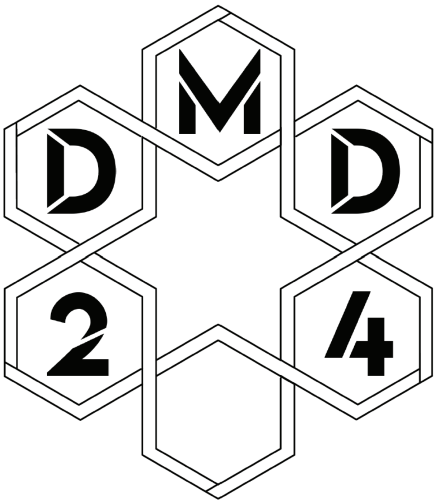
- Pascal Schweitzer, Technische Universität Darmstadt.
- María Serna, Universitat Politècnica de Catalunya.
- Maya Stein, Universidad de Chile.
- Julia Wolf, University of Cambridge.
- Öznur Yaşar, Kadir Has University.

Finally, we are grateful to the institutions that supported this edition:

- *Universidad de Alcalá* and its *Departamento de Física y Matemáticas*.
- *Universitat Politècnica de Catalunya* and its *Departament de Matemàtiques*.
- *Discrete and Algorithmic Mathematics Network*, project RED2022-134947-T funded by MCIN/ AEI /10.13039/501100011033.

July 2024,
Alcalá de Henares.

Delia Garijo,
David Orden,
Francisco Santos.



Invited talks

Graph universality

Julia Böttcher*[1]

[1]Department of Mathematics, London School of Economics and Political Sciences, Houghton Street, London WC2A 2AE, UK

Abstract

Given a class $\mathcal{G}$ of n-vertex graphs, how can we construct a host graph H that contains them all as subgraphs? Graphs H with this property are called universal for $\mathcal{G}$, and the question gets interesting when we put certain restrictions on H. For example, we might be interested in a graph H with as few edges as possible, or a graph H which has only n vertices itself and still only few edges. Or we might ask when certain random graphs are universal for $\mathcal{G}$. This all leads to a variety of interesting and challenging problems. In the talk, I will explain what is known and what is open for some classes of graphs $\mathcal{G}$. I will also detail some techniques that I recently used with my co-authors Peter Allen and Anita Liebenau for progress when $\mathcal{G}$ consists of all D-degenerate graphs for a fixed D.

*Email: j.boettcher@lse.ac.uk

Coboundary expansion, codes, and agreement tests

Irit Dinur*[1]

[1]Dept. of Computer Science and mathematics, Weizmann Institute of Science, Rehovot, Israel

Abstract

High dimensional expansion is a generalization of expansion in graphs to hypergraphs, simplicial complexes, and more general poset structures. Two main notions are studied: the first is a spectral notion that is related to random walks and mixing, and the second is a cohomological notion called coboundary expansion. Coboundary expansion was introduced by Linial and Meshulam, and by Gromov that combines combinatorics, topology, and linear algebra. Kaufman and Lubotzky observed its relation to "Property testing", and in recent years it has found several applications in theoretical computer science, including for error correcting codes (both classical and quantum), for PCP agreement tests, and even for studying polarization in social networks. In the talk I will introduce this notion and some of its applications. No prior knowledge is assumed, of course.

*Email: irit.dinur@weizmann.ac.il

Counting polytopes

Arnau Padrol*[1]

[1]Universitat de Barcelona and Centre de Recerca Matemàtica

Abstract

This talk will be an overview of the classical problem of estimating the number of combinatorial types of d-dimensional convex polytopes with n vertices, and its interactions with some of the milestones of combinatorial polytope theory. While in dimensions up to 3 we have a very good understanding on the asymptotic growth of the number of polytopes with respect to the number of vertices, in higher dimensions we only have coarse estimates. Upper bounds arise from results of Milnor and Thom from real algebraic geometry, whereas lower bounds are obtained with explicit constructions. I will present a recent construction giving the current best lower bounds for the number of polytopes, found in collaboration with Eva Philippe and Francisco Santos.

*Email: arnau.padrol@ub.edu. Research of A. P. supported by grants PID2022-137283NB-C21 of MCIN/AEI/10.13039/501100011033, CLaPPo (21.SI03.64658) of Universidad de Cantabria and Banco Santander, PAGCAP ANR-21-CE48-0020 of the French National Research Agency ANR, and SGR GiT-UB (2021 SGR 00697) from the Departament de Recerca i Universitats de la Generalitat de Catalunya.

Recent work on the Erdős-Hajnal Conjecture

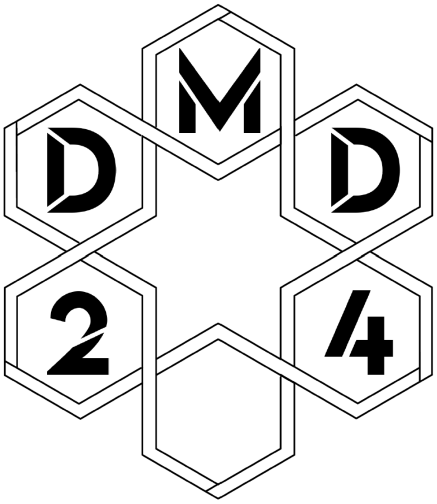
Alex Scott*[1]

[1]Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

A typical graph contains cliques and independent sets of no more than logarithmic size. The Erdős-Hajnal Conjecture asserts that if we forbid some induced subgraph H then we can do much better: the conjecture claims that there is some $c = c(H) > 0$ such that every H-free graph G contains a clique or independent set of size at least $|G|^c$. The conjecture looks far out of reach, and is only known for a small family of graphs. We will discuss some recent progress.

Joint work with Tung Nguyen and Paul Seymour.

*Email: scott@maths.ox.ac.uk. Research supported by EPSRC grant EP/X013642/1.



Ramon Llull Prize talk

Hamilton cycles in random graphs

Alberto Espuny Díaz*[1]

[1]Institut für Informatik, Universität Heidelberg, 69120 Heidelberg, Germany.

Abstract

Hamiltonicity (that is, the property of containing a cycle which covers all vertices of a graph) is among the simplest and most well-studied properties of graphs. It is well known that the associated decision problem is NP-complete, so we do not expect to find a nice characterisation of Hamiltonian graphs, which is why so much effort has been devoted to understanding conditions which are sufficient for Hamiltonicity. In parallel to this, however, a great deal of research has gone into understanding the "average case" behaviour, by considering probability distributions on different sets of graphs.

The most classical model of random graphs is the model of *binomial* random graphs, where edges appear independently with probability p. The Hamiltonicity of graphs in this model (and its closely related *uniform* model) has been well understood for decades. Other models of interest include random *regular* graphs or different models of random *geometric* graphs (though many other models have been studied as well).

More recently, a host of problems inspired by extremal graph theory have been considered in random graphs. They can broadly be classified into different subcategories. In one direction, given a graph G which is not Hamiltonian, one wishes to understand the "average case" behaviour of the supergraphs of G — this has led to the study of so-called *randomly perturbed* graphs. In the opposite direction, given a graph G which is Hamiltonian, we wish to understand the "average case" behaviour of its subgraphs — this relates to the *robustness* of Hamiltonicity in G. As a third direction, one may consider extremal questions on random graphs: does every subgraph of a random graph with some constraint (say, number of edges, or minimum degree) contain a Hamilton cycle? This direction has been dubbed the study of the *resilience* of Hamiltonicity. Other natural directions include counting, packings or coverings.

In this talk, we will survey some results about Hamiltonicity in each of these models and in each of the three main directions mentioned above, and discuss how they compare to one another.

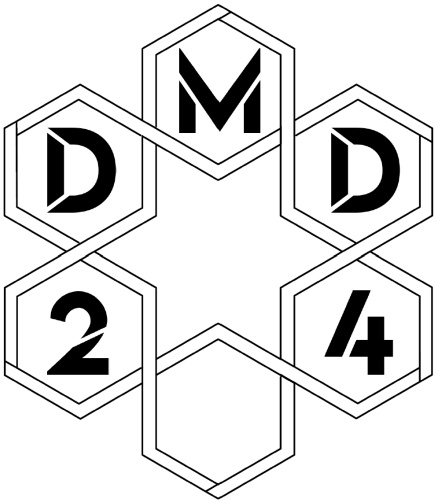
*Email: espuny-diaz@informatik.uni-heidelberg.de

AWARD MINUTES

The committee is happy to recommend that the first Ramon Llull Prize in Discrete Mathematics be awarded to Alberto Espuny Díaz for his PhD dissertation "Hamiltonicity problems in random graphs".

The thesis contains several excellent results, most notably the solution of a 40-year old conjecture of Bollobás regarding the threshold for the existence of a Hamilton cycle in the percolated hypercube, a truly outstanding achievement.

Committee:
Frédéric Havet (Université Côte d'Azur, CNRS, Inria)
Peter Keevash (Oxford)
Gabor Lugosi (ICREA-Universitat Pompeu Fabra, chair)



Contributed talks

Switching methods for the construction of cospectral graphs*

Aida Abiad[†1], Nils van de Berg[‡1], and Robin Simoens[§2]

[1]Dept. of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands
[2]Dept. of Mathematics: Analysis, Logic and Discrete Mathematics, Ghent University, Belgium;
Dept. of Mathematics, Universitat Politècnica de Catalunya, Spain

1 Introduction

An important problem in algebraic graph theory is to decide whether a graph is determined by the spectrum of its adjacency matrix (see the surveys [10, 11]). In 2003, van Dam and Haemers [10] conjectured that almost all graphs are uniquely determined by their spectrum. While the conjecture is still open, Brouwer and Spence [8] provided computational evidence by enumerating all graphs with up to 12 vertices and observing a decline in the fraction of *cospectral mates* (non-isomorphic graphs with the same spectrum) between 10 and 12 vertices. Recent work by Koval and Kwan [17] showed that an exponential number of graphs is determined by its spectrum. On the other side, Haemers and Spence [14] established an asymptotic lower bound for the number of cospectral mates. Their key ingredient is the notion of switching.

A *switching method* is an operation on a graph that results in a graph with the same spectrum. For such a method to work, the graph needs a special structure, called a *switching set*. This set of vertices makes it possible to swap some of the edges while preserving the spectrum of the adjacency matrix. While Godsil-McKay (GM) switching [13] is the oldest and most fruitful switching method in the literature (see e.g. [2, 3, 4]), new switching methods have recently been presented in the literature, most notably Wang-Qiu-Hu (WQH) switching [20] and Abiad-Haemers (AH) switching [4]. The latter captures all level 2 switching methods, and is motivated by the results of Wang and Xu [21], who suggested that almost all $\mathbb{R}$-*cospectral graphs* (cospectral graphs with cospectral complements) can be constructed using regular orthogonal matrices of level 2.

This work bridges a gap in the existing literature concerning the recently introduced switching methods of level 2. In particular, we present a combinatorial description of AH-switching that is more accessible than the algebraic description provided by Abiad and Haemers in [5]. We do this for switching sets of sizes 6, 8 and 10. Moreover, we show that the asymptotic lower bound on cospectral mates derived by Haemers and Spence [14] is tight for GM-switching. We also obtain analogous upper and lower bounds on the number of cospectral mates obtained via WQH-switching.

2 Preliminaries

In this work, graphs are considered to be simple and loopless. The *(adjacency) spectrum* of a graph is the multiset of eigenvalues of its adjacency matrix. Graphs are *cospectral* if they have the same spectrum. Two graphs are said to be *cospectral mates* if they are cospectral and non-isomorphic. Let I

*The full version of this work will be published elsewhere.

[†]Email: a.abiad.monge@tue.nl. Supported by the Dutch Research Council (NWO) through the grant VI.Vidi.213.085.

[‡]Email: n.p.v.d.berg@tue.nl. Supported by the Dutch Research Council (NWO) through the grant VI.Vidi.213.085.

[§]Email: Robin.Simoens@UGent.be. Supported by Research Foundation Flanders (FWO) through the grant 11PG724N.

denote the identity matrix and J the all-one matrix. Two graphs with adjacency matrices A and A' are called $\mathbb{R}$-*cospectral* if $A+rJ$ and $A'+rJ$ are cospectral for every $r \in \mathbb{R}$. An orthogonal matrix is *regular* if it has a constant row sum. Johnson and Newman [16] showed that two graphs are $\mathbb{R}$-cospectral if and only if their adjacency matrices are conjugated with a regular orthogonal matrix.

The *level* of a matrix is the smallest positive integer ℓ such that ℓ times the matrix is an integral matrix, or ∞ if it has irrational entries. A matrix is *decomposable* if it can be written as a non-trivial block-diagonal matrix after a certain permutation of the rows and columns. Otherwise, it is *indecomposable*.

3 Switching methods to construct cospectral graphs

The construction of cospectral graphs has multiple purposes: to disprove the conjecture stating that almost all graphs can be characterized by their spectrum for certain graph classes (see e.g. [3, 12]), to show which properties of a graph cannot be deduced from the spectrum (see e.g. [1, 7, 18]), or to construct new strongly regular and distance-regular graphs (see e.g. [6, 19]), among others. In what follows, we provide an overview of the existing switching methods, and present some new results concerning AH-switching.

3.1 Godsil-McKay switching

The following method for finding cospectral graphs was introduced by Godsil and McKay [13] in 1982.

Theorem 1 (GM-switching [13])**.** *Let Γ be a graph and let $\{C_1, \dots, C_t, D\}$ be a partition of its vertices such that, for all $i, j \in \{1, \dots, t\}$:*

(i) Every vertex in C_i has the same number of neighbours in C_j.

(ii) Every vertex in D has 0, $\frac{1}{2}|C_i|$ or $|C_i|$ neighbours in C_i.

For all $i \in \{1, \dots, t\}$ and every $v \in D$ that has exactly $\frac{1}{2}|C_i|$ neighbours in C_i, swap the adjacencies between v and C_i. The resulting graph is $\mathbb{R}$-cospectral with Γ.

The GM-switching operation corresponds to a conjugation of the adjacency matrix with the orthogonal matrix $\mathrm{diag}(R_1, \dots, R_t, I)$, where R_i equals the $|C_i| \times |C_i|$ matrix $\frac{2}{|C_i|}J - I$. Note that any C_i of order 2 only gives a permutation matrix and is therefore trivial. The simplest nontrivial case has one switching block of size four. This case has actually been the most fruitful in the literature, see e.g. [2, 3, 4]. Larger switching sets give more conditions on the graph, which intuitively explains the relative effectiveness of small switching sets. In Section 4, we give an asymptotic formula for the number of graphs with a switching set of size four. Note that the level of the corresponding matrix is 2 in that case, and the lowest common multiple of $\frac{1}{2}|C_i|$, $1 \le i \le t$, in general.

3.2 Wang-Qui-Hu switching

In 2019, Wang, Qiu and Hu [20] presented another switching method, which corresponds to a conjugation of the adjacency matrix with the orthogonal matrix $\mathrm{diag}(R_1, \dots, R_t, I)$, where each R_i is of the form

$$R_i = \begin{pmatrix} I - \frac{2}{|C_i|}J & \frac{2}{|C_i|}J \\ \frac{2}{|C_i|}J & I - \frac{2}{|C_i|}J \end{pmatrix}.$$

As illustrated in [3, 12, 15], WQH-switching is also a powerful tool for constructing cospectral graphs in cases where GM-switching fails. In combinatorial terms, the method can be described as follows.

Theorem 2 (WQH-switching [20])**.** *Let Γ be a graph and let $\{C_1^{(1)}, C_1^{(2)}, \dots, C_t^{(1)}, C_t^{(2)}, D\}$ be a partition of its vertices such that, for all $i, j \in \{1, \dots, t\}$:*

(i) $|C_i^{(1)}| = |C_i^{(2)}|$.

(ii) *The number* $\begin{cases} |N(v) \cap C_j^{(1)}| - |N(v) \cap C_j^{(2)}| & \text{if } v \in C_i^{(1)} \\ |N(v) \cap C_j^{(2)}| - |N(v) \cap C_j^{(1)}| & \text{if } v \in C_i^{(2)} \end{cases}$ *is the same for every* $v \in C_i^{(1)} \cup C_i^{(2)}$.

(iii) Every vertex in D *has either:*

- *(a)* $|C_i^{(1)}|$ *neighbours in* $C_i^{(1)}$ *and* 0 *neighbours in* $C_i^{(2)}$,
- *(b)* 0 *neighbours in* $C_i^{(1)}$ *and* $|C_i^{(2)}|$ *neighbours in* $C_i^{(2)}$,
- *(c) the same number of neighbours in* $C_i^{(1)}$ *as in* $C_i^{(2)}$.

For all $i \in \{1, \dots, t\}$ *and every* $v \in D$ *for which (a) or (b) holds, swap the adjacencies between* v *and* $C_i^{(1)} \cup C_i^{(2)}$. *The resulting graph is* $\mathbb{R}$*-cospectral with* Γ.

If $t = 1$ and $|C_1^{(1)}| = |C_2^{(2)}| = 2$, then WQH-switching is equivalent to GM-switching on $C_1^{(1)} \cup C_1^{(2)}$. But in general, they are different operations. Note that the level of the corresponding matrix is equal to the lowest common multiple of $\frac{1}{2}|C_i|$, $1 \le i \le t$, just like for GM-switching.

3.3 Abiad-Haemers switching

In 2012, Abiad and Haemers [5] considered switching methods that correspond to a conjugation of the adjacency matrix with a regular orthogonal matrix of level 2. In particular, these methods can be used to construct $\mathbb{R}$-cospectral graphs (see the characterization of $\mathbb{R}$-cospectral graphs by Johnson and Newman [16]). Their starting point is the following classification of indecomposable regular orthogonal matrices of level 2, which follows from the classification of weighing matrices of weight 4 by Chan, Rodger and Seberry [9] and has been restated by Wang and Xu [21] in the form below. Note that any regular orthogonal matrix has row sum 1 or -1, but without loss of generality, we may assume the row sum to be 1.

Theorem 3 ([9])**.** *Up to a permutation of the rows and columns, an indecomposable regular orthogonal matrix of level* 2 *and row sum* 1 *is one of the following:*

$$(i)\ \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix},\quad (ii)\ \frac{1}{2}\begin{bmatrix} J & O & \cdots & \cdots & O & Y \\ Y & J & O & \cdots & \cdots & O \\ O & Y & J & O & \cdots & O \\ & \ddots & \ddots & \ddots & \ddots & \\ O & \cdots & O & Y & J & O \\ O & \cdots & \cdots & O & Y & J \end{bmatrix},$$

$$(iii)\ \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & -1 \end{bmatrix},\quad (iv)\ \frac{1}{2}\begin{bmatrix} -I & I & I & I \\ I & -Z & I & Z \\ I & Z & -Z & I \\ I & I & Z & -Z \end{bmatrix},$$

where I, J, O, $Y = 2I - J$ *and* $Z = J - I$, *are square matrices of order* 2.

The matrix in Theorem 3(i) corresponds to GM-switching. The one in Theorem 3(ii) is an infinite family of matrices of even order, starting from order 6. In the following, we focus on this infinite family. The remaining matrices in Theorem 3(iii)-(iv) were studied by Abiad and Haemers in [5, Section 5 and Section 6].

It was already noticed by Abiad and Haemers [5] that sometimes, the six vertex AH-switching can be obtained by GM-switching twice. We make this notion concrete in the following new definition.

Definition 4. *Let Q be a regular orthogonal matrix of level 2 and let A be an adjacency matrix with the property that Q^TAQ is again an adjacency matrix. Then A is called* reducible *with respect to Q if there exist regular orthogonal matrices Q_1 and Q_2 of level 2 whose largest indecomposable block is smaller than that of Q, such that $Q = Q_1Q_2$ and $Q_1^TBQ_1$ is also an adjacency matrix. Otherwise, it is called* irreducible.

In what follows, we describe only the irreducible adjacency matrices for the AH-switching set, since the reducible ones can be obtained by repeated GM- and AH-switching on smaller sets.

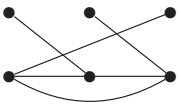
3.3.1 Six vertex switching

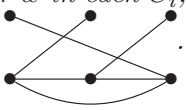
We present a combinatorial description of the switching on 6 vertices that was established by Abiad and Haemers [5, Section 4]. Recall that this switching corresponds to a conjugation of the adjacency matrix with the matrix in Theorem 3(ii) of order 6.

Theorem 5 (AH6-switching)**.** *Let Γ be a graph and let $\{C_1, C_2, C_3, D\}$ be a partition of its vertices such that:*

(i) $|C_1| = |C_2| = |C_3| = 2$.

(ii) Every vertex in D has the same number of neighbours in C_1, C_2 and C_3 modulo 2.

(iii) The induced subgraph on $C_1 \cup C_2 \cup C_3$ is  *(in that order, from left to right).*

Let π be the permutation on $C_1 \cup C_2 \cup C_3$ that shifts the vertices cyclically to the right. For every $v \in D$ that has exactly one neighbour w in each C_i, replace each edge $\{v, w\}$ by $\{v, \pi(w)\}$. Replace the induced subgraph on $C_1 \cup C_2 \cup C_3$ by  *. The resulting graph is $\mathbb{R}$-cospectral with Γ.*

Among the seven possible adjacency matrices for an AH-switching set of size 6, obtained by Abiad and Haemers in [5, Lemma 6], only two are irreducible. However, they are actually equivalent, and correspond to the induced subgraph in the statement of Theorem 5. In other words:

Theorem 6. *AH6-switching is the only switching that corresponds to a regular orthogonal matrix of level 2 with one indecomposable block of size 6 and that cannot be obtained by repeated GM-switching.*

3.3.2 Eight vertex switching

Surprisingly, all matrices that describe an AH-switching set of order 8 (corresponding to the matrix of order 8 in the infinite family of Theorem 3(ii)) are reducible:

Theorem 7. *Every switching that corresponds to a conjugation with the matrix* $\frac{1}{2}\begin{bmatrix} J & O & O & Y \\ Y & J & O & O \\ O & Y & J & O \\ O & O & Y & J \end{bmatrix}$ *can be obtained by repeated GM- and AH6-switching.*

3.3.3 Ten vertex switching

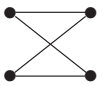
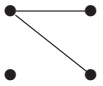
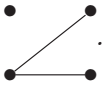
In contrast with the eight vertex case, there are $3 \cdot 2^{10} = 3072$ possibilities for an (irreducible) AH-switching set of size 10.

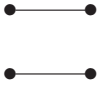
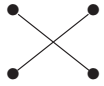
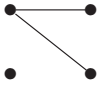
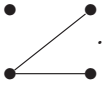
Theorem 8 (AH10-switching)**.** *Let Γ be a graph and let $\{C_1, C_2, C_3, C_4, C_5, D\}$ be a partition of its vertices such that:*

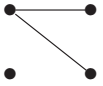
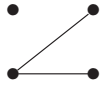
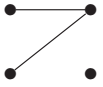
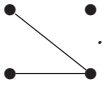
(i) $|C_1| = |C_2| = |C_3| = |C_4| = |C_5| = 2$.

(ii) *Every vertex in* D *has the same number of neighbours in* C_1, C_2, C_3, C_4 *and* C_5 *modulo* 2.

(iii) *One of the following holds (vertices are ordered, from left to right):*

(a) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the induced subgraph on* $C_i \cup C_{i+1}$ *is either*  *or* *and the induced subgraph on* $C_i \cup C_{i+2}$ *is either*  *or*  .

(b) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the induced subgraph on* $C_i \cup C_{i+1}$ *is either*  *or*  *and the induced subgraph on* $C_i \cup C_{i+2}$ *is either*  *or*  .

(c) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the induced subgraph on* $C_i \cup C_{i+1}$ *is either*  *or*  *and the induced subgraph on* $C_i \cup C_{i+2}$ *is either*  *or*  .

Let π *be the permutation on* $C_1 \cup \cdots \cup C_5$ *that shifts the vertices cyclically to the right. For every* $v \notin C$ *that has exactly one neighbour* w *in each* C_i, *replace each edge* $\{v, w\}$ *by* $\{v, \pi(w)\}$. *Replace the induced subgraph on* $C_1 \cup \cdots \cup C_5$ *by the unique graph such that, according to the cases above:*

(a) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the induced subgraph on* $C_i \cup C_{i+1}$ *remains invariant, and the new induced subgraph on* $C_i \cup C_{i+2}$ *is the former induced subgraph on* $C_i \cup C_{i+3}$.

(b) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the new induced subgraph on* $C_i \cup C_{i+1}$ *is the former induced subgraph on* $C_{i+1} \cup C_{i+2}$ *and the new induced subgraph on* $C_i \cup C_{i+2}$ *is the former induced subgraph on* $C_i \cup C_{i+3}$.

(c) *For every* $i \in \mathbb{Z}/5\mathbb{Z}$, *the new induced subgraph on* $C_i \cup C_{i+1}$ *is the former induced subgraph on* $C_{i-1} \cup C_{i+1}$ *and the new induced subgraph on* $C_i \cup C_{i+2}$ *is the former induced subgraph on* $C_{i+1} \cup C_{i+2}$.

The resulting graph is $\mathbb{R}$*-cospectral with* Γ.

Similar to the six vertex case, we have the following result:

Theorem 9. *AH10-switching is the only switching that corresponds to a regular orthogonal matrix of level* 2 *with one indecomposable block of size* 10 *and that cannot be obtained by repeated GM- and AH6-switching.*

4 Asymptotic bounds

Let g_n denote the number of graphs on n (unlabelled) vertices. In 2005, Haemers and Spence [14] established a lower bound on the number of graphs on n vertices that have a cospectral mate.

Theorem 10 ([14, Theorem 3])**.** *There are at least* $n^3 g_{n-1}(\frac{1}{24} - o(1))$ *graphs on* n *vertices with a cospectral mate.*

This bound was derived by counting the number of cospectral mates by GM-switching with respect to a switching set of size 4. Therefore, it is also a lower bound on the number of graphs which have a cospectral mate via GM-switching. From their proof, we can also deduce a matching upper bound.

Theorem 11. *There are* $n^3 g_{n-1}\left(\frac{1}{24} + o(1)\right)$ *non-isomorphic graphs on* n *vertices with a GM-switching set of size* 4.

Analogous bounds can be obtained for WQH-switching. Intuitively, there are less graphs with a WQH-switching set of size 6, because they require more conditions than a switching set of size 4.

Theorem 12. *There are between*

$$n^4 g_{n-2}(\frac{1}{72} - o(1)) \qquad \text{and} \qquad n^4 g_{n-2}(\frac{11}{8})^{n-6}(2^9 + o(1))$$

non-isomorphic graphs on n vertices with a cospectral mate that can be obtained via WQH-switching on 6 vertices.

References

[1] A. Abiad, B. Brimkov, J. Breen, T. R. Cameron, H. Gupta and R. Villagran, Constructions of cospectral graphs with different zero forcing numbers, *Electron. J. Linear Algebra* **38** (2022).

[2] A. Abiad, S. Butler and W. H. Haemers, Graph switching, 2-ranks, and graphical Hadamard matrices, *Discrete Math.* **342** (2019), 2850–2855.

[3] A. Abiad, J. D'haeseleer, W. H. Haemers and R. Simoens, Cospectral mates for generalized Johnson and Grassmann graphs, *Linear Algebra Appl.* **678** (2023), 1–15.

[4] A. Abiad and W. H. Haemers, Switched symplectic graphs and their 2-ranks, *Des. Codes, Cryptogr.* **81** (2016), 35–41.

[5] A. Abiad and W. H. Haemers, Cospectral Graphs and Regular Orthogonal Matrices of Level 2, *Electron. J. Comb.* **19** (2012), P13.

[6] S. G. Barwick, W.-A. Jackson and T. Penttila, New families of strongly regular graphs, *Australas. J. Comb.* **67** (2016), 486–507.

[7] Z. Blazsik, J. Cummings and W.H. Haemers, Cospectral regular graphs with and without a perfect matching, *Discrete Math.* **338** (2015), 199–201.

[8] A. E. Brouwer and E. Spence, Cospectral graphs on 12 vertices, *Electron. J. Comb.* **16** (2009), N20.

[9] H. C. Chan, C. A. Rodger and J. Seberry, On inequivalent weighing matrices, *Ars Comb.* **21A** (1986), 299–333.

[10] E. R. van Dam and W. H. Haemers, Which graphs are determined by their spectrum?, *Linear Algebra Appl.* **373** (2003), 241–272.

[11] E. R. van Dam and W. H. Haemers, Developments on spectral characterizations of graphs, *Discrete Math.* **309** (2009), 576–586.

[12] R. J. Evans, S. Goryainov, E. V. Konstantinova and A. D. Mednykh, A general construction of strictly Neumaier graphs and a related switching, *Discrete Math.* **346** (2023), 113384.

[13] C. Godsil and B. McKay, Constructing cospectral graphs, *Aeq. Math.* **25** (1982), 257–268.

[14] W. H. Haemers and E. Spence, Enumeration of cospectral graphs, *Eur. J. Comb.* **25** (2004), 199–211.

[15] F. Ihringer, F. Pavese and V. Smaldore, Graphs cospectral with $\mathrm{NU}(n+1, q^2)$, $n \neq 3$, *Discrete Math.* **344** (2021), 112560.

[16] C. R. Johnson and M. Newman, A note on cospectral graphs, *J. Comb. Theory B* **28** (1980), 96–103.

[17] I. Koval and M. Kwan, Exponentially many graphs are determined by their spectrum, preprint, 2023, `arXiv:2309.09788`.

[18] F. Liu, W. Wang, T. Yu and H.-J. Lai, Generalized cospectral graphs with and without Hamiltonian cycles, *Linear Algebra Appl.* **585** (2020), 199–208.

[19] A. Munemasa, Godsil–McKay switching and twisted Grassmann graphs, *Des. Codes Cryptogr.* **84** (2015), 173–179.

[20] W. Wang, L. Qiu and Y. Hu, Cospectral graphs, GM-switching and regular rational orthogonal matrices of level p, *Linear Algebra Appl.* **563** (2019), 154–177.

[21] W. Wang, C.-X. Xu, On the asymptotic behavior of graphs determined by their generalized spectra, *Discrete Math.* **310** (2010), 70–76.

The Four-Color Ramsey Multiplicity of Triangles*

Aldo Kiem[1,2], Sebastian Pokutta[1,2], and Christoph Spiegel[1,2]

[1]Technische Universität Berlin, Institute of Mathematics
[2]Zuse Institute Berlin, Department AIS2T, *lastname*@zib.de

Abstract

We study a generalization of a famous result of Goodman and establish that asymptotically at least a 1/256 fraction of all triangles needs to be monochromatic in any four-coloring of the edges of a complete graph. We also show that any large enough extremal construction must be based on a blow-up of one of the two $R(3,3,3)$ Ramsey-colorings of K_{16}. This result is obtained through an efficient flag algebra formulation by exploiting problem-specific combinatorial symmetries that also allows us to study some related problems.

1 Introduction

In 1959, Goodman [17] established precisely how few monochromatic triangles any two-edge-coloring of the complete graph on n vertices can contain, implying that asymptotically at least 1/4 of all triangles need to be monochromatic as n tends to infinity. Subsequently, in [18], he also asked for an answer to the natural generalization of this problem to more than two colors.[1] It took over 50 years and the advent of flag algebras for even the case of three colors to be settled: Cummings et al. [7] showed that asymptotically at least a 1/25 fraction of all triangles need to be monochromatic in any three-edge-coloring of K_n. For n large enough they also precisely characterize the set of extremal constructions, showing that the problem is closely linked to the Ramsey Number $R(3,3) = 6$ as previously noted by Fox [12, Theorem 5.2]. The purpose of this paper is to study the next iteration of this problem, in particular establishing an answer in the affirmative to Question 4 in [7] for the case of four colors.

Theorem 1. *Asymptotically at least a* 1/256 *fraction of all triangles are monochromatic in any four-edge-coloring of K_n and any sufficiently large extremal coloring must be based on one of the two $R(3,3,3)$ Ramsey-colorings of K_{16}.*

The proof of this result relies on the flag algebra framework of Razborov [32, 6]. This allows one to apply a formalized double counting and Cauchy-Schwarz-type argument to obtain bounds for classic problems in Turán and Ramsey theory by solving a concrete semidefinite programming (SDP) formulation. Broadly speaking, the larger this formulation, the better the derived bound becomes.

The major hurdle in establishing Theorem 1 therefore consisted of deriving an efficient formulation by identifying and exploiting combinatorial symmetries through a parameter-dependent notion of automorphisms. The resulting proof likely constitutes the largest *exact* flag algebra calculation done to date. The methods developed to derive it strengthen the previous approach of modifying the underlying notion of isomorphism and generalize Razborov's invariant-anti-invariant decomposition [33]. They

*The full version of this work is available as `arXiv:2312.08049` and will be published elsewhere. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

[1]He in fact calls the three-color version of this question "an old and difficult problem" and raises the question of more than three colors in Section 6 of [18]. The precise origin of this problem is unclear.

are applicable whenever the object we are minimizing has previously ignored symmetries and we hope that they will therefore find further applications. Accompanying these computational improvements, we also give an extension of the stability argument previously developed for the three-color case in [7]. We generalize it to the case of an arbitrary number of colors and establish a strong link between the problem of determining the Ramsey number and the Ramsey multiplicity problem.

2 The Ramsey Multiplicity Problem

We are studying the family of c-colorings of the edges between a finite number of vertices, that is maps $G : \big\{\{u,v\} \mid u,v \in V,\, u \neq v\big\} \to [c] = \{1,\dots,c\}$ where V is any finite set, but we will use common graph notation throughout. Let $\mathcal{G}^{(c)}$ denote the set of all such colorings and $\mathcal{G}_n^{(c)}$ the set of all colorings of order n. Given colorings $H \in \mathcal{G}_k^{(c)}$ and $G \in \mathcal{G}_n^{(c)}$, we write $p(H;G) = |\{S \subseteq V(G) \mid G[S] \simeq H\}| / \binom{n}{k}$ for the *density* of H in G. Note that $p(H;G) = 0$ if $n < k$. Denoting the monochromatic coloring of the edges between vertices in $[t]$ with color $i \in [c]$ by K_t^i, a multi-color version of Ramsey's theorem states that for any $t_1,\dots,t_c \in \mathbb{N}$ the number $R(t_1,\dots,t_c) = \min\big(\big\{n \in \mathbb{N} \mid \{G \in \mathcal{G}_n^{(c)} \mid p(K_t^1;G) + \ldots + p(K_t^c;G) = 0\} = \emptyset\big\}\big)$ is in fact finite. For the diagonal case, where $t_1 = \ldots = t_c$, we write $R_c(t) = R(t,\dots,t)$. The study of the parameter

$$m_c(t;n) = \min_{G \in \mathcal{G}_n^{(c)}} p(K_t^1;G) + \ldots + p(K_t^c;G)$$

is known as the *Ramsey multiplicity problem* for cliques. A simple double-counting argument establishes that $m_c(t;n)$ is monotonically increasing, so that the limit $m_c(t) = \lim_{n\to\infty} m_c(t;n)$ is well defined and satisfies $m_c(t) \geq m_c(t;n)$ for any $n \in \mathbb{N}$. Note that $m_c(t;n) > 0$ as long as $n \geq R_c(t)$ and therefore $m_c(t) > 0$ by Ramsey's theorem.

Concerning upper bounds for $m_c(t)$, coloring the edges uniformly at random with the c colors establishes that

$$m_c(t) \leq c^{1-\binom{t}{2}}. \tag{1}$$

Another way to obtain an upper bound is by blowing up a coloring of the edges of a *looped* complete graph, that is a map $C : \big\{\{u,v\} \mid u,v \in V\big\} \to [c]$. We use the same notation concerning the vertex and edge set as we did for unlooped colorings and write $\mathcal{L}^{(c)}$ for the set of all such colorings as well as $\mathcal{L}_n^{(c)}$ for colorings of order n. A coloring $H \in \mathcal{G}^{(c)}$ *embeds* into a given $C \in \mathcal{L}_k^{(c)}$, if there exists a (not necessarily injective) map $\varphi : V(H) \to V(C)$ satisfying $H(\{u,v\}) = C(\{\varphi(u),\varphi(v)\})$ for all $u,v \in V(H)$. We now let $\mathcal{B}(C) = \{H \in \mathcal{G}^{(c)} \mid H \text{ embeds into } C\}$ denote the *family of blow-up colorings* of C. Note that $\mathcal{B}(C)$ contains graphs of arbitrarily large order. Letting

$$\hat{p}(H;C) = |\{\varphi \text{ embeds } H \text{ into } C\}| \,/\, v(C)^{v(H)}$$

denote the *embedding density*, we have the following result

Lemma 2. *Given any $C \in \mathcal{L}_k^{(c)}$, we have $m_c(t) \leq \hat{p}(K_t^1;C) + \ldots + \hat{p}(K_t^c;C)$.*

In our case, the most relevant candidates for colorings C are obtained by considering a Ramsey-coloring on $r = R_{c-1}(t) - 1$ vertices avoiding cliques of size t in any of the $c-1$ colors. Coloring the loops with the additional c-th, this implies an upper bound of

$$m_c(t) \leq (R_{c-1}(t) - 1)^{1-t}, \tag{2}$$

see also Theorem 5.2 in [12]. The result of Goodman [17] implies that $m_2(3) = 1/4$. This aligns both with the probabilistic upper bound stated in Equation (1) as well as the Ramsey upper bound stated in Equation (2), where for the latter we are relying on the trivial case of Ramsey's theorem, that is $R_1(t) = R(t) = t$.

Given that the former bound dominates when t grows as long as $c = 2$, Erdős suggested [10] that the probabilistic upper bound should always be tight in this case. This was disproven by Thomason [39] for any $t \geq 4$ and a significant number of results since then have tried to either determine improved asymptotics for $m_2(t)$ or specific values of it for small t [5, 9, 11, 14, 15, 16, 20, 22, 27, 35, 38, 40, 41, 29]. The problem also links to Sidorenko's famous open conjecture and the search for a characterization of common graphs, cf. [4, 36]. As of now, even $m_2(4)$ remains open, with the best current lower and upper bounds of $0.0296 \leq m_2(4) \leq 0.03014$ respectively due to Grzesik et al. [20] as well as Parczyk et al. [29]. Note that we also obtained a slight improvement of $m_2(4) \geq 0.02961$.

For the asymptotic values, there has likewise been scarce progress, with the current best lower bound of $C^{-t^2(1+o(1))} \leq m_2(t)$ for $C \approx 2.18$ due to Conlon [5] and the best upper bound of $m_2(t) \leq 0.835 \cdot 2^{1-\binom{t}{2}}$ for $t \geq 7$ due to Jagger, Šťovíček, and Thomason [22]. Given the lack of progress on the two-color, diagonal version, there are two obvious directions to explore: the case of more colors, where $c > 2$, as well as the off-diagonal case, where $t_1 \neq t_2$.

3 Increasing the number of colors

Studying monochromatic triangles for more than two colors was, as already mentioned in the introduction, suggested by Goodman [18] and resolved for $c = 3$ by Cummings et al. [7], whose result aligns with Equation (2) since $R_2(3) = 6$. In order to state their result in its fullest strength, let $C_{R(3,3)}$ denote the coloring in $\mathcal{L}_5^{(3)}$ obtained by taking the unique Ramsey 2-coloring of a complete graph on five vertices that avoids monochromatic triangles and coloring the loops with the third color, that is $E_1(C_{R(3,3)})$ and $E_2(C_{R(3,3)})$ both are 5-cycles and $E_3(C_{R(3,3)})$ contains all five loops. Let $\mathcal{G}_{\text{ex}}^{(3)} \subset \mathcal{G}^{(3)}$ now consist of all colorings that can be obtained by (i) selecting an element in $\mathcal{B}(C_{R(3,3)})$, (ii) recoloring some of the edges from the first or second color to instead use the third color without creating any additional monochromatic triangles, and (iii) applying any permutation of the colors. Note that the second step implies that the recolored edges must form a matching between any of the five 'parts', though not every such recoloring avoids additional triangles.

Theorem 3 (Cummings et al. [7])**.** *There exists an $n_0 \in \mathbb{N}$ such that any element in $\mathcal{G}_n^{(3)}$ of order $n \geq n_0$ minimizing the number of monochromatic triangles must be in $\mathcal{G}_{ex}^{(3)}$.*

The result characterizes extremal constructions for large enough n, though more recently there has been increasing interest in deriving stability results based on flag algebra calculations [30]. Let $C'_{R(3,3,3)}$ and $C''_{R(3,3,3)}$ denote the two colorings in $\mathcal{L}_{16}^{(4)}$ obtained in a similar way to the previously defined $C_{R(3,3)}$ by respectively taking the two Ramsey 3-coloring of a complete graph on 16 vertices that avoid monochromatic triangles [19, 23, 24, 31] and coloring the vertices with the fourth color. Mirroring the construction of $\mathcal{G}_{\text{ex}}^{(3)}$, we let $\mathcal{G}_{\text{ex}}^{(4)} \subset \mathcal{G}^{(4)}$ consist of all colorings that can be obtained by

(i) selecting an element in $\mathcal{B}(C'_{R(3,3,3)})$ or $\mathcal{B}(C''_{R(3,3,3)})$,

(ii) recoloring some of the edges from any of the first, second or third color to instead use the fourth color without creating any additional monochromatic triangles,

(iii) applying any permutation of the four colors.

Theorem 4. *There exists an $n_0 \in \mathbb{N}$ such that for any $\varepsilon > 0$ there exists $\delta > 0$ such that any $G \in \mathcal{G}_n^{(4)}$ of order $n \geq n_0$ with $\sum_{i=1}^{c} p(K_3^i; G) \leq m_4(3; n) + \delta$ can be turned into an element of $\mathcal{G}_{ex}^{(4)}$ by recoloring at most $\varepsilon\binom{n}{2}$ edges.*

Note that this implies that any large enough element in $\mathcal{G}_n^{(4)}$ minimizing the number of monochromatic triangles must be in $\mathcal{G}_{\text{ex}}^{(4)}$. Our results in fact show that it likewise can be obtained for the case of three colors.

4 The off-diagonal case

The second of the previously suggested directions, that is considering the off-diagonal case, has recently started to receive some attention [29, 3, 26, 21] with two competing notions of off-diagonal Ramsey multiplicity having been suggested. The first is due to Parczyk et al. [29] and is concerned with determining

$$m(t_1,\dots,t_c;n) = \min_{G\in\mathcal{G}_n^{(c)}} p(K_{t_1}^1;G) + \dots + p(K_{t_c}^c;G).$$

This generalizes the previously defined $m_c(t;n)$ but does not consider the inherent imbalance when for example $c=2$ and $t_1 \ll t_2$; minimizing $p(K_{t_1}^1;G)+p(K_{t_2}^2;G)$ in this case will be equivalent to enforcing $p(K_{t_1}^1;G)=0$ and minimizing $p(K_{t_2}^2;G)$, a related problem previously suggested by Erdős [10, 28, 8, 29]. This issue was already noted in [29] and subsequently addressed by Moss and Noel [26], who instead suggested determining

$$m_s(t_1,\dots,t_c;n) = \min_{G\in\mathcal{G}_n^{(c)}} \max_{\substack{\lambda_1,\dots,\lambda_c\geq 0\\ \lambda_1+\dots+\lambda_c=1}} \lambda_1\, p(K_{t_1}^1;G) + \dots + \lambda_c\, p(K_{t_c}^c;G).$$

We will use $m(t_1,\dots,t_c)$ as well as $m_s(t_1,\dots,t_c)$ to respectively denote the limits of both of these functions as n tends to infinity. Both notions generalize the previous diagonal definition and clearly $m_s(t_1,\dots,t_c) \geq m(t_1,\dots,t_c)$. Unsurprisingly, determining $m_s(t_1,\dots,t_c)$ has proven much more difficult, with $m(3,4)$ and $m(3,5)$ having been settled in [29] and $m_s(3,4)$ still remaining open. Here we derive the following result for the weaker of the two notions.

Proposition 5. *We have* $m(3,3,4) = 1/125$.

The upper bound follows immediately by generalizing Equation (2) to the off-diagonal case, that is by noting that

$$m(t_1,\dots,t_c) \leq (R(t_1,\dots,t_{c-1})-1)^{1-t_c} \tag{3}$$

and inserting $R(3,3)=6$. The lower bound was derived using the same improvements to the flag algebra calculus that we developed to derive our main result.

5 Discussion and Outlook

The proposed computational improvements were crucial in order to derive a certificate for the upper bound and stability statement in Theorem 4. They are applicable whenever the problem studied exhibits symmetries with respects to the colors, with the reduction of the number of constraints essentially factorial in the number of colors. We therefore hope that they find further use for other problems, for example for improved upper bounds on Ramsey numbers through flag algebras, as recently done by Lidicky and Pfender [25]. The improvements however are largely not applicable when there are no previously ignored symmetries in the problem statement, as is for example the case with the famous $(3,4)$-Turán conjecture. They may also not be helpful for applications beyond graphs [1, 2, 37, 34], where there can be more drastic jumps on the numbers of constraints as N is increased.

Besides these computational improvements, it is notable that our generalization of the stability argument from [7] no longer requires explicit knowledge of the Ramsey colorings underlying the extremal construction. While in our case the colorings were both known and crucial in order to derive an exact rather than a floating point-based flag algebra certificate, our main stability result draws a connection between the Ramsey number $R_{c-1}(3)$ and the Ramsey multiplicity problem $m_c(3)$, in theory opens up an avenue to establish a sort of equivalence of the two problems without first explicitly solving both or even either problem:

1) We could derive a flag algebra certificate for a particular $c>4$ establishing $m_c(3)$ and meeting the necessary requirements without explicit knowledge of the $R_{c-1}(3)$-Ramsey colorings. Note that this would imply the exact value of $R_{c-1}(3) = m_c(3)^{-1/2}-1$.

2) We could show that the Ramsey multiplicity problem satisfies the necessary requirements for arbitrary $c \geq 3$, in particular that $\hat{K}_{3,1}$ and $\hat{K}_{3,3}$ have zero density in an extremal construction, through a purely theoretical argument not relying on the semidefinite programming method and without explicitly determining $m_c(3)$. This would imply that $m_c(3) = (R_{c-1}(3) - 1)^{-2}$ without giving us explicit knowledge of either value.

It should be noted that Fox and Wigerson [13] somewhat recently characterized an infinite family of 2-colorings for which an upper bound equivalent to the one given by Equation (2) is tight, i.e., Turán graphs determine the extremal constructions for the respective Ramsey multiplicity problem. They also obtained results for the case of $c = 3$ colors that are conditioned the conjecturec bound $R(t, \lceil t/2 \rceil) \leq 2^{-31} R(t, t)$. At the risk of extrapolating from a sample size of two, this fact motivates us to go so far as to conjecture the following to be true.

Conjecture 6. *For any $c \geq 3$, we have $m_c(3) = (R_{c-1}(3) - 1)^{-2}$ and the only extremal constructions are derived from $R_{c-1}(3)$-Ramsey colorings.*

References

[1] Baber, R.: Turán densities of hypercubes. arXiv preprint arXiv:1201.3587 p. 161171 (2012)

[2] Balogh, J., Hu, P., Lidickỳ, B., Liu, H.: Upper bounds on the size of 4-and 6-cycle-free subgraphs of the hypercube. European Journal of Combinatorics **35**, 75–85 (2014)

[3] Behague, N., Morrison, N., Noel, J.A.: Common pairs of graphs. arXiv preprint arXiv:2208.02045 (2022)

[4] Burr, S.A., Rosta, V.: On the Ramsey multiplicities of graphs—problems and recent results. Journal of Graph Theory **4**(4), 347–361 (1980)

[5] Conlon, D.: On the Ramsey multiplicity of complete graphs. Combinatorica **32**(2), 171–186 (2012). https://doi.org/10.1007/s00493-012-2465-x, `https://doi.org/10.1007/s00493-012-2465-x`

[6] Coregliano, L.N., Razborov, A.A.: Semantic limits of dense combinatorial objects. Russian Mathematical Surveys **75**(4), 627 (2020)

[7] Cummings, J., Král', D., Pfender, F., Sperfeld, K., Treglown, A., Young, M.: Monochromatic triangles in three-coloured graphs. Journal of Combinatorial Theory, Series B **103**(4), 489–503 (2013)

[8] Das, S., Huang, H., Ma, J., Naves, H., Sudakov, B.: A problem of erdős on the minimum number of k-cliques. Journal of Combinatorial Theory, Series B **103**(3), 344–373 (2013)

[9] Deza, A., Franek, F., Liu, M.J.: On a conjecture of Erdős for multiplicities of cliques. Journal of Discrete Algorithms **17**, 9–14 (2012)

[10] Erdős, P.: On the number of complete subgraphs contained in certain graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl **7**(3), 459–464 (1962)

[11] Even-Zohar, C., Linial, N.: A note on the inducibility of 4-vertex graphs. Graphs and Combinatorics **31**(5), 1367–1380 (2015)

[12] Fox, J.: There exist graphs with super-exponential Ramsey multiplicity constant. Journal of Graph Theory **57**(2), 89–98 (2008)

[13] Fox, J., Wigderson, Y.: Ramsey multiplicity and the Turán coloring. Advances in Combinatorics (2023)

[14] Franek, F.: On Erdős' conjecture on multiplicities of complete subgraphs lower upper bound for cliques of size 6. Combinatorica **22**(3), 451–454 (2002)

[15] Franek, F., Rödl, V.: 2-colorings of complete graphs with a small number of monochromatic K_4 subgraphs. Discrete mathematics **114**(1-3), 199–203 (1993)

[16] Giraud, G.: Sur le probleme de Goodman pour les quadrangles et la majoration des nombres de Ramsey. Journal of Combinatorial Theory, Series B **27**(3), 237–253 (1979)

[17] Goodman, A.W.: On sets of acquaintances and strangers at any party. The American Mathematical Monthly **66**(9), 778–783 (1959)

[18] Goodman, A.: Triangles in a complete chromatic graph with three colors. Discrete mathematics **57**(3), 225–235 (1985)

[19] Greenwood, R.E., Gleason, A.M.: Combinatorial relations and chromatic graphs. Canadian Journal of Mathematics **7**, 1–7 (1955)

[20] Grzesik, A., Lee, J., Lidickỳ, B., Volec, J.: On tripartite common graphs. arXiv preprint arXiv:2012.02057 (2020)

[21] Hyde, J., Lee, J.b., Noel, J.A.: Turán colourings in off-diagonal Ramsey multiplicity. arXiv preprint arXiv:2309.06959 (September 2023), https://arxiv.org/abs/2309.06959, 48 pages, 2 figures

[22] Jagger, C., Šťovíček, P., Thomason, A.: Multiplicities of subgraphs. Combinatorica **16**, 123–141 (1996)

[23] Kalbfleisch, J., Stanton, R.: On the maximal triangle-free edge-chromatic graphs in three colors. Journal of Combinatorial Theory **5**(1), 9–20 (1968)

[24] Laywine, C., Mayberry, J.: A simple construction giving the two non-isomorphic triangle-free 3-colored k16's. Journal of Combinatorial Theory, Series B **45**(1), 120–124 (1988)

[25] Lidicky, B., Pfender, F.: Semidefinite programming and ramsey numbers. SIAM Journal on Discrete Mathematics **35**(4), 2328–2344 (2021)

[26] Moss, E., Noel, J.A.: Off-diagonal Ramsey multiplicity. arXiv preprint arXiv:2306.17388 (2023)

[27] Nieß, S.: Counting monochromatic copies of K_4: a new lower bound for the Ramsey multiplicity problem. arXiv:1207.4714 (2012)

[28] Nikiforov, V.: On the minimum number of k-cliques in graphs with restricted independence number. Combinatorics, Probability and Computing **10**(4), 361–366 (2001)

[29] Parczyk, O., Pokutta, S., Spiegel, C., Szabó, T.: New Ramsey multiplicity bounds and search heuristics. arXiv preprint arXiv:2206.04036 (2022)

[30] Pikhurko, O., Sliačan, J., Tyros, K.: Strong forms of stability from flag algebra calculations. Journal of Combinatorial Theory, Series B **135**, 129–178 (2019)

[31] Radziszowski, S.: Small Ramsey numbers. The electronic journal of combinatorics **1000**, DS1–Aug (2011)

[32] Razborov, A.A.: Flag algebras. The Journal of Symbolic Logic **72**(4), 1239–1282 (2007)

[33] Razborov, A.A.: On 3-hypergraphs with forbidden 4-vertex configurations. SIAM Journal on Discrete Mathematics **24**(3), 946–963 (2010)

[34] Rué, J., Spiegel, C.: The rado multiplicity problem in vector spaces over finite fields. arXiv preprint arXiv:2304.00400 (2023)

[35] Sawin, W.: An improved lower bound for multicolor Ramsey numbers and the half-multiplicity Ramsey number problem. arXiv preprint arXiv:2105.08850 (2021)

[36] Sidorenko, A.: A correlation inequality for bipartite graphs. Graphs and Combinatorics **9**(2), 201–204 (1993)

[37] Sliacan, J., Stromquist, W.: Improving bounds on packing densities of 4-point permutations. Discrete Mathematics & Theoretical Computer Science **19**(Permutation Patterns) (2018)

[38] Sperfeld, K.: On the minimal monochromatic K_4-density. arXiv preprint arXiv:1106.1030 (2011)

[39] Thomason, A.: A disproof of a conjecture of Erdős in Ramsey theory. Journal of the London Mathematical Society **2**(2), 246–255 (1989)

[40] Thomason, A.: Graph products and monochromatic multiplicities. Combinatorica **17**(1), 125–134 (1997)

[41] Wolf, J.: The minimum number of monochromatic 4-term progressions in $\mathbb{Z}_p$. Journal of Combinatorics **1**(1), 53–68 (2010)

Speed and size of dominating sets in domination games*

Ali Deniz Bagdas[†1], Dennis Clemens[‡1], Fabian Hamann[§1], and Yannick Mogge[¶1]

[1]Institute of Mathematics, Hamburg University of Technology, Hamburg, Germany

Abstract

We consider Maker-Breaker domination games, a variety of positional games, in which two players (Dominator and Staller) alternately claim vertices of a given graph. Dominator's goal is to fully claim all vertices of a dominating set, while Staller tries to prevent Dominator from doing so, or at least tries to delay Dominator's win for as long as possible.

We prove a variety of results about domination games, including the number of turns Dominator needs to win and the size of a smallest dominating set that Dominator can occupy, when considering e.g. random graphs, powers of paths, and trees. We could also show that speed and size can be far apart, and we prove further non-intuitive statements about the general behaviour of such games.

We also consider the Waiter-Client version of such games.

1 Introduction

Let a hypergraph $\mathcal{H} = (X, \mathcal{F})$ and two integers $m, b \geq 1$ be given. The $(m : b)$ *Maker-Breaker game* on $(X, \mathcal{F})$ is played as follows. Maker and Breaker alternate in moves, where in a move Maker claims up to m unclaimed elements of the *board* X, and Breaker claims up to b unclaimed elements of X. Maker wins if during the course of the game she manages to claim all elements of a *winning set*, i.e. a hyperedge from $\mathcal{F}$, while Breaker wins otherwise. Surely, this outcome can depend on who makes the first move. Therefore, whenever it makes a difference in the following, we will state clearly whom we assume to be the first player. If $m = b = 1$, the game is called *unbiased*; and otherwise it is called *biased*. For a nice overview about positional games in general we recommend the monograph [15] as well as the survey [18].

Let $G = (V, E)$ be a graph. We denote the set of vertices of G by $V(G)$ and let $v(G) = |V(G)|$. We will mostly focus on *Maker-Breaker domination games*, a certain variety of Maker-Breaker games, which were recently introduced by Duchêne et al. [9]. While most Maker-Breaker games are played on the edge set of some graph, domination games are played on the vertex set of a given graph G instead. Two players, who are called Dominator and Staller, alternately claim vertices of G, and Dominator (who is playing as Maker) wins if and only if she manages to occupy all vertices of a *dominating set*, which is a subset of $V(G)$ such that every $v \in V(G)$ is either a neighbour of the dominating set or part of this subset itself. Note that the renaming of the players Maker and Breaker as Dominator and Staller is done to be consistent with the usual domination games; for an overview on these games we recommend the book [4].

*The full version of this work will be published elsewhere. This research of the second and fourth author is supported by Deutsche Forschungsgemeinschaft (Project CL 903/1-1).

[†]Email: ali.bagdas@tuhh.de

[‡]Email: dennis.clemens@tuhh.de

[§]Email: fabian.hamann@tuhh.de

[¶]Email: yannick.mogge@tuhh.de

Let $\gamma_{MB}(G, m : b)$ denote the smallest number of rounds in which Dominator can always win the $(m : b)$ Maker-Breaker domination game on G, provided that she starts the game, and where we set $\gamma_{MB}(G, m : b) = \infty$ if Dominator does not have a winning strategy. Similarly, let $\gamma'_{MB}(G, m : b)$ denote the smallest number of rounds for the case when Staller starts the game, and for short, let $\gamma_{MB}(G) := \gamma_{MB}(G, 1 : 1)$ and $\gamma'_{MB}(G) := \gamma'_{MB}(G, 1 : 1)$.

2 Known results

In their paper which introduced Maker-Breaker domination games, Duchêne et al. [9] proved that deciding who wins an unbiased Maker-Breaker domination game is PSPACE-complete. On the other hand, Gledel et al. [13] put a focus on the number of rounds which Dominator needs to win, and determined $\gamma_{MB}(G)$ and $\gamma'_{MB}(G)$ precisely when G is a tree or a cycle. Partial results for Cartesian products and paths [8, 11] as well as Corona products [7] of graphs were obtained afterwards as well. Additionally, Gledel et al. [13] provided examples which show that the domination number $\gamma(G)$ of a graph G and the Maker-Breaker domination numbers $\gamma_{MB}(G)$ and $\gamma'_{MB}(G)$ can take arbitrary values with the obvious restriction that $\gamma(G) \leq \gamma_{MB}(G) \leq \gamma'_{MB}(G)$. In particular, all these three values can be arbitrarily far apart from each other.

Theorem 1 (Theorem 3.1 in [13])**.** *For any integers $2 \leq r \leq s \leq t$, there exists a graph G such that $\gamma(G) = r$, $\gamma_{MB}(G) = s$ and $\gamma'_{MB}(G) = t$.*

3 Our results I: unbiased games

First we are interested in the general behaviour of domination games in the unbiased setting and, in particular, to find necessary or sufficient conditions for Dominator to win in given time. In this regards, Gledel et al. [13] already proved the following proposition. For this, for a graph G, let $X_\gamma(G)$ denote the number of dominating sets of size γ of G.

Proposition 2 (Proposition 3.3 in [13])**.** *If G is a graph and $X_\gamma(G) < 2^{\gamma(G)-1}$, then $\gamma_{MB}(G) > \gamma(G)$.*

Instead of looking at the number of smallest possible dominating sets, we proved the following minimum degree condition for Dominator to have a winning strategy.

Theorem 3. *Let n be a positive integer and let $\delta(G)$ denote the minimum degree of G. If G is a graph on n vertices with $\delta(G) > \log_2(n) - 1$, then Dominator wins the $(1 : 1)$ Maker-Breaker domination game on G.*

Moreover, the bound on $\delta(G)$ is asymptotically best possible. For infinitely many n, there is a graph G on n vertices and with $\delta(G) > \log_2(n) - 2$ such that Staller wins the $(1 : 1)$ Maker-Breaker domination game on G.

Next to this, when asking for the existence of winning strategies, it seems natural to study the behaviour of the Maker-Breaker domination game when played on a randomly chosen graph. Let $G \sim G_{n,p}$ denote a graph sampled from the binomial random graph model, where each edge of a graph with n vertices is present with probability p. When we play a Maker-Breaker domination game on such a graph with constant probability p, we have the following bound on the number of turns that Dominator needs to win.

Theorem 4. *If $p \in (0, 1)$ is constant and $G \sim G_{n,p}$, then a.a.s.*

$$\gamma_{MB}(G) = (1 + o(1)) \log_{1/(1-p)}(n).$$

Although the proofs of both Theorem 3 and Theorem 4 can be done with fairly standard methods from positional games theory, we believe that these statements are important for getting a general intuition for domination games and for predicting the outcome of such games.

4 Our results II: biased games

A lot of research in positional games considers games with a bias, yet we do not know of any paper considering biased versions of Maker-breaker domination games. As a first step, we extend [13] by proving results for biased game in which Dominator wants to dominate all vertices of the power of any path, or all vertices of a tree. Let P_n^k denote the k-th power of a path with n vertices. Then the following theorem holds which can be proven with an inductive argument that mainly involves ad-hoc winning strategies with case distinctions.

Theorem 5. *For all integers $b, k \leq n$ it holds that*

$$\gamma_{MB}(P_n^k, b:1) = \left\lceil \frac{n-1}{b(2k+1)-1} \right\rceil.$$

For a given a tree T and a bias b, let us say that T is b-good if we can recursively delete vertices, which have exactly b leaf neighbours, and also delete these leaf neighbours, until we reach a forest where every vertex has at most $b-1$ leaf neighbours. Then the following holds.

Theorem 6. *Let T be a tree with $v(T) \geq 2$. Then the following are equivalent:*

(i) Dominator wins the $(b:1)$ game on T when Staller is the first player.

(ii) T is b-good.

While the proof of the implication (i)⇒(ii) is a simple exercise, the other direction is less trivial. Here, we do an induction for a slightly stronger statement which considers games in which Dominator's goal is to dominate only a certain subset of vertices of T. Moreover, by having this more general statement we are also able to give an analogue theorem for the case when Dominator is the first player. We skip the details here, and will soon make them available on arXiv.

Additionally, in the case that a given tree T is not b-good, we are still able to prove the following quantitative statement.

Theorem 7. *For every tree T it holds that Dominator can dominate at least $\left(1 - \frac{1}{(b+1)^2}\right) v(T)$ vertices in the $(b:1)$ game on T. Moreover, the bound is sharp.*

Note that for the above games involving trees, we do not consider Staller's bias to be larger than 1 due to the fact that with bias 2, Staller can already win the game within one round, by claiming a leaf and its neighbour. Still, for other graphs, it makes sense to increase Staller's bias and in fact, interesting (and maybe surprising) behaviours can be shown, see Theorem 8.

Before stating this theorem, note that a nice property of $(m:b)$ Maker-Breaker games is that these are monotone with respect to each of the biases m and b. That is, roughly speaking, increasing the bias of one of the players can never be a disadvantage for this player; see e.g. [2, 15]. This observation leads to the natural definition of *threshold biases*, which in many cases have proven to be related to properties of random graphs, see e.g. [12, 17]. When both biases get increased simultaneously, we however cannot expect monotonicity in general. For an example, Balogh et al. [1] considered the 2-diameter game in which Maker's goal is to occupy a spanning subgraph of K_n with diameter 2, and they proved that Breaker wins the $(1:1)$ variant of this game, while Maker has a winning strategy for the $(2:b)$ variant even when $b \leq \frac{1}{9} n^{1/8} (\log n)^{-3/8}$, provided n is large enough. In particular, the $(b:b)$ variant is won by Maker for every constant $b \geq 2$ if n is large.

Now, looking only at these fair $(b:b)$ games, the above example could still be considered to be monotone for all $b \geq 1$, since increasing b never worsens Maker's chances of winning. So, one could wonder whether such a behaviour always holds for fair Maker-Breaker games. With our next result we show that this is not the case, even for Maker-Breaker domination games.

Theorem 8. *Let $B \subset \mathbb{N}$ be any finite set. Then there exists a graph G, such that Dominator wins the $(b:b)$ Maker-Breaker domination game on G (when Dominator starts) if and only if $b \notin B$.*

5 Our results III: speed and size

Another interesting question is the following: when playing the $(m : b)$ Maker-Breaker domination game on a graph G, of what size is the smallest dominating set which Dominator can always claim? Let $s_{MB}(G, m : b)$ and $s_{MB}(G) = s_{MB}(G, 1 : 1)$ denote this values when Dominator starts the game, and let $s'_{MB}(G, m : b)$ and $s'_{MB}(G) = s_{MB}(G, 1 : 1)$ denote this value when Staller starts, where again we set such a value to ∞ if Dominator does not have a winning strategy. A priori it is not clear why the minimal number of rounds and the minimal size of a dominating set that Dominator can achieve should be different. In fact, from the proofs in [13] it can be deduced easily that $\gamma_{MB}(G) = s_{MB}(G)$ and $\gamma'_{MB}(G) = s'_{MB}(G)$ hold when G is a tree or a cycle. However, in contrast to this, we can prove the following statement which, roughly speaking, says that all these parameters in questions can take almost arbitrary values and hence can be arbitrarily far apart. Note that this statement is a strengthening of Theorem 1 from [13].

Theorem 9. *For any biases $m \leq b$ and any integers r, s, s', t, t' such that $m+1 \leq r, \max\{2m+1, r\} \leq s \leq s', t \leq t'$, $s \leq m \cdot t$, $s' \leq m \cdot t'$, there exists a graph G such that*

$$\begin{aligned} &\gamma(G) = r, \\ &s_{MB}(G, m : b) = s \quad \text{and} \quad \gamma_{MB}(G, m : b) = t \\ &s'_{MB}(G, m : b) = s' \quad \text{and} \quad \gamma'_{MB}(G, m : b) = t'. \end{aligned}$$

One step in the proof of this theorem is to provide a construction which allows us to transfer constructive results from general Maker-Breaker games to domination games. We can also use this transference construction to prove the following rather non-intuitive result:

Theorem 10. *For any biases m, b with $m \leq b$ and any integers $t > s \geq 2m + 1$, there exists a graph G such that $s_{MB}(G : m : b) = s$, but Dominator cannot occupy a dominating set of size s before she has occupied another minimal dominating set of size t.*

Moreover, with similar arguments, we can show the following result which roughly states that claiming a smallest possible dominating set can take Dominator arbitrarily much longer than claiming an arbitrarily large dominating set which can be claimed in optimal time. Hence, as already supported by Theorem 9, studying the parameters s_{MB} and γ_{MB} can be two very different problems which may require very different tools when proving exact results.

Theorem 11. *For any biases m, b with $m \leq b$ and any integers $t' \geq t \geq s' \geq s \geq 2m + 1$, there exists a graph G such that $\gamma_{MB}(G, m : b) = t$ and $s_{MB}(G, m : b) = s$, but in the $(m : b)$ Maker-Breaker domination game on G we have that*

- *s' is the smallest size of a dominating set that Dominator can get within t rounds,*
- *t' is smallest number of rounds that Dominator needs to claim a dominating set of size s.*

6 Our results IV: Waiter-Client domination games

Another variety of Maker-Breaker games are *Waiter-Client games* (earlier called Picker-Chooser games, see e.g. [2]), which have received increasing attention lately, ranging from results on fast winning strategies [5, 10] over biased games [3, 14, 19] to games played on random graphs [6, 16]. In the following we will stick to the case of unbiased Waiter-Client games. On a given hypergraph $\mathcal{H} = (X, \mathcal{F})$, these games are played almost the same way as Maker-Breaker games with the following difference: In every round, Waiter chooses two unclaimed elements of the board X and then Client decides which of these elements goes to Waiter while the other one goes to Client. Waiter wins if and only if she manages to claim all elements of a winning set from $\mathcal{F}$.

So far, domination games have not been studied in this setting. So, we also aim to give first results for Waiter-Client domination games and on the relation of Waiter-Client and Maker-Breaker domination games. Given a graph G, we define the *Waiter-Client domination game* on G in the obvious way: Dominator (playing as Waiter) offers two unclaimed vertices of G and then Staller (playing as Client) picks one of these vertices for himself and the other goes to Dominator. In accordance with previous notation, we denote with $\gamma_{WC}(G)$ the smallest number of rounds in which Dominator can always occupy a dominating set in the Waiter-Client domination game on G, and we let $s_{WC}(G)$ denote the size of the smallest dominating set that Waiter can always claim. For our first results in this setup, we can prove that for cycles and trees the game behaves the same way as in the Maker-Breaker setting [13] (when Breaker starts).

Theorem 12. *For every $n \geq 3$,*

$$\gamma_{WC}(C_n) = s_{WC}(C_n) = \left\lfloor \frac{n}{2} \right\rfloor.$$

Theorem 13. *Let T be a tree on n vertices. If T has a perfect matching, then*

$$\gamma_{WC}(T) = s_{WC}(T) = \frac{n}{2}.$$

In all other cases, Dominator does not win the $(1:1)$ Waiter-Client domination game on T.

Due to these results and due to the fact that in the literature, Waiter most of the time can play at least as good as Maker can do in the analogue game with same winning sets, it seems natural to wonder whether relations such as $\gamma_{WC}(G) \leq \gamma_{MB}(G)$ can be proven for arbitrary graphs G. As our last result we negate this with the following theorem which states that the parameters $\gamma_{WC}(G)$ and $\gamma_{MB}(G)$ can take almost arbitrary values and, in particular, $\gamma_{WC}(G)$ can be much larger than $\gamma_{MB}(G)$. The proof again uses our transference argument from the previous section together with suitable hypergraphs on which either Maker (in the usual Maker-Breaker game) or Waiter (in the usual Waiter-Client game) can win fast, while the other player does not have a strategy that ensures a fast win.

Theorem 14. *For all integers $s, t \geq 7$ there is a graph G such that $\gamma'_{MB}(G) = s$ and $\gamma_{WC}(G) = t$.*

References

[1] J. Balogh, R. Martin, and A. Pluhár, The diameter game, *Random Structures & Algorithms* **35.3** (2009), 369–389.

[2] J. Beck, *Combinatorial games: Tic-Tac-Toe theory*, Volume 114, Cambridge University Press, Cambridge (2008).

[3] M. Bednarska-Bzdega, D. Hefetz, M. Krivelevich, and T. Łuczak, Manipulative waiters with probabilistic intuition, *Combinatorics, Probability and Computing* **25.6** (2016), 823–849.

[4] B. Brešar, M. A. Henning, S. Klavžar, and D. F. Rall, *Domination games played on graphs*, Springer, Cham (2021).

[5] D. Clemens, P. Gupta, F. Hamann, A. Haupt, M. Mikalački, and Y. Mogge, Fast strategies in Waiter-Client games, *The Electronic Journal of Combinatorics* **27.3** (2020), 1–35.

[6] D. Clemens, F. Hamann, Y. Mogge, and O. Parczyk, Waiter-Client Games on Randomly Perturbed Graphs, in: *Extended Abstracts EuroComb 2021: European Conference on Combinatorics, Graph Theory and Applications*, Springer (2021), 397–403.

[7] A. Divakaran, T. James, S. Klavžar, and L. S. Nair, Maker–Breaker domination game played on Corona products of graphs, preprint (2024), `arXiv:2402.13581`.

[8] P. Dokyeesun, Maker–Breaker domination game on Cartesian products of graphs, preprint (2023), `arXiv:2310.04103`.

[9] E. Duchêne, V. Gledel, A. Parreau, and G. Renault, Maker–Breaker domination game, *Discrete Mathematics* **343.9** (2020), 111955.

[10] V. Dvořák, Waiter–Client triangle-factor game on the edges of the complete graph, *European Journal of Combinatorics* **96** (2021), 103356.

[11] J. Forcan and J. Qi, Maker–Breaker domination number for Cartesian products of path graphs P_2 and P_n, preprint (2020), `arXiv:2004.13126`.

[12] H. Gebauer and T. Szabó, Asymptotic random graph intuition for the biased connectivity game, *Random Structures & Algorithms* **35.4** (2009), 431–443.

[13] V. Gledel, V. Iršič, and S. Klavžar, Maker–Breaker domination number, *Bulletin of the Malaysian Mathematical Sciences Society* **42** (2019), 1773–1789.

[14] D. Hefetz, M. Krivelevich, and W. E. Tan, Waiter-Client and Client-Waiter planarity, colorability and minor games, *Discrete Mathematics* **339** (2016), 1525–1536.

[15] D. Hefetz, M. Krivelevich, M. Stojaković, and T. Szabó, *Positional games*, Volume 44, Birkhäuser, Basel (2014).

[16] D. Hefetz, M. Krivelevich, and W. E. Tan, Waiter–Client and Client–Waiter Hamiltonicity games on random graphs, *European Journal of Combinatorics* **63** (2017), 26–43.

[17] M. Krivelevich, The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln n$, *Journal of the American Mathematical Society* **24.1** (2011), 125–131.

[18] M. Krivelevich, Positional games, preprint (2014), `arXiv:1404.2731`.

[19] R. Nenadov, Probabilistic intuition holds for a class of small subgraph games, *Proceedings of the American Mathematical Society* **151.04** (2023), 1495–1501.

On the solutions of linear systems over additively idempotent semirings*

Álvaro Otero Sánchez†[1], Daniel Camazón Portela‡[1], and Juan Antonio López Ramos§[1]

[1]Department of Mathematics, University of Almería

Abstract

The aim of this paper is to address the system $AX = Y$, where $A = (a_{ij}) \in M_{m\times n}(S)$, $Y \in S^m$, and X represents an unknown vector of size n, with S being an additively idempotent semiring. Should the system possess solutions, we aim to comprehensively characterize a particular solution as it is the so-called maximal solution with respect to an order that is induced by the addition of the semiring. Additionally, in the specific scenario where S is what we call a generalized tropical semiring, we offer a thorough characterization of its solutions along with an explicit estimation of the computational cost involved in its computation.

1 Introduction

A semiring $(S, +, \cdot)$ is a set S with two internal operations, $+, \cdot$ where $(S, +)$ is a commutative monoid, and $(S, \cdot)$ is a monoid, being both internal operations connected by a ring-like distributivity. We also assume that for both operations, there exists an identity element; 0 for $+$ and 1 for $\cdot$. In addition, a semiring $(S, +, \cdot, 0, 1)$ is said to be additively idempotent if $x + x = x$ for all $x \in S$.

One of the most important examples of semirings are the tropical semirings. The semiring $(\mathbb{R}, min, +)$ appeared in optimization problems such as Floyd's algorithm for finding shortest paths in a graph [5]. However, a systematic study of the tropical semiring began only after the Simon's work (see [3]) and since then the study has significantly increased due to the huge number of applications.

The first paper [4] about linear algebra on such a semirings appeared in 2005. However, solving linear systems was a major task from the beginning of tropical algebras, but it was not until the work of Viro [6] that the problem actually took a most present role in mathematics. Moreover, this problem has already proved to be very interesting from the algorithmic point of view as it is known to be in $NP \cap coNP$ [7].

Letting $(S, +, \cdot)$ be an additively idempotent semiring, we want to solve the system $AX = Y$, where $A = (a_{ij}) \in M_{m\times n}(S)$, $Y \in S^m$ and X is an unknown vector of size n. In the context where the system $AX = Y$ admits solutions, we can compute the maximal one. Moreover, within the specific framework where S is a generalized tropical semiring (see Definition 1.1.1), we present a complete characterization of all its solutions, with an explicit polynomial computational cost.

2 Preliminars

We will recall some basic background and introduce the notation we will use through this work.

*The full version of this work can be found in [10] and has been submited to Funzy set and Systems. This research is supported by the Department of Mathematics, University of Almería.

†Email: aos073@inlumine.ual.es. Research of supported by Department of Mathematics, University of Almería.

‡Email: danielcp@ual.es. Research supported by Ministerio de Ciencia e Innovación PID2022-138906NB-C21..

§Email: jlopez@ual.es. Research supported by Ministerio de Ciencia e Innovación PID2020-113552GB-I00 and FQM 0211 Junta de Andalucía..

Definition 1. *A semiring $(R, +, \cdot)$ is a non-empty set R together with two operations $+$ and $\cdot$ such that $(R, +)$ is a commutative monoid, $(R, \cdot)$ is a monoid and the distributive laws hold:*

$$\begin{aligned} a(b+c) &= ab + ac \\ (a+b)c &= ac + bc \end{aligned} \tag{1}$$

We say that $(R, +, \cdot)$ is additively idempotent if $a + a = a$ for all $a \in R$.

Example 2. *From the work of J. Zumbrägel [13], the following additively idempotent semiring with 5 elements can be obtained:*

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	1	1	1	1	5
2	2	1	2	1	2	5
3	3	1	1	3	3	5
4	4	1	2	3	4	5
5	5	5	5	5	5	5

$\cdot$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	2	0	0	5
3	0	3	4	3	4	3
4	0	4	4	0	0	3
5	0	5	2	5	2	5

Example 3. *In [14], a classification of all additively commutative semirings with two elements is presented. In that article, we can see that the set $\{0, 1\}$ endowed with the following operations results in an additively idempotent semiring:*

$+$	0	1
0	0	0
1	0	1

$\cdot$	0	1
0	0	1
1	1	1

Definition 4. *Let R be a semiring and $(M, +)$ be a commutative semigroup with identity 0_M. M is a right semimodule over R if there is an external operation $\cdot : M \times R \to M$ such that*

$$\begin{aligned} (m \cdot a) \cdot b &= m \cdot (a \cdot b) \\ m \cdot (a+b) &= m \cdot a + m \cdot b \\ (m+n) \cdot a &= m \cdot a + n \cdot a \\ 0_M \cdot a &= 0_M \end{aligned} \tag{2}$$

for all $a, b \in R$ and $m, n \in M$. We will denote $m \cdot a$ by the concatenation ma.

In an additively idempotent semiring $(R, +, \cdot)$, an order can be induced by the addition operation, by:

$$a \leq b \text{ if and only if } a + b = b. \tag{3}$$

This order respects the operation in R and enables us to define a partial order in R^n for every positive integer n.

$$X = (x_1, \dots x_n) \geq Y = (y_1, \dots, y_n) \text{ if and only if } x_i \geq y_i \ \forall i = 1, \dots, n. \tag{4}$$

In addition, note that this order also respect the multiplication by a square matrices of order n whose entries are in R.

Let $AX = Y$ be the system of linear equations in R with indeterminates $x_1, \dots, x_n$,

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{(m-1)1} \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{(m-1)2} \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{(m-1)n} \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m-1} \\ y_m \end{pmatrix}, \tag{5}$$

with $a_{i,j}, y_j \in R$ for all $i = 1, \ldots, n$ $j = 1, \ldots, m$. Let A_j be the $j-th$ column of A, $A_j = (a_{1j}, a_{2j}, \ldots, a_{mj})$, then, the previous system can be written as

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = Y. \tag{6}$$

Definition 5. *Let R be an additively idempotent semiring, and let $AX = Y$ be a linear system of equations. We say that $\hat{X}$ is the maximal solution of the system if and only if the two following conditions are satisfied*

1. *$\hat{X} \in R^n$ is a solution of the system, i.e. $A\hat{X} = Y$,*
2. *if $Z \in R^n$ is any other solution of the system, then $Z \leq \hat{X}$*

The following result depicts a method to compute the maximal solution of such a system of equations.

Theorem 6. *Given $(R, +, \cdot)$ an additively idempotent semiring, let $W_i = \{x \in R : xA_i + Y = Y\}$ $\forall i = 1, \ldots, n$. Suppose that these subsets have a maximum with respect to the order induced in R*

$$C_i = \max W_i. \tag{7}$$

If $XA = Y$ has as a solution, then $\hat{X} = (C_1, \ldots C_n)$ is the maximal solution of the system.

Proof. If there is a solution $Z = (z_1, \ldots, z_n)$, then, it is enough to proof that $z_k \cdot A_k + Y = Y$ for all $k = 1, \ldots, n$, and therefore we can show that $z_k \in W_k$. As a consequence, $\hat{X} \geq Z$. Finally, we show that $\hat{X}$ is a solution, and therefore, it is the maximal solution. □

In [10, example 3.19], an example of a direct application of this theorem can be found.

3 Particular cases

An important example of the considered semirings is the so-called tropical semiring, which is the semiring given by $(\mathbb{R} \cup \{\infty\}, \max, +)$. The following definition is a generalization of this concept.

Definition 7. *Let $(R, +, \cdot)$ be a semiring. We say that R is a generalized tropical semiring if*

$$a + b = a \text{ or } a + b = b \text{ of all } a, b \in R.$$

It is straightforward that the tropical semiring is the tropicalized of $\mathbb{R}$ with the usual operations.

Using the argument given in the proof of the preceding theorem to this specific case, allows us to provide the following result. A complete proof of theorems 8 and 9 can be found in [10, Theorem 3.6] and [10, Theorem 3.12] respectively.

Theorem 8. *Let $(R, +, \cdot)$ be a generalized tropical semiring where $(R, \cdot)$ is a group. Then the linear system $A \cdot X = Y$ has at least one solution.*

Tropical lineal algebra over tropical semirings appears naturally in several problems of graph theory (c.f. [12] or [11]). The following result shows a characterization of all solutions of the linear system $AX = Y$.

Theorem 9. *Let R be a generalized tropical semiring, and let $AX = Y$ be a system of equations with $Y = (y_i) \in R^m$ and $A = (a_{i,j}) \in Mat_{n\times m}(R)$. $X = (x_1, x_2, ..., x_n)$ is solution of the system if and only if*

1. $a_{j,i} \cdot x_i + y_j = y_j$,$\forall j = 1, \ldots, m$,
2. $\forall j = 1, \ldots, m$ $\exists h \in \{1, \ldots, n\}$ *such that* $a_{j,h} \cdot x_h = y_j$.

Another significant case is that of finite idempotent semirings, which has garnered renewed interest in the scientific community due to its potential applications in cryptography. As an example, [8] provides a characterization of all finite commutative simple semirings, among which one of the five possible cases is the additively idempotent semiring.

Then, due to the finiteness of the semiring we get that

Theorem 10. *Let R be an additively idempotent finite semiring, and let $AX = Y$ be a system of equations, with $Y \in R^m$ and $A = (a_{i,j}) \in Mat_{n\times m}(R)$. Then, the system is compatible, $W_i = \{x \in R : x \cdot A_i + Y = Y\}$ is finite and*

$$X = (x_1, \ldots, x_n) \text{ such that } x_i = \sum_{x\in W_i} x \tag{8}$$

is the maximal solution of the system.

An important consequence of this result is that we are able to provide a cryptanalysis of the key exchange over finite semirings that are congruence simple and that is introduced in [2] and that it is published in [9].

References

[1] J. S. Golan, Semirings and their applications, Kluwer Academic Publishers, Dordrecht, 1999, xii+381.

[2] G. Maze, C. Monico, J. Rosenthal, Public key cryptography based on semigroup actions, *Adv. Math. Commun.* **1** (2007), 489–507.

[3] I. Simon, Limited subsets of a free monoid, in: *19th Annual Symposium on Foundations of Computer Science (Ann Arbor, Mich., 1978)*, IEEE, Long Beach, CA, 1978, pp. 143–150.

[4] M. Develin, F. Santos, B. Sturmfels, On the rank of a tropical matrix, in: *Combinatorial and computational geometry*, Math. Sci. Res. Inst. Publ., vol. 52, Cambridge Univ. Press, Cambridge, 2005, pp. 213–242.

[5] R. W. Floyd, Algorithm 97: Shortest Path, *Commun. ACM* **5(6)** (1962), 345. `https://doi.org/10.1145/367766.368168`

[6] O. Viro, Dequantization of Real Algebraic Geometry on Logarithmic Paper, in: *European Congress of Mathematics*, eds. Carles Casacuberta, Rosa Maria Miró-Roig, Joan Verdera, Sebastià Xambó-Descamps, Birkhäuser Basel, Basel, 2001, pp. 135–146. ISBN: 978-3-0348-8268-2.

[7] D. Grigoriev, Complexity of solving tropical linear systems, *comput. complex.* **22** (2013), 71–88, DOI 10.1007/s00037-012-0053-5.

[8] C. Monico, On finite congruence-simple semirings, *J. Algebra* **271** (2004) 846-854.

[9] A. Otero Sánchez and J. A. López Ramos, Cryptanalysis of a key exchange protocol based on a congruence-simple semiring action, Journal of Algebra and Its Applications, Online Ready No Access, https://doi.org/10.1142/S0219498825502299

[10] A. Otero Sánchez, D. Camazón, J. A. López Ramos, "On the solutions of linear systems over additively idempotent semirings", arXiv:2404.03294 [cs.IT], https://doi.org/10.48550/arXiv.2404.03294

[11] B. M. E. Moret and H. D. Shapiro, An Empirical Assessment of Algorithms for Constructing a Minimum Spanning Tree, Computational Support for Discrete Mathematics, 15(1):99–117, 1992.

[12] D. Speyer and B. Sturmfels, Tropical Mathematics, Mathematics Magazine, 82(2):163–173, April 2009.

[13] J. Zumbrägel, Classification of finite congruence-simple semirings with zero, Journal of Algebra and Its Applications, 7:363–377, 2008.

[14] R. El Bashir, J. Hurt, A. Jančařík, and T. Kepka, Simple commutative semirings, Journal of Algebra, 236:277–306, 2001.

Rainbow loose Hamilton cycles in Dirac hypergraphs

Amarja Kathapurkar*[1], Patrick Morris†[2], and Guillem Perarnau‡[2,3]

[1]University of Birmingham, United Kingdom.
[2]Universitat Politècnica de Catalunya (UPC), Barcelona, Spain.
[3]Centre de Recerca Matemàtica, Bellaterra, Spain.

Abstract

A meta-conjecture of Coulson, Keevash, Perarnau and Yepremyan [6] states that above the extremal threshold for a given spanning structure in a (hyper-)graph, one can find a rainbow version of that spanning structure in any suitably bounded colouring of the host (hyper-)graph. We solve one of the most pertinent outstanding cases of this conjecture, by showing that if G is an n-vertex k-uniform hypergraph with $\delta_{k-1}(G) \geq \left(\frac{1}{2(k-1)} + o(1)\right) n$, then any bounded colouring of G contains a rainbow loose Hamilton cycle.

1 Introduction

A famous theorem of Dirac [11] states that any n-vertex graph G with $\delta(G) \geq n/2$ contains a Hamilton cycle. This inspired many further results exploring the optimal minimum degree conditions for certain spanning structures in a host (hyper-)graph. This area, sometimes referred to as 'Dirac theory', is a cornerstone of modern extremal combinatorics and has flourished in recent decades due to powerful tools being developed to tackle these questions, such as the regularity method [31] and absorption [27]. In graphs, this has led to a deep understanding of the full picture with celebrated results including the minimum degree threshold for F-factors [24] (vertex disjoint copies of F covering the vertex set of the host graph) for arbitrary graphs F and the so-called Bandwidth Theorem [3] of Böttcher, Taraz and Schacht.

In hypergraphs, the situation is considerably more complex. This is, in part, due to the various ways in which one can generalise the graph case. For example, when generalising Dirac's theorem to hypergraphs, one has a range of choices as to which minimum degree condition is considered and what type of Hamilton cycle is desired. Indeed, for a k-uniform hypergraph G (k-graph for short), one can consider

$$\delta_j(G) := \min\left\{|\{e \in E(G) : T \subset e\}| : T \in \binom{V(G)}{j}\right\},$$

for $1 \leq j \leq k-1$. The case $j = k-1$ is often called the *codegree* of the k-graph G. Likewise with Hamilton cycles, one can consider a cyclic ordering of the vertices of G and require that each edge of the Hamilton cycle occupies k consecutive vertices in the ordering and every pair of consecutive edges intersect in precisely ℓ vertices for some $1 \leq \ell \leq k-1$. Such a Hamilton cycle is called a Hamilton ℓ-cycle and when $\ell = 1$, we refer to it as a *loose* cycle, whilst the case $\ell = k-1$ is referred to as a *tight*

*Email: amarja.kathapurkar@gmail.com. Research of A. K. supported by EPSRC Research grant EP/R034389/1.

†Email: pmorrismaths@gmail.com. Research of P. M. supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Walter Benjamin program - project number 504502205.

‡Email: guillem.perarnau@upc.edu. Research of G.P. supported by the Grant PID2020-113082GB-I00, the Grant RED2022-134947-T and the Programme Severo Ochoa y María de Maeztu por Centros y Unidades de Excelencia en I&D (CEX2020-001084-M), all of them funded by MICIU/AEI/10.13039/501100011033.

cycle. Note that if an n-vertex k-graph G has a loose Hamilton cycle, then one necessarily has that $(k-1)|n$ and similar divisibility conditions hold for the other Hamilton ℓ-cycles.

In hypergraphs our understanding of minimum degree thresholds is far from complete despite a wealth of results. Indeed, even in the case that the spanning structure is a perfect matching, there are unanswered questions. We refer the reader to the survey [32] on the matter.

Whilst establishing minimum degree thresholds can be a considerable challenge, the lower bounds often follow from simple constructions that are derived to force the non-existence of the spanning structure in question. For example, the minimum codegree threshold for a loose Hamilton cycle in a k-graph is (asymptotically) $\frac{n}{2(k-1)}$ and the following example establishes the lower bound. Take n divisible by $2(k-1)$, partition $V(G) = A \cup B$ such that $|A| = \frac{n}{2(k-1)} - 1$ and take any set of k vertices that intersects A as an edge of G. Then $\delta_{k-1}(G) = |A|$ and if there was a loose Hamilton cycle in G, then in the cyclic order defining the cycle, there cannot be $2(k-1)$ consecutive vertices from B as this would contain an edge of the Hamilton cycle but no such edge exists in B. Thus there are at least $\frac{n}{2(k-1)}$ vertices in A, contradicting the size of A.

The fact that these constructions are contrived and atypical, for example having large independent sets, suggests that although one cannot weaken the respective minimum degree condition, perhaps one can strengthen the conclusion of the degree threshold. That is to say, when we are above the minimum degree threshold (we will informally refer to such (hyper-)graphs as being 'Dirac') with respect to a given spanning structure, the Dirac (hyper-)graph is in fact *robust* with respect to containing that spanning structure. Various results of this flavour have been established, in particular in the context of Dirac's condition for Hamilton cycles, see the nice survey of Sudakov [30]. For example, it has been shown that there are in fact *many* Hamilton cycles above the extremal threshold [10], as well as many edge-disjoint Hamilton cycles [9]. In this paper, we will consider a notion of robustness related to finding *rainbow* spanning structures in any bounded edge colouring of the Dirac (hyper-)graph. This is motivated by the classical study of rainbow spanning structures in certain colourings of graphs.

1.1 Rainbow spanning structures

A subgraph H of an edge coloured graph G is said to be *rainbow* if each of the edges of H is a different colour. Rainbow subgraphs appeared early on in combinatorics via connections with design theory. Indeed, already Euler [15] was interested in transversals in Latin squares, which is a collection of entries in the Latin square with distinct rows, columns and symbols. Viewing an $n \times n$ Latin square as an edge colouring of a complete bipartite graph, with parts corresponding to columns and rows and colour classes corresponding to symbols, a transversal becomes a rainbow matching. Several beautiful conjectures were posed in design theory, that are only now being solved by heavily utilising connections to rainbow spanning subgraphs. Indeed, perhaps the most famous such conjecture, known as the Ryser-Brualdi-Stein conjecture [4, 28, 29] states that every $n \times n$ Latin square has a transversal of size at least $n-1$ and one of size n when n is odd. The first part of this (establishing the existence of transversals of size $n-1$) has only recently been solved by Montgomery [25]. Translating to colourings of graphs, the Ryser-Brualdi-Stein conjecture asserts that one can always find an (almost) perfect rainbow matching. Here, the conditions of the Latin square are equivalent to the colouring of $K_{n,n}$ having n colours and being *proper*, that is, there are no two edges of the same colour at a vertex.

From a graph theoretic perspective one can ask more generally what conditions on a colouring of a host graph guarantee the existence of a rainbow (almost) spanning structure of interest. The Ryser-Brualdi-Stein conjecture, as well as a host of other conjectures inspired by design theory, suggest that the colouring being proper is enough. In search for other conditions, researchers noted that a colouring being proper is equivalent to saying that the colouring is *locally bounded*, that is, at each vertex we see every colour at most once, or more generally, a bounded number of times. One can also then consider *globally bounded* conditions where we bound the size of each colour class.

An early example of interest in rainbow structures under global bounded conditions on colouring

was due to Erdős and Stein (see [13]) who asked whether there is some constant $c > 0$ such that any colouring of K_n with at most cn edges of each colour contains a rainbow Hamilton cycle. This was then explicitly conjectured by Hahn and Thomassen [18] and, after several results towards the conjecture, was solved by Albert, Frieze and Reed [1]. A generalisation to hypergraph Hamilton cycles was then given by Dudek, Frieze and Ruciński [12]. There has been a wealth of similar results studying different spanning structures.

One may wonder how optimal these results are. For example, note that the result of Albert, Frieze and Reed is tight up to the choice of constant $c > 0$; a value of $c < 1/2$ is certainly necessary as otherwise there may not be enough colours to have a rainbow Hamilton cycle. In the setting of perfect matchings in complete bipartite graphs, Stein [29] boldly conjectured that the condition of being proper could be dropped and replaced by each colour class simply having size n. This turned out to be false with Pokrovskiy and Sudakov [26] recently giving a construction with n edges of each colour and no rainbow transversal bigger than $n - \Omega(\log n)$. This shows that in this setting, a colouring having a global bound of n edges of each colour is not enough to guarantee the desired rainbow matching of size $n - 1$. However, in what was a hugely influential paper and the first in this area of finding rainbow structures in globally bounded colourings, Erdős and Spencer [14] showed that any colouring of $K_{n,n}$ with at most $\frac{n}{16}$ edges of each colour contains a rainbow perfect matching.

1.2 Rainbow structures in Dirac (hyper-)graphs

The vast majority of results concerning rainbow spanning substructures in bounded (and proper) colourings have focused on the case where the host graph is a complete (hyper-)graph or complete bipartite graph. When considering other possible host graphs, Dirac graphs arise naturally. Indeed, in order to contain a rainbow copy of a desired spanning subgraph in any bounded colouring, the host graph certainly needs to contain copies of that subgraph and so imposing the existence of such subgraphs through minimum degree conditions gives a natural class of candidate host graphs. This perspective was first considered by Cano, Perarnau and Serra [5] who showed that one can find a rainbow Hamilton cycle in any globally $o(n)$-bounded colouring of G when G is either an n-vertex graph or a balanced bipartite graph with n vertices in each part, and such that G has minimum degree $\delta(G) \geq (1 + o(1))\frac{n}{2}$. The asymptotic minimum degree condition was then replaced to give an exact minimum degree condition $\delta(G) \geq \frac{n}{2}$ by Coulson and Perarnau, first in the bipartite case [7] and then in the non-bipartite case [8] as in Dirac's original theorem. These results thus give evidence of robustness for the extremal thresholds for Hamilton cycles. Note also that in the bipartite case, these results can be seen as a direct strengthening of the result of Erdős and Spencer [14], allowing for host graphs that are not complete (at the expense of a potentially worse constant for the boundedness).

Further examples of these types of results came from Coulson, Keevash, Perarnau and Yepremyan [6] who proved that (asymptotically) above the minimum degree for a given (hyper-)graph F-factor, one finds a rainbow F-factor in any suitably bounded colouring, and from Glock and Joos [16] who gave a rainbow version of the famous blow-up lemma [22], which allowed them to give results of this flavour in considerable generality for graphs, in particular providing a rainbow version of the bandwidth theorem [3]. We remark that a nice feature of the work of [6] is that they could establish such a result, even in cases where the minimum degree threshold has not yet been determined.

All of these results provide evidence of a general phenomenon and caused Coulson, Keevash, Perarnau and Yepremyan [6] to explicitly give the "meta-conjecture" that once one is above the extremal threshold for a given spanning structure, rainbow copies of that structure can be found in any suitably bounded colouring of the Dirac graph. Our main result provides further evidence for this conjecture, by establishing that this is the case for loose Hamilton cycles in hypergraphs.

Theorem 1. *For any $2 \leq k \in \mathbb{N}$ and $\varepsilon > 0$, there exists $\mu > 0$ such that for any sufficiently large $n \in (k-1)\mathbb{N}$, the following holds. If G is a k-graph with $\delta_{k-1}(G) \geq (1+\varepsilon)\frac{n}{2(k-1)}$ and $\chi : E(G) \to \mathbb{N}$ is colouring of G with at most μn^{k-1} edges of each colour and at most μn edges of each colour containing*

any given $(k-1)$-set of vertices, then there exists a rainbow loose Hamilton cycle.

Theorem 1 provides a first generalisation of the result of Coulson and Perarnau [8] to the hypergraph setting. Note that the minimum degree condition is asymptotically tight due to the construction discussed above. The fact that hypergraphs with minimum codegree at least $(1+o(1))\frac{n}{2(k-1)}$ contain loose Hamilton cycles was proven originally by Kühn and Osthus [23] for $k=3$ and for general k by Hàn and Schacht [19] and independently by Keevash, Kühn, Mycroft and Osthus [20]. Our result can thus be seen as a direct strengthening of these results, providing robustness. We remark that the tight minimum codegree threshold (without the $o(1)$ factor) for the existence of a loose Hamilton cycle is unknown and seems to be a considerable challenge.

Note also that the global bound in Theorem 1 is also tight, up to the choice of the constant μ. Indeed, some global bound of the order of n^{k-1} is needed to guarantee enough colours. The local bound in Theorem 1, requiring each $(k-1)$-set to be in at most μn edges of any given colour, is rather weak in comparison to requiring a colouring to be proper, for example. It is unclear whether this local bound is in fact necessary. Indeed this condition arises as somewhat of a technicality within the proof which nonetheless seems hard to bypass. This local boundedness condition was also present in the previous result of Coulson, Keevash, Perarnau and Yepremyan [6] on rainbow factors and it can be shown to be necessary when dealing with clique factors or tight Hamilton cycles (for which it remains an open question to prove an analogue of Theorem 1). At the cost of this extra local bound, Theorem 1 strengthens the previously mentioned work of Dudek, Frieze and Ruciński [12] who proved Theorem 1 in the case that the host hypergraph G is complete. Finally, we mention a result of Antoniuk, Kamčev and Ruciński [2] who showed that under the same assumption that $\delta_{k-1}(G) \geq (1+o(1))\frac{n}{2(k-1)}$, any colouring in which each vertex is contained in at most $o(n^{k-1})$ edges of the same colour results in a Hamilton loose cycle that is properly coloured, that is, the Hamilton cycle does not contain incident edges of the same colour. Our result strengthens the conclusion by guaranteeing a rainbow loose Hamilton (which is in particular proper) at the cost of adopting both a local bound and a global bound for the colouring, the latter being necessary for the rainbow setting, as previously discussed.

2 A proof overview

The *lopsided local lemma*, originally introduced by Erdős and Spencer [14] in the context of rainbow perfect matchings in $K_{n,n}$, provides a general tool for finding rainbow spanning structures in bounded colourings of host graphs. The setup works by taking a uniformly random copy of the desired spanning structure and defining bad events based on two edges of the same colour appearing in this random sample. This setting does *not* have limited dependence between our bad events and so the original local lemma cannot be used to show that the uniform copy is rainbow with some positive probability. Nonetheless, Erdős and Spencer showed that the desired conclusion of the local lemma indeed holds if we can bound the amount of *negative dependence* between bad events. In the setting of complete (bipartite) graphs, one can carefully count copies of the desired spanning structure subject to certain bad events not taking place, allowing calculations of conditional probabilities necessary to show such negative dependence.

When the host graph is no longer complete, precise counts of spanning structures are no longer accessible. The key idea in the initial works [5, 7, 8] in Dirac host graphs, is that one can still estimate the required conditional probabilities necessary, by applying a "switching method". Here one locally alters some fixed copy of the spanning structure in a way that maintains some fixed events that we want to condition on. If we can find many ways of performing valid switchings, we can provide upper bounds on conditional probabilities to show that there is enough negative dependence in the collection of bad events for the lopsided local lemma. This switching approach was then used again in the work of Coulson, Keevash, Perarnau and Yepremyan [6] finding rainbow F-factors. Their key innovation was that one can find many switchings via probabilistic methods. They take a random sample of the vertex

set (in fact, a random sample of copies of F in the factor we are switching from) and show that with probability bounded away from 0 one can perform the switch within this random set, obtaining a new factor where some copies have been reshuffled. This translates to having many subsets providing valid switches and opens up the power of the probabilistic method to prove the existence of valid switchings. Indeed, with high probability, the sampled vertex set will inherit many nice properties of the host graph, in particular the minimum degree condition. After some work (to ensure the switching is valid), this allows the authors of [6] to apply the existence of a sub-F-factor in the random vertex set as a black box, using that the minimum degree condition is satisfied.

Our proof again follows this template and we will again use random samples to provide many switchings, setting up an application of the lopsided local lemma. There is one major hurdle in our setting as opposed to F-factors though, which comes from the fact that we are now dealing with *connected* spanning structures. This means that we cannot locally adjust our copy within the random set independently of the rest of the spanning structure. This hurdle was noted also in [2] and means that one can no longer use black box results in the random set of vertices. In order to overcome this, we use absorption techniques to rebuild the loose Hamilton cycle in the random set in such a way that it provides a valid switching. In more detail, we use an absorbing strategy due to Hàn and Schacht [19] which gives an absorbing structure as well as a connecting lemma that we can use to piece back together our loose Hamilton cyle.

To our knowledge, this is a first example of absorption being used in the context of the local lemma and we find it a nice feature of our proof that it simultaneously incorporates two of the most powerful methods in modern extremal and probabilistic combinatorics.

3 Further directions

We believe our method of using the lopsided local lemma in conjunction with absorption techniques has the potential to prove more results in the setting of robustness via rainbow structures in bounded colourings. In particular, for different Hamilton ℓ-cycles in hypergraphs under different minimum j-degree conditions, whenever there is an existing proof for the existence of the cycle that appeals to absorption techniques, there is a hope to apply our framework. This is reminiscent of recent work in the setting of transversal spanning structures [17] and robustness via percolation [21], where they provide certain 'absorption-necessary' conditions in order to give general results that follow from the previous work in establishing extremal thresholds, in particular covering many different types of Hamilton cycle and minimum degree conditions. The full power of our approach will be explored in a forthcoming journal version of this extended abstract.

References

[1] M. Albert, A. Frieze, and B. Reed. Multicoloured Hamilton cycles. *The Electronic Journal of Combinatorics*, 2(1):R10, 1995.

[2] S. Antoniuk, N. Kamčev, and A. Ruciński. Properly colored hamilton cycles in dirac-type hypergraphs. *The Electronic Journal of Combinatorics*, pages P1–44, 2023.

[3] J. Böttcher, M. Schacht, and A. Taraz. Proof of the bandwidth conjecture of Bollobás and Komlós. *Mathematische Annalen*, 343(1):175–205, 2009.

[4] R. A. Brualdi, H. J. Ryser, et al. *Combinatorial matrix theory*, volume 39. Springer, 1991.

[5] P. Cano, G. Perarnau, and O. Serra. Rainbow spanning subgraphs in bounded edge–colourings of graphs with large minimum degree. *Electronic Notes in Discrete Mathematics*, 61:199–205, 2017.

[6] M. Coulson, P. Keevash, G. Perarnau, and L. Yepremyan. Rainbow factors in hypergraphs. *Journal of Combinatorial Theory, Series A*, 172:105184, 2020.

[7] M. Coulson and G. Perarnau. Rainbow matchings in Dirac bipartite graphs. *Random Structures & Algorithms*, 55(2):271–289, 2019.

[8] M. Coulson and G. Perarnau. A rainbow Dirac's theorem. *SIAM Journal on Discrete Mathematics*, 34(3):1670–1692, 2020.

[9] B. Csaba, D. Kühn, A. Lo, D. Osthus, and A. Treglown. Proof of the 1-factorization and Hamilton decomposition conjectures. *Memoirs of the American Mathematical Society*, 244(1154), 2016.

[10] B. Cuckler and J. Kahn. Hamiltonian cycles in Dirac graphs. *Combinatorica*, 29:299–326, 2009.

[11] G. A. Dirac. Some theorems on abstract graphs. *Proceedings of the London Mathematical Society*, 3(1):69–81, 1952.

[12] A. Dudek, A. Frieze, and A. Ruciński. Rainbow Hamilton cycles in uniform hypergraphs. *The Electronic Journal of Combinatorics*, page P46, 2012.

[13] P. Erdős, J. Nešetril, and V. Rödl. Some problems related to partitions of edges of a graph. *Graphs and other combinatorial topics, Teubner, Leipzig*, 5463, 1983.

[14] P. Erdős and J. Spencer. Lopsided Lovász local lemma and Latin transversals. *Discrete Applied Mathematics*, 30(151-154):10–1016, 1991.

[15] L. Euler. Recherches sur un nouvelle espéce de quarrés magiques. *Verhandelingen uitgegeven door het zeeuwsch Genootschap der Wetenschappen te Vlissingen*, pages 85–239, 1782.

[16] S. Glock and F. Joos. A rainbow blow-up lemma. *Random Structures & Algorithms*, 56(4):1031–1069, 2020.

[17] P. Gupta, F. Hamann, A. Müyesser, O. Parczyk, and A. Sgueglia. A general approach to transversal versions of Dirac-type theorems. *Bulletin of the London Mathematical Society*, 55(6):2817–2839, 2023.

[18] G. Hahn and C. Thomassen. Path and cycle sub-Ramsey numbers and an edge-colouring conjecture. *Discrete Mathematics*, 62(1):29–33, 1986.

[19] H. Hàn and M. Schacht. Dirac-type results for loose Hamilton cycles in uniform hypergraphs. *Journal of Combinatorial Theory, Series B*, 100(3):332–346, 2010.

[20] P. Keevash, D. Kühn, R. Mycroft, and D. Osthus. Loose Hamilton cycles in hypergraphs. *Discrete Mathematics*, 311(7):544–559, 2011.

[21] T. Kelly, A. Müyesser, and A. Pokrovskiy. Optimal spread for spanning subgraphs of Dirac hypergraphs. *arXiv preprint arXiv:2308.08535*, 2023.

[22] J. Komlós, G. N. Sárközy, and E. Szemerédi. Blow-up lemma. *Combinatorica*, 17:109–123, 1997.

[23] D. Kühn and D. Osthus. Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree. *Journal of Combinatorial Theory, Series B*, 96(6):767–821, 2006.

[24] D. Kühn and D. Osthus. The minimum degree threshold for perfect graph packings. *Combinatorica*, 29(1):65–107, 2009.

[25] R. Montgomery. A proof of the Ryser-Brualdi-Stein conjecture for large even n. *arXiv preprint arXiv:2310.19779*, 2023.

[26] A. Pokrovskiy and B. Sudakov. A counterexample to Stein's Equi-n-square conjecture. *Proceedings of the American Mathematical Society*, 147(6):2281–2287, 2019.

[27] V. Rödl, A. Ruciński, and E. Szemerédi. A Dirac-type theorem for 3-uniform hypergraphs. *Combinatorics, Probability and Computing*, 15(1-2):229–251, 2006.

[28] H. J. Ryser. Neuere probleme der kombinatorik. *Vorträge über Kombinatorik, Oberwolfach*, 69(91):35, 1967.

[29] S. K. Stein. Transversals of latin squares and their generalizations. *Pacific Journal of Mathematics*, pages 567–575, 1975.

[30] B. Sudakov. Robustness of graph properties. *Surveys in Combinatorics 2017*, 440:372, 2017.

[31] E. Szemerédi. Regular partitions of graphs. In *Problémes Combinatoires et Théorie des Graphes Colloques Internationaux CNRS 260*, pages 399–401. 1978.

[32] Y. Zhao. Recent advances on Dirac-type problems for hypergraphs. *Recent trends in combinatorics*, pages 145–165, 2016.

d-regular graph on n vertices with the most k-cycles

Gabor Lippner[1] and Arturo Ortiz San Miguel*[1]

[1]Dept. of Mathematics, Northeastern University, Boston, MA, USA

Abstract

We construct the unique d-regular graph G with the maximum number of k-cycles for $k = 5, 6$ with a fixed number $n = c(d+1)$ of vertices for $k = 5$ and $n = 2cd$ vertices for even $k = 6$. Using a Möbius inversion relation between graph homomorphism numbers and injective homomorphism numbers, we reframe the problem as a continuous optimization problem on the eigenvalues of G by leveraging the fact that the number of closed walks of length k is $\mathrm{tr}(A^k)$. For $k = 5$ and $d > 3$, we show G is a collection of disjoint K_{d+1} graphs. For $d = 3$, disjoint Petersen graphs emerge. For $k = 6$ and d large enough, G consists of copies of $K_{d,d}$. We conjecture that for odd k and sufficiently large d, the optimal G is a collection of K_{d+1}, while for even k with sufficiently large d, the optimal G consists of $K_{d,d}$.

Additionally, we introduce and give formulas for non-backtracking homomorphism numbers and backtracking homomorphism numbers, respectively. Moreover, we find the unique d-regular graph on n vertices with the most non-backtracking closed walks of length k by considering an optimization problem on the non-backtracking spectrum of G. We also solve the same problem, but for backtracking closed walks. Lastly, a corollary gives formulas for the number of 4-cycles and 5-cycles of a graph with respect to its spectrum, regardless of regularity.

1 Introduction

For given positive integers d, n, k we consider the d-regular graph G on n vertices that maximizes the number of k-cycles. For convenience, throughout this paper, n will be a multiple of $c(d+1)$ or of $2cd$ as we are ultimately interested in asymptotic behavior similar to [5]. Note that uniqueness of an optimizer is not true for all n, d, k. However, if there are no 'remainder vertices,' then the optimizer is unique. Here are some preliminary and elementary results from [5].

1. For $k = 3$ and $n = c(d+1)$, c copies of K_{d+1} is optimal.
2. For $k = 4$ and $n = 2cd$, c copies of $K_{d,d}$ is optimal.
3. Let $n = c(d+1)$. The d-regular graph with the most K_k subgraphs is c copies of K_{d+1}.

First, we give a technical lemma that will be used for the continuous optimization problems that follow. Furthermore, every graph that we call "optimal" or "maximal" is the *unique* graph that maximizes the objective. Furthermore, all optimizers given are graphs that are determined by their spectra [7].

Lemma 1. *Let p be a degree k polynomial with a positive leading coefficient. For d large enough, the constrained optimization problem,*

$$\textit{maximize } \sum_{i=1}^{n} p(\lambda_i), \quad \textit{subject to } \sum_{i=1}^{n} \lambda_i = 0, \sum_{i=1}^{n} \lambda_i^2 = nd, \lambda_{\max} = d, |\lambda_i| \le d,$$

is uniquely solved by

$$\begin{cases} \lambda_1 = ... = \lambda_{n/c} = d, \quad \lambda_{n/c+1} = ... = \lambda_n = -1, \ \textit{if } k \textit{ odd, } n = c(d+1) \\ \lambda_1 = ... = \lambda_{n/c} = d, \quad \lambda_{n/c+1} = ... = \lambda_{2n/c} = -d, \quad \lambda_{2n/c+1} = ... = \lambda_n = 0, \ \textit{if } k \textit{ even, } n = 2cd \end{cases}.$$

*Email: ortizsanmiguel.a@northeastern.edu.

Proof for odd k. We will show this in two steps. First, that there must be exactly n/c variables with a value of d, and then that the rest of the variables must be equal to each other.
Step 1: Exactly n/c variables are equal to d.
Case 1: Suppose $\lambda_1, ..., \lambda_n$ satisfy the constraints and that $\ell < n/c$ of them are equal to d. Then, for large enough d, it suffices to consider $p(x) = x^k$. Then, since k is odd, $\lambda_i \leq d$, and the fact that $|\sum x_i|^k \geq \left|\sum x_i^k\right|$, we have

$$\ell d^k + \sum_{i=\ell+1}^{n} \lambda_i^k < (\ell+1)d^k + \sum_{i=\ell+2}^{n} \left(\lambda_i - \frac{d - \lambda_{\ell+1}}{n-\ell-1}\right)^k.$$

Case 2: Suppose there are $m > n/c$ variables that are equal to d. Then, for some $\epsilon > 0$,

$$md^k + \sum_{i=m+1}^{n} \lambda_i^k < (m-1)d^k + (d-\epsilon)^k + \sum_{i=m+1}^{n} \left(\lambda_i + \frac{\epsilon}{n-m}\right)^k$$
$$= md^k + \sum_{i=m+1}^{n} \lambda_i^k + \sum_{i=m+1}^{n} \left[\sum_{j=1}^{k} \binom{k}{j} d^{k-j}(-\epsilon)^j + \sum_{j=1}^{k} \binom{k}{j} \lambda_i^{k-j} \left(\frac{\epsilon}{n-m}\right)^j\right].$$

Thus, the optimizer has $\lambda_1, ..., \lambda_{n/c} = d$. **Step 2: The rest of the variables are equal.**
Suppose that they are not equal. Then, without loss of generality, by the constraints, we can assume that $\lambda_{n/c+1} > -1 > \lambda_{n-s} \geq ... \geq \lambda_n$ so that $|\lambda_{c/n+1} + 1| \leq |\lambda_{n-s} + 1|$. Note that if this is not true then the same holds in the reverse direction. Then,

$$\sum_{i=m+1}^{n} \lambda_i^k = \lambda_{n/c+1}^k + \left(\frac{n(c-1)}{c} - 1 - s\right)(-1)^k + \sum_{i=n-s}^{n} \lambda_i^k < \left(\frac{n(c-1)}{c} - s\right)(-1)^k + \sum_{i=n-s}^{n} (\lambda_i^k + g(\lambda_i))^k,$$

for some function g such that the constraints are still satisfied. Thus, $\lambda_1 = ... = \lambda_{n/c} = d$, $\lambda_{n/c+1} = ... = \lambda_n = -1$ is a maximizer. □
Proof for even k. A similar argument is used to find the solutions. For sufficiently large d, it suffices to consider $p(x) = x^k$. Suppose there are $\ell < 2n/c$ variables that have magnitude d, that $|\lambda_{\ell+1}| \geq ... \geq |\lambda_n|$, and that the constraints are satisfied. Without loss of generality, let $\lambda_{\ell+1} > 0$. Then, for $\epsilon > 0$,

$$\ell d^k + \sum_{i=\ell+1}^{n} \lambda_i^k < \ell d^k + (\lambda_{\ell+1} + \epsilon)^k + \sum_{i=\ell+2} \left(\lambda_i^k - \frac{\epsilon}{n-m}\right)^k.$$

Thus, exactly $2n/c$ variables have magnitude d. The constraints force the optimizer to be what is claimed in the statement. □

Lemma 2. *For odd k and $n = c(d+1)$, the graph with the maximal number of closed walks of length k is the graph consisting of c copies of K_{d+1}.*

Proof. If there is a graph with adjacency matrix A with eigenvalues λ_i that solve the optimization problem,

$$\text{maximize } \sum_{i=1}^{n} \lambda_i^k, \quad \text{subject to } \sum_{i=1}^{n} \lambda_i = 0, \sum_{i=1}^{n} \lambda_i^2 = nd, \lambda_{\max} = d, |\lambda_i| \leq d,$$

then it is an optimizer. By Lemma 1, $\lambda_1 = ... = \lambda_{n/c} = d$, $\lambda_{n/c+1} = .. = \lambda_n = -1$ is optimal. The graph consisting of c copies of K_{d+1} uniquely has this spectrum and is thus optimal. □

Lemma 3. *For even k and $n = 2cd$ the graph with the maximal number of closed walks of length k is the graph consisting of c copies of $K_{d,d}$.*

Proof. The problem is equivalent to the optimization problem in the previous lemma. By Lemma 1, $\lambda_1 = ... = \lambda_{n/c} = d, \lambda_{n/c+1} = ... = \lambda_{2n/c} = -d, \lambda_{2n/c+1} = ... = \lambda_n = 0$ is a maximizer. The graph with c copies of $K_{d,d}$ uniquely has this spectrum. □

It is remarkable that there exist graphs whose spectra are the solutions to these optimization problems, which is not *a priori* the case. This happens for every optimization problem we consider. In the language of graph homomorphisms we just found $\max \hom(C_k, G)$ over all $d-$regular G with n vertices. For k-cycles instead of closed walks, the problem becomes $\max \mathrm{inj}(C_k, G)$. The following equations relate these quantities using the Möbius inverse of the partition lattice.

Lemma 4.

$$\hom(H, G) = \sum_P \mathrm{inj}(H/P, G).$$

$$\mathrm{inj}(H, G) = \sum_P \mu_p \cdot \hom(H/P, G), \text{ with } \mu_p = (-1)^{v(G)-|P|} \prod_{S \in P} (|S| - 1)!$$

where P ranges over all partitions of $V(H)$ and where $|P|$ is the number of classes in the partition and S are the classes of P. [6]

It is important to note that some of the resulting quotient graphs will have self-loops. Since G is simple, these terms vanish. We will use these formulas to find an eigenvalue optimization problem corresponding to finding the d-regular graph with the most k-cycles. When the context is clear we will write $\hom(H) = \hom(H, G)$. We will now display the type of results that can be achieved by using spectral theory and Lemma 4 by giving a formula for the number of 4-cycles of a graph.

Proposition 5. *Given a graph G, with adjacency matrix A and eigenvalues $\lambda_1, ..., \lambda_n$, the number of 4-cycles in G is*

$$\frac{1}{8}\left(\left[\sum_{i=1}^{n} \lambda_i^4\right] - 2 \cdot \mathbf{1}^T A^2 \mathbf{1} + \mathbf{1}^T A \mathbf{1}\right),$$

where $\mathbf{1}$ is the all ones vector. In particular, if G is d-regular, then the number of 4-cycles is

$$\frac{1}{8}\left(\sum_{i=1}^{n} \lambda_i^4 - nd^2 + nd\right).$$

Proof. Notice that the number of 4-cycles is exactly $\frac{1}{8}\mathrm{inj}(C_4, G)$. Then, by Lemma 4,

$$\mathrm{inj}(C_4, G) = \hom(C_4, G) - 2 \cdot \hom(P_3, G) + \hom(K_2, G).$$

The first term is the number of closed walks of length 4, the second the number of walks of length 2, and the last being the number of walks of length 1. □

Theorem 6. *For $d > 3$ and $n = c(d+1)$, the d-regular graph on n vertices with the most 5-cycles is c copies of K_{d+1}. For $d = 3$ and $n = 10c$, then the optimal graph is c copies of the Petersen graph.*

Proof. By Lemma 4, we calculate: $\mathrm{inj}(C_5, G) = \hom(C_5) - 5 \cdot \hom(K_3 + e) + 5 \cdot \hom(K_3)$,
where the "$+e$" means with an antenna. Then, since G is d-regular, we have $\hom(K_3{+}e) = d \cdot \hom(K_3)$. Thus, we consider the optimization problem,

$$\text{maximize} \sum_{i=1}^{n} \lambda_i^5 + (5 - 5d)\lambda_i^3, \quad \text{subject to } \sum_{i=1}^{n} \lambda_i = 0, \sum_{i=1}^{n} \lambda_i^2 = nd, \lambda_{\max} = d, |\lambda_i| \leq d.$$

By Lemma 1, for $d > 3$, this is solved when $\lambda_i = ... = \lambda_c = d$ and $\lambda_{c+1}, ..., \lambda_n = -1$. The graph consisting of c copies of K_{d+1} has this spectrum. For $d = 3$, the solution is the spectrum of the Petersen graph. □

Alternate proof for Petersen graphs. Consider a 3-regular graph with an edge containing the maximal number of 5-cycles going through it, which is 11. Then, there is an edge with no 5-cycles going through it. Thus, the Petersen graph, with 10 5-cycles going through every edge, is optimal. □

This alternate method, which was originally used to find maximal graphs in [5], becomes impractical for large d and k. The new spectral method works for all d and can be used to obtain formulas for the number of 4-cycles and 5-cycles of a graph, regardless of regularity.

Corollary 7. *Given a graph G with adjacency matrix A with eigenvalues $\lambda_1, ..., \lambda_n$, the number of 5-cycles in G is*

$$\frac{1}{10}\left(\left[\sum_{i=1}^{n}\lambda_i^5+5\lambda_i^3\right]-5\cdot\mathrm{tr}(\mathrm{diag}(A^3)D)\right),$$

where D is the diagonal degree matrix and the diag *operator sets the non-diagonal entries to zero. In particular, if G is d-regular then, the number of 5-cycles is*

$$\frac{1}{10}\left(\sum_{i=1}^{n}\lambda_i^5+(5-5d)\lambda_i^3\right).$$

Proof. The only term that is not immediately clear is $-5\cdot\mathrm{tr}(\mathrm{diag}(A^3)D)$. This is the number of homomorphisms $\phi: K_3+e\to G$. Without loss of generality, this equals the number of homomorphisms $\psi: K_3\to G$ with $\psi(1)=v$ times the degree of v summed over all $v\in G$ as the antenna may map to any of the neighbors of v. □

Furthermore, the new method also works for $k=6$. When using Lemma 4, we notice that the terms where H/P is a tree are constant for d-regular G. So, we can omit them. We get

$$\mathrm{inj}(C_6,G)=\left[\sum_{i=1}^{n}\lambda_i^6+(6-6d)\lambda_i^4-6\lambda_i^3\right]-3\cdot\hom(B,G)+9\cdot\hom(K_4\setminus e,G)+C, \qquad (*)$$

for some $C\in\mathbb{Z}$ where B is the 'bowtie' graph. We note that $\hom(B)$ and $\hom(K_4\setminus e)$ cannot be expressed using eigenvalues. Thus, we need the following.

Lemma 8. *For any G,* $\hom(B)\geq 4\cdot\hom(K_4\setminus e)$.

Proof. $\hom(B)=\mathrm{inj}(B)+4\cdot\hom(K_4\setminus e)+2\cdot\hom(K_3)$. □

Theorem 9. *For $n=2cd$ and d large enough, the d-regular graph on n vertices with the most 6-cycles is c copies of $K_{d,d}$.*

Proof. Since $\hom(B,G)\geq 4\cdot\hom(K_4\setminus e,G)$, we have that the non-constant terms outside of the sum in $(*)$ are non-positive and thus maximized when they are zero. Note that if G is bipartite, then these terms are zero. By Lemma 1, the spectral part of $(*)$ is maximized when $\lambda_1=...=\lambda_c=d, \lambda_{c+1}=...=\lambda_{2c}=-d, \lambda_{2c+1}=...=\lambda_n=0$. Thus the upper bound given by,

$$\mathrm{maxinj}(C_6,G)\leq\max\left(f(\lambda)\right)+\max\left(-3\cdot\hom(B,G)+9\cdot\hom(K_4\setminus e,G)\right),$$

where $f(\lambda)$ is the spectral term in $(*)$ is attained. □

As k grows, more non-spectral terms appear and more inequalities between homomorphsim numbers are needed. As a result, it is hard to come up with a scheme that does this for all k. In fact, in [4], it was shown that any linear inequality between homomorphism densities, which are defined using homomorphism numbers, can be shown using a (possibly infinite) number of Cauchy-Schwarz inequalities. However, deciding whether such an inequality is true is indeterminable. Thus, we introduce the notion of a non-backtracking, respectively backtracking, homomorphism number and use non-backtracking spectral theory developed in [1, 2, 3].

2 Non-backtracking Homomorphisms

A homomorphism $\phi : V(H) \to V(G)$ is a non-backtracking homomorphism if for each vertex $u \in G$, each neighbor of u has distinct images. That is, ϕ is a non-backtracking homomorphism if

$$\forall u \in V(H), \forall v, w \in N(u), \phi(v) \neq \phi(w).$$

Denote the number of non-backtracking homomorphisms from H to G as $\mathrm{nob}(H, G)$. We see that, $\hom(H, G) \geq \mathrm{nob}(H, G) \geq \mathrm{inj}(H, G)$. We give a relation between these quantities.

Proposition 10. *For graphs H, G, we have*

$$\mathrm{nob}(H, G) = \sum_{Q} \mathrm{inj}(H/Q, G),$$

where Q ranges over all partitions of G where each part is an independent set with no common neighbors.

Proof. For any partition Q and any injective homomorphism of H/Q, we get exactly one non-backtracking homomorphism of H. It is exactly the one where the vertices $v_i \in H$ that are in the part of Q represented by $v \in H/Q$ are mapped to the same vertex that v is mapped to in the injective homomorphism. There are no other non-backtraccking homomorphisms because any partition where some part has a common neighbor, the resulting homomorphism from $H \to G$ will be backtracking. □

Similarly, we can denote $\mathrm{bac}(H, G)$ as the number of backtracking homomorphisms from H to G. A backtracking homomorphsism is a homomorphsism that is not non-backtracking. Clearly, we have,

$$\hom(H, G) = \mathrm{nob}(H, G) + \mathrm{bac}(H, G) = \sum_{P} \mathrm{inj}(H/P, G) \implies \mathrm{bac}(H, G) = \sum_{S} \mathrm{inj}(H/S, G),$$

where S ranges over partitions of $V(H)$ with common neighbors in some part. Note that a Möbius inversion relation like in Lemma 4 does not hold. However, if we label the vertices and edges of the quotients and define neighbors in a way that considers the labeling, then such an inversion holds.

2.1 Maximizing Non-backtracking and Backtracking

We can count the number of closed non-backtracking walks of length k of G with the following results from [1, 2, 3]. In the latter, it was shown that there is a closed form for the non-backtracking spectrum in terms of the ordinary spectrum of G. Consider the directed graph $\tilde{G} = (\tilde{V}, \tilde{E})$ where $|\tilde{V}| = 2|E|$, where each vertex is represented by $(u, v) \subset E$. Then, we have $\tilde{E} = \{(u, v), (x, y) : v = x, u \neq y\}$. The non-backtracking matrix of G, denoted B, is the adjacency matrix of $\tilde{G}$, which is given by,

$$B_{(u,v),(x,y)} = \begin{cases} 1, & \text{if } v = x, u \neq y \\ 0, & \text{otherwise} \end{cases}.$$

Note that each distinct directed closed walk of $\tilde{G}$ corresponds to a unique non-backtracking walk of G of the same length. Thus, the number of closed non-backtracking walks of length k of G is equal to $\mathrm{tr}(B^k)$. Furthermore, we have the following result.

Proposition 11. *Let G be a d-regular graph. Then, the eigenvalues of B are*

$$\pm 1, \frac{\lambda_i \pm \sqrt{\lambda_i^2 - 4(d-1)}}{2},$$

where λ_i are the eigenvalues of A and ± 1 each have multiplicity $m - n$, where m is the number of edges in G [3].

In general, by the binomial theorem, the problem of finding the graph with the most non-backtracking closed walks of length k becomes:

$$\max_{\lambda} \sum_{i=1}^{n} \left(\frac{\lambda_i + \sqrt{\lambda_i^2 - 4(d-1)}}{2} \right)^k + \left(\frac{\lambda_i - \sqrt{\lambda_i^2 - 4(d-1)}}{2} \right)^k$$

$$= \max_{\lambda} \sum_{j=1}^{n} \sum_{i=1}^{\lfloor k/2 \rfloor} \binom{k}{2i} 4 \left(\frac{\lambda_j}{2} \right)^{k-2i} \frac{(\lambda_j^2 - 4(d-1))^i}{2^{2i}}, \text{ s.t. } \sum_{i=1}^{n} \lambda_i = 0, \sum_{i=1}^{n} \lambda_i^2 = nd, \lambda_{\max} = d, |\lambda_i| \leq d.$$

Note that the objective function above is always real as $z^k + \overline{z}^k = z^k + \overline{(z^k)} = 2\mathrm{Re}(z^k)$. This is a sum of polynomials with positive leading coefficient and satisfies the assumptions of Lemma 1.

Theorem 12. *For odd k, sufficiently large d, and $n = c(d+1)$, the d-regular graph on n vertices with the most non-backtracking closed walks of length k is c copies of K_{d+1}.*

For even k, sufficiently large d and $n = 2cd$, the d-regular graph on n vertices with the most non-backtracking closed walks of length k is c copies of $K_{d,d}$.

Proof. By Lemma 1, the solution of the optimization problem is the spectrum for these graphs. □

We now find $\max_G \mathrm{bac}(C_k, G)$. The number of backtracking walks of length k is the sum of those that backtrack once, those that backtrack twice, thrice, and so on. Denote $\mathrm{bac}_{i_1,i_2,\ldots,i_\ell}(C_k, G)$ as the number of backtracking homomorphisms that backtrack $i = \sum_j^\ell i_j$ times where i_j denotes the length of the jth consecutive backtracking streak. We compute,

$$\mathrm{bac}(C_k) = \sum_{i=1}^{n} \sum_{i_1+\ldots i_\ell = i} \mathrm{bac}_{i_1,i_2,\ldots,i_\ell}(C_k) = \sum_{i=1}^{n} \sum_{i_1+\ldots+i_\ell=i} \hom(H_{i_1,\ldots,i_\ell}) = \sum_{i=1}^{n} \sum_{i_1+\ldots+i_\ell=i} d^a \sum_{j=1}^{n} \lambda_j^{k-i},$$

where $H_{i_1,\ldots,i_\ell}$ is C_{k-i} with a antennas with $a = (\#$ odd length streaks in $i_1, i_2, \ldots, i_\ell)$. This is maximized at the desired spectrum by Lemma 1 because every coefficient is positive. Thus, we have the following result.

Proposition 13. *For d sufficiently large, the d-regular graph on $n = c(d+1)$ vertices with the most closed backtracking walks of odd length k is c copies of K_{d+1}. Similarly, for sufficiently large d, if $n = 2cd$ and k is even, then the optimal graph is c copies of $K_{d,d}$.*

Proof. Using the above equation as the objective function with the same constraints as before, by Lemma 1, gives the spectra of K_{d+1}, or $K_{d,d}$ respectively, as the optimizer. □

References

[1] Noga Alon et al. "Non-backtracking random walks mix faster". In: *Communications in Contemporary Mathematics* 9.04 (2007), pp. 585–603.

[2] Ewan Davies et al. "Independent sets, matchings, and occupancy fractions". In: *Journal of the London Mathematical Society* 96.1 (2017), pp. 47–66.

[3] Cory Glover and Mark Kempton. "Spectral properties of the non-backtracking matrix of a graph". In: *Linear Algebra and its Applications* 618 (2021), pp. 37–57.

[4] Hamed Hatami and Serguei Norine. "Undecidability of linear inequalities in graph homomorphism densities". In: *Journal of the American Mathematical Society* 24.2 (2011), pp. 547–565.

[5] Pim van der Hoorn, Gabor Lippner, and Elchanan Mossel. "Regular graphs with linearly many triangles are structured". In: *The Electronic Journal of Combinatorics* 29.1 (2022).

[6] László Lovász. *Large networks and graph limits*. Vol. 60. American Mathematical Soc., 2012.

[7] Edwin R Van Dam and Willem H Haemers. "Which graphs are determined by their spectrum?" In: *Linear Algebra and its applications* 373 (2003), pp. 241–272.

The weight spectrum of the Reed-Muller codes $RM(m-5,m)$ *

Claude Carlet†

Dept. of Informatics, University of Bergen, 5005 Bergen, Norway
Dept. of Mathematics, University of Paris 8, 93526 Saint-Denis, France

Abstract

The weight spectra (i.e. the lists of all possible weights) of the Reed-Muller codes $RM(r,m)$, of length 2^m and order r, are unknown for $r \in \{3,\ldots,m-5\}$ (and m large enough). Those of $RM(m-4,m)$ and $RM(m-3,m)$ have been determined very recently (but not the weight distributions, giving the number of codewords of each weight, which seem out of reach). We determine the weight spectrum of $RM(m-5,m)$ for every $m \geq 10$. We proceed by first determining the weights in $RM(5,10)$. To do this, we construct functions whose weights are in the set $\{62,74,78,82,86,90\}$, and functions whose weights are all the integers between 94 and $2^9-2=510$ that are congruent with 2 modulo 4 (those weights that are divisible by 4 are easier to determine and they are indeed known). This allows us to determine completely the weight spectrum, thanks to the well-known result due to Kasami, Tokura and Azumi, which precisely determines those codeword weights in Reed-Muller codes which lie between the minimum distance d and 2.5 times d, and thanks to the fact the weight spectrum is symmetric with respect to 2^9. Then we use this particular weight spectrum for determining that of $RM(m-5,m)$, by an induction on m.

This extended abstract is an excerpt of the full paper [3].

1 Introduction

Given $0 \leq r \leq m$, the Reed-Muller code $RM(r,m)$, of length 2^m and order r, is made of all m-variable Boolean functions f of algebraic degree at most r (or more precisely of the binary vectors of length 2^m that are the lists of values of $f(\mathbf{x})$ when $\mathbf{x}=(x_1,\ldots,x_m)$ ranges over $\mathbb{F}_2^m$ in some fixed order). All codeword weights in the Reed-Muller codes of length 2^m and orders $0,1,2,m-2,m-1,m$ are known (as well as the weight distributions of these codes). They are recalled for instance in [7] and in [4].
The low Hamming weights are also known in all Reed-Muller codes: Kasami and Tokura [5] have shown that, for $r \geq 2$, the only Hamming weights in $RM(r,m)$ occurring in the range $[2^{m-r};2^{m-r+1}[$ are of the form $2^{m-r+1}-2^{m-r+1-i}$ where $i \leq \max(\min(m-r,r),\frac{m-r+2}{2})$.

Kasami, Tokura and Azumi determined later in [6] all the weights lying between the minimum distance $d=2^{m-r}$ and 2.5 times d. The functions having such weights are characterized in this reference (all weights are described at pages 392 and following of the reference, and the corresponding functions are described under some conditions in its Table I).

The weight spectra (i.e., the sets of all possible codeword weights) of the codes $RM(r,m)$ are unknown for $3 \leq r \leq m-5$ (and therefore, their weight distributions are also unknown) but they have been recently determined in [4] for $r=m-4,m-3$, thanks to the fact that there is a simple way to determine many weights in $RM(r,m)$ from the weights in $RM(r-1,m-1)$; the weight spectra of $RM(m-c,m)$

*The full version of this work is to appear in IEEE Transactions on Information Theory. It has never been presented in a conference (only in a local seminar) .This research is supported by the Norwegian Research Council.

†Email: claude.carlet@gmail.com.

were then deduced for $c = 3, 4$, thanks to the Kasami-Tokura's results [5], which allowed to know that the numbers missing in the obtained lists could not be weights in these codes.

Reference [4] could not address the cases $c \geq 5$, mainly because the weights that are not divisible by 4 in $RM(5, 10)$ could not be determined. In the present paper, we solve the case $c = 5$, by constructing codewords in $RM(5, 10)$ achieving all the weights allowed by [6] and all those that are larger than $2.5d$ and smaller than 2^{m-1}, and thanks to an induction on m.

2 Preliminaries

The *Hamming weight* (in brief, the weight) of a binary vector $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ is the size of its support $\{i \in \{1, \ldots, n\}; x_i \neq 0\}$. The *Hamming distance* between two vectors in $\mathbb{F}_2^n$ is the weight of their difference (that is, of their sum). Hence, since m-variable Boolean functions can be identified with binary vectors of length $n = 2^m$, the *Hamming weight* of an m-variable Boolean function f is the size of its support $\{\mathbf{x} \in \mathbb{F}_2^m; f(\mathbf{x}) \neq 0\}$, and the *Hamming distance* between two Boolean functions is the weight of their sum. A binary linear code of *length* n is an $\mathbb{F}_2$-subspace of $\mathbb{F}_2^n$. This allows to define its *dimension* (as an $\mathbb{F}_2$-vector space). Its *minimum distance* is the minimum Hamming distance between distinct codewords, that is (thanks to the linearity of the code) the minimum Hamming weight of the nonzero codewords.

The set of the codeword weights of a given linear code C will be called the *weight spectrum* of C, and for simplicity, we will sometimes write "the weights of C" instead of "the weights of the codewords in C" for the elements of its weight spectrum. The *weight distribution* of the code is the list of the numbers A_i, where A_i equals the number of codewords of weight i for $i \in \{0, \ldots, n\}$.

Given two integers m and $r \in \{0, \ldots, m\}$, the *Reed Muller* code $RM(r, m)$ of length $n = 2^m$ and order r is defined in terms of Boolean functions (see [7]): each m-variable Boolean function $f : \mathbb{F}_2^m \mapsto \mathbb{F}_2$ admits a unique representation as a polynomial in $\mathbb{F}_2[x_1, \ldots, x_m]/(x_1^2 + x_1, \ldots, x_m^2 + x_m)$, called the algebraic normal form (ANF) of f. We choose an order on $\mathbb{F}_2^m$, that is, we write $\mathbb{F}_2^m = \{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_n\}$, and we denote by ev the evaluation map from the space of Boolean functions to $\mathbb{F}_2^n$ by the rule $ev(f) = (f(\mathbf{P}_1), \ldots, f(\mathbf{P}_n))$. Then $RM(r, m)$ equals $\{ev(f) \mid f \in B_m \text{ and } \deg(f) \leq r\}$, where B_m is the vector space of all m-variable Boolean functions and $\deg(f)$, called the algebraic degree of f, is the (global) degree of the ANF of f. Boolean function f has an odd Hamming weight if and only if it has (maximal) algebraic degree m.

The dimension of $RM(r, m)$ equals $\sum_{i=0}^{r} \binom{m}{i}$ and its minimum distance equals 2^{m-r}. The minimum weight codewords are the indicators of the $(m - r)$-dimensional affine subspaces of $\mathbb{F}_2^m$; up to affine equivalence, they equal $\prod_{i=1}^{r} x_i$ (two Boolean functions are called affine equivalent if one equals the composition of the other by an affine permutation).

The McEliece theorem gives a divisibility lower bound on the weights in $RM(r, m)$:

Theorem 1 (McEliece divisiblity theorem)**.** *[8] The weights in $RM(r, m)$ are multiples of $2^{\lfloor \frac{m-1}{r} \rfloor}$.*

This bound is tight, as shown in [1]; more precisely, for each pair (r, m), there is at least one codeword of $RM(r, m)$ with weight equal to $2^{\lfloor \frac{m-1}{r} \rfloor}$ times an odd integer.

Another important result on Reed-Muller codes is the following (already evoked in the introduction):

Theorem 2 (Kasami-Tokura)**.** *[5] Let w be a weight of some nonzero codeword in $RM(r, m)$ in the range $2^{m-r} \leq w < 2^{m-r+1}$. Let $\alpha = \min(r, m - r)$, and $\beta = \frac{m-r+2}{2}$. The weight w is of the form $w = 2^{m-r+1} - 2^{m-r+1-i}$, for i in the range $1 \leq i \leq \max(\alpha, \beta)$. Conversely, for any such i, there is a w of that form in the range $2^{m-r} \leq w < 2^{m-r+1}$.*

This result has been extended in [6] into the characterization of all the weights of $RM(r, m)$ that are in the range $2^{m-r} \leq w < 2^{m-r+1} + 2^{m-r-1}$ (i.e. that lie between the minimum distance of the code

and 2.5 times the minimum distance). It is impossible to summarize these results; we shall refer below to the pages in this reference where the results that we shall need can be found.

Notation: for every n, we denote respectively by $\mathbf{0}_n$ and $\mathbf{1}_n$ the all-0 and all-1 vectors of length n.

3 The weights of the Reed-Muller codes of length 2^m and order $m-5$

It is well-known that we obtain all the codewords in $RM(r,m)$ by concatenating any codeword u of $RM(r,m-1)$ and the sum of u and of a codeword v of $RM(r-1,m-1)$ (this is called the $(u,u+v)$ construction of $RM(r,m)$, see [7]). If we take u also in $RM(r-1,m-1)$, then u and $u+v$ range freely and independently in $RM(r-1,m-1)$. Hence, $RM(r,m)$ contains the concatenations of any two codewords of $RM(r-1,m-1)$ (which can also be seen directly by considering functions of the form $u(\mathbf{x}')+x_m v(\mathbf{x}')$, where u and v are two $(m-1)$-variable Boolean functions of algebraic degrees at most $r-1$ and $\mathbf{x}' \in \mathbb{F}_2^{m-1}$). This implies that the sums of two weights in $RM(r-1,m-1)$ are weights in $RM(r,m)$. This allowed in [4] to determine the weights of $RM(3,6)$ and $RM(4,8)$ and deduce by induction the weights of $RM(m-c,m)$ when $c \leq 4$.

But the weights in $RM(m-5,m)$ could not be determined. This would have needed to determine the weights in $RM(5,10)$. Indeed, determining the weights in the codes $RM(m-c,m)$ for a given $c>0$ needs in practice, for starting an induction, to determine the weights in the code $RM(m-c,m)$ for which m is the smallest such that $\left\lfloor \frac{m-1}{m-c} \right\rfloor$ (in the McEliece divisiblity theorem) has value 1, that is, $m=2c=2r$ (in which case the condition $i \leq \max(\min(m-r,r), \frac{m-r+2}{2})$ of Kasami-Tokura writes $i \leq c$). Taking m smaller than $2c$ allows by computing sums of two weights in $RM(m-c,m)$ to obtain only weights that are divisible by 4 in $RM(m+1-c,m+1)$. And only a half of the weights of $RM(5,10)$ could be determined in [4] (almost all weights that are not divisible by 4 missing).

For the reasons presented above, determining the weights in $RM(r,2r)$ that are divisible by 4 is easier than determining those which are not divisible by 4 (and divisible by 2): many of the former can be obtained by adding two weights from $RM(r-1,2r-1)$ if these weights are known, or from $RM(r-j,2r-j)$ where $j>1$ is the smallest value for which the weights are known. This is how they have been determined in [4] for $RM(5,10)$.
Let us then work on the most difficult part: the weights that are not divisible by 4.

3.1 The weights in $RM(5,10)$ that are congruent with 2 mod 4

Since using a computer for obtaining the weight spectrum of $RM(5,10)$ seems out of reach, we need to mathematically construct Boolean functions in 10 variables and of algebraic degree at most 5, whose Hamming weights can be determined and cover as many values allowed by [6] as possible (and are congruent with 2 mod 4). Of course, we only need to determine the weights up to $2^{m-1}-2$, since Reed-Muller codes being invariant by the complementation of their codewords to the all-one vector, their weight spectra are invariant by complement to 2^m.
We shall use the structure of the so-called Maiorana-McFarland functions (see e.g. [2]). Let m be a positive integer. An m-variable Boolean function is Maiorana-McFarland if there exist $2 \leq k \leq m$, $\phi : \mathbb{F}_2^{m-k} \mapsto \mathbb{F}_2^k$ and $g : \mathbb{F}_2^{m-k} \mapsto \mathbb{F}_2$ such that:

$$f(\mathbf{x},\mathbf{y}) = \mathbf{x} \cdot \phi(\mathbf{y}) + g(\mathbf{y}); \quad \mathbf{x} \in \mathbb{F}_2^k, \quad \mathbf{y} \in \mathbb{F}_2^{m-k},$$

where $(\mathbf{x},\mathbf{y})$ is the concatenation of the vectors $\mathbf{x}=(x_1,\ldots,x_k)$ and $\mathbf{y}=(y_1,\ldots,y_{m-k})$ and "$\cdot$" is an inner product in $\mathbb{F}_2^k$ (for instance the so-called usual inner product $\mathbf{x}\cdot\mathbf{x}' = x_1x_1' + \cdots + x_kx_k'$, where of course $\mathbf{x}'=(x_1',\ldots,x_k')$). We assume $k \geq 2$ because for $k=1$, the corresponding Maiorana-McFarland functions are all m-variable Boolean functions, and the Maiorana-McFarland structure is then weak and does not help the study.

Such function f belongs to $RM(r,m)$ if and only if ϕ has algebraic degree at most $r-1$ (that is, all its coordinate functions have algebraic degree at most $r-1$) and g has algebraic degree at most r.

Considering the value $W_f(\mathbf{0}_k, \mathbf{0}_{m-k})$ of the Walsh transform W_f of function f (see e.g. [2]), we have:

$$2^m - 2w_H(f) = W_f(\mathbf{0}_k, \mathbf{0}_{m-k}) := \sum_{\mathbf{x}\in\mathbb{F}_2^k, \mathbf{y}\in\mathbb{F}_2^{m-k}} (-1)^{\mathbf{x}\cdot\phi(\mathbf{y})+g(\mathbf{y})} =$$

$$\sum_{\mathbf{y}\in\mathbb{F}_2^{m-k}} \left((-1)^{g(\mathbf{y})} \sum_{\mathbf{x}\in\mathbb{F}_2^k} (-1)^{\mathbf{x}\cdot\phi(\mathbf{y})} \right) = 2^k \sum_{\mathbf{y}\in\phi^{-1}(\mathbf{0}_k)} (-1)^{g(\mathbf{y})},$$

where $\phi^{-1}(\mathbf{0}_k)$ denotes the pre-image by ϕ of the zero vector in $\mathbb{F}_2^k$. Hence:

$$w_H(f) = 2^{m-1} - 2^{k-1} \sum_{\mathbf{y}\in\phi^{-1}(\mathbf{0}_k)} (-1)^{g(\mathbf{y})}. \tag{1}$$

We want this number to be congruent with 2 mod 4, which obliges to take $k=2$.

Let ϕ_1, ϕ_2 be the two coordinate functions of ϕ. We have $\phi^{-1}(\mathbf{0}_2) = \{\mathbf{y} \in \mathbb{F}_2^{m-2}; \phi_1(\mathbf{y}) = \phi_2(\mathbf{y}) = 0\}$. The indicator function of $\phi^{-1}(\mathbf{0}_2)$ equals then $(\phi_1(\mathbf{y})+1)(\phi_2(\mathbf{y})+1)$. According to what we recalled in Section 2, a Boolean function in $m-2$ variables has an odd Hamming weight if and only if it has (maximal) algebraic degree $m-2$. Hence, $\phi^{-1}(\mathbf{0}_2)$ has an odd size if and only if $\phi_1\phi_2 + \phi_1 + \phi_2$ has algebraic degree $m-2$.

We fix now $m=10$ and $r=5$ ($c=5$). The fact that $\phi_1\phi_2$ has algebraic degree $m-2=8$ implies that ϕ_1 and ϕ_2 both have algebraic degree 4 exactly.

We wish that $\phi^{-1}(\mathbf{0}_2)$ is as large as possible (then we can try to reach as many weights as possible with f by visiting as many Boolean functions g as possible). For this, we wish that the co-support of ϕ_1 (that is, the complement of its support) is as large as possible. We take then for ϕ_1 a minimum weight codeword in $RM(4,8)$. Up to affine equivalence, we can take $\phi_1(\mathbf{y}) = \prod_{j=1}^4 y_j$ (see [7, 2]). This ϕ_1 being chosen, we want that $\phi_1\phi_2$ has the algebraic degree 8 and that $\phi^{-1}(\mathbf{0}_2)$ has a maximum size. Let us then take $\phi_2(\mathbf{y}) = \prod_{j=5}^8 y_j$.

3.1.1 The weights achievable by f when $m=10$, $k=2$, $\phi_1(\mathbf{y}) = \prod_{j=1}^4 y_j$ and $\phi_2(\mathbf{y}) = \prod_{j=5}^8 y_j$

With such choices, we have:

$$\begin{aligned}\phi^{-1}(\mathbf{0}_2) &= \Big\{\mathbf{y} \in \mathbb{F}_2^8; \prod_{j=1}^4 y_j = \prod_{j=5}^8 y_j = 0\Big\} \\ &= (\mathbb{F}_2^4 \setminus \{\mathbf{1}_4\}) \times (\mathbb{F}_2^4 \setminus \{\mathbf{1}_4\}).\end{aligned}$$

Then, according to (1), denoting by g' the restriction of g to $(\mathbb{F}_2^4 \setminus \{\mathbf{1}_4\})^2$, by g_1 the restriction of g to $\{\mathbf{1}_4\} \times \mathbb{F}_2^4$ and by g_2 the restriction of g to $\mathbb{F}_2^4 \times \{\mathbf{1}_4\}$, we have:

$$\begin{aligned} w_H(f) &= 2^9 - 2 \sum_{\mathbf{y}\in(\mathbb{F}_2^4\setminus\{\mathbf{1}_4\})^2} (-1)^{g(\mathbf{y})} \\ &= 2^9 - 2\Big(15^2 - 2w_H(g')\Big) \\ &= 62 + 4w_H(g') \qquad (2) \\ &= 62 + 4w_H(g) - 4w_H(g_1) - 4w_H(g_2) + 4\,g(\mathbf{1}_8). \end{aligned}$$

The detailed explanations on how we obtained all the possible weights of g' when g belongs to $RM(5,8)$ can be found at URL:
`https://d197for5662m48.cloudfront.net/documents/publicationstatus/171039/preprint_pdf/5e3b1a34b6f649e6b532796b16033485.pdf`

The weights congruent with 2 mod 4 between 62 and 94 Considering the case where g has minimum nonzero weight 8 (i.e. g is the indicator of a 3-dimensional affine space A), and considering all possible cases, we have:

Lemma 3. *Let:*

$$f(\mathbf{x},\mathbf{y}) = x_1 \prod_{j=1}^{4} y_j + x_2 \prod_{j=5}^{8} y_j + g(\mathbf{y}); \quad \mathbf{x} \in \mathbb{F}_2^2, \quad \mathbf{y} \in \mathbb{F}_2^8,$$

where g is any minimum weight codeword in $RM(5,8)$. Then the set of weights of such codewords of $RM(5,10)$ includes $\{62, 74, 78, 82, 86, 90, 94\}$ and covers all the weights in $RM(5,10)$ that are congruent with 2 modulo 4 and between 62 and 94.

The weights congruent with 2 mod 4 between 96 and 126 Choosing now for g a codeword of $RM(5,8)$ having the three weights that come immediately after 8 when visiting the weight spectrum in ascending order, that is $16-4=12$, $16-2=14$ and 16 itself, we obtain:

Lemma 4. *Let f be defined as in Lemma 3, where g is the sum of two minimum weight codewords in $RM(5,8)$. Then the set of weights of such codewords of $RM(5,10)$ includes additionally to Lemma 3, the numbers: $98, 102, 106, 110, 114, 118, 122, 126$, and covers then all the weights in $RM(5,10)$ that are congruent with 2 modulo 4 and which lie between 98 and 126.*

The weights congruent with 2 mod 4 between 130 and 226 We now need to take a function g such that the weight w of g' is between 17 and 41. We have:

Lemma 5. *Let f be defined as in Lemma 3, where g is the sum of three to six minimum weight codewords in $RM(5,8)$ with disjoint supports. Then the set of weights of such codewords of $RM(5,10)$ includes additionally to Lemmas 3 and 4, all the numbers congruent with 2 modulo 4 and lying between 130 and 226.*

All remaining weights congruent with 2 mod 4

Lemma 6. *Let g be the 8-variable Maiorana-McFarland function:*

$$g(\mathbf{z},\mathbf{t}) = \mathbf{z} \cdot \psi(\mathbf{t}) + h(\mathbf{t}); \quad \mathbf{z},\mathbf{t} \in \mathbb{F}_2^4,$$

where ψ is any function from $\mathbb{F}_2^4$ to $\mathbb{F}_2^4$ and h is any Boolean function over $\mathbb{F}_2^4$. Let:

$$f(\mathbf{x},\mathbf{z},\mathbf{t}) = x_1 \prod_{j=1}^{4}(z_j+1) + x_2 \prod_{j=1}^{4}(t_j+1) + g(\mathbf{z},\mathbf{t});$$

$$\mathbf{x} \in \mathbb{F}_2^2, \quad \mathbf{z},\mathbf{t} \in \mathbb{F}_2^4.$$

Then the algebraic degree of any such 10-variable Boolean function f is at most 5 and the set of the weights of such functions includes all those integers between 230 and 510 that are congruent with 2 modulo 4.

3.2 The weight spectrum of $RM(5,10)$

Proposition 7. *The set of all weights in $RM(5,10)$ equals $\{0, 32, 48, 56, 60, 62,$
$64, 68, 72+2i, 2^{10}-68, 2^{10}-64, 2^{10}-62, 2^{10}-60, 2^{10}-56, 2^{10}-48, 2^{10}-32, 2^{10}\}$, where i ranges over the set of consecutive integers from 0 to 2^9-72.*

Proof. The result is deduced from Lemmas 3,4,5,6, the results of [5], and the facts that the spectrum is symmetric with respect to 512 and that, according to [4], all the numbers divisible by 4 between 56 and $2^{10}-56=968$ are weights in $RM(5,10)$. □

3.3 The weight spectrum of every code $RM(m-5,m)$ for $m \geq 10$

Theorem 8. *For every $m \geq 10$, the set of all weights in $RM(m-5,m)$ equals $\{0, 32, 48, 56, 60, 62, 64, 68, 72+2i, 2^m-68, 2^m-64, 2^m-62, 2^m-60, 2^m-56, 2^m-48, 2^m-32, 2^m\}$, where i ranges over the set of consecutive integers from 0 to $2^{m-1}-72$.*

The proof by an induction on $m \geq 10$ is omitted because of length limitation.

Open question: *Let c be any positive integer. For $m \geq 2c$, is the weight spectrum of $RM(m-c,m)$ of the form:*

$$\{0\} \cup A \cup B \cup C \cup \overline{B} \cup \overline{A} \cup \{2^m\}?$$

where:

- $A \subseteq [2^c, 2^{c+1}]$, *is given by Kasami and Tokura [5],*
- $B \subseteq [2^{c+1}, 2^{c+1}+2^{c-1}]$, *is given by Kasami, Tokura, and Azumi in [6, Page 392 and foll.],*
- $C \subseteq [2^{c+1}+2^{c-1}, 2^m-2^{c+1}-2^{c-1}]$, *consists of all consecutive even integers,*
- $\overline{A}$ *stands for the complement to 2^m of A, and $\overline{B}$ stands for the complement to 2^m of B.*

References

[1] Y.L. Borissov, On McEliece's result about divisibility of the weights in the binary Reed-Muller codes, *Seventh International Workshop on Optimal Codes and Related Topics, September 6–12, 2013, Albena, Bulgaria* pp. 47–52. `http://www.moi.math.bas.bg/oc2013/a7.pdf`

[2] C. Carlet. *Boolean Functions for Cryptography and Coding Theory*. Cambridge University Press, 2021.

[3] C. Carlet. The weight spectrum of the Reed-Muller codes $RM(m-5,m)$. To appear in *IEEE Transactions on Information Theory*. 2024. 10.1109/TIT.2023.3343697

[4] C. Carlet and P. Solé. The weight spectrum of two families of Reed-Muller codes. *Discrete Mathematics* 346 (10), 113568, 2023. See also http://arxiv.org/abs/2301.13497.

[5] T. Kasami and N. Tokura. On the weight structure of the Reed-Muller codes, *IEEE Transactions on Information Theory* 16, pp. 752-759, 1970.

[6] T. Kasami, N. Tokura, and S. Azumi. On the Weight Enumeration of Weights Less than $2.5d$ of Reed-Muller Codes. *Information and Control*, 30:380–395, 1976.

[7] F. J. MacWilliams and N. J. Sloane. *The theory of error-correcting codes*, North Holland. 1977.

[8] R. J. McEliece. Weight congruence for p-ary cyclic codes. *Discrete Mathematics*, 3, pp. 177-192, 1972.

[9] List of weight distributions [from The On-Line Encyclopedia of Integer Sequences (OEIS)] https://oeis.org/wiki/List_of_weight_distributions

Separating cycle systems*

Fábio Botler[†1] and Tássio Naia[‡2]

[1]Instituto de Matemática e Estatística, São Paulo, Brazil
[2]Centre de Recerca Matemàtica, 2.7182 Sabadell, Spain

Abstract

A *separating system* of graph G is a family $\mathcal{F}$ of subgraphs of G such that, for all distinct edges $e, f \in E(G)$, some element in $\mathcal{F}$ contains e but not f. Recently, it has been shown that every n-vertex graph admits a separating system of paths of size $O(n)$ [*Separating the edges of a graph by a linear number of paths*, M. Bonamy, F. Botler, F. Dross, T. Naia, J. Skokan. Advances in Combinatorics, October 2023]. This result improved an almost linear bound of $O(n \log^\star n)$ found by Letzter in 2022, and settled a conjecture independently posed by Balogh, Csaba, Martin, and Pluhár and by Falgas-Ravry, Kittipassorn, Korándi, Letzter, and Narayanan. We extend this result, showing that every n-vertex graph admits a separating system consisting of $O(n)$ edges and cycles.

1 Introduction

Given a set Ω and a family $\mathcal{F} \in 2^\Omega$ of subsets of Ω, we say that $\mathcal{F}$ *separates* Ω if for all distinct $\omega, \rho \in \Omega$ there exist $A_\omega, A_\rho \in \mathcal{F}$ such that $A_\omega \cap \{\omega, \rho\} = \{\omega\}$ and $A_\rho \cap \{\omega, \rho\} = \{\rho\}$. The study of separating systems dates back to the work of Rényi in 1961 [9]. The particular setting where $\Omega = E(G)$ is the edge set of a graph G and only certain subgraphs are allowed in $\mathcal{F}$ has also been investigated multiple times in the Computer Science literature, in the context of fault detection in networks (see, e.g., [5, 6] and the references therein). A generic problem in the area is the following.

Question 1. Let $\mathcal{G}$ be a (possibly infinite) family of graphs, and let H be an n-vertex graph. What is the smallest size of a collection $\mathcal{F} \subseteq \mathcal{G}$ of H-subgraphs such that $\{E(H) : H \in \mathcal{F}\}$ separates $E(H)$?

A *separating system* of a graph G is a collection of G-subgraphs such that their edge sets separate $E(G)$. Recently, Bonamy, Dross, Skokan and the two authors showed that every n-vertex graph admits a separating system consisting of at most $19n$ paths [2], improving a previous bound of $O(n \log^\star n)$ found by Letzter in 2022 [7], and settling a conjecture independently posed by Balogh, Csaba, Martin, and Pluhár [1] and by Falgas-Ravry, Kittipassorn, Korándi, Letzter, and Narayanan [4].

A natural follow-up question is to ask whether every graph G admits a *cycle separating system* of size $O(|V(G)|)$, that is, a collection of cycles and edges of G which separate $E(G)$. (Note that cycles alone are not enough in general, since G might contain a cycle-free component.) This question was independently posed by Girão and Pavez-Signé[1]. Here we answer their question in the affirmative.

*This research has been partially supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil – CAPES – Finance Code 001. This work is also supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M). F. Botler was partially supported by CNPq (Proc. 304315/2022-2), FAPERJ (Proc. 201.334/2022), and CAPES (Proc. 88881.878881/2023-01). T. Naia was partially supported by the Grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

[†]Email: fbotler@ime.usp.br
[‡]Email: tnaia@member.fsf.org

[1]Personal communication.

Theorem 2. *Every graph on n vertices admits a separating cycle system of size $41n$.*

Note that any cycle separating system of K_n contains at least $(n-1)/2$ elements, since each of the $\binom{n}{2}$ edges must be covered and any cycle contains at most n edges, so the bound in Theorem 2 is optimal apart from the leading constant. We do not believe that 41 is the correct multiplicative constant, and we wonder whether every graph of order n admits a separating cycle system of order $n + o(n)$.

Our proof uses a combination of properties of Pósa's rotation-extension method, a covering result due to Pyber, combined with algebraically-constructed edge covers of Hamiltonian graphs.

1.1 Pósa rotation-extension.

Given a graph G and $S \subseteq V(G)$, we denote by $N_G(S)$ the set of vertices in $V(G) \setminus S$ which are adjacent (in G) to some vertex in S. We omit subscripts when clear from the context. Let $P = u \cdots v$ be a path from u to v. If $x \in V(P)$ is a neighbor of u in G and x^- is the vertex preceding x in P, then $P' = P - xx^- + ux$ is a path in G for which $V(P') = V(P)$. We say that P' *has been obtained from* P by an *elementary exchange* fixing v (see Figure 1). A path obtained from P by a (possibly empty) sequence of elementary exchanges fixing v is said to be a path *derived* from P. The set of endvertices of paths derived from P distinct from v is denoted by $S_v(P)$. Since all paths derived from P have the same vertex set as P, we have $S_v(P) \subseteq V(P)$. When P is a longest path ending at v, we obtain the following (for a proof see [2]).

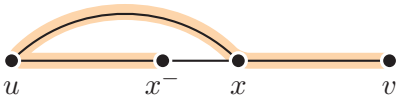


Figure 1: a path (highlighted) obtained by an elementary exchange fixing v.

Lemma 3 ([3])**.** *If $P = u \cdots v$ is a longest path of a graph G, then $\left|N_G(S)\right| \leqslant 2\left|S_v(P)\right|$.*

We also use the following property of Pósa rotations.

Lemma 4. *If $P = u \cdots v$ be a longest path of a graph G and $S = S_v(P)$, then G contains a subgraph C which is either and edge or a cycle and moreover $S \cup N(S) \subseteq C$.*

Proof. Consider the vertex $z \in V(P) \cap N(S)$ which lies closest to v in P, and let $P' = u' \cdots v$ be a path obtained from P by elementary exchanges fixing v so that P' starts with a neighbor u' of z. Note that C is an edge when $S = \{u\}$ and $|N(S)| = 1$. Since the section $P[z, v]$ of P from z to v intersects $S \cup N(S)$ precisely in v, and $P' \cup u'z$ has at most one cycle, we conclude that $C = (P' + uv) \setminus E(P[w, v])$ is either an edge or a cycle that contains $S \cup N(S)$ □

2 Separating into cycles

For the sake of clarity, we make no attempt to optimize multiplicative constants in the argument. This allows us to better highlight its main ideas. It also seems unlikely that the optimal multiplicative constant can be reached by this approach alone.

The following theorem of Pyber is useful in our proof.

Theorem 5 (Pyber [8])**.** *Every graph G contains $|V(G)| - 1$ cycles and edges covering $E(G)$.*

Given a graph G, a collection $\mathcal{J}$ of subgraphs of G, and $e, f \in E(G)$, we say that $\mathcal{J}$ *separates e from f* if there exists $J \in \mathcal{J}$ such that $E(J) \cap \{e, f\} = \{e\}$. Similarly, given $\mathcal{E}, \mathcal{F} \subseteq E(G)$, we say that $\mathcal{J}$ *separates $\mathcal{E}$ from $\mathcal{F}$* if $\mathcal{J}$ separates e from f for all distinct $e \in \mathcal{E}$ and $f \in \mathcal{F}$.

Proof of Theorem 2. We proceed by induction on n. Let G be a graph with n vertices. If G is empty, the result trivially holds. Let $P = u \cdots v$ be a longest path of G and let $S = S_v(P)$. By Lemma 4, there exists $C \subseteq G$ which is either an edge or a cycle and which contains $S \cup N(S)$.

Let H be the subgraph of G induced by the edges incident to at least a vertex in S, let $h = |V(H)|$, and let $G' = G \setminus S$. Then $G = H \cup G'$ and $V(H) = S \cup N(S)$, so $h \leqslant 3|S|$ by Lemma 3.

Note that S is not empty (because G is not empty). By the induction hypothesis, there is a cycle separating system $\mathcal{C}'$ of G' of size at most $41(n - |S|)$. Note that $\mathcal{C}'$ separates $E(G')$ from $E(H)$. In what follows, we construct a set $\mathcal{C}$ of at most $41|S|$ edges and cycles which separates $E(H)$ from $E(G)$, i.e., separates edges in H and also separates $E(H)$ from $E(G')$. This set $\mathcal{C}$ is the union of three collections of cycles and edges ($\mathcal{D}, \mathcal{E}$ and $\mathcal{H}$) which we next define.

Let $\mathcal{D}$ be a collection of at most $h-1 \leqslant 3|S|-1$ edges and cycles which covers $E(H) \setminus E(C)$ (such $\mathcal{D}$ exists by Lemma 5), and let $\mathcal{E} = E(C) \cap E(H)$ be the collection of edges of C which contain a vertex in S. Note that $|\mathcal{E}| \leqslant 2|S|$, and that $\mathcal{E}$ separates the edges of $E(C) \cap E(H)$ among themselves and from all other edges of G. Moreover, $\mathcal{D}$ separates the edges of $E(H) \setminus E(C)$ from all other edges. The final component of $\mathcal{C}$ will separate the edges of $E(H) \setminus E(C)$ from one another.

Note that every edge in $E(H) \setminus E(C)$ has both endvertices in $V(H) = S \cup N(S)$. Let $v_1, \ldots, v_h$ denote the vertices in $V(H)$, labeled following the cyclic order in which they appear in C. From this point onward, whenever we refer to an edge v_iv_j, we will always assume that $i < j$. We say that edges v_iv_j and v_rv_s *cross* each other if either $i < r < j < s$ or $r < i < s < j$. For given integers k and ℓ, consider the two matchings

$$M_k = \{v_iv_j \in E(H) \setminus E(C) : j - \ \ i = k\}$$
$$N_\ell = \{v_iv_j \in E(H) \setminus E(C) : j - 2i = \ell\}$$

Note that at most $3h \leqslant 9|S|$ of these matchings are nonempty, because M_k is empty whenever $k < 2$ or $k > h - 1$, and N_ℓ is empty if $\ell < -h + 2$ or $\ell > h - 2$. We claim that the nonempty matchings separate the edges in $E(H) \setminus E(C)$. Pick two edges v_iv_j and v_rv_s. If $j - i \neq s - r$, then M_{j-i} separates v_iv_j from v_rv_s and, moreover, M_{s-r} separates v_rv_s from v_iv_j. Similarly, if $j - 2i \neq s - 2r$, then N_{j-2i} separates v_iv_j from v_rv_s and N_{s-2r} separates v_rv_s from v_iv_j. Finally, it is easy to check that $j-i = s-r$ and $i - 2j = s - 2r$ if and only if $i = r$ and $j = s$, that is, if and only if $v_iv_j = v_rv_s$. We conclude that every pair of distinct edges in $E(H) \setminus E(C)$ is separated by these matchings.

To construct $\mathcal{H}$ we shall cover each nonempty M_k (respectively, N_ℓ) using at most 4 cycles in $M_k \cup C$ (respectively, $N_\ell \cup C$) each. A trivial yet crucial observation we shall use here is that if $M \subseteq M_k$ (or $M \subseteq N_\ell$) is a set of pairwise crossing edges and $|M|$ is odd, then $M \cup C$ contains a cycle which covers M (see Figure 2). More generally, if M admits a partition $\bigcup_\alpha S_\alpha^{(M)}$ such that

(i) each $S_\alpha^{(M)}$ is formed by odd number of pairwise crossing edges, and

(ii) each pair of distinct edges $v_iv_j \in S_\alpha^{(M)}$ and $v_rv_s \in S_\beta^{(M)}$ cross if and only if $\alpha = \beta$,

then $M \cup C$ contains a cycle which covers M (see Figure 2).

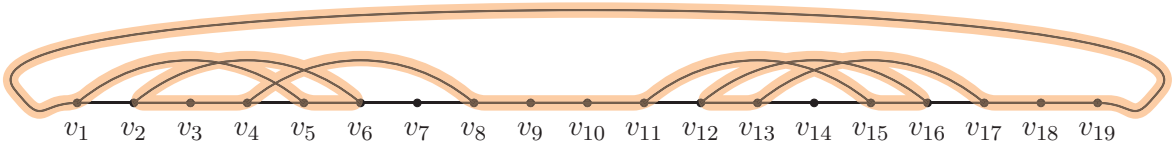


Figure 2: A cycle covering an odd number of pairwise crossing edges (v_2v_9, v_4v_{11} and v_8v_{15}).

Indeed, it turns out that each of the matchings M_k and N_ℓ admits a 4-piece partition such that each part satisfies both *(i)* and *(ii)*. Consequently, each nonempty M_k and each N_ℓ can be covered by at

most four cycles using only edges in the matching and in C. The required partitions will be obtained by splitting each matching into two, twice. Given positive integers u, let $f(u)$ be the largest integer such that $2^q(\ell+1)-\ell \leqslant u$. For all k, all ℓ and all $\pi \in \{0,1\}$, put

$$M_{k,\pi} = \{v_iv_j \in E(H)\setminus E(C) : \lfloor i/k \rfloor \equiv \pi \pmod 2\}$$
$$N_{\ell,\pi} = \{v_iv_j \in E(H)\setminus E(C) : \ f(i) \equiv \pi \pmod 2\}.$$

Let M be an arbitrary $M_{k,\pi}$ or $N_{\ell,\pi}$. We claim that M admits a partition $\bigcup_\alpha S_\alpha^{(M)}$ such that distinct edges of M cross if and only if they belong to the same part.

Proof of claim (partition of $M_{k,\pi}$). We begin with the case $M = M_{k,\pi}$. Let v_iv_j and v_rv_s be distinct edges in $M_{k,\pi}$. Without loss of generality, we assume $i < r$. By definition, $j = i+k$ and $s = r+k$. Since $\lfloor i/k \rfloor$ and $\lfloor r/k \rfloor$ have the same parity, then either $\lfloor i/k \rfloor = \lfloor r/k \rfloor$ or $\lfloor r/k \rfloor - \lfloor i/k \rfloor \geqslant 2$. In the former case, we have that

$$r < \left(\left\lfloor \frac{r}{k} \right\rfloor + 1\right)k = \left(\left\lfloor \frac{i}{k} \right\rfloor + 1\right)k = i+k = j,$$

so v_iv_j and v_rv_s cross, while in the latter they do not, since

$$j = i+k < \big(\lfloor i/k \rfloor + 1\big)k + k \leqslant \big(\lfloor r/k \rfloor - 1\big)k + k \leqslant r.$$

Hence the crossing relation defines equivalence classes among the edges in $M_{k,\pi}$, and thus the a partition of M satisfying *(ii)* exists.

Proof of claim (partition of $N_{\ell,\pi}$). The case $M = N_{\ell,\pi}$ is similar. Consider distinct $v_{i'}v_{j'}$ and $v_{r'}v_{s'}$ in $N_{\ell,\pi}$, where without loss of generality we assume $i' < r'$. By definition, either $f(i') = f(r')$ or $f(r') - f(i') \geqslant 2$. In the former case $v_{i'}v_{j'}$ and $v_{r'}v_{s'}$ must cross, since

$$r' \leqslant 2^{f(r')+1}(\ell+1) - \ell = 2^{f(i')+1}(\ell+1) - \ell \leqslant 2\big(2^{f(i')}(\ell+1) - \ell\big) + \ell \leqslant 2i' + \ell = j'.$$

On the other hand, if $f(r') - f(i') \geqslant 2$, then

$$j' = 2i' + \ell \leqslant 2\big(2^{f(i')+1}(\ell+1) - \ell\big) + \ell \leqslant 2^{f(i')+2}(\ell+1) - \ell \leqslant 2^{f(r')+2}(\ell+1) - \ell \leqslant r',$$

and consequently $v_{i'}v_{j'}$ and $v_{r'}v_{s'}$ do not cross. As before we conclude that M admits a partition $\bigcup_\alpha S_\alpha^{(M)}$ which satisfies *(ii)*.

Returning to the proof of the theorem, we complete our partitioning by refining each one of the nonempty matchings $M_{k,\pi}$ and $N_{\ell,\pi}$ (for each k, ℓ and π) further into two pieces each, so that any matching after the refinement also satisfies *(i)* (note that partition refinement does not break *(ii)*). More precisely, by *(ii)*, $M_{k,\pi}$ has a natural partition $\bigcup_\alpha S_\alpha^{(M_{k,\pi})}$ into equivalence classes such that edges in the same class are pairwise crossing and edges in distinct classes do not cross. Form $M^1_{k,\pi}$ by selecting arbitrarily one edge from each even-sized equivalence class, and let $M^2_{k,\pi} = M_{k,\pi} \setminus M^1_{k,\pi}$ be the remaining edges of M_k (i.e., $M^1_{k,\pi}$ contains at most one edge from each equivalence class, and $M^2_{k,\pi}$ contains an odd number of edges from each equivalence class). We use the same criterion for partitioning $N_{\ell,\pi}$ into $N^1_{\ell,\pi} \cup N^2_{\ell,\pi}$.

Note that each part resulting from this refinement satisfies both *(i)* and *(ii)*. It follows that there exists a collection $\mathcal{H}$ of at most $4 \cdot 9|S| = 36|S|$ cycles such that each nonempty M_k (respectively, N_ℓ) is covered by a at most 4 cycles in $M_k \cup E(C)$ (respectively, $N_\ell \cup E(C)$), as desired.

Note that $\mathcal{H}$ separates the edges in $E(H)\setminus E(C)$ from $E(G)$, and contains at most $36|S|$ elements. Since $|\mathcal{E}| \leqslant 2|S|$ and $|\mathcal{D}| \leqslant 3|S|$, we have that $\mathcal{C} = \mathcal{D}\cup\mathcal{E}\cup\mathcal{H}$ has at most $41|S|$ edges and paths. Hence, $\mathcal{C}' \cup \mathcal{C}$ is a cycle separating system of G with cardinality at most $41(n-|S|) + 41|S| = 41n$ as desired. This completes the proof. □

3 Concluding remarks

In this article, we have shown that every n-vertex graph admits a separating system consisting of $O(n)$ edges and cycles (which we call cycle separating systems for short). This is, in at least two ways, a natural extension of previous results about the existence of path separating systems. On the one hand, a cycle separating system immediately yields a path separating system (obtained by breaking each cycle into two paths). On the hand, since paths and cycles are, respectively, subdivisions of K_2 and K_3, the following question immediately suggests itself.

Question 6. Is it true that for every natural $t \geqslant 2$, every n-vertex graph admits a separating system consisting of $O_t(n)$ edges and subdivisions of K_t?

Note that edges are necessary in the separating systems in Question 6, because a union of disjoint K_{t-1} cliques has linearly many edges and no K_t subdivision. This follows in more generality from a classical result of Mader, stating that for every t there exists $f(t)$ such that every graph free of a K_t subdivision has average degree at most $f(t)$.

Our Theorem 2 and the results in [2] confirm the conjecture for $t \leqslant 3$. In a forthcoming article, the authors extend this for $t = 4$ as well, but to the best of our knowledge no further cases have been settled.

References

[1] J. Balogh, B. Csaba, R. R. Martin, and A. Pluhár. On the path separation number of graphs. *Discrete Appl. Math.*, **213** (2016), 26–33.

[2] M. Bonamy, F. Botler, F. Dross, T. Naia, and J. Skokan. Separating the edges of a graph by a linear number of paths. *Advances in Combinatorics* (2023), October.

[3] S. Brandt, H. Broersma, R. Diestel, and M. Kriesell. Global connectivity and expansion: long cycles and factors in f-connected graphs. *Combinatorica*, **26** (2006), 17–36.

[4] V. Falgas-Ravry, T. Kittipassorn, D. Korándi, S. Letzter, and B. P. Narayanan. Separating path systems. *J. Comb.*, **5** (2014), 335–354.

[5] C. G. Fernandes, G. O. Mota, and N. Sanhueza-Matamala. Separating path systems in complete graphs. In *Latin American Symposium on Theoretical Informatics, Springer* (2024), 98–113.

[6] G. Kontogeorgiou and M. Stein. An exact upper bound for the minimum size of a path system that weakly separates a clique. *arXiv preprint 2403.08210*.

[7] S. Letzter. Separating paths systems of almost linear size, 2022. *Trans. Amer. Math. Soc.* (to appear).

[8] L. Pyber. An Erdős–Gallai conjecture. *Combinatorica*, **5** (1985), 67–79.

[9] A. Rényi. On random generating elements of a finite boolean algebra. *Acta Sci. Math. Szeged*, **22** (1961), 75–81.

The algorithmic Fried Potato Problem in two dimensions*

Francisco Criado†[1] and Francisco Santos‡[2]

[1]Departamento de Métodos Cuantitativos, CUNEF Universidad, Madrid, Spain
[2]Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, 39005 Santander, Spain

Abstract

Conway's Fried Potato Problem seeks to determine the best way to cut a convex body in n parts by $n-1$ hyperplane cuts (with the restriction that the i-th cut only divides in two one of the parts obtained so far), in a way as to minimize the maxuimum of the inradii of the parts. It was shown by Bezdek and Bezdek that the solution is always attained by $n-1$ parallel cuts. But the algorithmic problem of finding the best direction for these parallel cuts remains.

In this note we show that for a convex m-gon P, this direction (and hence the cuts themselves) can be found in time $O(m \log^4 m)$, which improves on a quadratic algorithm proposed by Cañete-Fernández-Márquez (DMD 2022). Our algorithm first preprocesses what we call the *dome* (closely related to the medial axis) of P using a variant of the Dobkin-Kirkpatrick hierarchy, so that linear programs in the dome and in slices of it can be solved in polylogarithmic time.

1 From fried potatoes to baker's potatoes

Conway's fried potato problem is stated in [2] (problem C1) as follows: "In order to fry it as expeditiously as possible Conway wishes to slice a given convex potato into n pieces by $n-1$ successive plane cuts (just one piece being divided by each cut) so as to minimize the greatest inradius of the pieces."

The problem was solved by A. Bezdek and K. Bezdek [1] who showed that, no matter what convex potato you start with, the best solution is to cut it with $n-1$ parallel and equally spaced hyperplanes. Let us formalize this a little bit:

Definition 1. *Let $C \subseteq \mathbb{R}^d$ be a convex body (that is, a compact convex subset with nonempty interior).*

1. *The* directional width *of $C \subseteq \mathbb{R}^d$ in a direction $v \in \mathbb{S}^{d-1}$ is the distance between two parallel supporting hyperplanes of C with normal vector v:*

$$\mathrm{width}_v(C) = \max_{x \in C} v^T x - \min_{x \in C} v^T x.$$

The width *of C is its minimum directional width:*

$$\mathrm{width}(C) = \min_{v \in \mathbb{S}^{d-1}} \mathrm{width}_v(C).$$

*This research is supported by Grants PID2019-106188GB-I00 and PID2022-137283NB-C21 of MCIN/AEI/10.13039/501100011033 / FEDER, UE and by project CLaPPo (21.SI03.64658) of Universidad de Cantabria and Banco Santander.

†Email: francisco.criado@cunef.edu
‡Email: francisco.santos@unican.es

2. *The* inner parallel body *of C at a distance $\rho \geq 0$ [8, p. 134] is the set of points of C that are centers of balls of radius ρ contained in C.*

$$\mathrm{inn}_\rho(C) = \{x \in C : B(x, \rho) \subseteq C\}.$$

The ρ-rounded body C^ρ is the union of all closed ρ-balls contained in C:

$$C^\rho = \bigcup_{x \in \mathrm{inn}_\rho(C)} B(x, \rho).$$

3. *The inradius $I(C)$ of C is the maximum radius of a ball contained in C. Equivalently, it is the maximum ρ for which* $\mathrm{inn}_\rho(C) \neq \emptyset$.

Observe that $C^\rho = \mathrm{inn}_\rho(C) + B(0, \rho)$. Also if $C = \{x \in \mathbb{R}^d : Ax \leq b\}$ is a polyhedron with $||A_i|| = 1$ for each i, then $\mathrm{inn}_\rho(C) = \{x \in \mathbb{R}^d : Ax \leq b - \rho\}$, where $b - \rho$ is shorthand for $(b_1 - \rho, \ldots, b_m - \rho)$.

The statement and solution of Conways's fried potato problem can now be stated as follows:

Theorem 2 (Bezdek-Bezdek [1])**.** *Let C be a convex body in $\mathbb{R}^d$ and $n \in \mathbb{N}$. Let P be a division of C into n subsets $C_1, \ldots, C_n$ given by $n-1$ successive hyperplane cuts. These cuts of P do not extend beyond previously made cuts, therefore $(n-1)$ cuts produce n pieces.*

Then,

$$\max_{i \in [n]} I(C_i) \geq \rho,$$

where $\rho > 0$ is the unique number satisfying

$$\mathrm{width}(C^\rho) = 2n\rho. \tag{1}$$

Furthermore, equality holds for the division of C given by $n-1$ parallel and equally spaced hyperplanes normal to the direction attaining $\mathrm{width}(C^\rho)$.

The solution to the fried potato problem raises the algorithmic question of how to find ρ, v and the cuts in the statement. We suggest calling this the *baker's potato problem.*[1]

Clearly, the difficult part is to find ρ and the direction $v \in \mathbb{S}^{d-1}$ such that $\mathrm{width}_v(C^\rho) = \mathrm{width}(C^\rho)$. Cañete, Fernández and Márquez [3, 4] have proposed a quadratic algorithm to do this for a convex polygon in the plane. We here describe a quasi-linear one:

Theorem 3. *Let $P = \{x \in \mathbb{R}^2 : Ax \leq b\}$ be a polygon with non-empty interior, where $A \in \mathbb{R}_{m \times 2}$ and $b \in \mathbb{R}^m$. We can compute the ρ of Theorem 2 and a direction $v \in \mathbb{S}^2$ satisfying* $\mathrm{width}_v(P^\rho) = 2n\rho$ *in $O(m \log^4 m)$ time.*

2 The Dobkin-Kirkpatrick hyerarchy

Equation 1 suggests to formalize the Baker's potato problem adding one dimension to it. If, for a given convex body $C \in \mathbb{R}^d$, we define

$$\overline{C} = \{(x, t) \in \mathbb{R}^d \times [0, \infty) : x \in C^t\} \subset \mathbb{R}^{d+1}\},$$

the problem to solve is to find the ρ such that $\mathrm{width}(\overline{C} \cap \{t = \rho\}) = 2n\rho$.

We solve this using the Dobkin-Kirkpartrick hierarchy, which allows to do linear programming queries in a 3-dimensional polytope in logarithmic type per query. The classical version (which we do not use but state for completeness) is the following statement in which an *extreme-point query* in a set S of m points has as input a linear functional $c \in \mathbb{R}^3$ and as output the point p (or one of the points) of S maximizing $c^T p$.

[1] Baker's potatoes (*pommes boulangère* in French and *patatas panaderas* in Spanish) are potatoes cut in parallel slices of 2–3 mm. and cooked in the oven.

Theorem 4 (Dobkin-Kirkpatrick Hierarchy [5, 6], see also [7, Theorem 7.10.4])**.** *After $O(m \log^2 m)$ time and space preprocessing, extreme-point queries in 3 dimensions can be solved in $O(\log m)$ time each.*

The version we need works in the dual. In what follows we will assume that the facet hyperplanes in our polytopes are generic, that is, no $d+1$ of them have a common point. This implies the polytopes to be simple and is not a loss of generality since it can be achieved by a symbolic perturbation of the input matrix.

Definition 5. *Let $A \in \mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$ be the half-space description of a polytope P in $\mathbb{R}^d$. We call* Dobkin-Kirkpatrick hierarchy *on P a data structure consisting of:*

1. *The face poset of P, in which each face F of codimension k is represented by the subset of size k of $[m]$ consisting of facets containing F.*
2. *A stratification of the set $[m]$ as*

$$[m] = I_0 \supset I_1 \supset \cdots \supset I_k$$

   with the property that the facets labelled by each $I_l \setminus I_{l+1}$ are independent (i.e., mutually non-adjacent) in the polytope P_l defined by the inequalities I_l.
3. *For each vertex x of each P_{l+1} the following information: either the fact that x is still a vertex in P_l or the label of the unique facet inequality of P_l that is violated at x.*

We call k the depth *of the hierarchy and $|I_k|$ the* core size*.*

Observe that in part (3) uniqueness of the facet follows from the fact that the facets labelled by $I_l \setminus I_{l+1}$ are independent in P_l.

Lemma 6. *Let $P = \{x \in \mathbb{R}^3 : Ax \leq b\}$ be a bounded 3-polyhedron defined by $A \in \mathbb{R}^{m \times 3}$, $b \in \mathbb{R}^m$. Then, a Dobkin-Kirkpatrick hierarchy on P of depth $O(m \log m)$ and base size $O(1)$ can be computed in time $O(\log m)$.*

Proof. First, it is well-known that the face poset of a 3-polytope can be fully computed in the way we require in time $O(m \log m)$.

Let $I = [m]$ be the row indices of A. Let I' be a subset of $[m]$ of size at most six and that defines a bounded polyhedron. This is, $\{x \in \mathbb{R}^3 : A_i x \leq b_i \ \ \forall i \in I'\}$ is a bounded set. [2]

We now define the subsets $I = I_0, I_1, \ldots, I_k$ of I in the following recursive manner: Given I_l, we compute the face poset of the polyhedron P_l defined by the rows of $Ax \leq b$ with indices in I_l. We then compute a coloring of the facets of P_l with at most 6 colors, which can be done in linear time because the dual graph of P_l is planar, so that the graph and all its subgraphs contain vertices of degree at most 5.

We choose a color $C \subset I_l$ with

$$|C \cap (I_l \setminus I')| \geq \frac{1}{6}|I_l \setminus I'|$$

and let $I_{l+1} = I_l \setminus C$. Eventually we reach an I_k with $I_k = I'$, hence $|I_k| \leq 6 = O(1)$. Since each time we remove at least 1/6th of the original inequalities (not in I'), $|I_l| \leq |I'| + \lceil \left(\frac{5}{6}\right)^l |I_0 \setminus I'| \rceil$. Thus, we

[2] Such an I', of size at most $2d$, can be found in any facet-described d-polytope as follows: by inductive hypothesis assume that you know how to find such facets in polytopes of dimension smaller than d. To find them for P, start with any facet of P, say I_1 and solve the linear program $\min A_1^T x$ on P. If the program has a unique minimum (a vertex) then let I' be the original facet plus the d containing that vertex. If the program is minimized at a face F of dimension $0 < d' < d$, then let I' equal the original facet plus the $d - d'$ containing F plus the at most $2d'$ that you can find by recursion. This gives $1 + d + d' \leq 2d$.

To find the facets, in the worst case you need to solve d linear programs in dimension $\leq d$, which can be done on $O(m)$ time (with a hidden constant depending on d).

have at most $\log_{5/6}(m)+1$ steps in the hierarchy, so $k \in O(\log m)$. The whole computation needs time proportional to

$$\begin{aligned}\sum_{l=0}^{k} |I_l| \log(|I_l|) \leq \sum_{l=0}^{k} |I_l| \log(m) \leq \\ \leq \log(m) \sum_{l=0}^{k} |I_l| \leq \\ \leq \log(m) \left(6 + \sum_{l=0}^{k} (5/6)^l m\right) \leq \\ \leq \log(m)\,(6+6m)) \leq O(m \log m).\end{aligned}$$

At each step we can easily identify which vertices appear and disappear, and what facet of I_l is violated at each new vertex. When doing this each facet is only considered at one of the levels (the one in which it dissappears) and the total number of vertices in all layers is linear, since the sizes of the polytopes in the layers are bounded by a geometric sequence of ratio 5/6. □

Lemma 7. *Let P be a facet-described polytope and P' be a section of P obtained by intersecting with a linear system of independent inequalities. Then, any Dobkin-Kirkpatrick hierarchy on P is also a Dobkin-Kirkpatrick hierarchy on P'.*

Proof. By induction on $\dim P - \dim P'$ it is enough to consider the case $\dim P - \dim P' = 1$, so that P' is obtained from P by adding one inequality, that is, intersecting with a hyperplane H.

The first observation is that of a set of facets are mutually non-adjacent on P_l then they are also mutually non-adjacent on $P'_l := P_l \cap H$, so the stratification of $[m]$ in the hyerarchy of P works also in P'. We need only to show how the hierarchy on P allows to find the information of which facets remove which vertices on P'. For this, observe that a vertex x of a P'_l is an edge of the corresponding P_l. Let u and v be the end-points of that edge. Then, x can only be eliminated by the facet of P'_{l+1} that eliminates one (or both) of u and v in P_{l+1}, and this can be checked in logarithmic time (the time needed to find the vertices u and v). □

Theorem 8. *Let P be a d-polytope with m facets and suppose that we have a Dobkin-Kirkpatrick hyerarchy on P of depth k and base $O(1)$. Then, a linear program on P can be solved in time $O(k^d \log m)$.*

Proof. In order to solve the linear program with objective function $c^T x$ we traverse the hierarchy in reverse. In the last polytope P_k we need constant time since it has at most $O(1)$ facets. Once we have the maximizer x^*_{l+1} in P_{l+1} we find the maximizer in P_l as follows: if x^*_{l+1} is in P_l (that is, if it satisfies the inequalities with indices in $I_l \setminus I_{l+1}$) then we set $x^*_l = x^*_{l+1}$.

If x^*_{l+1} is not in P_l then by construction of the hierarchy, there is a unique inequality in $I_l \setminus I_{l+1}$ violated by x^*_{l+1} (this is because no two facets indexed by $I_l \setminus I_{l+1}$ are adjacent in P_l). We solve the linear program on that facet to find x^*_{l+1}. By inductive hypothesis this step requires $O(k^{d-1} \log m)$, and we need to do this at most k times. □

Corollary 9. *Any facet described 3-polytope with m facets can be preprocessed in time $O(m \log m)$ so that linear programs on P can be solved in time $O(\log^4 m)$ and linear programs on planar sections of it in time $O(\log^3 m)$*

3 Proof of Theorem 3

Without loss of generality let us assume $||A_i|| = 1$ for every i. We also assume that the given description of P is irredundant (every row of A is a facet), which we can check with a (dual) convex hull computation in $O(m \log m)$ time.

We want to compute the value $\rho > 0$ for which

$$\text{width}(P^\rho) = 2n\rho.$$

Since $P^\rho = \text{inn}_\rho(P) + B(0,\rho)$, we have that

$$\text{width}(P^\rho) = \text{width}(\text{inn}_\rho(P)) + 2\rho.$$

Hence, by definition of width, the ρ and v we are looking for must satisfy:

$$\min_{v\in\mathbb{S}^2} \left(\text{width}_v\left(\text{inn}_\rho(P)\right) - 2(n-1)\rho\right) = 0.$$

The direction minimizing width in the polygon $\text{inn}_\rho(P)$ is normal to an edge of $\text{inn}_\rho(P)$,[3] hence to an edge of P. Thus, we do not need to check for all v, only those normal to edges of P. We use the outwards normals without loss of generality; that is, v must be a row of A.

Now let $r > 0$ be the inradius of P; observe that $0 < \rho < r$. For each $i \in [m]$, let $f_i : [0, r] \to \mathbb{R}$ be defined as

$$f_i(t) = \text{width}_{A_i}\left(\text{inn}_t(P)\right) - 2(n-1)t,$$

so that the equation that characterizes ρ is:

$$\min_{i\in[m]} f_i(\rho) = 0. \tag{2}$$

Each f_i is well defined (as $\text{inn}_t(P)$ is not empty for $0 \le t < r$), continuous, piece-wise linear, and monotonically decreasing. At $t = 0$ every f_i is positive and at $t = r$ some f_i is negative because the width of $\text{inn}_r(P)$ is 0 in some direction.

Then, (2) implies that ρ is exactly:

$$\rho = \min\{0 < t < r : \exists i \in [m] : f_i(t) = 0\}.$$

Indeed, some f_i is guaranteed to have a root by continuity, and ρ is a root for some f_i. If ρ were not the minimum root, then some other root is smaller and by the f_i being strictly decreasing, some f_i is negative at ρ.

So, in order to find ρ we need only to compute the minimum of the roots of the f_i. This is not trivial, since the definition of each f_i is quite implicit. However we need not verify all of them. For each $i \in [m]$, let M_i be the maximum t such that $A_i x \le b - t$ still defines an edge of $\text{inn}_t(P)$. Then, for $t > M_i$ the minimum width of $\text{inn}_t(P)$ cannot be attained at the direction A_i, since it needs to be attained at the normal to an edge of $\text{inn}_t(P)$. Thus,

$$\rho = \min\{0 < t < r : \exists i \in [m] : f_i(t) = 0, t_i \le M_i\}.$$

Equivalently, by continuity and monotonicity,

$$\rho = \min\{0 < t < r : \exists i \in [m] : f_i(t) = 0, f_i(M_i) \le 0\}.$$

We claim that (after preprocessing), we can compute each M_i in polylogarithmic time. For this, consider the following three-dimensional polytope that we call the *dome* of P:

$$\widetilde{P} := \{(x, y, t) \in \mathbb{R}^3 : Ax \le b - t, t \ge 0\}.$$

That is, $\widetilde{P} \cap \{t = 0\}$ equals P and, apart from the horizontal facet, $\widetilde{P}$ has a facet with normal $(A_i, 1) \in \mathbb{R}^3$ for each $i \in [m]$. Assume that we have preprocessed $\widetilde{P}$ as required in Corollary 9. The

[3] A version of this is true in any dimension: by the Karush-Kuhn-Tucker conditions, the v minimizing width in any polytope must be the common normal to two faces of P with sum of dimensions $\ge d - 1$.

value M_i equals the maximum of the t coordinate of the i-th facet of $\widetilde{P}$ and, by the corollary, it can be computed in time $O(\log^3 m)$.

Thus, we can compute all the M_i in time $O(m \log^3 m)$. Once this is done we evaluate all $f_i(M_i)$ also in total time $O(m \log^3 m)$ by solving the linear programs with objective functions A_i in the horizontal slices $P_i := \widetilde{P} \cap \{t = M_i\}$ of the dome.

The motivation for adding the restrictions $t_i \leq M_i$ is that in this range the functions f_i are easier to compute. Recall that

$$f_i(t) = \max_{x:Ax \leq b-t} A_i^T x - \min_{x:Ax \leq b-t} A_i^T x - (2n-2)t.$$

Now, in the range $0 \leq t \leq M_i$ we have

$$\max_{x:Ax \leq b-t} A_i^T x = b_i - t,$$

so we can rewrite

$$f_i(t) = b_i - \min_{x:Ax \leq b-t} A_i^T x - (2n-1)t = \max_{x:Ax \leq b-t} \left(b_i - A_i^T x - (2n-1)t\right).$$

We want to solve $f_i(t) = 0$, for $0 \leq t \leq M_i$ and $f_i(M_i) < 0$. Equivalently, we want the unique t such that:

$$0 = \max_{x:Ax \leq b-t} \left(b_i - A_i^T x - (2n-1)t\right).$$

This is the same as finding:

$$\arg \max_{t:0 \leq t \leq M_i} \max_{\substack{x:Ax \leq b-t \\ A_i^T x \geq b_i - (2n-1)t}} (b_i - A_i^T x - (2n-1)t).$$

This is a linear program on the dome, except we have an extra constraint $A_i^T x \geq b_i - (2n-1)t$. In order to solve it we solve it first without the constraint. If the optimum satisfies the extra constraint we are done, and if not the optimum we want is obtained solving the linear program in the section $\widetilde{P} \cap \{A_i^T x \geq b_i - (2n-1)t\}$. So, this linear program is solved in time $O(\log^4 m)$. The minimum of the solutions of these programs for the different choices of i is the value of ρ we are looking for.

References

[1] A. Bezdek, K. Bezdek, A solution of Conway's fried potato problem. *Bull. London Math. Soc.* 27, (1995) 492–496. doi:10.1112/blms/27.5.492

[2] H. Croft, K. Falconer and R. Guy, *Unsolved Problems in Geometry*. Berlin–Heidelberg–New York, 1991. doi:10.1007/978-1-4612-0963-8

[3] Antonio Cañete, Isabel Fernández and Alberto Márquez. Conway's fried potato problem: a (quadratic) algorithm leading to an optimal division for convex polygons. In: (L. Tabera Alonso (ed) *Discrete Math Days, Santander, July 4–6, 2022*. Editorial Universidad de Cantabria, Santander, 2022, pp 71–76. doi:doi.org/10.22429/Euc2022.016

[4] Antonio Cañete, Isabel Fernández and Alberto Márquez. Optimal divisions of a convex body. *Mathematical Inequalities & Applications*, 26:2 (2023), 315–342. doi:10.7153/mia-2023-26-21.

[5] D.P. Dobkin and D.G. Kirkpatrick, Fast detection of polyhedral intersections, In: Nielsen, M., Schmidt, E.M. (eds) *Automata, Languages and Programming. ICALP 1982*. Lecture Notes in Computer Science, vol 140. Springer, Berlin, Heidelberg, pp. 154–165. doi:10.1007/BFb0012765

[6] D.P. Dobkin and D.G. Kirkpatrick, Determining the separation of preprocessed polyhedra – A unified approach. In: Paterson, M.S. (eds) *Automata, Languages and Programming. ICALP 1990*. Lecture Notes in Computer Science, vol 443. Springer, Berlin, Heidelberg, pp. 400–413. https://doi.org/10.1007/BFb0032047

[7] Joseph O'Rourke. *Computational geometry in C*. Cambridge University Press, 1998.

[8] R. Schneider, *Convex bodies: the Brunn-Minkowski theory*. Second edition, Cambridge University Press, Cambridge, 2014.

Three-term arithmetic progressions in two-colorings of the plane*

Gabriel Currier[†1], Kenneth Moore[‡1], and Chi Hoi Yip[§1]

[1]Dept. of Mathematics, University of British Columbia, Canada

Abstract

We show that in any two-coloring of the plane, there exists a monochromatic congruent copy of any arithmetic progression of length 3. This problem lies at the intersection of two longstanding but active research projects. The first is the study of Ramsey problems for arithmetic progressions in colorings of euclidean space, for which there are many results dating back over 50 years, but about which much is still not known. The second is centered around a conjecture of Erdős, Graham, Montgomery, Rothschild, Spencer and Straus, which posits that any two-coloring of the plane must contain a monochromatic congruent copy of every non-equilateral three-point configuration. Our result confirms one of the most natural open cases of this conjecture.

1 Introduction

We let $\mathbb{E}^n$ denote n-dimensional Euclidean space, that is, $\mathbb{E}^n$ equipped with the Euclidean norm. The field of Euclidean Ramsey theory is concerned with what types of configurations (monochromatic, rainbow, etc.) must exist in *any* coloring of $\mathbb{E}^n$ using a prescribed number of colors. One of the most commonly studied configurations is denoted ℓ_m, and consists of m collinear points with consecutive points at distance 1 apart. In other words, ℓ_m is an m-term arithmetic progression with common difference 1. Our main result is the following.

Theorem 1. *In any two-coloring of $\mathbb{E}^2$, there exists a monochromatic congruent copy of ℓ_3.*

Thus, by scaling, there naturally exists a monochromatic 3-term arithmetic progression with any common difference. The classical question in this area, known as the Hadwiger-Nelson (HN) problem, is one of the most famous open problems in combinatorics. The HN problem, first discussed by Nelson (not in print) in 1950, asks how many colors one would need to color $\mathbb{E}^2$ so that there is no monochromatic copy of ℓ_2; i.e. two points of unit distance apart. This quantity is known sometimes as the *chromatic number* $\chi(\mathbb{E}^2)$ of the plane. It was known that the answer is between 4 and 7 for a long time, and a 2018 breakthrough by de Grey [7] showed that one needs at least 5 colors. In general, it is known that $(1.239 + o(1))^n \leq \chi(\mathbb{E}^n) \leq (3 + o(1))^n$ as $n \to \infty$ [12, Section 11.1].

After the introduction of the HN problem, the area was further developed by Erdős, Graham, Montgomery, Rothschild, Spencer, and Straus in a series of papers [8, 9, 10]. In these papers, they ask if, for any non-equilateral three-point configuration K, there must be a monochromatic congruent copy of K in any 2-coloring of $\mathbb{E}^2$. The conjecture was confirmed when the coloring is assumed to be polygonal [14], but it is still widely open in general. As noted in [2, Section 6.3], Theorem 1 gives perhaps

*The full version of this work can be found in [6] and will be published elsewhere.

[†]Email: currierg@math.ubc.ca Research of G. C. supported by a Killam Doctoral Scholarship and Four Year Fellowship from the University of British Columbia

[‡]Email: kjmoore@math.ubc.ca Research of K. M. supported a Four Year Fellowship from the University of British Columbia

[§]Email: kyleyip@math.ubc.ca

the most natural open case of this conjecture. This problem was discussed as well in the concluding remarks of a very recent paper of Führer and Tóth [11, Page 12].

To discuss further known results, we introduce some standard notation. If we have configurations $K_1, \dots, K_r$ in $\mathbb{E}^n$, we say that $\mathbb{E}^n \to (K_1, \dots, K_r)$ if, for any coloring of $\mathbb{E}^n$ with r colors, there exists a monochromatic (congruent) copy of K_i in color i, for some i. If there exists a coloring where this does not hold, we say $\mathbb{E}^n \not\to (K_1, \dots, K_r)$. For simplicity, if $K_i = K$ for all i and $\mathbb{E}^n \to (K_1, \dots, K_r)$ or $\mathbb{E}^n \not\to (K_1, \dots, K_r)$, we say simply $\mathbb{E}^n \xrightarrow{r} K$ (resp. $\mathbb{E}^n \not\xrightarrow{r} K$).

Using the above terminology, our Theorem 1 says that $\mathbb{E}^2 \xrightarrow{2} \ell_3$. The question of for which $n, r, s_1, \dots, s_r$ we have $\mathbb{E}^n \to (\ell_{s_1}, \dots, \ell_{s_r})$ also has a rich history, so we collect here the known results. Perhaps the most relevant results to this manuscript are that $\mathbb{E}^2 \not\xrightarrow{3} \ell_3$, that $\mathbb{E}^3 \xrightarrow{2} \ell_3$, and that there exists m such that $\mathbb{E}^n \not\to (\ell_3, \ell_m)$ for all n. The first of these results was shown by Graham and Tressler [13] using a simple hexagonal grid construction. In [8, Theorem 8] Erdős et. al. proved that $\mathbb{E}^3 \xrightarrow{2} T$ for any triangle[1] T; in particular, the second result, that $\mathbb{E}^3 \xrightarrow{2} \ell_3$. The third result, that there exists m such that $\mathbb{E}^n \not\to (\ell_3, \ell_m)$ for all n, is a recent result of Conlon and Wu [4]. They were able to show a bound of $m = 10^{50}$, and in a recent paper, Führer and Tóth [11] were able to improve this to $m = 1177$. Some other relevant results in the area are as follows.

- $\mathbb{E}^2 \to (\ell_2, K)$ for any K with 4 points (Juhász [15])
- $\mathbb{E}^2 \to (\ell_2, \ell_5)$ (Tsaturian [17])
- There is a set K with 8 points, such that $\mathbb{E}^2 \not\to (\ell_2, K)$ (Csizmadia and Tóth [5])
- $\mathbb{E}^3 \to (\ell_2, \ell_6)$ (Arman and Tsaturian [1])
- $\mathbb{E}^n \not\to (\ell_2, \ell_{2^{cn}})$ for some constant $c > 0$ (Conlon and Fox [3])
- $\mathbb{E}^n \not\xrightarrow{2} \ell_6$ (Erdős et. al. [8, Theorem 12])

An (a, b, c) *triangle* is a triangle with side lengths a, b, c. The following theorem is due to Erdős et. al. [10, Theorem 1].

Theorem 2 (Erdős et. al.)**.** *A given 2-coloring admits a monochromatic (a, b, c) triangle if and only if it admits a monochromatic equilateral triangle of side a, b, or c.*

Note that ℓ_3 is a $(1, 1, 2)$ triangle. Thus, by scaling, Theorem 1 and Theorem 2 imply the following corollary.

Corollary 3. *If $n \geq 2$, then $\mathbb{E}^n \xrightarrow{2} T$ for an $(\alpha, 2\alpha, x\alpha)$ triangle T for any $\alpha > 0$ and $x \in [1, 3]$.*

This verifies another interesting case of the aforementioned conjecture of Erdős et. al. from [10]. We refer to [10, 16] and [12, Theorem 11.1.4 (a)] for a collection of known families of triangles T such that $\mathbb{E}^2 \xrightarrow{2} T$. In particular, Erdős et. al. [10] showed that $\mathbb{E}^2 \xrightarrow{2} T$ if T has a ratio between two sides equal to $2\sin(\theta/2)$ with $\theta \in \{30°, 72°, 90°, 120°\}$. Our result handles the case that $\theta = 180°$.

In the next section, we will give an outline of the proof of Theorem 1.

2 Sketch of proof

In this section, we will discuss the main ideas of the proof of Theorem 1. We start with a simple corollary of Theorem 2.

[1] Throughout, degenerate triangles (that is, three collinear points) are also regarded as triangles.

Corollary 4. *If a coloring of $\mathbb{E}^2$ does not contain a monochromatic ℓ_3, then it also does not contain a monochromatic equilateral triangle of side-length 1 or 2.*

Our proof of Theorem 1 will proceed in two parts, both with the same general outline. Each will proceed by contradiction, starting with the assumption that there exists a coloring of $\mathbb{E}^2$ that has no monochromatic ℓ_3. Then, we begin with a small set of starting points, and show that all possible colorings of those starting points will result another point that must be colored both blue and red; that is, a contradiction. So far, we have the following "rules" at our disposal that will allow us to execute this proof: we can take two points of the same color at distance 1 or 2 apart, and do one of the following.

- Add a third point of the opposite color to form an ℓ_3 (as a midpoint if the points are distance 2 apart), or
- add a third point of the opposite color to create an equilateral triangle (of side-length 1 or 2).

Where the second option follows directly from corollary 4. Visually, we can think of these rules as follows.

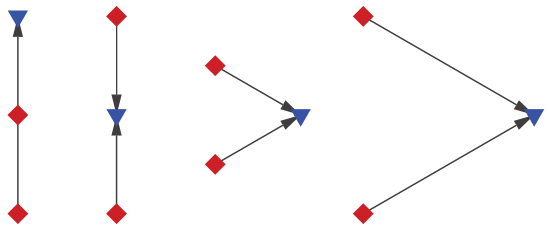


Figure 1: The color implication steps

These two rules are not quite enough to establish Theorem 1. Thus, the first part of the proof, detailed in Section 2.1, will be to establish one more useful rule. Then, in section 2.2 we will describe how to use our rules to prove Theorem 1.

2.1 Another rule

The goal of this section is to describe the following result.

Lemma 5. *In any two-coloring of $\mathbb{E}^2$ containing no monochromatic ℓ_3, any unit triangle colored blue-blue-red has a blue centroid.*

By a symmetric argument, under the same assumptions any red-red-blue triangle has a red centroid. To outline the proof of this result, we need an efficient way to describe the coordinates of our point sets. If a, b, c, d are integers, then all points we use will be of the following form:

$$[a, b, c, d] := \left(\frac{a\sqrt{3} + b\sqrt{11}}{12}, \frac{c + d\sqrt{33}}{12} \right). \tag{1}$$

The proof will proceed as follows: we will start with a basic pointset, containing a unit equilateral triangle and its centroid. Then, we will assume the triangle is colored blue-blue-red but has a red centroid, and use the rules from the previous section (as in Figure 1) to derive a contradiction. The pointset we will use is drawn in Figure 2, with the following explicit coordinates.

$$\begin{array}{llll} p_1 = [-4,0,0,0], & p_2 = [0,0,0,0], & p_3 = [2,0,-6,0], & p_4 = [2,0,6,0], \\ q_1 = [-1,-3,3,-1], & q_2 = [-1,-3,-3,1], & q_3 = [2,0,0,2], & q_3' = [2,0,0,-2], \\ q_4 = [-3,-3,-3,-1], & q_5 = [-3,-3,3,1]. & & \end{array}$$

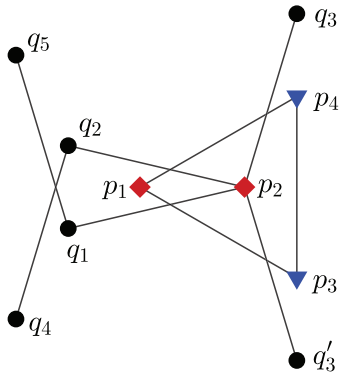


Figure 2: The base points needed to verify the lemma

We will need to consider all possible colorings of this pointset, which results in some case work. However, the symmetries present allow us to limit this to only 6 cases, and the color implications we end up with are simple enough to be verified (somewhat tediously) by hand. Alternatively, we provide a method to quickly verify the result computationally. We again refer to [6] for a complete description of these results, but for the moment we use Figure 3 to visualize the simplest case - that is, where q_1 and q_2 have different colors.

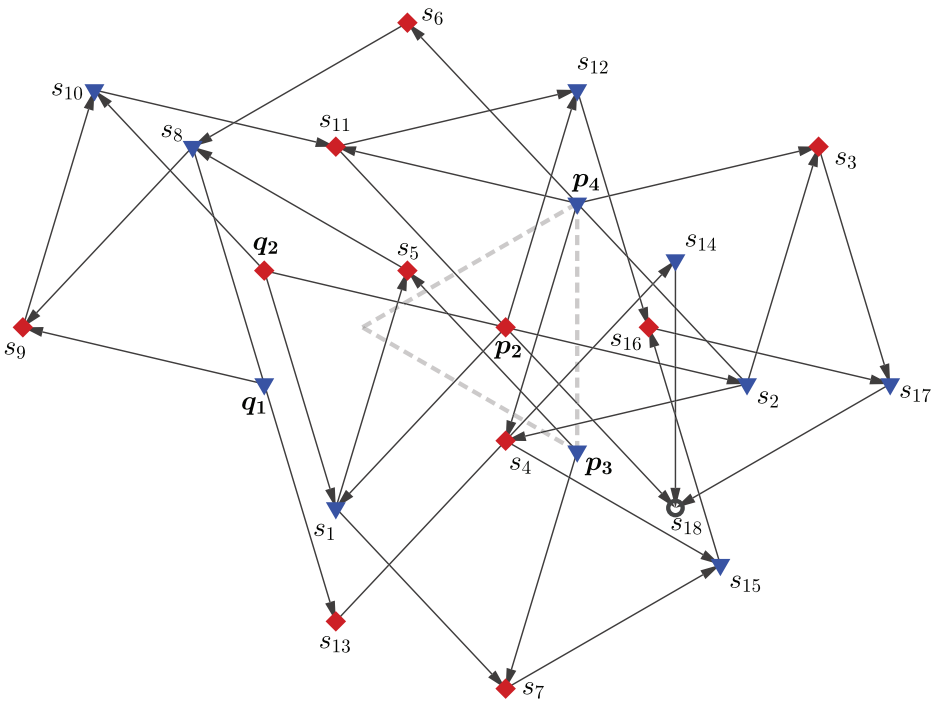


Figure 3: Case 1: q_1 is red and q_2 is blue (or vice versa)

The contradiction comes from the fact that the point s_{18} must be colored both red and blue; the blue coloring comes from the ℓ_3 created with red points s_{11} and p_2, and the red coloring comes from the equilateral triangle created with blue points s_{14} and s_{17}. We note, finally, that not all points from Figure 2 are used in this case. However, the remaining cases will make use of all of the p_i and q_j.

2.2 Putting it all together

Using the rules described in section 2 as well as the one established in section 2.1, we can now complete the proof of Theorem 1. We'll deal with a $\frac{1}{\sqrt{3}}$-scaled hexagonal grid - that is, where the smallest triangles are scaled to have edge-length $\frac{1}{\sqrt{3}}$. A straightforward argument (detailed in [6]) using these rules shows that if we assume there is no monochromatic ℓ_3, then there is only one coloring of this grid up to isometry - that is, the one pictured in Figure 4.

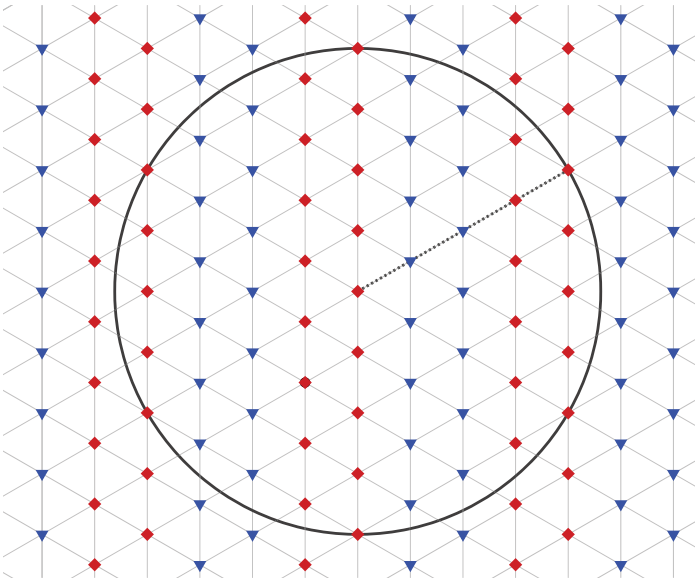


Figure 4: A circle with radius $\frac{4}{\sqrt{3}}$ in the colored grid

To finish the proof is straightforward. We pick two points in our coloring that are less than distance $\frac{8}{\sqrt{3}}$ from one another and have different colors; call them p_1 and p_2, and let them be red and blue respectively. As illustrated in Figure 4, any point on the hexagonal grid at distance $\frac{4}{\sqrt{3}}$ from p_1 must be red as well. By rotating the grid, we can actually show that all points at distance $\frac{4}{\sqrt{3}}$ from p_1 are red, and symmetrically all points at distance $\frac{4}{\sqrt{3}}$ from p_2 are blue. However, since p_1 and p_2 are of distance less than $\frac{8}{\sqrt{3}}$ from one another there must be a point that is distance $\frac{4}{\sqrt{3}}$ from both of these points, which provides our contradiction.

References

[1] A. Arman and S. Tsaturian. A result in asymmetric Euclidean Ramsey theory. *Discrete Math.*, 341(5):1502–1508, 2018.

[2] P. Brass, W. Moser, and J. Pach. *Research problems in discrete geometry.* Springer, New York, 2005.

[3] D. Conlon and J. Fox. Lines in Euclidean Ramsey theory. *Discrete Comput. Geom.*, 61(1):218–225, 2019.

[4] D. Conlon and Y.-H. Wu. More on lines in Euclidean Ramsey theory. *C. R. Math. Acad. Sci. Paris*, 361:897–901, 2023.

[5] G. Csizmadia and G. Tóth. Note on a Ramsey-type problem in geometry. *J. Combin. Theory Ser. A*, 65(2):302–306, 1994.

[6] G. Currier, K. Moore, and C. H. Yip. Any two-coloring of the plane contains monochromatic 3-term arithmetic progressions, 2024. arXiv:2402.14197.

[7] A. D. N. J. de Grey. The chromatic number of the plane is at least 5. *Geombinatorics*, 28(1):18–31, 2018.

[8] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus. Euclidean Ramsey theorems. I. *J. Combinatorial Theory Ser. A*, 14:341–363, 1973.

[9] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus. Euclidean Ramsey theorems. II. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III*, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 529–557. North-Holland, Amsterdam-London, 1975.

[10] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. Spencer, and E. G. Straus. Euclidean Ramsey theorems. III. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III*, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 559–583. North-Holland, Amsterdam-London, 1975.

[11] J. Führer and G. Tóth. Progressions in Euclidean Ramsey theory, 2024. arXiv:2402.12567.

[12] J. E. Goodman, J. O'Rourke, and C. D. Tóth, editors. *Handbook of discrete and computational geometry*. Discrete Mathematics and its Applications (Boca Raton). CRC Press, Boca Raton, FL, third edition, 2018.

[13] R. Graham and E. Tressler. Open problems in Euclidean Ramsey theory. In *Ramsey theory*, volume 285 of *Progr. Math.*, pages 115–120. Birkhäuser/Springer, New York, 2011.

[14] V. Jelínek, J. Kynčl, R. Stolař, and T. Valla. Monochromatic triangles in two-colored plane. *Combinatorica*, 29(6):699–718, 2009.

[15] R. Juhász. Ramsey type theorems in the plane. *J. Combin. Theory Ser. A*, 27(2):152–160, 1979.

[16] L. E. Shader. All right triangles are Ramsey in E^2! *J. Combinatorial Theory Ser. A*, 20(3):385–389, 1976.

[17] S. Tsaturian. A Euclidean Ramsey result in the plane. *Electron. J. Combin.*, 24(4):Paper No. 4.35, 9, 2017.

Bounding the balanced upper chromatic number *

Gabriela Araujo-Pardo[†1], Silvia Fernández-Merchant[‡2], Adriana Hansberg[§1], Dolores Lara[¶3], Amanda Montejano[‖4], and Déborah Oliveros[**1]

[1]Instituto de Matemáticas, Universidad Nacional Autónoma de México, Campus Juriquilla, Mexico.
[2]California State University, Northridge, 18311 Nordhoff St, Northridge, CA, 91330, USA.
[3]Departamento de Computación, Cinvestav, Mexico City, Mexico.
[4]UMDI, Facultad de Ciencias, UNAM Juriquilla, Querétaro, México

Abstract

In this paper, we provide a general upper bound on the *balanced upper chromatic number of any linear hypergraph*, that is, the largest size of a vertex coloring of any linear hypergraph in which all color-class sizes differ by at most one (balanced) and each hyperedge contains at least two vertices of the same color (rainbow-free). We are particularly interested in understanding this parameter for the n-dimensional cube on t elements due to its close connection to the unexistence of a rainbow Ramsey version of the Hales-Jewett Theorem. We improve the lower and upper bounds for this hypergraph and (except for four cubes) completely determine this parameter in dimensions 2 and 3.

1 Introduction

Many classical Ramsey theory results that deal with the existence of a monochromatic object have a rainbow counterpart: a theorem that guarantees the existence of certain rainbow subset (i.e. such that no color is repeated) provided that the color set is sufficiently large and that all colors are well represented. An example of this is van der Waerden's theorem, whose rainbow counterpart for three colors was studied in [8]. The novelty in [8] was to notice that rainbow structures can be forced to appear not only by letting the number of colors grow, but also by fixing the number of colors and letting all chromatic classes be large enough. This is because the more balanced the color classes are the higher the number of rainbow substructures is. For instance, in k-colorings of the set of vertices of a t-uniform hypergraph H, the number of rainbow t-sets of vertices is higher as the coloring becomes more balanced. Therefore, when looking for a rainbow-free k-coloring of H, it is in principle harder to find one among balanced k-colorings (this, of course, depends on the structure of H). Consequently, we consider ~~here~~ *balanced colorings* of the vertices of a given hypergraph H, that is, colorings in which the cardinalities of all color classes differ in at most one. In this setting, we aim to maximize the number k of colors for which there is a balanced k-coloring of H without *rainbow hyperedges*, i.e. hyperedges where all colors appear at most once. To avoid inconsistencies with this definition, we assume that all hyperedges have size at least two.

*Part of this research was performed during a Simons Laufer Mathematical Sciences Institute (SLMath, formerly MSRI) summer research program, supported by the National Science Foundation (Grant No. DMS-1928930) in partnership with the Mathematics Institute of the National Autonomous University of Mexico (UNAM).

[†]Email: garaujo@im.unam.mx. Research of G. A-P. supported by DGAPA PAPIIT IN113324

[‡]Email: silvia.fernandez@csun.edu. Research of S. F-M. supported by a 2024 RSP CSUN Campus Funding Initiative.

[§]Email: ahansberg@im.unam.mx. Research of A. H. supported by DGAPA PAPIIT IG100822.

[¶]Email: dolores.lara@cinvestav.mx

[‖]Email: amandamontejano@ciencias.unam.mx. Research of A. M. supported by DGAPA PAPIIT IG100822.

[**]Email: doliveros@im.unam.mx. Research of D. O. supported by DGAPA PAPIIT 35-IN112124

This maximum value, defined originally in [3], is called the *balanced upper chromatic number of* H, and it is denoted by $\overline{\chi}_b(H)$. Clearly, this parameter is related to the *upper chromatic number* $\overline{\chi}(H)$ defined as the maximum number of colors in a coloring of H (not necessarily balanced) without rainbow hyperedges, which has been the subject of study in many papers, see for instance [4, 5, 10]. Observe that, if $E(H)$ is the set of hyperedges and n is the order of H, then $\min\{|e| : e \in E(H)\} - 1 \leq \overline{\chi}_b(H) \leq \overline{\chi}(H) \leq n$. Often, lower bounds for $\overline{\chi}(H)$ are obtained by colorings with one very large color class and all other classes of size one. Evidently, such colorings do not provide lower bounds for $\overline{\chi}_b(H)$ which requires more involved constructions.

The upper balanced chromatic number of C_t^n, the n-dimensional cube over t-elements, is in close connection with the unexistence of a rainbow Ramsey version of the Hales-Jewett theorem, a central result in Ramsey theory that establishes the existence of monochromatic combinatorial lines in any finite coloring of C_t^n provided that n is sufficiently large, it was shown in [9] that, except for the case $(t, n) = (3, 2)$, for every $t \geq 3$ and every $n \geq 2$, there are balanced t-colorings of C_t^n without rainbow lines. Moreover, for every even $t \geq 4$ and every n, there are balanced $\left(\frac{t}{2}\right)^n$-colorings of C_t^n without rainbow lines [9]. This shows that $\overline{\chi}_b(C_t^n) \geq \left(\frac{t}{2}\right)^n$. In a recent work by the authors of this paper [2], it is proved that

$$\overline{\chi}_b(C_t^n) \leq \frac{3t^n - (t+2)^n}{2}, \tag{1}$$

for $t \geq \frac{2}{\sqrt[n]{2}-1}$, and that the bound is attained when $t \geq 4n - 2$.

In Section 2, we generalize the idea leading into Inequality (1), providing a general upper bound for the balanced upper chromatic number of any *linear hypergraph*, that is, a hypergraph where every two edges intersect in at most one vertex, Theorem 1. Moreover, we complete the upper bound for the case that $v < 2e$, showing that $\overline{\chi}_b(H) \leq \left\lceil 2(v+2e) - 4\sqrt{e^2+ev} \right\rceil - 1$. This bound comes very close to the known upper bounds for the upper balanced chromatic number of finite projective planes, which constitute a special family of linear hypergraphs, that were given in [3, 7].

In Section 3, we provide bounds for the balanced chromatic number of the n-dimensional cube. We present a general lower bound that follows from Theorem 1 and sharper bounds for specific values of t (improving those in [9]). Finally, Section 4 refines our results for the plane and the space.

2 General upper bound

We start presenting an upper bound for the balanced upper chromatic number of a linear hypergraph.

Theorem 1. *Let H be a linear hypergraph with v vertices and e hyperedges. Then*

$$\overline{\chi}_b(H) \leq \begin{cases} v - e & \text{if } v \geq 2e, \\ \left\lceil 2(v+2e) - 4\sqrt{e^2+ev} \right\rceil - 1 & \text{if } v < 2e. \end{cases}$$

Proof. Consider a balanced c-coloring of H for some integer $2 \leq c \leq v$. Then there is an integer $1 \leq k < v$ such that all color classes are of size k or possibly $k+1$. Let $c_k \geq 1$ and $c_{k+1} \geq 0$ be the number of classes of size k and $k+1$, respectively. Then $c = c_k + c_{k+1}$ and $v = kc_k + (k+1)c_{k+1} = ck + c_{k+1}$. Since $0 \leq c_{k+1} < c$, then k and c_{k+1} are the quotient and the remainder, respectively, when v is divided by c. That is, $k = \lfloor \frac{v}{c} \rfloor$ and $c_{k+1} = v - c\lfloor \frac{v}{c} \rfloor$. Let $\varepsilon = \frac{c_{k+1}}{c} = \frac{v}{c} - \lfloor \frac{v}{c} \rfloor$. Thus $c_{k+1} = \varepsilon c$ and $k = \frac{v}{c} - \varepsilon$, where $0 \leq \varepsilon < 1$. We say that a color *blocks* a hyperedge if at least two vertices of the hyperedge receive that color. So an *unblocked* edge is a rainbow edge. Note that at most $\binom{k}{2}c_k + \binom{k+1}{2}c_{k+1} = \binom{k}{2}c + kc_{k+1}$ hyperedges can be blocked by the distinct colors in the c-coloring. Hence, if

$$e > \binom{k}{2}c + kc_{k+1} = \binom{\frac{v}{c} - \varepsilon}{2}c + \left(\frac{v}{c} - \varepsilon\right)\varepsilon c = \frac{1}{2}\left(\frac{v}{c} - \varepsilon\right)(v + \varepsilon c - c) = \frac{1}{2c}(v - \varepsilon c)(v + \varepsilon c - c), \tag{2}$$

then there is at least one rainbow hyperedge.

First, assume that $v \geq 2e$ and let $c = v - e + 1$. To prove that any c-coloring of H has rainbow hyperedges, it is enough to verify Inequality (2). In this case, $v = (v-e+1)\cdot 1 + (e-1) = c\cdot 1 + (e-1)$ and the assumption $v \geq 2e$ implies $0 \leq e-1 < v-e+1 = c$. Thus $\varepsilon c = e-1$ and

$$\frac{1}{2c}\left(v - \varepsilon c\right)\left(v + c\varepsilon - c\right) = \frac{1}{2(v-e+1)}\left(v-(e-1)\right)\left(v+(e-1)-(v-e+1)\right) = e-1 < e.$$

Similarly, assume that $v < 2e$ and let $c = \left\lceil 2(v+2e) - 4\sqrt{e^2+ev} \right\rceil$. It is enough to show that Inequality (2) holds. Note that (rationalizing)

$$c > 2(v+2e) - 4\sqrt{e^2+ev} = \frac{2v^2}{v+2e+2\sqrt{(e^2+ev)}} = \frac{2v^2}{v+2e+\sqrt{(v+2e)^2-v^2}}. \tag{3}$$

Because $(v+2e)^2 - v^2 < (v+2e)^2$, it follows that $c > \frac{v^2}{v+2e}$. This implies that $e > \frac{v(v-c)}{2c}$, which is precisely Inequality (2) when $\varepsilon = 0$.

Assume now that $0 < \varepsilon < 1$. Thus $0 < \varepsilon(1-\varepsilon) \leq \frac{1}{4}$. Since $2e > v$ and $v \geq c$, we have that

$$\frac{v+2e}{2\varepsilon(1-\varepsilon)} > \frac{v}{\varepsilon(1-\varepsilon)} \geq 4v > v \geq c. \tag{4}$$

Also, $v^2 \geq 4v^2\varepsilon(1-\varepsilon)$. By Inequality (3) and rationalizing, we obtain

$$c > \frac{2v^2}{v+2e+\sqrt{(v+2e)^2-4v^2\varepsilon(1-\varepsilon)}} = \frac{v+2e-\sqrt{(v+2e)^2-4v^2\varepsilon(1-\varepsilon)}}{2\varepsilon(1-\varepsilon)}. \tag{5}$$

Inequalities (4) and (5) imply that

$$\frac{\sqrt{(v+2e)^2-4v^2\varepsilon(1-\varepsilon)}}{2\varepsilon(1-\varepsilon)} > \frac{v+2e}{2\varepsilon(1-\varepsilon)} - c > 0.$$

Multiplying by $2\varepsilon(1-\varepsilon)$ and squaring, we obtain $(v+2e)^2 - 4v^2\varepsilon(1-\varepsilon) > (v+2e-2\varepsilon(1-\varepsilon)c)^2$, which is equivalent to

$$e > \frac{1}{2c}\left(v^2 - vc + \varepsilon c^2 - \varepsilon^2 c^2\right) = \frac{1}{2c}\left(v-\varepsilon c\right)\left(v+\varepsilon c - c\right).$$

This is again Inequality (2) and thus there must be a rainbow hyperedge. □

The finite projective planes Π_q of order q have the same number of lines (of size $q+1$) and points, namely $v = e = q^2+q+1$. It has been shown that $\overline{\chi}_b(\Pi_q) \leq v/3$ [3, 7]. The upper bound above, which in this setting falls into the second case, gives $\overline{\chi}_b(\Pi_q) \leq (6-4\sqrt{2})v \approx 0.34v$.

3 Geometric lines in hypercubes

In this section, we consider the *n-cube over t elements*, denoted by C_t^n as an application of Theorem 1, where the set of *vertices* is the set of points in $\mathbb{R}^n$ with entries in $\{0,1,\ldots,t-1\}$ and the hyperedges are the geometric lines of C_t^n, that is, all the lines parallel to the axes and the main diagonals of maximal hyperplanes. More precisely, $C_t^n = \{\mathbf{x} = (x_1, x_2, \ldots, x_n) : 0 \leq x_i \leq t-1, x_i \in \mathbb{Z}\}$; and a set of t points in C_t^n is a geometric line if there is a labeling of the points $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{t-1}$ such that if $\mathbf{x}_i = (x_{i,1}, x_{i,2}, x_{i,3}, \ldots, x_{i,n-2}, x_{i,n-1}, x_{i,n})$ for all $0 \leq i \leq t-1$, then for every $1 \leq j \leq n$ it holds that the entries of $(x_{0,j}, x_{1,j}, x_{2,j}, \ldots, x_{t-2,j}, x_{t-1,j})$ are all equal to some fixed value $a \in \{0,1,\ldots,t-1\}$; appear in increasing order $0,1,2,\ldots,t-1$; or appear in decreasing order $t-1, t-2, \ldots, 1, 0$.

It is clear, that the number of vertices of this hypergraph is $v = |(C_t^n)| = t^n$ and it is known that the number of hyperedges is $e = |\mathcal{L}(C_t^n)| = \frac{1}{2}((t+2)^n - t^n))$ [6]. Since any two points in the cube C_2^n are in a line, then $\overline{\chi}_b(C_2^n) = 1$. The general lower bound $\overline{\chi}_b(C_t^n) \geq \left(\frac{t}{2}\right)^n$ for any even $t \geq 4$ was proved in [9]. The following result is a direct application of Theorem 1 with $v = |C_t^n| = t^n$ and $e = |\mathcal{L}(C_t^n)| = ((t+2)^n - t^n)/2$.

Corollary 2. *Let t and n be positive integers. Then*

$$\overline{\chi}_b(C_t^n) \leq \begin{cases} \dfrac{3t^n-(t+2)^n}{2} & \text{if } t \geq \dfrac{2}{\sqrt[n]{2}-1}, \\ \left\lceil 2(t+2)^n - 2\sqrt{(t+2)^{2n}-t^{2n}} \right\rceil - 1 & \text{if } 2 \leq t < \dfrac{2}{\sqrt[n]{2}-1}. \end{cases}$$

It can be checked that when $t=2$, this result implies $\overline{\chi}_b(C_2^n)=1$. Recently, we have proved that this upper bound is tight when $t \geq 4n-2$ (see Theorem 3) by giving an intricate construction that uses the Hall's Marriage Theorem.

Theorem 3 ([2])**.** *For integers $n \geq 2$ and $t \geq 4n-2$, the balanced upper chromatic number of C_t^n is $\overline{\chi}_b(C_t^n) = \frac{3t^n-(t+2)^n}{2}$. This identity also holds for $(t,n)=(5,2),(8,3)$, and $(9,3)$.*

We conjecture that Theorem 3 remains true for $2/(\sqrt[n]{2}-1) \leq t \leq 4n-2$. To confirm this, a rainbow-free coloring with classes of sizes 1 and 2 needs to be found. The inclusion of the cases $(t,n)=(5,2),(8,3)$, and $(9,3)$ confirms this conjecture for dimensions 2 and 3. We now focus on the small values of t, namely, $2 < t < 2/(\sqrt[n]{2}-1)$. First, we present a *recursive* lower bound that will allow us to improve the lower bound in [9] and the best-known colorings in the space.

Theorem 4. *Let t and n be positive integers and suppose that t has a proper divisor $1 < d < t$. Then $\overline{\chi}_b(C_t^n) \geq \left(\frac{t}{d}\right)^n \overline{\chi}_b(C_d^n)$.*

Proof. (Sketch) Partition C_t^n into $(t/d)^n$ n-cubes over d elements. Color each of these smaller n-cubes with $\overline{\chi}_b(C_d^n)$ different colors for a total of $(t/d)^n\,\overline{\chi}_b(C_d^n)$. To prove that this coloring has no rainbow lines, we argue that any geometric line of C_t^n completely contains a geometric line of one of the $\left(\frac{t}{d}\right)^n$ copies of C_d^n. Since none of these smaller lines is rainbow (i.e., each of them has at least two points of the same color), then the larger line is not rainbow. □

When $t=4$ and any n, the idea used in Theorem 4 can improve the lower bound in [9] for this case.

Proposition 5. *For $n \geq 2$, $\overline{\chi}_b(C_4^n) \geq 2^n+1$.*

Proof. (Sketch) Partition the cube C_4^n into 2^n n-cubes over 2 elements $C_1, C_2, \ldots C_{2^n}$ and denote by C_0 the centered n-cube over 2 elements, that is $C_0 = \{\mathbf{x}=(x_1,x_2,\ldots,x_n) \in C_4^n, :\ x_i \in \{1,2\}\}$. Consider the sets $R_i = C_i - C_0$. Assign color 0 to all vertices in C_0 and color i to every vertex in R_i, $1 \leq i \leq 2^n$. Note that this is a balanced partition of C_4^n into 2^n+1 parts, 2^n of size 2^n-1 and one of size 2^n, (see Figure 1 for this coloring of C_4^n for $n=3$). This coloring contains no rainbow lines. □

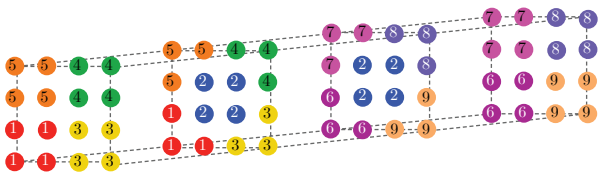


Figure 1: An illustration of the coloring in Proposition 5 for $n=3$.

A direct application of Theorem 4 and Proposition 5 gives a lower bound that improves when t is a multiple of 4. Moreover, we were able to adapt the construction for every even t.

Theorem 6. *Let $n \geq 2$. If $2 \leq t \leq n$ and t is even, then $\overline{\chi}_b(C_t^n) \geq \left(\frac{t}{2}\right)^n + \left\lfloor \frac{t}{4} \right\rfloor^n$.*

Proof. (Sketch) If $t \equiv 0 \pmod 4$ the result follows directly by Theorem 4 and Proposition 5. If $t \equiv 2 \pmod 4$ then we use the construction for $t-2$ adapted as follows. For every $1 \leq i \leq n$ consider the two central $(n-1)$-dimentional hypercubes $C_{i,1}$ and $C_{i,2}$ defined as $C_{i,1} = \{(x_1, x_2, \ldots, x_n) \in C_t^n : x_i = \frac{t-2}{2}\}$ and $C_{i,2} = \{(x_1, x_2, \ldots, x_n) \in C_t^n : x_i = \frac{t}{2}\}$. Use the coloring provided by Theorem 4 and Proposition 5 for the set $C_t^n \setminus \cup_{i=1}^n (C_{i,1} \cup C_{i,2})$, which can be seen as a copy of C_{t-2}^n contained in C_t^n (see the top part of Figure 2(e) to visualize this copy of C_{t-2}^n contained in C_t^n when $(t,n) = (6,3)$, ignore the colors.) Then, only the lines contained in $\cup_{i=1}^n (C_{i,1} \cup C_{i,2})$ are not yet blocked by this partial coloring, but we can block them using copies of C_2^n, each of a different color. □

Note that this lower bound is nonsignificant for $t \geq 4n-2$ due to Theorem 3, but provides the best known bound for the remaining even values of t. The other modular classes of t mod n would require a more detailed analysis that strongly depends on the dimension to expand the coloring of C_{t-1}^n to C_t^n. We illustrate this approach in Figure 2(d) for the case $(t,n) = (5,3)$.

4 A summary of exact results and best bounds

In this section, we summarize the best-known bounds for the cases $n = 2$ and $n = 3$. In the 2-dimentional case, the balanced upper chromatic number is completely determined. In the 3-dimentional case, four cases $4 \leq t \leq 7$ remain open. Table 1 shows the best bounds we know for these values.

Theorem 7. *For $n = 2$, $\overline{\chi}_b(C_3^2) = 2$, $\overline{\chi}_b(C_4^2) = 7$, and $\overline{\chi}_b(C_t^2) = t^2 - 2t - 2$, for $t \geq 5$. For $n = 3$, $\overline{\chi}_b(C_3^3) = 3$, and $\overline{\chi}_b(C_t^3) = t^3 - 3t^2 - 6t - 4$, for $t \geq 8$.*

Proof. In the plane ($n = 2$), the cases $t \geq 5$ are covered by Theorem 3. The balanced rainbow-free 3-coloring shown in Figure 2(a) shows that the upper bound in Theorem 1 is tight for $t = 4$. It is known that in any balanced 3-coloring of C_3^2, there is a rainbow line, which means that $\overline{\chi}_b(C_3^2) \leq 2$ [9]. Since 2 colors are not enough to block a line in this cube, then $\overline{\chi}_b(C_3^2) = 2$. In the space ($n = 3$), the cases $t \geq 8$ are covered by Theorem 3. For $t = 3$, the balanced rainbow-free 3-coloring shown in Figure 2(b) shows that $\overline{\chi}_b(C_3^3) \geq 3$. We ran a computer program to check that all balanced 4-colorings of C_3^3 contain a rainbow line showing that $\overline{\chi}_b(C_3^3) = 3$. Our program searched among a reduced set of $O(13!)$ colorings. Such a coloring would have 3 colors that are used 7 times and one color that is used 6 times. We reduced the number of possible colorings to be checked by fixing the color of the point in the center of the cube and using the fact that the other two points in any line through the center are either the same color or one of them is the same color as the center. □

The question of determining $\overline{\chi}_b(C_t^n)$ for $3 \leq t < 4n-2$ remains open for higher dimensions.

t	lower bound Th. 6	lower bound Fig. 2	upper bound Th. 2
4	4	12	18*
5	-	26	47
6	28	40	95
7	-	64	171

Table 1: Best bounds for $\overline{\chi}_b(C_t^3)$ when $n = 3$ and $4 \leq t \leq 7$. * The bound resulting from Theorem 2 is 19 but we have improved it to 18 [2].

References

[1] M. Aigner, Lexicographic matching in Boolean algebras, *J. Comb. Theory Ser.B*, 14:187–194, 1973.

[2] G. Araujo-Pardo, S. Fernández-Merchant, A. Hansberg, D. Lara, A. Montejano, D. Oliveros, The exact balanced upper chromatic number of the n-cube over t elements, 36th Canadian Conference on Computational Geometry, 2024, to appear.

Figure 2: Rainbow-free colorings of (a) C_4^2, (b) C_3^3, (c) C_4^3. (d) C_5^3, (e) C_6^3, (f) C_7^3. The shaded regions highlight a coloring of a smaller size (e.g. the shaded region in (e) corresponds to the coloring in (c).) All colorings in (d)-(f) expand the one in (c), so (c)-(f) have color classes of sizes 5 and 6.

[3] G. Araujo-Pardo, G. Kiss, A. Montejano, On the balanced upper chromatic number of cyclic projective planes and projective spaces, Discrete Mathematics 338:12, 2562–2571, 2015.

[4] G. Bacsó, Z. Tuza. Upper chromatic number of finite projective planes. *J. Combinatorial Designs*, 16(3):221–230, 2008.

[5] S. Bhandari, V. Voloshin. Upper chromatic number of n-dimensional cubes, *Alabama Journal of Mathematics* 43, 2019.

[6] J. Beck, W. Pegden, and S. Vijay. "The Hales-Jewett number is exponential: game-theoretic consequences". In: *Analytic Number Theory: Essays in Honour of Klaus Roth*. Ed. by W.W.L. Chen et al. Cambridge University Press, 2009.

[7] Z. L. Blázsik, A. Blokhuis, Š. Miklavič, Z. L. Nagy, and T. Szőnyi, On the balanced upper chromatic number of finite projective planes. *Discrete Mathematics*, 344(3):112266, 2021.

[8] V. Jungić, J. Licht (J. Fox), M. Mahdian, J. Nešetřil, and R. Radoičić. Rainbow arithmetic progressions and anti-Ramsey results. *Combinatorics, Probability and Computing*, 12(5–6):599–620, 2003.

[9] A. Montejano, Rainbow considerations around the Hales-Jewett theorem, preprint, 2024, arXiv:2403.13726.

[10] V. I. Voloshin. On the upper chromatic number of a hypergraph. *Australas. J. Comb.*, 11:25–46, 1995.

Creating trees with high maximum degree *

Grzegorz Adamski[†1], Sylwia Antoniuk[‡1], Małgorzata Bednarska-Bzdęga[§1], Dennis Clemens [¶2], Fabian Hamann [‖2], and Yanick Mogge [**2]

[1]Department of Discrete Mathematics, Faculty of Mathematics and CS, Adam Mickiewicz University, Poznań, Poland
[2]Institute of Mathematics, Hamburg University of Technology, Hamburg, Germany.

Abstract

We consider positional games where the winning sets are edge sets of copies of fixed spanning trees or tree universal graphs. We prove that in Maker-Breaker games on the edges of a complete graph K_n, Maker has a strategy to occupy the edges of a graph which contains copies of all spanning trees with almost linear maximum degree, and we give a similar result for Waiter-Client games. By this, it follows that both Maker and Waiter can play at least as good as predicted by the so-called random graph intuition. Moreover, our results improve on special cases of earlier results by Johannsen, Krivelevich, and Samotij as well as Han and Yang. Additionally, when the target of the building player is a copy of only one fixed spanning tree, then we show that in the Waiter-Client game on K_n, Waiter can do even better than suggested by the random graph intuition, while the same is not true for Client in the similarly looking Client-Waiter game.

1 Introduction

Tree embedding problems have a long history, ranging from the embedding of a fixed tree (e.g. [13, 18, 19]) over universality results (e.g. [11, 17, 22]) to packing problems (e.g. [3, 16, 23]). Research in this branch of combinatorics was influenced by many beautiful problems, including the appearance of particular subgraphs in the binomial random graph $G(n,p)$, as well as challenging conjectures, such as the well known Ringel's Conjecture from 1968 and Gyárfás Tree Packing Conjecture from 1978, just to mention a few. For an overview on general graph embedding problems we recommend the survey [6].

In our paper, we want to take a look at such tree embedding problems from a game theoretic perspective, as it has been started already in a series of papers, see e.g. [5, 7, 8, 10, 12, 17, 21]. In general, given any hypergraph $\mathcal{H} = (X, \mathcal{F})$, a *positional game* on $\mathcal{H}$ is played as follows. Two players claim the elements of the *board* X in rounds according to some predefined rule; and the winner is determined according to some rule that involves the *winning sets* in $\mathcal{F}$. Specifically, we will be interested in the following three types of such games.

- **Maker-Breaker** games: Maker and Breaker alternatingly claim one element of X which was not claimed before. Maker wins if she occupies all elements of a winning set, and Breaker wins otherwise.

*The full version of this work can be found in [1, 2] and will be published elsewhere. The research of the fourth and sixth author is supported by Deutsche Forschungsgemeinschaft (Project CL 903/1-1).

[†]Email: grzegorz.adamski@amu.edu.pl
[‡]Email: sylwia.antoniuk@amu.edu.pl
[§]Email: mbed@amu.edu.pl
[¶]Email: dennis.clemens@tuhh.de
[‖]Email: fabian.hamann@tuhh.de
[**]Email: yannick.mogge@tuhh.de

- **Waiter-Client** games: In each round, Waiter offers two elements of X to Client, and then Client decides which element is claimed by him, and which element goes to Waiter. Client wins if he avoids to claim a full winning set, and otherwise Waiter wins. (If in the last round there is only one unclaimed element in X, then it is given to Waiter.)

- **Client-Waiter** games: The elements of X are claimed in the same way as in Waiter-Client games, but this time Client wins if at some point he occupies a winning set, and Waiter wins otherwise.

We note that the above games, when played on the edges of the complete graph K_n, often but not always show to have some strong connection to properties of random graphs, referred to as *random graph intuition*, which roughly speaking suggests that the outcome of a game between perfect players can be predicted by looking at the typical behaviour of a randomly played game in which each player creates a random graph. Prominent examples for such a relation between positional games and random graphs are e.g. the Maker-Breaker clique game [4], the Maker-Breaker Hamiltonicity game [20], and the Waiter-Client H-game [24]. For a general overview on positional games we refer to the monograph [15].

In the following we will stick to games on $X = E(K_n)$, the edge set of a complete graph K_n on n vertices. For any spanning tree T of K_n, we will consider the family $\mathcal{F}_T$ consisting of all copies of T in K_n. Moreover, we will be interested in the family $T(n, \Delta)$ of all graphs which are *universal* for trees on n vertices with maximum degree at most Δ, i.e. graphs which contain a copy of every such tree.

Starting with games in which Maker wants to claim a copy of a fixed tree, Ferber, Hefetz and Krivelevich [10] asked for the largest value $d = d(n)$ such that in a Maker-Breaker game on the edges of K_n, Maker has a strategy to claim a copy of any tree T provided that the maximum degree satisfies $\Delta(T) \leq d$ and n is large enough. An analogue question for Waiter-Client games has then been asked in [8], and related questions regarding tree universality were studied in [5, 17]. We note that in all cases the random graph intuition would suggest that the largest value for the maximum degree $\Delta(T)$ such that the building player (i.e. the player who aims for a winning set) wins should be of the order $\frac{n}{\log(n)}$, see e.g. [19] for the case when a tree T is fixed. However, all previously known results are quite far away from this desired bound on $\Delta(T)$: Hefetz et al. [14] proved that Maker can claim a Hamilton path within $n-1$ rounds. With a tiny worsening in the number of rounds, this was extended to trees of constant maximum degree [7] and trees with $\Delta(T) \leq n^{0.05}$ [10]. Not aiming for a fast winning strategy, Johannsen, Krivelevich, and Samotij [17] further improved the bound on the maximum degree, where their result is much more general as it also considers games played on expander graphs and it gives a Maker's winning strategy for tree universality, i.e. for $T(n, \Delta)$, when $\Delta \leq \frac{cn^{1/3}}{\log(n)}$. Recently, the latter was further improved to $\Delta \leq \frac{cn^{1/2}}{\log(n)}$ by Han and Yang [12]. Moreover, all of the above results stay true when considered in the Waiter-Client context, see [5, 8].

2 Tree Universality

As our first contribution to positional games involving spanning trees, we show that for the tree universality game $T(n, \Delta)$, Maker and Waiter can play at least as good as predicted by the random graph intuition.

Theorem 1 (Tree Universality, Maker-Breaker version, Theorem 1.1 in [2])**.** *There exists a constant $c > 0$ such that the following holds for every large enough integer n. In the Maker-Breaker game on K_n, Maker has a strategy to occupy a graph which contains a copy of every tree T with n vertices and maximum degree $\Delta(T) \leq \frac{cn}{\log(n)}$.*

Theorem 2 (Tree Universality, Waiter-Client version, Theorem 1.2 in [2])**.** *There exists a constant $c > 0$ such that the following holds for every large enough integer n. In the Waiter-Client game on*

K_n, Waiter has a strategy to force Client to claim a graph which contains a copy of every tree T with n vertices and maximum degree $\Delta(T) \leq \frac{cn}{\log(n)}$.

For the proofs of Theorem 1 and Theorem 2 we combine many different tools, including properties of expander graphs, simple absorption and random embedding arguments as well as winning criteria for positional games. While most of our tools are rather standard, the difficulty and novelty in our proof, when compared with the earlier results in [12, 17], lies in finding a suitable list of structural properties which (a) help to embed every tree of the mentioned maximum degree and (b) can be achieved by Maker and Waiter, respectively. Note that the more structural properties are added to such a list, the easier (a) can be proven, but the more difficult (b) gets. The following theorem provides such a list.

Theorem 3 (Theorem 3.1 in [2])**.** *Let $\alpha \in (0,1)$, and $C_0 > 0$ be any constants. There exist constants $\gamma', c > 0$ and a positive integer n_0 such that the following is true for every $\gamma \in (0, \gamma')$ and every integer $n \geq n_0$.*

Let $G = (V, E)$ be a graph on n vertices with a partition $V = V_1 \cup V_2$ of its vertex set such that the following properties hold:

(1) Partition size: $|V_2| = 500\lfloor \gamma n \rfloor$.

(2) Suitable star: *There are a vertex x^* and disjoint sets $R^*, S^* \subset V_1$ such that the following holds:*

- *(a)* $|S^*| = \lfloor 25C_0 \log(n) \rfloor$ *and* $S^* \subset N_G(x^*)$.
- *(b)* $|R^*| \leq 25$ *and for each $v \in R^*$ the following holds: If v is not adjacent with x^*, then v is adjacent with a vertex $s_v \in S^*$, such that $s_v \neq s_w$ if $v \neq w$.*
- *(c) For all $w \in V \setminus (R^* \cup S^*)$, we have $d_G(w, S^*) \geq 2C_0 \log(n)$.*

(3) Pair degree conditions: *For every $v \in V(G)$ there are at most $\log(n)$ vertices $w \in V(G)$ such that $|N_G(v) \cap N_G(w) \cap V_1| < \alpha n$.*

(4) Edges between sets: *Between every two disjoint sets $A \subset V_1$ and $B \subset V$ of size $\lfloor C_0 \log(n) \rfloor$ there is an edge in G.*

(5) Suitable clique factor: *In $G[V_2]$ there is a collection $\mathcal{K}$ of $100\lfloor \gamma n \rfloor$ vertex-disjoint K_5-copies such that the following holds:*

- *(a) There is a partition $\mathcal{K} = \mathcal{K}_{good} \cup \mathcal{K}_{bad}$ such that $|\mathcal{K}_{bad}| = \lfloor \gamma n \rfloor$.*
- *(b) Every vertex $v \in V$ which is not in a clique of $\mathcal{K}_{good}$ satisfies $d_G(v, V_2) \geq 40\lfloor \gamma n \rfloor$.*
- *(c) For every clique $K \in \mathcal{K}_{good}$ there are at most γn cliques $K' \in \mathcal{K}_{good}$ such that G does not have a matching of size 3 between $V(K)$ and $V(K')$.*

Then G contains a copy of every tree T on n vertices with maximum degree $\Delta(T) \leq \frac{cn}{\log(n)}$.

The proof of Theorem 3 can be found in [2], and its overall idea can be summarized as follows. We make a case distinction depending on whether the given tree T contains many bare paths of suitable length (i.e. paths such that all inner vertices have degree 2 in the given tree) or many leaves. In the first case, we embed T minus the bare paths into V_1, by using the properties (3) and (4) together with a criterion by Haxell [13] that helps to embed almost spanning trees into expander graphs. Then, with property (5), we manage to embed all the remaining bare paths to complete a copy of T, and at the same time absorb all leftover vertices from V_1 into our embedding. In the second case, we proceed similarly and embed the leaves at the end of our embedding procedure. However, in order to succeed with this final embedding step, we slightly modify the first step involving Haxell's criterion as follows: If there is a vertex x in T which is adjacent to many neighbours of leaves, we modify Haxell's criterion to make sure that x can be embedded onto x^* (see property (2)) and that we can use S^* exclusively for

the embedding of leaf neighbours. Otherwise, if such a vertex x does not exist, we make sure that in the application of Haxell's criterion a small subtree of T, which itself contains many leaf neighbours, is embedded in a suitable (i.e. random) way into V_1. In both cases, also using the properties (2)–(4), we then obtain suitable properties for our partial embedding that help to finish the embedding of T with a generalization of Hall's Theorem.

For the final proofs of Theorem 1 and Theorem 2, i.e. for giving strategies that create a graph satisfying the properties (1)–(5), we combine many standard tools from positional games, including results on degree games, clique factor games plus the well-known Erdős-Selfridge Criterion and variants of it. A novelty in our proof is that we also play a pair degree game which is necessary for applying our random embedding argument above. While for Maker-Breaker games property (3) cannot be improved in the sense that each pair of vertices gets a large common neighbourhood, for Waiter-Client games we can prove the following more general statement which allows to obtain large common neighbourhoods for all sets of at most logarithmic size.

Lemma 4 (Lemma 6.2 in [1]). *Let $\beta \in (0,1)$. Then for every large enough integer n and every $t \in \mathbb{N}$ such that $t \leq 0.1 \log_2(n)$ the following holds. Suppose G is a graph on n vertices and for every set A of t vertices we have a set $Y_A \subset N_G[A]$ of at least βn common neighbours. Then in the Waiter-Client game on G, Waiter has a strategy such that at the end of the game, Client's graph C satisfies the following:*

$$|N_C[A] \cap Y_A| \geq \frac{\beta n}{200^{t+1}} \quad \text{for every } A \subset V(G) \text{ such that } |A| = t.$$

We believe that the above lemma could be of independent interest, as it may be helpful for other games in which Waiter's goal is to claim complex spanning structures.

3 Results on fixed trees

We believe that the bound of $\frac{n}{\log n}$ in Theorem 1 is best possible and pose this as a conjecture. One reason for believing in this conjecture is that Maker-Breaker games often behave as predicted by the random graph intuition, or Maker performs even worse than this prediction. Indeed, for the randomly played game it follows from [18] that there are fixed trees of maximum degree $\Theta(\frac{n}{\log n})$ which with high probability are not contained in Maker's random graph.

For Waiter-Client games, the situation is completely different, and in fact, we can prove that for any fixed tree T of not too large but linear maximum degree, Waiter has a winning strategy for the Waiter-Client game with winning sets $\mathcal{F}_T$. It then becomes natural to ask for the largest constant c such that Waiter can always win if $\Delta(T) \leq cn$ and n is large enough. With the following two theorems we give a small window for the size of c.

Theorem 5 (Theorem 1.2 in [1]). *For every $\varepsilon \in \left(0, \frac{1}{3}\right)$ there exist positive constants b and n_0 such that the following holds. Let T_n be a tree on $n \geq n_0$ vertices with $\Delta(T_n) < \left(\frac{1}{3} - \varepsilon\right) n$. Then Waiter has a strategy to force a copy of T_n in the Waiter-Client game on K_n within at most $n + b\sqrt{n}$ rounds.*

Theorem 6 (Theorem 1.3 in [1]). *There are positive constants γ and n_0 such that the following holds for every $n \geq n_0$. There exists a tree T_n with n vertices and $\Delta(T_n) < \left(\frac{1}{2} - \gamma\right) n$ such that Client can avoid claiming a copy of T_n in the Waiter-Client game on K_n.*

The proof of Theorem 5, which is given in [1], is an involved study of ad-hoc winning strategies for Waiter consisting of several cases and stages, depending on the existence and distribution of large degree vertices in T, the structure of the tree after all such vertices get deleted, and the existence of suitable bare paths as well as matchings incident with leaves. We skip the details here.

In contrast to this, Theorem 6 is obtained by analysing a partially randomized strategy for Client. We prove this theorem with $\gamma = 0.001$ but do no effort to optimize it, as we believe that our randomized strategy is not optimal. We also note that it is easy to find trees with maximum degree close to $\frac{n}{2}$ that

Client can avoid. Although this improvement by the constant γ in Theorem 6 may seem cosmetic, we believe that it is important for determining a best possible constant c for which Waiter can always win if $\Delta(T) \leq cn$ and n is large enough.

Finally, we consider the Client-Waiter version of the above game. In contrast to the above results, it turns out that Client, who is the building player now, cannot do better than predicted by the random graph intuition. Indeed, the following statement can be obtained as a corollary of Lemma 4.

Theorem 7 (Theorem 1.3 in [1])**.** *There are positive constants c and n_0 such that the following holds. For every $n \geq n_0$ there exists a tree T_n with n vertices and $\Delta(T_n) \leq \frac{cn}{\log(n)}$ such that in a Client-Waiter game on K_n, Waiter can prevent Client from claiming a copy of T_n.*

4 Open problems

As already stated, we believe that Theorem 1 is optimal up to the constant factor c, but we think that Waiter can do better. Therefore, we state the following two conjectures.

Conjecture 8. *There exists a constant $C > 0$ such that the following holds for every large enough integer n. In the Maker-Breaker game on K_n, Breaker has a strategy such that Maker cannot build a graph which contains a copy of every tree T with n vertices and maximum degree $\Delta(T) \leq \frac{Cn}{\log(n)}$.*

Conjecture 9. *There exists a constant $c > 0$ such that the following holds for every large enough integer n. In the Waiter-Client game on K_n, Waiter has a strategy to force Client to claim a graph which contains a copy of every tree T with n vertices and maximum degree $\Delta(T) \leq cn$.*

Similarly and based on other known results on Client-Waiter games we suspect that the Client-Waiter game with winning sets $T(n, \Delta)$ behaves according to the random graph intuition. Due to Theorem 7 it remains to prove the following conjecture.

Conjecture 10. *There exists a constant $c > 0$ such that the following holds for every large enough integer n. In the Client-Waiter game on K_n, Client has a strategy to build a graph which contains a copy of every tree T with n vertices and maximum degree $\Delta(T) \leq \frac{cn}{\log(n)}$.*

Last but not least, recall that in our strategy for Theorem 2 it was beneficial to know that Waiter can force a spanning graph where every pair of vertices has a common neighbourhood of linear size. We wonder how large this pair degree can be made.

Problem 11. *Find the maximum α such that for every large enough n Waiter has a strategy in the Waiter-Client game on K_n to force Client to claim a spanning subgraph C with the following property: for any two vertices v, w we have $|N_C(v) \cap N_C(w)| \geq \alpha n$.*

References

[1] G. Adamski, S. Antoniuk, M. Bednarska-Bzdęga, D. Clemens, F. Hamann and Y. Mogge, Creating spanning trees in Waiter-Client games, preprint, 2024, `arXiv:2403.18534`.

[2] G. Adamski, S. Antoniuk, M. Bednarska-Bzdęga, D. Clemens, F. Hamann and Y. Mogge, Tree universality in positional games, preprint, 2023, `arXiv:2312.00503`.

[3] P. Allen, J. Böttcher, D. Clemens, J. Hladký, D. Piguet and A. Taraz, The tree packing conjecture for trees of almost linear maximum degree, preprint, 2021, `arXiv:2106.11720`.

[4] J. Beck, On two theorems of positional games, *Periodica Mathematica Hungarica* **78.1** (2019), 1–30.

[5] M. Bednarska-Bzdęga, On weight function methods in Chooser-Picker games, *Theoretical Computer Science* **475** (2013), 21–33.

[6] J. Böttcher, Large-scale structures in random graphs, *Surveys in Combinatorics 2017, London Mathematical Society Lecture Note Series* **440** (2017), 87–140.

[7] D. Clemens, A. Ferber, R. Glebov, D. Hefetz and A. Liebenau, Building spanning trees quickly in Maker-Breaker games, *SIAM Journal on Discrete Mathematics* **29.3** (2015), 1683–1705.

[8] D. Clemens, P. Gupta, F. Hamann, A. Haupt, M. Mikalački and Y. Mogge, Fast strategies in Waiter-Client games, *The Electronic Journal of Combinatorics* **27.3** (2020), 1–35.

[9] P. Erdős and J. L. Selfridge, On a combinatorial game, *Journal of Combinatorial Theory, Series A* **14.3** (1973), 298–301.

[10] A. Ferber, D. Hefetz and M. Krivelevich, Fast embedding of spanning trees in biased Maker-Breaker games, *European Journal of Combinatorics* **33.6** (2012), 1086–1099.

[11] J. Friedman and N. Pippenger, Expanding graphs contain all small trees, *Combinatorica* **7** (1987), 71–76.

[12] J. Han and D. Yang, Spanning trees in sparse expanders, preprint, 2022, `arXiv:2211.04758`.

[13] P. Haxell, Tree embeddings, *Journal of Graph Theory* **36** (2001), 121–130.

[14] D. Hefetz, M. Krivelevich, M. Stojaković and T. Szabó, Fast winning strategies in Maker-Breaker games, *Journal of Combinatorial Theory, Series B* **99.1** (2009), 39–47.

[15] D. Hefetz, M. Krivelevich, M. Stojaković and T. Szabó, Positional games, *Oberwolfach Seminars* **44**, Birkhäuser, 2014.

[16] B. Janzer and R. Montgomery, Packing the largest trees in the tree packing conjecture, preprint, 2024, `arXiv:2403.10515`.

[17] D. Johannsen, M. Krivelevich and W. Samotij, Expanders are universal for the class of all spanning *Combinatorics, Probability and Computing* **22.2** (2013), 253–281.

[18] J. Komlós, G. Sárközy and E. Szemerédi, Spanning trees in dense graphs, *Combinatorics, Probability and Computing* **10.5** (2001), 397–416.

[19] M. Krivelevich, Embedding spanning trees in random graphs, *SIAM Journal on Discrete Mathematics* **24.4** (2010), 1495–1500.

[20] M. Krivelevich, The critical bias for the Hamiltonicity game is $(1 + o(1))n/\ln n$, *Journal of the American Mathematical Society* **24.1** (2011), 125–131.

[21] A. Lehman, A solution of the Shannon switching game, *Journal of the Society for Industrial and Applied Mathematics* **12.4** (1964), 687–725.

[22] R. Montgomery, Spanning trees in random graphs, *Advances in Mathematics* **356** (2019), 106793, 1–92.

[23] R. Montgomery, A. Pokrovskiy and B. Sudakov, A proof of Ringel's conjecture, *Geometric and Functional Analysis* **31.3** (2021), 663–720.

[24] R. Nenadov, Probabilistic intuition holds for a class of small subgraph games, *Proceedings of the American Mathematical Society* **151.04** (2023), 1495–1501.

Random lifts of very high girth and their applications to frozen colourings*

Guillem Perarnau[†1,2] and Giovanne Santos[‡3]

[1]Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona.
[2]Centre de Recerca Matemàtica, Bellaterra, Spain.
[3]Departamento de Ingeniería Matemática, Universidad de Chile, Santiago, Chile.

Abstract

We study the cycle distribution of a random n-lift of a fixed d-regular graph on m vertices, deriving an asymptotic formula for the probability that it has girth at least $g = g(n)$, provided that $g(n)$ grows sufficiently slowly with respect to m, d and n. As a consequence of the existence of lifts with high girth, we construct graphs with very large girth that admit frozen colourings, and graphs with moderately large girth where typical colourings are partially-frozen. The latter result shows the tightness on the girth condition of a recent theorem on graph colouring rigidity by Hurley and Pirot [STOC, 2023].

1 Introduction

An *n-lift* of a graph G is a graph $L = L_n(G)$ with vertex set $V(L) := V(G) \times [n]$ and edge set obtained as follows: for every edge $uv \in E(G)$, we place a perfect matching between the sets $\{u\} \times [n]$ and $\{v\} \times [n]$. Let $\mathcal{L}_n(G)$ be the set of all n-lifts of a graph G.

A *random n-lift of G*, denoted by $\mathbb{L}_n(G)$, is an n-lift of G chosen uniformly at random from $\mathcal{L}_n(G)$. It is worth noticing that we may generate $\mathbb{L}_n(G)$ by choosing each perfect matching corresponding to an edge in G, independently and uniformly at random.

Random lifts of graphs were introduced by Amit and Linial in 2002 [2, 3] and since then, they have attracted a lot of interest in the area. Among other works, we highlight the results of Amit, Linial and Matoušek [4] on their independent and chromatic numbers, the results of Greenhill, Janson and Ruciński [10] on their number of perfect matchings, and the work of Bordenave [6] on their spectral properties.

Fortin and Rudinsky [8] studied the distribution of short cycles in random lifts. Given a subgraph $H \subseteq L$, its *pattern* is the multigraph on $V(G)$ obtained by adding an edge (u, v) for every edge $(u, x)(v, y) \in E(H)$. The following observation is key to study the cycles of random lifts

If C is a k-cycle of $L \in \mathcal{L}_n(G)$, then the pattern of C is a closed non-backtracking k-walk in G.

*This project was initiated during a research stay of GS at Universitat Politècnica de Catalunya supported by the Research and Innovation Staff Exchange (Horizon 2020) *RandNET: Randomness and learning in networks (MSCA-RISE-2020-101007705).*

[†]Email: guillem.perarnau@upc.edu. Research of GP supported by the Grant PID2020-113082GB-I00, the Grant RED2022-134947-T and the Programme Severo Ochoa y María de Maeztu por Centros y Unidades de Excelencia en I&D (CEX2020-001084-M), all of them funded by MICIU/AEI/10.13039/501100011033.

[‡]Email: gsantos@dim.uchile.cl. Research of GS supported by ANID Becas/Doctorado Nacional 21221049.

Let $w_k(G)$ be the number of closed non-backtracking k-walks in G. For all $k \geq 3$, we let

$$\lambda_k(G) := \frac{w_k(G)}{2k}.$$

and

$$\mu_k(G) := \sum_{\ell=3}^{k-1} \lambda_\ell(G).$$

Theorem 1 ([8])**.** *Let $n \in \mathbb{N}$ and $d \geq 3$, and let G be a d-regular graph. For any $k \geq 3$, let $Z_{k,n}$ be the number of cycles of length k in $\mathbb{L}_n(G)$. Let Z_k be independent random variables with Poisson distribution of parameter $\lambda_k(G)$ respectively. Then $(Z_{k,n})_{k\geq 3} \to (Z_k)_{k\geq 3}$ in distribution as $n \to \infty$.*

For any graph H, let $g(H)$ be its *girth*; the length of a shortest cycle. The previous result implies that, for every $g_0 \geq 3$,

$$\lim_{n\to\infty} \mathbb{P}(g(\mathbb{L}_n(G)) \geq g_0) = e^{-\mu_{g_0}(G)} > 0. \tag{1}$$

The qualitative behaviour of short cycles in random lifts is the same as for other random graph models, such as Erdős-Rényi random graphs or random regular graphs (see e.g. [9]). In the case of random regular graphs, McKay, Wormald and Wysocka [12] went a step further and studied the distribution of long cycles. As a corollary, they obtained an enumeration formula for d-regular graphs on n vertices with girth at least g, provided that $(d-1)^{2g-3} = o(n)$.

Our main result extends (1) in the line of [12], allowing for the girth $g(n)$ to tend to infinity when $n \to \infty$, provided it does it sufficiently slowly with respect to the other parameters.

Theorem 2. *Let $n \in \mathbb{N}$, $d = d(n) \geq 3$, $m = m(n)$ and $g = g(n)$ such that $m(d-1)^{2g-4} = o(n)$. If G is a d-regular graph on m vertices, then,*

$$\mathbb{P}(g(\mathbb{L}_n(G)) \geq g(n)) \sim e^{-\mu_{g(n)}(G)}. \tag{2}$$

An immediate corollary of our main theorem is the existence of lifts of any fixed regular graph G with very high girth.

Corollary 3. *Let $n \in \mathbb{N}$, $d = d(n) \geq 3$, $m = m(n)$ and $g = g(n)$ such that $m(d-1)^{2g-4} = o(n)$. If G is a d-regular graph on m vertices, then, for any sufficiently large n, there exists $L \in \mathcal{L}_n(G)$ with $g(L) \geq g(n)$.*

The condition on the parameters is not far from optimal. Recall Moore's bound for odd girth: the number of vertices of any d-regular graph with girth at least $g = 2s+1$ is

$$n \geq 1 + d\sum_{i=1}^{s-1}(d-1)^i \geq (d-1)^{(g-1)/2}$$

and thus the restriction on the girth is tight up to a constant factor.

1.1 Ideas of the proof

To exemplify the ideas behind the proof of Theorem 2, we give a sketch of a direct proof of Corollary 3 that does not use the theorem. The approach is a combination of the second moment technique and the switching method, that we now detail.

One of the most important aspects is to control the expected number of appearances of k-cycles (and other subgraphs) in random lifts, which is done using the two lemmas below.

Lemma 4. *For every $k \geq 3$ satisfying $m(d-1)^{2k-4} = o(n)$, if X_k is the number of k-cycles in $\mathbb{L}_n(G)$, then*

$$\mathbb{E}(X_k) \sim \lambda_k(G).$$

A graph H is *feasible* if its vertex set is a subset of $V(G) \times [n]$ and there exists $L \in \mathcal{L}_n(G)$ with $H \subseteq L$.

Lemma 5. *Let H be a connected feasible graph on h vertices and e edges. Let Y_H be the number of subgraphs isomorphic to H in $\mathbb{L}_n(G)$. If $e = o(n^{1/2})$, then*

$$\mathbb{E}(Y_H) = O(md^{h-1}n^{h-e}).$$

We use these two results in the next lemma, which is proved by an application of the second moment method. Crucially, we also use an upper bound on the number of subgraphs H that can be obtained from the union of two cycles, that depends on the number of components and the number of edges of the intersection graph of the two cycles, that was derived in [12].

Lemma 6. *Let $s_k := \max\{2\lambda_k(G), \log^2 n\}$. Then,*

$$\mathbb{P}(X_k > s_k, \text{ for some } 3 \leq k < g) = o(1).$$

Moreover, the probability that there are two cycles of length shorter than g that share at least one edge is $o(1)$.

With the previous lemma in hand, we give a proof of the existence of a lift of G with no cycles of length less than g (short cycles), that we now sketch.

Let L_0 be a graph that satisfies the conclusions of Lemma 6 (few short cycles and all of them edge-disjoint). A key property is that the number of vertices participating in short cycles is at most

$$\sum_{k=3}^{g-1} ks_k = o(n).$$

In the classical argument of Erdős to find graphs with large girth and large chromatic number (see e.g. [1]), an arbitrary vertex of each short cycle is deleted, which enforces the girth to be at least g. However, since we want to maintain the property that the final graph is a lift of G, we need to find an alternative way to get rid of the short cycles of L_0. The idea will be to use a switching-type argument to destroy them one by one, while keeping the structure of a lift. In doing so, we will strongly use that the cycles are edge-disjoint.

A *switch* on $L \in \mathcal{L}_n(G)$ is a local transformation defined as follows: Let $uv \in E(G)$ and $x_1, x_2, y_1, y_2 \in [n]$ such that $(u, x_1)(v, y_1)$ and $(u, x_2)(v, y_2)$ are edges of L, and $(u, x_1)(v, y_2)$ and $(u, x_2)(v, y_1)$ are not. Then, we delete the former two edges from L and add the latter two. Observe that the resulting graph is also in $\mathcal{L}_n(G)$.

Given an edge $e = (u, x_1)(v, y_1)$ and a cycle C of L with $e \in E(C)$, we say that $f = (u, x_2)(v, y_2)$ is (e, C)-*good* if and only if:

(i) f is not in a short cycle of L, and

(ii) $\text{dist}(f, c) \geq g$ for every $c \in V(C)$.

Starting with L_0, we construct a sequence of lifts $L_0, L_1, L_2, \ldots$ such that every lift has less short cycles than the previous one. To do so, at step i we choose any cycle C_i of L_i and any edge $e_i \in E(C_i)$. Then, we choose f_i to be a (e_i, C_i)-good edge of L_i and we switch e_i and f_i. The rest of the proof consists on showing that: (1) after the switch the number of short cycles in the resulting graph has decreased, and (2) every pair (e_i, C_i) has at least one good edge f_i.

2 Existence of large girth graphs with frozen and partially-frozen colourings

Beyond the existence of lifts with large girth, our result have implications in Graph Colouring. Let G be a graph and $m \in \mathbb{N}$. The *m-recolouring graph* of G, denoted by $R_m(G)$, is the graph whose vertices are the proper m-colourings of G and two colourings are adjacent if they differ at exactly one vertex. An isolated vertex in $R_m(G)$ is called a *frozen colouring* of G, and can be understood as a colouring that admits no single-vertex recolouring that keeps its properness. The frozen terminology comes from Glauber dynamics on colourings, a Markov chain with state space $R_m(G)$ used to sample almost uniform m-colourings of G. In the dynamics, frozen states correspond to absorbing states of the chain, and impede the chain to converge to the uniform distribution. It is thus interesting to study under which conditions, such colourings may appear.

Let us first review some structural properties of $R_m(G)$. If G has maximum degree Δ, a necessary condition for the existence of frozen colourings is $m \leq \Delta + 1$. In particular, if $m = \Delta + 1$, then the graph G must be d-regular, where $d = \Delta$. In this abstract, we will restrict ourselves to the case where G is a d-regular graph and $m = d + 1$.

Feghali, Johnson and Paulusma [7] proved that $R_{d+1}(G)$ is composed of a unique connected component of size at least 2 and a number of isolated vertices (frozen colourings). Bonamy, Bousquet, and the first author [5] studied the fraction of vertices that are isolated in $R_{d+1}(G)$: when G is a large connected graph, the number of frozen colourings is exponentially smaller than the total number of colourings. This justifies that, even though the Glauber dynamics might not be irreducible, it can still be used to sample almost uniform $(d+1)$-colourings of G.

Observe that a d-regular graph G on N vertices has a frozen colouring if and only if G is isomorphic to an n-lift of K_{d+1}, the complete graph on $d+1$ vertices, for $n(d+1) = N$. For the "if" part, one can obtain a frozen colouring of G by colouring each vertex with the corresponding vertex from K_{d+1}. For the "only if" part, any frozen colouring splits the vertex set of G into $d+1$ independent sets of equal size, in this case n. By a simple counting argument, there are n edges within any two sets, and by the frozen condition, they form a matching. Together with (1), this shows the existence of graphs of large girth that admit a frozen colouring. As a consequence of Corollary 3, we obtain the following.

Corollary 7. *Let $n \in \mathbb{N}$, $d = d(n) \geq 3$ and $g = g(n)$ such that $(d-1)^{2g-3} = o(n)$. Then there exists a d-regular graph on n vertices and girth at least g that admits a frozen $(d+1)$-colouring.*

Recently, Hurley and Pirot [11] studied uniformly random proper m-colourings of sparse graphs with maximum degree d in the regime $d < m \log m$. Sparsity in this setting is controlled by the girth: the larger the girth, the less density of edges in local neighbourhoods. The main concern of their paper is to understand the *shattering threshold*, the minimum number of colours that are needed for $R_m(G)$ to resemble $R_m(\mathbb{G}_{n,d})$, where $\mathbb{G}_{n,d}$ is a random d-regular graph. In this direction, they proved that a typical m-colouring of a large girth graph is not "rigid" in the following sense.

Theorem 8 ([11])**.** *Let $\epsilon > 0$ and $m \in \mathbb{N}$ large enough such that $d < (1-\epsilon)m \ln m$. If G is a graph on n vertices, maximum degree d and girth at least $\ln \ln n$, then a uniformly random proper m-colouring σ of G satisfies w.h.p.[1], for all $v \in V(G)$*

(i) for all $j \in [m]$, there exists a colouring τ with $\tau(v) = j$, that differs at $O(\log^2 n)$ vertices with σ.

(ii) for all $j \in [m]$, the component of σ in $R_m(G)$ contains a colouring τ with $\tau(v) = j$.

Properties (i) and (ii) deal with the geometry of the solution space (of colourings) and are also shared with colourings of random graphs. In this direction, a natural problem is to determine which are the minimum sparsity requirements on G that ensure such properties hold. Hurley and Pirot showed that the condition $g(G) \geq \ln \ln n$ cannot be replaced by $g(G) \geq C$, for any constant $C > 0$. Here, indeed, we show that lower bound on the girth required in Theorem 8 is essentially optimal, even for $m = d + 1$.

[1] We say that a property holds *with high probability (w.h.p.)* if the probability it holds tends to 1 as $n \to \infty$.

Given an m-colouring σ of G and $v \in V(G)$, following [11], we say that v is *frozen* in σ if $\tau(v) = \sigma(v)$ for all τ in the same component of $R_m(G)$ as σ. Note that if v is frozen, then condition (ii) is not satisfied.

Proposition 9. *For every $\gamma > 0$, $d \geq 3$ and sufficiently large n_0, there exists $n \geq n_0$ and a d-regular graph G on n vertices of girth at least $\left(\frac{1}{2\ln(d-1)} - \gamma\right) \ln\ln n$ with the following property: if σ is a uniformly random proper $(d+1)$-colouring of G, w.h.p. σ has at least $n^{1-o(1)}$ frozen vertices.*

We include the proof of this proposition which is a simple application of our previous results.

Proof of Proposition 9. Let $\delta > 0$ be sufficiently small with respect to γ and d. Let $g = (1/2 - \delta)\log_{d-1} n_0$. By Corollary 7, there exists a d-regular graph G_0 on n_0 vertices of girth at least g that has a frozen colouring.

Let $\epsilon > 0$ be sufficiently small and fix n the smallest multiple of n_0 such that $n^\epsilon \geq (d+1)^{n_0}$. As δ has been chosen small enough, we have that

$$g(G_0) \geq \left(\frac{1}{2\ln(d-1)} - \gamma\right) \ln\ln n.$$

Let $k = n/n_0$ and let G be the graph composed of k vertex-disjoint copies of G_0. The uniform probability space over $(d+1)$-colourings of G is a product space of k uniform and independent probability spaces over $(d+1)$-colourings of G_0. Since G_0 admits at least one frozen $(d+1)$-colouring, the probability a uniform random $(d+1)$-colouring of G_0 is frozen is at least $p := (d+1)^{-n_0}$. It follows that the number of frozen $(d+1)$-colourings in the copies of G_0 in G stochastically dominates a Binomial random variable with k trials and probability p. By Chernoff inequality, w.h.p. the number of copies of G_0 where σ induces a frozen colourings is at least

$$\frac{k}{2(d+1)^{n_0}},$$

and the number of frozen vertices is at least

$$n_0 \cdot \frac{k}{2(d+1)^{n_0}} \geq \frac{n^{1-\epsilon}}{2}.$$

Since the choice of $\epsilon > 0$ is arbitrary, we conclude the proof of the proposition. □

Acknowledgements

The authors would like to thank François Pirot for suggesting the question on partially-frozen graphs with very large girth, that motivated the study of large long cycles in random lifts. The authors also thank the two anonymous reviewers for their comments that helped improving the article.

References

[1] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016.

[2] Alon Amit and Nathan Linial. Random graph coverings. I. General theory and graph connectivity. *Combinatorica*, 22(1):1–18, 2002.

[3] Alon Amit and Nathan Linial. Random lifts of graphs II: Edge expansion. *Combinatorics Probability and Computing*, 14:317–332, 2006.

[4] Alon Amit, Nathan Linial, and Jiří Matoušek. Random lifts of graphs: independence and chromatic number. *Random Structures Algorithms*, 20(1):1–22, 2002.

[5] Marthe Bonamy, Nicolas Bousquet, and Guillem Perarnau. Frozen $(\Delta + 1)$-colourings of bounded degree graphs. *Combin. Probab. Comput.*, 30(3):330–343, 2021.

[6] Charles Bordenave. A new proof of friedman's second eigenvalue theorem and its extension to random lifts. In *Annales Scientifiques de l'École Normale Supérieure*, volume 4, pages 1393–1439, 2020.

[7] Carl Feghali, Matthew Johnson, and Daniël Paulusma. A reconfigurations analogue of Brooks' theorem and its consequences. *J. Graph Theory*, 83(4):340–358, 2016.

[8] Jean-Philippe Fortin and Samantha Rudinsky. Asymptotic eigenvalue distribution of random lifts. *The Waterloo Mathematics Review*, pages 20–28, 2013.

[9] Alan Frieze and Michal Karonski. *Introduction to Random Graphs*. Cambridge University Press, Cambridge, 2016.

[10] Catherine Greenhill, Svante Janson, and Andrzej Ruciński. On the number of perfect matchings in random lifts. *Combin. Probab. Comput.*, 19(5-6):791–817, 2010.

[11] Eoin Hurley and François Pirot. Uniformly random colourings of sparse graphs. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. ACM, jun 2023.

[12] Brendan D. McKay, Nicholas C. Wormald, and Beata Wysocka. Short cycles in random regular graphs. *Electron. J. Combin.*, 11(1):Research Paper 66, 12, 2004.

A covering problem for zonotopes and Coxeter permutahedra

Gyula Károlyi*[1,2]

[1]HUN–REN Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13–15, H–1053 Budapest, Hungary
[2]Department of Algebra and Number Theory, Eötvös University, Pázmány P. sétány 1/C, H–1117 Budapest, Hungary

Abstract

An almost cover of a finite set in the affine space is a collection of hyperplanes that together cover all points of the set except one. According to the Alon-Füredi theorem, every almost cover of the vertex set of an n-dimensional cube requires at least n hyperplanes. Here we investigate a possible generalization of this result to Coxeter permutahedra: convex polytopes whose vertices form the orbit of a generic point under the action of a finite reflection group.

1 Introduction

An almost cover of a finite set in the affine space is a collection of hyperplanes that together cover all points of the set except one. According to a classical result of Jamison [11], an almost cover of the n-dimensional affine space over the q-element finite field requires at least $(q-1)n$ hyperplanes. Equivalently, to pierce every affine hyperplane in $\mathbb{F}_q^n$ one needs at least $(q-1)n+1$ points, see [5]. See also [4] for further results in finite geometries. Another example is the Alon–Füredi theorem [2]: *Every almost cover of the vertex set of an n-dimensional cube requires at least n hyperplanes.*

Consider those points in the n-dimensional space whose coordinates form a permutation of the first n positive integers. The elements of this set P_n are the vertices of a convex $(n-1)$-dimensional polytope called the permutahedron (spelled also as permutohedron) Π_{n-1}. For $n=3$ it is a regular hexagon, for $n=4$ a truncated octahedron. This polytope has many fascinating properties and can be used to illustrate various concepts in geometry, combinatorics and group theory, see [13]. Our starting point is the following analogue of the Alon–Füredi theorem observed by Hegedüs, see [8].

Theorem 1. *Every almost cover of the vertices of Π_{n-1} consists of at least $\binom{n}{2}$ hyperplanes. This bound is sharp.*

A zonotope is a convex polytope that can be represented as the Minkowski sum of a finite number of line segments. A collection of line segments is called nondegenerate if no two of the segments are parallel to each other. Each zonotope Z can be written as the Minkowski sum of a nondegenerate collection of line segments, unique up to translations. The number of the summands, denoted by $\mathrm{rk}(Z)$, we call the rank of Z. In [8] we suggested that the above result and the Alon–Füredi theorem must be representatives in a more general famework.

Conjecture 2. *Every almost cover of the vertices of a zonotope Z consists of at least $\mathrm{rk}(Z)$ hyperplanes.*

*Email: karolyi.gyula@renyi.hu

Apart from some small examples, all zonotopes for which we were able to verify this hypothesis turned out to be Coxeter permutahedra. Our purpose here is to initiate a systematic study of the almost covers of their vertex sets based on a polynomial method colloquially referred to as the application of the Combinatorial Nullstellensatz.

We express our gratitude to Günter M. Ziegler for identifying one of our first examples as a permutahedron of type B, and to Francesco Santos for drawing the beautifully illuminating paper [6] of Fomin and Reading to our attention. For additional background information we refer to [9, 10].

2 Two elementary examples

The 2-dimensional zonotopes of rank r are exactly the centrally symmetric convex $2r$-gons, and every almost cover of such a polygon with lines requires at least r lines. There are two types of them that occur as zonotopal Coxeter permutahedra: regular $2r$-gons and equiangular $2r$-gons (r even) with alternating edge lengths. (The vertices of) any prism over such polygons have almost covers of size $r+1$. An elementary argument using a simple modular invariant reveals that r planes do not suffice.

Theorem 3. *Let Z be a prism over a regular $2n$-gon. Then every almost cover of the vertices of Z consists of at least $\mathrm{rk}(Z) = n+1$ planes.*

Theorem 4. *Let Z be a prism over an equiangular $4n$-gon having alternating edge lengths. Then every almost cover of the vertices of Z consists of at least $\mathrm{rk}(Z) = 2n+1$ planes.*

3 The polynomial toolkit

The Combinatorial Nullstellensatz, formulated by Noga Alon in the late nineties, describes, in an efficient way, the structure of multivariate polynomials whose zero-set includes a Cartesian product over a field $\mathbb{F}$. This characterization immediately implies ([1]) the first part of the following theorem.

Theorem 5. *Let $S_1, \dots, S_n$ be subsets of $\mathbb{F}$, $|S_i| = k_i$, and let f be a polynomial in $\mathbb{F}[x] = \mathbb{F}[x_1, \dots, x_n]$ whose degree is at most $\sum_{i=1}^n (k_i - 1)$.*

(i) If $f(s) = 0$ for every $s \in S_1 \times \cdots \times S_n$, then the coefficient of the monomial $\prod_{i=1}^n x_i^{k_i-1}$ in f is zero.

(ii) If $f(s) = 0$ for all but one element $s \in S_1 \times \cdots \times S_n$, then the coefficient of the monomial $\prod_{i=1}^n x_i^{k_i-1}$ in f is not zero.

The second part can be derived directly from (i) rather easily and is contained implicitly in many works, e.g. it is a very special case of Corollary 4.2 in [3]. The result has innumerable variations with even more different proofs, see e.g. [12]. Apparently they all depend on two basic principles: reduction modulo a standard Gröbner basis and Lagrange interpolation. It also implies the following immediate consequence of Theorem 5 in [2] we find particularly useful for the present work.

Theorem 6. *Let $S_1, \dots, S_n$ be nonempty subsets of $\mathbb{F}$, $B = S_1 \times \cdots \times S_n$. If a polynomial $f \in \mathbb{F}[x_1, \dots, x_n]$ vanishes at every point of B except one, then its degree is at least $\sum_{i=1}^n (|S_i| - 1)$.*

For a polynomial $f \in \mathbb{R}[x_1, \dots, x_n]$ set $V(f) = \{a \in \mathbb{R}^n \mid f(a) = 0\}$; it is called a hypersurface of degree $\deg f$. Note that the union of m hyperplanes is a hypersurface of degree m. Thus an almost cover of $X \subseteq \mathbb{R}^n$ is a hypersurface satifying $X \setminus \{v\} \subseteq V(f)$, $v \notin V(f)$ for some $v \in X$ and a polynomial f that splits into linear factors over $\mathbb{R}$. For an arbitrary hypersurface $V(f)$ satisfying the above two conditions for X and v we say that it is an almost cover of X: it covers every point of X except v. Throughout this work we are going to employ the following consequence of Theorem 6.

Corollary 7. *Let $\emptyset \neq X \subseteq B = S_1 \times \cdots \times S_n \subseteq \mathbb{R}^n$, $f \in \mathbb{R}[x_1, \ldots, x_n]$ and $d = (\sum_{i=1}^n |S_i|) - n - \deg f$. If $X = B \setminus V(f)$, then every hypersurface which is an almost cover of X has degree at least d.*

For example, the Alon–Füredi theorem follows with the choice $S_i \equiv \{0, 1\}$, $X = B$, $f = 1$. For the first statement in Theorem 1 one can use $S_i \equiv \{1, 2, \ldots, n\}$, $X = P_n$, $f = \prod_{1 \leq i < j \leq n}(x_j - x_i)$.

4 Prisms over permutahedra

Here we demonstrate how Theorem 5 can be used via a polynomial invariant to verify Conjecture 2 for prisms over permutahedra. Because of affine invariance it is enough to prove it for the prism whose bases are Π_{n-1} and $-\Pi_{n-1} = \Pi_{n-1} - (n+1)(e_1 + \cdots + e_n)$, where $e_1, \ldots, e_n$ is the standard orthonormal basis for $\mathbb{R}^n$.

Theorem 8. *Every almost cover of $P_n \cup (-P_n)$ consists of at least $\binom{n}{2} + 1$ hyperplanes.*

Proof. Let $m = \binom{n}{2}$ and suppose that the hyperplanes H_i, $1 \leq i \leq m$ cover every point of $P_n \cup (-P_n)$ except v. By symmetry, we may assume that $v \in -P_n$. The hyperplane H_i is defined by an equation $f_i(x) = a_i$ where f_i is a linear form. Consider the Vandermonde polynomial $V(x) = \prod_{i<j}(x_j - x_i)$. The polynomial

$$f(x) = V(x) \prod_{i=1}^{m} (f_i(x) - a_i))$$

of degree $n(n-1)$ vanishes at every point of the Cartesian product $\{1, 2, \ldots, n\}^n$. By Theorem 5 (i), the coefficient of the monomial $\prod_{i=1}^n x_i^{n-1}$ in f must be zero.

On the other hand, the polynomial f attains the value 0 at every point of the Cartesian product $\{-1, -2, \ldots, -n\}^n$ except v. That is, the polynomial

$$g(x) = f(-x) = (-1)^{\binom{n}{2}} V(x) \prod_{i=1}^{m} (-f_i(x) - a_i)) = V(x) \prod_{i=1}^{m} (f_i(x) + a_i))$$

of degree $n(n-1)$ vanishes at every point of the Cartesian product $\{1, 2, \ldots, n\}^n$ except $-v$. By Theorem 5 (ii), the coefficient of the monomial $\prod_{i=1}^n x_i^{n-1}$ in g must be nonzero. Since the degree $n(n-1)$ parts of the polynomials f and g are identical, we arrive at a contradiction. □

5 Reflection groups, root systems and Coxeter permutahedra

Let V be an n-dimensional real euclidean space with orthonormal basis $e_1, \ldots, e_n$. Here and in what follows we identify the vectors of V with the points of $\mathbb{R}^n$. For a nonzero vector $\alpha \in V$ we denote by s_α the orthogonal reflection in the linear hyperplane H_α orthogonal to α. Thus, $s_\alpha(\alpha) = -\alpha$. A finite reflection group acting on V is any finite group generated by (a nonempty set of) such reflections. A root system Φ is a set of nonzero vectors satisfying $\Phi \cap \mathbb{R}\alpha = \{-\alpha, \alpha\}$ and $s_\alpha(\Phi) = \Phi$ for every $\alpha \in \Phi$. Crytallographic root systems satisfy an extra integrality condition. The group $W(\Phi)$ (called Weyl group in the crystallographic case) of orthogonal transformations generated by the reflections s_α, $\alpha \in \Phi$ is always a finite reflection group in which the reflections exhaust Φ. Thus, Φ is invariant under the action of W. Conversely, if W is a finite reflection group, then the unit vectors α for which $s_\alpha \in W$ form a root system Φ for which $W = W(\Phi)$. If the vectors in Φ form one orbit under the action of W, then $W = W(\Phi')$ if and only if $\Phi' = c\Phi$ for some $0 \neq c \in \mathbb{R}$. On the other hand, if Φ is the union of more than one orbits, then the common length of the vectors in an orbit may be scaled arbitrarily for each orbit. Thus, if $W = \mathrm{I}_2(m)$ is the symmetry group of a regular m-gon centered at the origin, then each corresponding root system has $2m$ elements, which form one orbit if m is odd and splits into two orbits of equal size if m is even.

Let $W = W(\Phi)$ be a finite reflection group. For any point $a \in \mathbb{R}^n$, consider its orbit $W(a)$. The point a is called *generic* with respect to W, if $|W(a)| = |W|$, or equivalently, $a \notin \bigcup_{\alpha\in\Phi} H_\alpha$. In this case $W(a)$ is the vertex set of a (not necessarily full dimensional) convex polytope $\Pi W(a)$, referred to as a *W-permutahedron*, or a Coxeter permutohedron of type W. Thus, a permutahedron of type $\mathrm{I}_2(m)$ is either a regular $2m$-gon, or an equiangular $2m$-gon with alternating edge lengths (the latter being a zonotope only for m even), and each such polygon centered at the origin can be obtained as a Coxeter permutahedron for an appropriate choice of Φ. All vertices except one can be covered by m, but not less lines.

A root system Φ is irreducible if it cannot be partitioned into two subsets lying in two nontrivial orthogonal complements of V, or equivalently, if $W(\Phi)$ is not the direct sum of two proper subgroups acting as reflection groups on two such subspaces. Theorems 3 and 4 thus read as follows: *Every almost cover of a zonotopal permutahedron of type* $\mathrm{I}_2(m) \oplus \mathrm{A}_1$ *requires at least* $m + 1$ *hyperplanes.* Note that the group contains exactly $m + 1$ reflections.

Next consider the reflection group A_{n-1} acting on $\mathbb{R}^n$, generated by the reflections in the hyperplanes of equation $x_{i+1} = x_i$, $i = 1, \dots, n-1$. It is isomorphic to the symmetric group S_n, and a point is generic if and only if all its coordinates are different. Thus we have $\Pi_{n-1} = \Pi\mathrm{A}_{n-1}(1, 2, \dots, n)$, and Thm 1 coupled with the remark following its proof in [8] can be read as follows: *Every almost cover of the vertices of a Coxeter permutahedron of type* A_{n-1} *consists of at least* $\binom{n}{2}$ *hyperplanes.* The bound is also sharp. Note that the vectors $e_i - e_j$ $(i \neq j)$ form a root system for A_{n-1}, so the bound equals the number of reflections contained in A_{n-1}. In general, for a reflection group $W = W(\Phi)$, the number of reflections contained in W is $N(W) = |\Phi|/2$.

It is not difficult to prove an analogue of Thm 1 for permutahedra of type B. The hyperoctahedral group B_n acting on $\mathbb{R}^n$ is generated by the reflections in the hyperplanes of equation $x_{i+1} = x_i$, $i = 1, \dots, n-1$, together with the reflection in the hyperplane $x_1 = 0$; it contains A_{n-1} as a subgroup. Altogether it contains n^2 reflections in the hyperplanes $x_i = \varepsilon x_j$ $(1 \leq i < j \leq n,\ \varepsilon = \pm 1)$ and $x_i = 0$ $(1 \leq i \leq n)$. Thus, $N(\mathrm{B}_n) = n^2$. A point $a = (a_1, \dots, a_n)$ is generic if and only if $a_i \neq 0$ for all i and $|a_i| \neq |a_j|$ for all $i \neq j$. Thus every orbit of a generic point is of the form

$$B_n(a) = \{\varepsilon_1 a_{\pi(1)} + \cdots + \varepsilon_n a_{\pi(n)} \mid \varepsilon_i = \pm 1, \pi \in S_n\}$$

for some $a \in \mathbb{R}^n$ with coordinates $0 < a_1 < \cdots < a_n$.

Theorem 9. *Every almost cover of the vertices of a Coxeter permutahedron of type* B_n *consists of at least* n^2 *hyperplanes. This bound is sharp.*

Proof. The vertex set of the permutahedron $\Pi\mathrm{B}_n(a)$ with $0 < a_1 < \cdots < a_n$ is contained in the Cartesian product $S_1 \times \cdots \times S_n$ where $S_i = \{a_i, -a_i \mid 1 \leq i \leq n\}$, and each point in $(S_1 \times \cdots \times S_n) \setminus \mathrm{B}_n(a)$ is a root of the polynomial

$$f(x) = \prod_{1\leq i<j\leq n} (x_j - x_i)(x_j + x_i)$$

of degree $n(n-1)$. According to Corollary 7, every almost cover of $\mathrm{B}_n(a)$ consists of at least

$$(\sum_{i=1}^{n} |S_i|) - n - \deg f = 2n^2 - n - n(n-1) = n^2$$

hyperplanes. To see that the bound cannot be improved, notice that the hyperplanes $x_i = a_j$ $(i < j)$, $x_i = -a_j$ $(i \leq j)$ cover every vertex but $a = (a_1, a_2, \dots, a_n)$. □

The study of almost covers of the vertices of permutahedra of type D is more subtle. The group D_n is the subgroup of index 2 in B_n generated by the reflections in the hyperplanes of equation $x_{i+1} = x_i$, $i = 1, \dots, n-1$, together with the reflection in the hyperplane $x_2 = -x_1$. Altogether it contains

$n(n-1)$ reflections in the hyperplanes $x_i = \varepsilon x_j$ ($1 \le i < j \le n$, $\varepsilon = \pm 1$). A point $a = (a_1, \dots, a_n)$ is generic if and only if $|a_i| \neq |a_j|$ for all $i \neq j$. Thus every orbit of a generic point is of the form

$$\mathrm{D}_n(a) = \{\varepsilon_1 a_{\pi(1)} + \cdots + \varepsilon_n a_{\pi(n)} \mid \pi \in S_n, \varepsilon \in E\}$$

for some $a \in \mathbb{R}^n$ with coordinates $-a_2 < a_1 < a_2 < \cdots < a_n$, where E is either of the two subsets of $\{-1, 1\}^n$ that consists of all vectors in which the number of -1 coordinates are the same modulo 2.

Theorem 10. *Every almost cover of the vertices of a Coxeter permutahedron of type* D_n *consists of at least* $n(n-1)$ *hyperplanes. This bound is sharp in the following sense: if* a *is a generic point one of whose coordinates is 0, then* $\mathrm{D}_n(a)$ *has an almost cover of size* $n(n-1)$*.*

Proof. It is very similar to the previous one if the vertices of the permutahedron have a 0 coordinate. Otherwise we may assume by symmetry that the vertex set is $\mathrm{D}_n(a)$ with $0 < a_1 < \cdots < a_n$. In this case we can apply Corollary 7 with the polynomial

$$f(x) = \prod_{1 \le i < j \le n} (x_j - x_i)(x_j + x_i) \left(\prod_{i=1}^{n} x_i + \prod_{i=1}^{n} a_i \right)$$

of degree n^2. □

These results suggest that the following might be true.

Conjecture 11. *For a finite reflection group* W*, every almost cover of the vertices of a permutahedron of type* W *consists of at least* $N(W)$ *hyperplanes.*

In contrast, all vertices of a Coxeter permutahedron are contained in a single hypersurface of degree 2, namely a sphere centered at the origin.

6 Zonotopal permutahedra

For the reflection group $W = \mathrm{A}_n$, the orbit of any generic point contains a unique point $a = (a_1, \dots, a_{n+1})$ with $a_1 < \cdots < a_{n+1}$. Similarly, for $W = \mathrm{B}_n$, the orbit of any generic point contains a unique point $a = (a_1, \dots, a_n)$ with $0 < a_1 < \cdots < a_n$. For such points it is known that the Coxeter permutahedron $\Pi W(a)$ is a zonotope if and only if the coordinates a_i form an arithmetic progression, see [7, Thm 4.13]. We can prove an analogous statement for permutahedra of type D, and in fact all these results can be viewed as special cases of a more general phenomenon. For a root system Φ, consider any set Φ^+ of positive roots. The Minkowski sum of the line segments $[-\alpha/2, \alpha/2]$, $\alpha \in \Phi^+$, independent of the choice of Φ^+ we denote by $Z(\Phi)$. Then $\mathrm{rk}(Z(\Phi)) = N(W(\Phi))$.

Theorem 12. *Let* W *be a finite reflection group with a corresponding root system* Φ*. Then* $Z(\Phi)$ *is a permutahedron of type* W*.*

The reflection group W is called essential if it acts on V without nonzero fixed points. In general, $V = U \oplus U'$, where W is essential relative to U and the orthogonal complement U' consists of all fixed points of W.

Theorem 13. *A permutahedron* Π *of type* W *is a zonotope if and only if there exists a root system* Φ *with* $W(\Phi) = W$ *and a vector* $u \in U'$ *such that* $\Pi = Z(\Phi) + u$*.*

Although it is not likely that Conjectures 2 and 11 for $Z(\Phi)$ in general can be attacked by our methods, it is possible to say something more for crystallographic root systems. We call a zonotope $Z \subset \mathbb{R}^n$ *special* if there exist finite sets $S_1, \dots, S_n \subset \mathbb{R}$ and a polynomial f such that the vertex set X of Z is $(S_1 \times \cdots \times S_n) \setminus V(f)$ and

$$\mathrm{rk}(Z) \le |S_1| + \cdots + |S_n| - n - \deg f.$$

According to Corollary 7, every almost cover of the vertices of a special zonotope Z consists of at least $\mathrm{rk}(Z)$ hyperplanes. Now for an irreducible crystallographic root system Φ, $Z(\Phi)$ is special if the type of Φ is $\mathrm{A}_n, \mathrm{B}_n, \mathrm{C}_n, \mathrm{D}_n$ or G_2. Moreover, if V is the sum of the orthogonal subspaces V_1, V_2 and $\Phi = \Phi_1 \cup \Phi_2$ with $\Phi_i = \Phi \cap V_i$, then $Z(\Phi)$ is the product polytope $Z(\Phi_1) \times Z(\Phi_2)$. In general, $\mathrm{rk}(Z_1 \times Z_2) = \mathrm{rk}(Z_1) + \mathrm{rk}(Z_2)$ holds for arbitrary zonotopes Z_1, Z_2. Then the following construction yields further examples for which these conjectures hold.

Theorem 14. *If $Z_1, \ldots, Z_k$ are special zonotopes, then so is $Z_1 \times \cdots \times Z_k$.*

For the crystallographic root system Φ of type F_4, the vertex set of $Z(\Phi)$ splits into three B_4-orbits. We can construct an almost cover of size $24 = \mathrm{rk}(Z(\Phi))$, but we do not see if our method suits a proof that this is best possible.

7 Conclusion

We investigated how the polynomial method can be used to study almost covers of vertex sets of zopotopes and Coxeter permutahedra. In the meantime, Conjecture 2 was refuted by Gábor Damásdi, whereas Conjecture 11 was verified by Péter Frenkel.

References

[1] N. Alon, Combinatorial Nullstellensatz, *Combinatorics, Probability and Computing* **8** (1999), 7–29.

[2] N. Alon and Z. Füredi, Covering the cube by affine hyperplanes, *European Journal of Combinatorics* **14** (1993), 79–83.

[3] S. Ball and O. Serra, Punctured combinatorial Nullstellensätze, *Combinatorica* **29** (2009), 511–522.

[4] A. Blokhuis, A.E. Brouwer, and T. Szőnyi, Covering all points except one, *Journal of Algebraic Combinatorics* **32** (2010,) 59–66.

[5] A.E. Brouwer and A. Schrijver, The blocking number of an affine space, *Journal of Combinatorial Theory, Series A* **24** (1978), 251–253.

[6] S. Fomin and N. Reading, Root systems and generalized associahedra, in: *Geometric Combinatorics*, IAS/Park City Mathematics Series, 13, American Mathematical Society, Providence, 2007, 63–131.

[7] T. Godland and Z. Kabluchko, Projections and angle sums of belt polytopes and permutohedra, *Results in Mathematics* **78** (2023), #140, 29 pp.

[8] G. Hegedüs and Gy. Károlyi, Covering the permutohedron by affine hyperplanes, *Acta Mathematica Hungarica*, to appear. See `arXiv:2305.06202v3`.

[9] C. Hohlweg, Permutahedra and Associahedra: Generalized associahedra from the geometry of finite reflection groups, in: *Associahedra, Tamari Lattices and Related Structures*, Progress in Mathematics, 299, Birkhäuser, Basel, 2012, 129–159.

[10] J.E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge Studies in Advanced Mathematics, 29, Cambridge University Press, Cambridge, 1990.

[11] R.E. Jamison, Covering finite fields with cosets of subspaces, *Journal of Combinatorial Theory, Series A* **22** (1977), 253–266.

[12] G. Rote, The generalized combinatorial Lasoń-Alon-Zippel-Schwartz Nullstellensatz lemma, preprint, 2023, `arXiv:2305.10900`.

[13] G.M. Ziegler, *Lectures on Polytopes*, Graduate Texts in Mathematics, 153, Springer, New York, 1995.

Classification of Edge-to-edge Monohedral Tilings of the Sphere

Ho Man Cheung*[1], Hoi Ping Luk†[2], and Min Yan‡[3]

[1,2,3]The Hong Kong University of Science & Technology, Hong Kong.

Abstract

The history of studies on tilings of the sphere can be traced back to Plato (5 Platonic solids) and Archimedes (13 Archimedean solids). We study edge-to-edge monohedral tilings of the sphere. The classification of such tilings was pioneered by D. Sommerville in 1923. Significant progress was made in the past decades. However, the remaining cases have been the most difficult to classify. They are also of the utmost importance as they give rise to the majority of the tilings. We have recently classified all of them and hence completed the whole classification celebrating its centenary. The process involved new techniques ranging from combinatorics, geometry, algebra and number theory. All the tilings can be classified into 3 types: Platonic type, earth map type, and sporadic type. The full classification gives us a comprehensive understanding of their structural relations.

1 Introduction

The *tilings* in our studies cover the surface of the sphere without holes and overlaps. A tiling is *monohedral* if all tiles are geometrically congruent to a fixed polygon. The polygon, assumed to have geodesic arcs as edges, is called the *prototile*. By [8, Lemma 1], the prototile of a monohedral tiling of the sphere must be simple, i.e., its boundary is a simple closed curve. The tilings are also *edge-to-edge*, which means that no vertex of a tile lies in the interior of an edge of another tile (for example, see Figure 1). We also assume that *the degree* of a vertex in a tiling is at least 3 to avoid trivial examples by artificially adding extra vertices to edges and the complications inflicted by that. For simplicity, by *tiling* we mean edge-to-edge monohedral tiling of the sphere satisfying the above assumptions.

Figure 1: Edge-to-edge v.s. non-edge-to-edge

By [11, Proposition 4], the prototile in a tiling is either a triangle, a quadrilateral, or a pentagon. We call the prototiles resulting in tilings the *admissible prototiles*. From [4, 10] and [12], they are shown in Figure 2) with notations for their edge combinations. For example, a^4b means 4 a-edges and 1 b-edge in a pentagon. Edges with different labels are assumed to have different lengths. In a^4, the notation $\bullet$ (and $\circ$) denotes the opposite angles of equal value, and $\bullet$ will be used in Figures 5 and 7.

D. Sommerville [9] first studied the tilings with triangle prototiles in 1923. H. L. Davies gave an outline for the classification [6], which was completed by Y. Ueno and Y. Agaoka [10] in 2002. H. H. Gao, N. Shi and M. Yan [8] classified the minimal case for pentagon prototiles in 2013 and significant progress has since been made by Y. Akama, E. X. Wang and M. Yan [1, 2, 12, 13] in the quadrilateral and the pentagon direction. The remaining and the hardest problems have prototiles with edge combinations a^2bc, a^3b and a^4b. By overcoming these challenges [3, 4, 5], we present the main result below.

*Email: hmcheungae@connect.ust.hk.

†Email: hoi@connect.ust.hk. Author was supported in part by the Li Po Chun Charitable Trust Fund scholarship.

‡Email: mamyan@ust.hk. Research was supported by Hong Kong RGC General Research Fund 16303515 and 16305920.

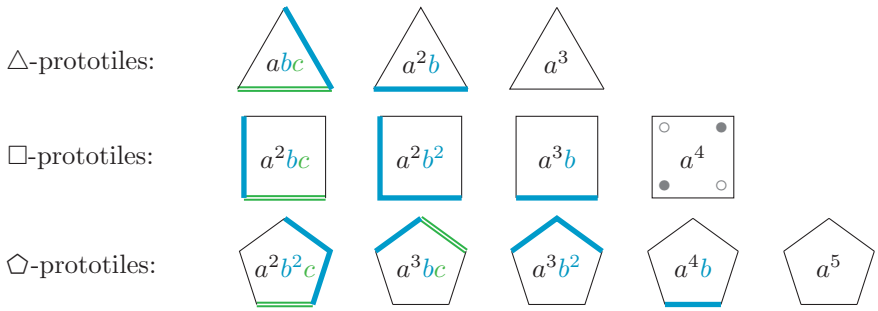


Figure 2: The admissible prototiles

2 Main result

Theorem 1. *The edge-to-edge monohedral tilings of the sphere are*

1. *Platonic type: Platonic solids* $P_* = P_4, P_6, P_8, P_{12}, P_{20}$ *and subdivisions on* P_* *below*
   - *Simple subdivision* S_iP_6 *of the cube for* $i = 1, ..., 7$*;*
   - *Triangular subdivision* TP_**;*
   - *Barycentric subdivision* BP_**;*
   - *Quadrilateral subdivision* QP_**;*
   - *Quadricentric subdivision* CP_**;*
   - *Pentagonal subdivision* PP_**;*
   - *Double pentagonal subdivision* DP_**;*
2. *Earth map type:*
   - 3 *infinite families of* △*-tilings:* $E_△1$ *(with reductions* $E^I_△1, E^J_△1$*),* $E_△2$ *and* $E_△3$*;*
   - 2 *infinite families of* □*-tilings:* $E_□1$ *(with reductions* $E^A_□1, E^K_□1, E^R_□1$*) and* $E_□2$*;*
   - 2 *infinite families of* ⬠*-tilings:* $E_⬠1$ *and* $E_⬠2$*;*
3. *Sporadic type:* $S_{12□}1, S_{16□}1, S_{16□}2, S_{16□}3$*(and* $FS_{16□}3$*),* $S_{16□}4, S_{36□}5, S_{36□}6, S_{16⬠}$*;*
4. *Modifications:*
   - *Flip* F*:*
     Platonic - $FBP_8, FQP_6, FQP_8, FPP_8, F_1PP_{20}, F_2PP_{20}$*;*
     Earth map △*-tilings -* $FE_△i$ *where* $i = 1, 2, 3$*;*
     Earth map □*-tilings -* $FE_□1, F_1E_□2, F_2E_□2$*;*
     Earth map ⬠*-tilings -* $F_1E_⬠i, F_2E_⬠i$ *for* $i = 1, 2$*, and* $F'_2E_⬠2, F''_2E_⬠2$*;*
     Sporadic - $FS_{16□}3$*;*
   - *Rearrangement* R*:* $RE_□1$*.*

The distinguishing features of tilings are best demonstrated in plane drawings. Platonic type tilings are shown in Figures 3, 4, 5 and 6, where the open ends of the outmost edges in a drawing converge to a single vertex. Earth map type tilings are shown in Figures 7 and 8, where the vertical edges in the top row of each drawing converge to a vertex (the "north pole") and those in the bottom converge to another (the "south pole"), and the left and right boundaries are identified. Sporadic tilings are shown in Figures 9 and 10. Two examples of modifications on QP_8 and on $E^A_□1$ are shown respectively in Figures 11 and 12. The readers are referred to [4] and [5] for detailed discussion on modifications, including the most sophisticated ones.

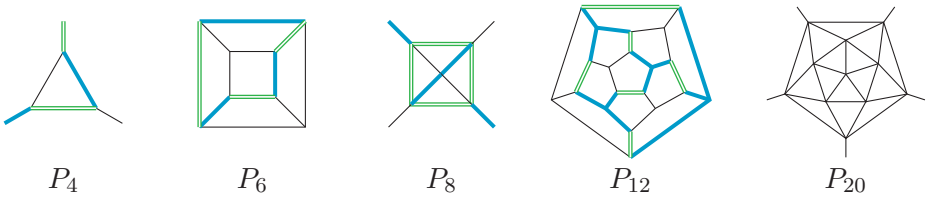


Figure 3: Platonic solids

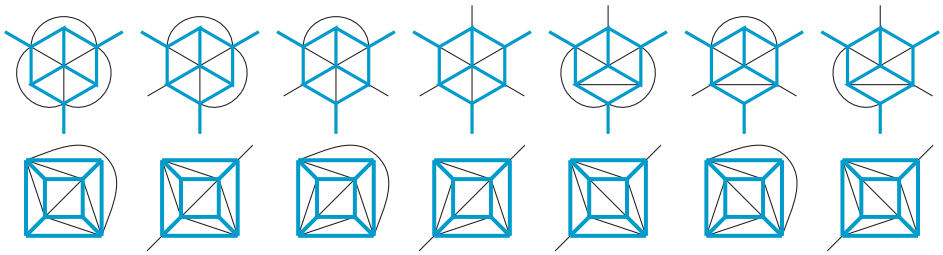


Figure 4: Simple triangular subdivisions of the cube P_6

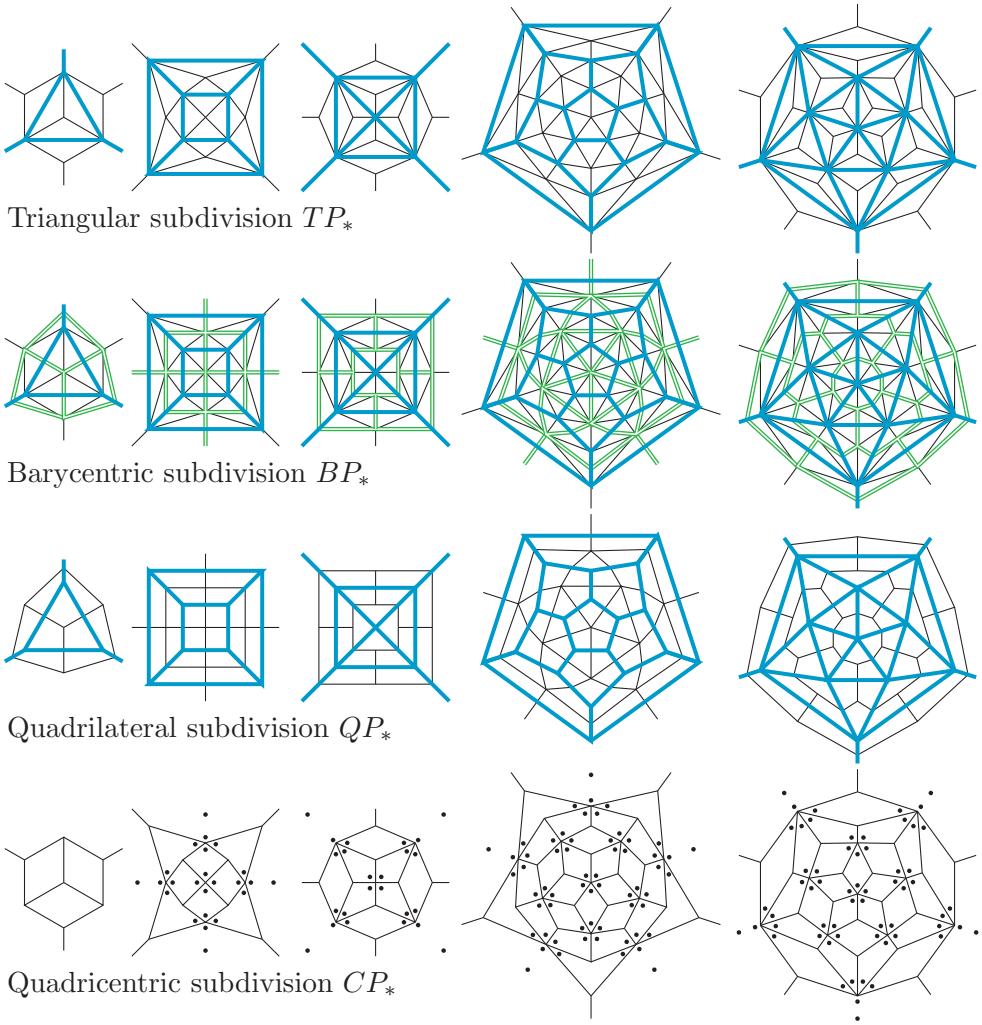


Figure 5: Subdivisions of Platonic solids TP_*, QP_*, BP_*, and CP_*

We highlight some interesting facts before the sketch of the proof. First, P_{20} is the only Platonic

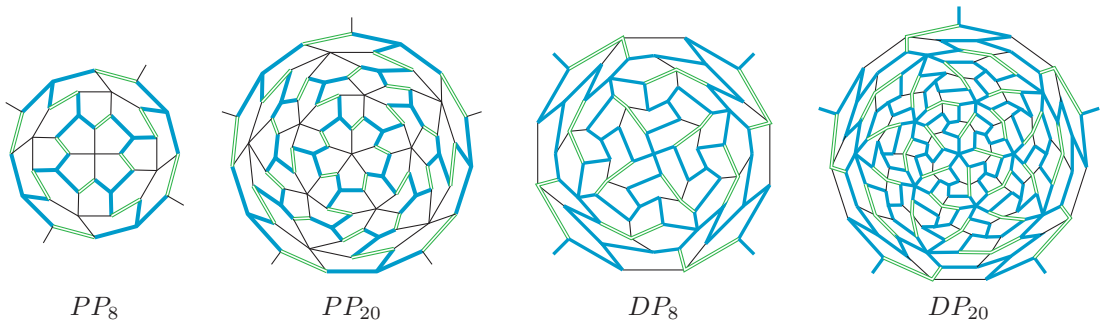


Figure 6: Pentagonal subdivisions and double pentagonal subdivisions of P_8 and P_{20}

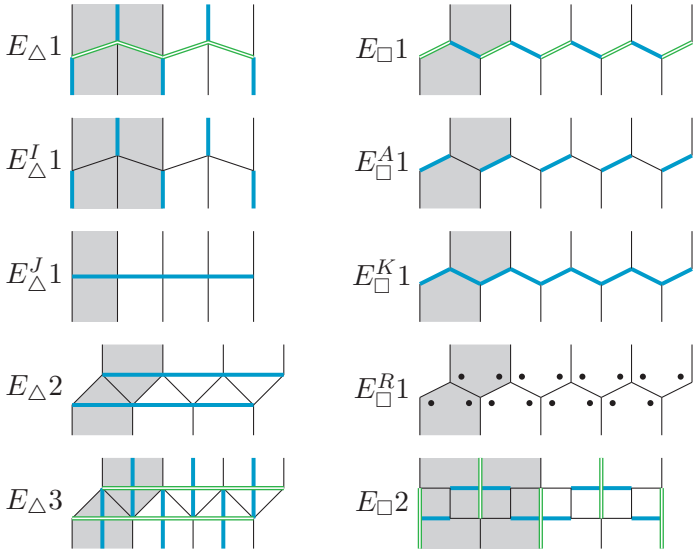


Figure 7: Earth map type $\triangle$-tilings and $\square$-tilings

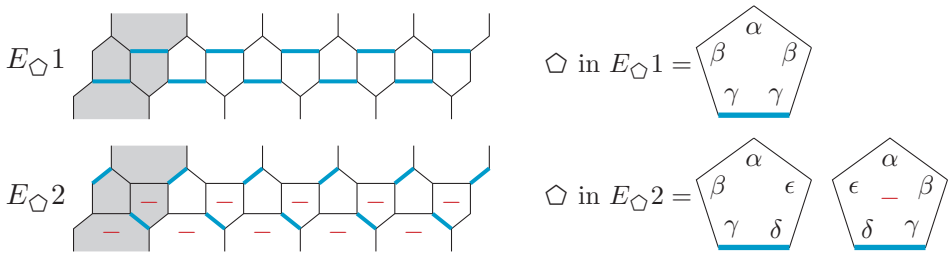


Figure 8: Earth map type $\pentagon$-tilings

solid that gives a rigid tiling. Second, the earth map type tilings (or earth map tilings) resemble the earth map – hence the name. Notably, the poles of earth map tilings are the vertices with negative combinatorial curvature (see definition in [7]). Between them, a tiling is formed by repeating copies of *a timezone* (shaded). Third, in $S_{16\square}3$ and $FS_{16\square}3$, one angle is actually π. Hence they are also non-edge-to-edge $\triangle$-tilings.

Sketch of proof. The complete classification is obtained by determining

1. the admissible prototiles, and

2. the corresponding admissible vertices in terms of angle combinations for each admissible prototile.

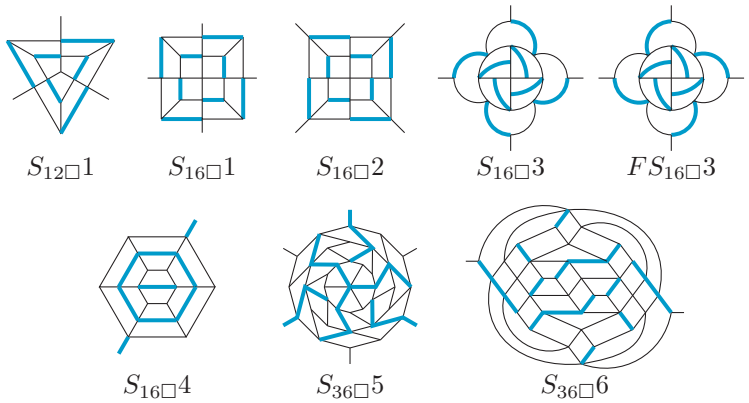


Figure 9: Sporadic $\square$-tilings $S_{12}\square 1, S_{16}\square 1, S_{16}\square 2, S_{16}\square 3, FS_{16}\square 3, S_{16}\square 4, S_{36}\square 5, S_{36}\square 6$

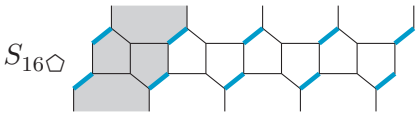


Figure 10: The sporadic $\pentagon$-tiling $S_{16}\pentagon$

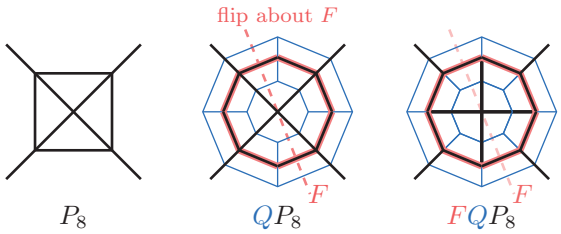


Figure 11: Platonic type tiling from subdivision to modification: $P_8 \to QP_8 \to FQP_8$

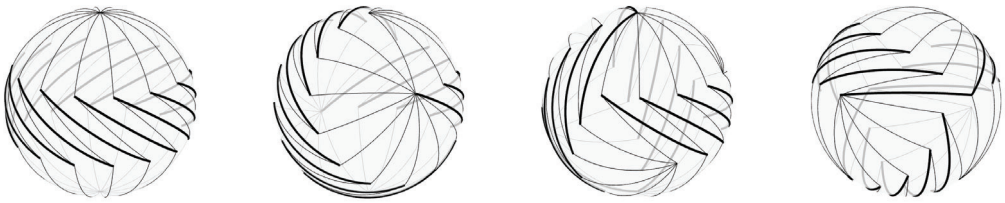


Figure 12: An example of modifications – earth map tiling $E^A_{28}\square 1$, two tilings from flip modification $FE^A_{28}\square 1$ and a rearrangement $RE^A_{28}\square 1$

Such a set of vertices satisfies various combinatorial and geometric contraints. We call it *anglewise-vertex combination* (or AVC for short). The tiling in the first picture of Figure 13 has AVC $= \{\alpha\gamma\delta, \beta^n\}$.

The knowledge of AVC is pivotal: it serves as the instruction of how to put the tiles together. For example, suppose that we have AVC $= \{\alpha\gamma\delta, \beta^3\}$ for the prototile a^3b. Then every vertex is $\alpha\gamma\delta$ or β^3. The notation $\alpha\gamma\delta$ means that a vertex has one α, one γ and one δ (see first picture, Figure 13) whereas β^3 means that a vertex has three β's. In the second picture, a vertex $\alpha\gamma\delta$ uniquely determines the three incident tiles ①, ②, ③. Similarly, we then determine $\alpha_3\gamma_1\cdots = \alpha\gamma\delta$ and $\gamma_3\delta_2\cdots = \alpha\gamma\delta$ and $\beta_3\cdots, \beta_1\beta_2\cdots = \beta^3$. Repeating such process, we uniquely determine the tiling given by the cube P_6 in the third picture. The same argument works for AVC $= \{\alpha\gamma\delta, \beta^n\}$ with any fixed integer $n \geq 3$. The tiling obtained is indeed $E^A_{\square}1$ in the first picture where $n = 3$ gives P_6 (shaded).

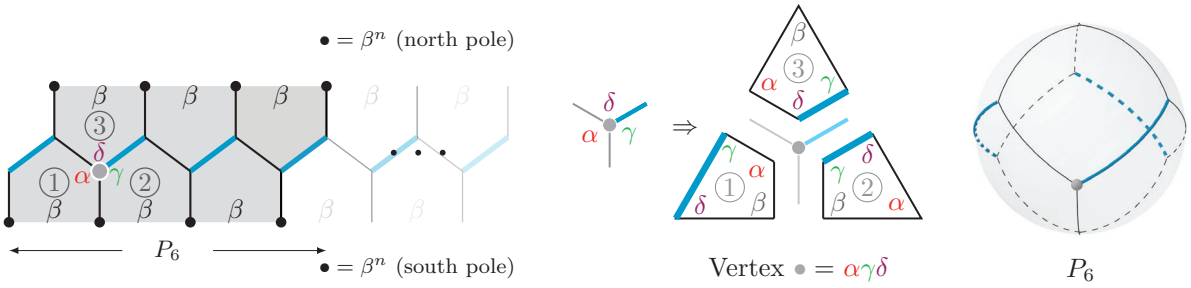


Figure 13: Construction of the tiling $E_{\square}1$ with prototile a^3b and AVC = $\{\alpha\gamma\delta, \beta^n\}$

By edge configurations and the existence of vertices of certain degrees, we obtain the prototiles in Figure 2. See [4, Lemma 1] and [12, Lemma 9] for further details.

For each admissible prototile, it takes both combinatorial and geometric arguments to determine the AVCs. It boils down to the study of the angles in a tiling. Powerful tools, such as discharging method, convexity analysis, spherical trigonometry, Gröbner basis, trigonometric Diophantine analysis and integer linear programming, are implemented for this purpose. □

The full classification of the △-tilings can be see in [4, 10], the full classification of the □-tilings can be seen in [4], and the full classification of ⬠-tilings is the collective effort of [1, 2, 5, 8, 12, 13]. An alternative classification of tilings with a^3b prototile via a noval approach is given in [3].

References

[1] Y. Akama, E. X. Wang, M. Yan, Tilings of the sphere by congruent pentagons III: edge combination a^5, *Adv. in Math.* **394** (2022), #107881.

[2] Y. Akama, M. Yan, On deformed dodecahedron tiling, *Australas. J. of Comb.* **85(1)** (2023), 1–14.

[3] H. P. Luk, H. M. Cheung, Rational angles and tilings of the sphere by congruent quadrilaterals, *Ann. Comb.* **28** (2024), 485–527.

[4] H. M. Cheung, H. P. Luk, M. Yan, Tilings of the sphere by congruent quadrilaterals or triangles, *preprint*, 2022, `arXiv:2204.02736`.

[5] H. M. Cheung, H. P. Luk, M. Yan, Tilings of the sphere by congruent pentagons IV: edge combination a^4b, *preprint*, 2023, `arXiv:2307.11453`.

[6] H. L. Davies, Packings of spherical triangles and tetrahedra, *Proceedings of the Colloquium on Convexity* **ed. W. Fenchel** (1967), Københavns Univ. Mat. Inst., Copenhagen, 42–51.

[7] Y. Higuchi, Combinatorial curvature for planar graphs, *J. Graph Theory* **38(4)** (2001), 220–229.

[8] H. H. Gao, N. Shi, M. Yan, Spherical tiling by 12 congruent pentagons, *J. Comb. Theory Ser. A* **120(4)** (2013), 744–776.

[9] D. M. Y. Sommerville, Division of space by congruent triangles and tetrahedra, *Proc. Royal Soc. Edinburgh* **43** (1923), 85–116.

[10] Y. Ueno, Y. Agaoka, Classification of tilings of the 2-dimensional sphere by congruent triangles, *Hiroshima Math. J.* **32(3)** (2002), 463–540.

[11] Y. Ueno, Y. Agaoka, Examples of spherical tilings by congruent quadrangles, *Math. Inform. Sci., Fac. Integrated Arts Sci., Hiroshima Univ.* **Ser. IV 27** (2001), 135–144.

[12] E. X. Wang, M. Yan, Tilings of sphere by congruent pentagons I: edge combinations a^2b^2c and a^3bc, *Adv. in Math.* **394** (2022), #107866.

[13] E. X. Wang, M. Yan, Tilings of sphere by congruent pentagons II: edge combination a^3b^2, *Adv. in Math.* **394** (2022), #107867.

Betti numbers of monomial curves*

Ignacio García-Marco[†1], Philippe Gimenez[‡2], and Mario González-Sánchez[§2]

[1]Instituto de Matemáticas y Aplicaciones (IMAULL), Sección de Matemáticas, Facultad de Ciencias, Universidad de La Laguna, 38200, La Laguna, Spain
[2]Instituto de Investigación en Matemáticas de la Universidad de Valladolid (IMUVA), Universidad de Valladolid, 47011 Valladolid, Spain

Abstract

In this work, we explore when the Betti numbers of the coordinate rings of a projective monomial curve and one of its affine charts are identical. Given an infinite field k and a sequence of relatively prime integers $a_0 = 0 < a_1 < \cdots < a_n = d$, we consider the projective monomial curve $\mathcal{C} \subset \mathbb{P}_k^n$ of degree d parametrically defined by $x_i = u^{a_i} v^{d-a_i}$ for all $i \in \{0, \ldots, n\}$ and its coordinate ring $k[\mathcal{C}]$. The curve $\mathcal{C}_1 \subset \mathbb{A}_k^n$ with parametric equations $x_i = t^{a_i}$ for $i \in \{1, \ldots, n\}$ is an affine chart of $\mathcal{C}$ and we denote by $k[\mathcal{C}_1]$ its coordinate ring. The main contribution of this paper is the introduction of a novel (Gröbner-free) combinatorial criterion that provides a sufficient condition for the equality of the Betti numbers of $k[\mathcal{C}]$ and $k[\mathcal{C}_1]$. Leveraging this criterion, we identify infinite families of projective curves satisfying this property.

Introduction

Let k be an infinite field, and $k[\mathbf{x}] := k[x_1, \ldots, x_n]$ and $k[\mathbf{t}] := k[t_1, \ldots, t_m]$ be two polynomial rings over k. Given $\mathcal{B} = \{b_1, \ldots, b_n\} \subset \mathbb{N}^m$, a set of nonzero vectors, each element $b_i = (b_{i1}, \ldots, b_{im}) \in \mathbb{N}^m$ corresponds to the monomial $\mathbf{t}^{b_i} := t_1^{b_{i1}} \cdots t_m^{b_{im}} \in k[\mathbf{t}]$. The affine toric variety $X_\mathcal{B} \subset \mathbb{A}_k^n$ determined by $\mathcal{B}$ is the Zariski closure of the set given parametrically by $x_i = u_1^{b_{i1}} \cdots u_m^{b_{im}}$ for all $i = 1, \ldots, n$. Consider

$$\mathcal{S}_\mathcal{B} := \langle b_1, \ldots, b_n \rangle = \{\alpha_1 b_1 + \cdots + \alpha_n b_n \,|\, \alpha_1, \ldots, \alpha_n \in \mathbb{N}\} \subset \mathbb{N}^m \,,$$

the affine monoid spanned by $\mathcal{B}$. The toric ideal determined by $\mathcal{B}$ is the kernel $I_\mathcal{B}$ of the k-algebra homomorphism $\varphi_\mathcal{B} : k[\mathbf{x}] \longrightarrow k[\mathbf{t}]$ induced by $x_i \mapsto \mathbf{t}^{b_i}$. Since k is infinite, one has that $I_\mathcal{B}$ is the vanishing ideal of $X_\mathcal{B}$ and, hence, the coordinate ring of $X_\mathcal{B}$ is (isomorphic to) the semigroup algebra $k[\mathcal{S}_\mathcal{B}] := \mathrm{Im}(\varphi_\mathcal{B}) \simeq k[\mathbf{x}]/I_\mathcal{B}$. The ideal $I_\mathcal{B}$ is an $\mathcal{S}_\mathcal{B}$-homogeneous binomial ideal, i.e., if one sets the $\mathcal{S}_\mathcal{B}$-degree of a monomial $\mathbf{x}^\alpha \in k[\mathbf{x}]$ as $\deg_{\mathcal{S}_\mathcal{B}}(\mathbf{x}^\alpha) := \alpha_1 b_1 + \cdots + \alpha_n b_n \in \mathcal{S}_\mathcal{B}$, then $I_\mathcal{B}$ is generated by $\mathcal{S}_\mathcal{B}$-homogeneous binomials. One can thus consider a minimal $\mathcal{S}_\mathcal{B}$-graded free resolution of $k[\mathcal{S}_\mathcal{B}]$ as $\mathcal{S}_\mathcal{B}$-graded $k[\mathbf{x}]$-module,

$$\mathcal{F} : 0 \longrightarrow F_p \longrightarrow \cdots \longrightarrow F_0 \longrightarrow k[\mathcal{S}_\mathcal{B}] \longrightarrow 0 \,.$$

The projective dimension of $k[\mathcal{S}_\mathcal{B}]$ is $\mathrm{pd}(k[\mathcal{S}_\mathcal{B}]) = \max\{i \,|\, F_i \neq 0\}$. The i-th Betti number of $k[\mathcal{S}_\mathcal{B}]$ is the rank of the free module F_i, i.e., $\beta_i(k[\mathcal{S}_\mathcal{B}]) = \mathrm{rank}(F_i)$; and the Betti sequence of $k[\mathcal{S}_\mathcal{B}]$ is

*This work is supported in part by the grant PID2022-137283NB-C22 funded by MCIN/AEI/10.13039/501100011033 and by ERDF "A way of making Europe."

[†]Email: iggarcia@ull.edu.es.
[‡]Email: pgimenez@uva.es.
[§]Email: mario.gonzalez.sanchez@uva.es. T. A. thanks financial support from European Social Fund, *Programa Operativo de Castilla y León*, and *Consejería de Educación de la Junta de Castilla y León*.

$(\beta_i(k[\mathcal{S}_\mathcal{B}])\,;\, 0 \leq i \leq \mathrm{pd}(k[\mathcal{S}_\mathcal{B}]))$. When the Krull dimension of $k[\mathcal{S}_\mathcal{B}]$ coincides with its depth as $k[\mathbf{x}]$-module, the ring $k[\mathcal{S}_\mathcal{B}]$ is said to be Cohen-Macaulay. By the Auslander-Buchsbaum formula, this is equivalent to $\mathrm{pd}(k[\mathcal{S}_\mathcal{B}]) = n - \dim(k[\mathcal{S}_\mathcal{B}])$. When $k[\mathcal{S}_\mathcal{B}]$ is Cohen-Macaulay, its (Cohen-Macaulay) type is the rank of the last nonzero module in the resolution, i.e., $\mathrm{type}(k[\mathcal{S}_\mathcal{B}]) := \beta_p(k[\mathcal{S}_\mathcal{B}])$ where $p = \mathrm{pd}(k[\mathcal{S}_\mathcal{B}])$.

Now consider $d \in \mathbb{Z}^+$ and $a_0 := 0 < a_1 < \cdots < a_n = d$ a sequence of relatively prime integers. Denote by $\mathcal{C}$ the projective monomial curve $\mathcal{C} \subset \mathbb{P}^n_k$ of degree d parametrically defined by $x_i = u^{a_i}v^{d-a_i}$ for all $i \in \{0,\ldots,n\}$, i.e., $\mathcal{C}$ is the Zariski closure of

$$\{(u^{a_0}v^{d-a_0} : \cdots : u^{a_i}v^{d-a_i} : \cdots : u^{a_n}v^{d-a_n}) \in \mathbb{P}^n_k \,|\, (u:v) \in \mathbb{P}^1_k\}.$$

Taking $\mathcal{A} = \{\mathbf{a}_0,\ldots,\mathbf{a}_n\} \subset \mathbb{N}^2$ with $\mathbf{a}_i = (a_i, d-a_i)$ for all $i = 0,\ldots,n$, one has that $I_\mathcal{A}$ is the vanishing ideal of $\mathcal{C}$, and the coordinate ring of $\mathcal{C}$ is the two-dimensional ring $k[\mathcal{C}] = k[x_0,\ldots,x_n]/I_\mathcal{A}$, where $\mathcal{S} = \mathcal{S}_\mathcal{A}$ denotes the monoid spanned by $\mathcal{A}$. The projective monomial curve $\mathcal{C}$ is said to be arithmetically Cohen-Macaulay if the ring $k[\mathcal{C}]$ is Cohen-Macaulay.

The monomial projective curve $\mathcal{C}$ has two affine charts, $\mathcal{C}_1 = \{(u^{a_1},\ldots,u^{a_n}) \in \mathbb{A}^n_k \,|\, u \in k\}$ and $\mathcal{C}_2 = \{(v^{d-a_0}, v^{d-a_1},\ldots,v^{d-a_{n-1}}) \in \mathbb{A}^n_k \,|\, v \in k\}$, associated to the sequences $a_1 < \cdots < a_n$ and $d-a_{n-1} < \cdots < d-a_1 < d-a_0$, respectively. The second sequence is sometimes called the dual of the first one. Denote by $\mathcal{S}_1 := \mathcal{S}_{\mathcal{A}_1}$ the numerical semigroup generated by $\mathcal{A}_1 = \{a_1,\ldots,a_n\}$. The vanishing ideal of $\mathcal{C}_1$ is $I_{\mathcal{A}_1} \subset k[x_1,\ldots,x_n]$, and hence, its coordinate ring is the one-dimensional ring $k[\mathcal{C}_1] = k[x_1,\ldots,x_n]/I_{\mathcal{A}_1}$. Moreover, $I_\mathcal{A}$ is the homogenization of $I_{\mathcal{A}_1}$ with respect to the variable x_0. Similarly, denoting by $\mathcal{S}_2 := \mathcal{S}_{\mathcal{A}_2}$ the numerical semigroup generated by $\mathcal{A}_2 := \{d-a_0, d-a_1,\ldots,d-a_{n-1}\}$, the vanishing ideal of $\mathcal{C}_2$ is $I_{\mathcal{A}_2} \subset k[x_0,\ldots,x_{n-1}]$, its coordinate ring is $k[\mathcal{C}_2] = k[x_0,\ldots,x_{n-1}]/I_{\mathcal{A}_2}$, and $I_\mathcal{A}$ is the homogenization of $I_{\mathcal{A}_2}$ with respect to x_n.

One has that $\beta_i(k[\mathcal{C}]) \geq \beta_i(k[\mathcal{C}_1])$ for all i, and the goal of this work is to understand when the Betti sequences of $k[\mathcal{C}]$ and $k[\mathcal{C}_1]$ coincide. A necessary condition is that $k[\mathcal{C}]$ is Cohen-Macaulay. Indeed, affine monomial curves are always arithmetically Cohen-Macaulay while projective ones may be arithmetically Cohen-Macaulay or not. Thus, $\mathrm{pd}(k[\mathcal{C}]) = \mathrm{pd}(k[\mathcal{C}_1])$ if and only if $\mathcal{C}$ is arithmetically Cohen-Macaulay. In Theorem 5, which is the main result of this work, we provide a combinatorial sufficient condition for having equality between the Betti sequences of $k[\mathcal{C}]$ and $k[\mathcal{C}_1]$ by means of the poset structures induced by $\mathcal{S}$ and $\mathcal{S}_1$ on the Apery sets of both $\mathcal{S}$ and $\mathcal{S}_1$. In Propositions 9 and 11, we use our main result to provide explicit families of curves where $\beta_i(k[\mathcal{C}]) = \beta_i(k[\mathcal{C}_1])$ for all i.

The motivation of this work comes from [7], where the authors obtain a sufficient condition in terms of Gröbner bases to ensure the equality of the Betti sequences.

The computations in the examples given in this paper are performed using Singular [4].

1 Apery sets and their poset structure

Let $d \in \mathbb{Z}^+$ and $a_0 := 0 < a_1 < \cdots < a_n = d$ be a sequence of relatively prime integers. For each $i = 0,\ldots,n$, set $\mathbf{a}_i := (a_i, d-a_i) \in \mathbb{N}^2$, and consider the three sets $\mathcal{A}_1 = \{a_1,\ldots,a_n\}$, $\mathcal{A}_2 = \{d, d-a_1,\ldots,d-a_{n-1}\}$ and $\mathcal{A} = \{\mathbf{a}_0,\ldots,\mathbf{a}_n\} \subset \mathbb{N}^2$. We denote by $\mathcal{C} \subset \mathbb{P}^n_k$ the projective monomial curve defined by $\mathcal{A}$ as defined in the introduction, and by $\mathcal{C}_1$ and $\mathcal{C}_2$ its affine charts. Consider $\mathcal{S}_1$ and $\mathcal{S}_2$ the numerical semigroups generated by $\mathcal{A}_1$ and $\mathcal{A}_2$ respectively, and $\mathcal{S}$ the monoid spanned by $\mathcal{A}$ that we call the homogenization of $\mathcal{S}_1$ (with respect to d).

As already mentioned, $k[\mathcal{S}_1]$ and $k[\mathcal{S}_2]$ are always Cohen-Macaulay, while $k[\mathcal{C}]$ can be Cohen-Macaulay or not. There are many ways to determine when a projective monomial curve is arithmetically Cohen-Macaulay; see, e.g., [2, Cor. 4.2], [3, Lem. 4.3, Thm. 4.6] or [6, Thm. 2.6]. We give

some of them in Proposition 1, but let us previously recall the notion of Apery set since it is involved in some of those charaterizations.

For $i = 1, 2$, the Apery set of $\mathcal{S}_i$ with respect to d is $\mathrm{Ap}_i := \{y \in \mathcal{S}_i \,|\, y - d \notin \mathcal{S}_i\}$. Since $\gcd(\mathcal{A}_1) = 1$, we know that Ap_i is a complete set of residues modulo d, i.e., $\mathrm{Ap}_1 = \{r_0 = 0, r_1, \ldots, r_{d-1}\}$ and $\mathrm{Ap}_2 = \{t_0 = 0, t_1, \ldots, t_{d-1}\}$ for some positive integers r_i and t_i such that $r_i \equiv t_i \equiv i \pmod{d}$ for all $i = 1, \ldots, d-1$. One can also define the Apery set of $\mathcal{S}$ as $\mathrm{AP}_{\mathcal{S}} := \{\mathbf{y} \in \mathcal{S} \,|\, \mathbf{y} - \mathbf{a}_0 \notin \mathcal{S}, \mathbf{y} - \mathbf{a}_n \notin \mathcal{S}\}$. Note that this set has at least d elements by [5, Lem. 2.5].

Proposition 1. *The following assertions are equivalent:*

(a) $\mathcal{C}$ *is arithmetically Cohen-Macaulay.*

(b) $\mathrm{AP}_{\mathcal{S}}$ *has exactly* d *elements.*

(c) $\mathrm{AP}_{\mathcal{S}} = \{(0,0)\} \cup \{(r_i, t_{d-i}) \,|\, 1 \leq i < d\}$.

(d) For all $i = 1, \ldots, d-1$, $(r_i, t_{d-i}) \in \mathcal{S}$. *In other words, if* $q_1 \in \mathrm{Ap}_1$, $q_2 \in \mathrm{Ap}_2$ *and* $q_1 + q_2 \equiv 0 \pmod{d}$, *then* $(q_1, q_2) \in \mathcal{S}$.

(e) If $\mathbf{s} \in \mathbb{Z}^2$ *satisfies* $\mathbf{s} + \mathbf{a}_0 \in \mathcal{S}$ *and* $\mathbf{s} + \mathbf{a}_n \in \mathcal{S}$, *then* $\mathbf{s} \in \mathcal{S}$.

In order to compare $\beta_i(k[\mathcal{C}])$ and $\beta_i(k[\mathcal{C}_1])$ for all i, we will relate in Theorem 5 the Apery sets Ap_1 and $\mathrm{AP}_{\mathcal{S}}$ with the natural poset structure that both have and that we now define. For $i = 1, 2$, $(\mathrm{Ap}_i, \leq_i)$ is a poset, where $\leq_i$ is given by $y \leq_i z \iff z - y \in \mathcal{S}_i$. Similarly, $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ is a poset for $\leq_{\mathcal{S}}$ defined by $\mathbf{y} \leq_{\mathcal{S}} \mathbf{z} \iff \mathbf{z} - \mathbf{y} \in \mathcal{S}$.

Since $\mathcal{S} \subset \mathcal{S}_1 \times \mathcal{S}_2$, it follows that if $(y_1, y_2) \leq_{\mathcal{S}} (z_1, z_2)$, then $y_i \leq_i z_i$ for $i = 1, 2$. Using Proposition 1, one can prove that the poset structure of $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ is completely determined by those of $(\mathrm{Ap}_1, \leq_1)$ and $(\mathrm{Ap}_2, \leq_2)$ when $\mathcal{C}$ is arithmetically Cohen-Macaulay.

Proposition 2. *If* $\mathcal{C}$ *is arithmetically Cohen-Macaulay, then for all* $(y_1, y_2), (z_1, z_2) \in AP_{\mathcal{S}}$,

$$(y_1, y_2) \leq_{\mathcal{S}} (z_1, z_2) \iff y_1 \leq_1 z_1 \text{ and } y_2 \leq_2 z_2.$$

Let us recall some notions about posets that will be needed in the sequel.

Definition 3. *Let* $(P, \leq)$ *be a finite poset.*

(a) For $y, z \in P$, *we say that* z *covers* y, *and denote it by* $y \prec z$, *if* $y < z$ *and there is no* $w \in P$ *such that* $y < w < z$.

(b) We say that P *is graded if there exists a function* $\rho : P \to \mathbb{N}$, *called rank function, such that* $\rho(z) = \rho(y) + 1$ *whenever* $y \prec z$.

As the following result shows, the poset $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ is always graded. Since $(\mathrm{Ap}_1, \leq_1)$ has a minimum, whenever it is graded, the corresponding rank function is completely determined by the value of the rank function in the minimum, which we will fix to be 0. In the following proposition, we characterize the covering relation in Ap_1 and $\mathrm{AP}_{\mathcal{S}}$ and describe the rank functions of $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$, and of $(\mathrm{Ap}_1, \leq_1)$ when it is graded.

Proposition 4. *(a) If* $y, z \in \mathrm{Ap}_1$, *then* $y \prec_1 z$ *if and only if* $z = y + a_i$ *for some minimal generator* a_i *of* $\mathcal{S}_1$ *such that* $a_i \neq d$. *Therefore, if* Ap_1 *is graded and* $\rho_1 : \mathrm{Ap}_1 \to \mathbb{N}$ *denotes the rank function, for any* $y \in \mathrm{Ap}_1$, $\rho_1(y)$ *is the number of elements involved in any writing of* y *in terms of minimal generators of* $\mathcal{S}_1$.

(b) If $\mathbf{y} = (y_1, y_2)$, $\mathbf{z} = (z_1, z_2)$ *and* $\mathbf{y}, \mathbf{z} \in \mathrm{AP}_{\mathcal{S}}$, *then* $\mathbf{y} \prec_{\mathcal{S}} \mathbf{z}$ *if and only if* $\mathbf{z} = \mathbf{y} + \mathbf{a}_i$ *for some* $i \in \{1, \ldots, n-1\}$. *Therefore,* $\mathrm{AP}_{\mathcal{S}}$ *is graded by the rank function* $\rho : \mathrm{AP}_{\mathcal{S}} \to \mathbb{N}$ *defined by* $\rho(y_1, y_2) := (y_1 + y_2)/d$.

2 Betti numbers of affine and projective monomial curves

Recall that $I_{\mathcal{A}_1} \subset k[x_1, \ldots, x_n]$ is the vanishing ideal of $\mathcal{C}_1$ and $I_{\mathcal{A}} \subset k[x_0, \ldots, x_n]$ is the vanishing ideal of $\mathcal{C}$. When $\mathcal{C}$ is arithmetically Cohen-Macaulay, $\mathrm{pd}(k[\mathcal{C}]) = \mathrm{pd}(k[\mathcal{C}_1])$. Moreover, by Proposition 1, in this case, one has that $|\mathrm{AP}_{\mathcal{S}}| = |\mathrm{Ap}_1| = d$. The main result in this section is Theorem 5 where we give a sufficient condition in terms of the poset structures of the Apery sets Ap_1 and $\mathrm{AP}_{\mathcal{S}}$ for the Betti sequences of $k[\mathcal{C}_1]$ and $k[\mathcal{C}]$ to coincide.

Theorem 5. *If* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}}) \simeq (\mathrm{Ap}_1, \leq_1)$*, then* $\beta_i(k[\mathcal{C}]) = \beta_i(k[\mathcal{C}_1])$ *for all* i*.*

Note that the converse of this result is far from being true, as shown in Example 6.

Example 6. *For the sequence* $1 < 2 < 4 < 8$*, one has that both* $k[\mathcal{C}_1]$ *and* $k[\mathcal{C}]$ *are complete intersections with Betti sequence* $(1, 3, 3, 1)$*. However, the posets* $(\mathrm{Ap}_1, \leq_1)$ *and* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ *are not isomorphic since* $\leq_1$ *is a total order on* Ap_1*, while* $\leq_{\mathcal{S}}$ *is not.*

In order to compare the two posets $\mathrm{AP}_{\mathcal{S}}$ and Ap_1, one can use the following result.

Proposition 7. *The following two claims are equivalent:*

(a) The posets $(\mathrm{Ap}_1, \leq_1)$ *and* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ *are isomorphic;*

(b) $k[\mathcal{C}]$ *is Cohen-Macaulay,* $(\mathrm{Ap}_1, \leq_1)$ *is graded, and* $\{a_1, \ldots, a_{n-1}\}$ *is contained in the minimal system of generators of* $\mathcal{S}_1$*.*

Note that Ap_1 can be a graded poset even if $(\mathrm{Ap}_1, \leq_1)$ and $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ are not isomorphic as the following example shows.

Example 8. *For the sequence* $a_1 = 5 < a_2 = 11 < a_3 = 13$*, the Apery set of the numerical semigroup* $\mathcal{S}_1 = \langle a_1, a_2, a_3 \rangle$ *is* $\mathrm{Ap}_1 = \{0, 27, 15, 16, 30, 5, 32, 20, 21, 22, 10, 11, 25\}$*. This Apery set is graded with the rank function* $\rho_1 : \mathcal{S}_1 \to \mathbb{N}$ *defined below (see Figure 1):*

- $\rho_1(0) = 0$,
- $\rho_1(5) = \rho(11) = 1$,
- $\rho_1(10) = \rho_1(16) = \rho_1(22) = 2$,
- $\rho_1(15) = \rho_1(21) = \rho_1(27) = 3$,
- $\rho_1(20) = \rho_1(32) = 4$,
- $\rho_1(25) = 5$,
- $\rho_1(30) = 6$.

Moreover, since $\mathrm{AP}_{\mathcal{S}}$ *has 16 elements,* $k[\mathcal{C}]$ *is not Cohen-Macaulay, and hence* $(\mathrm{Ap}_1, \leq_1)$ *and* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ *are not isomorphic by Proposition 7.*

3 Examples of application

In Propositions 9 and 11, we provide some sequences $a_1 < \cdots < a_n$ for which the condition in Theorem 5 is satisfied. Let us start with arithmetic sequences, i.e., sequences $a_1 < \cdots < a_n$ such that $a_i = a_1 + (i-1)e$ for some positive integer e with $\gcd(a_1, e) = 1$. For this family, we refine [7, Cor. 4.2] that considers $a_1 > n - 1$.

Proposition 9. *Let* $a_1 < \ldots < a_n$ *be an arithmetic sequence of relatively prime integers. Then,* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}}) \simeq (\mathrm{Ap}_1, \leq_1)$ *if and only if* $a_1 > n - 2$*. Therefore, if* $a_1 > n - 2$*, the Betti sequences of* $k[\mathcal{C}_1]$ *and* $k[\mathcal{C}]$ *coincide.*

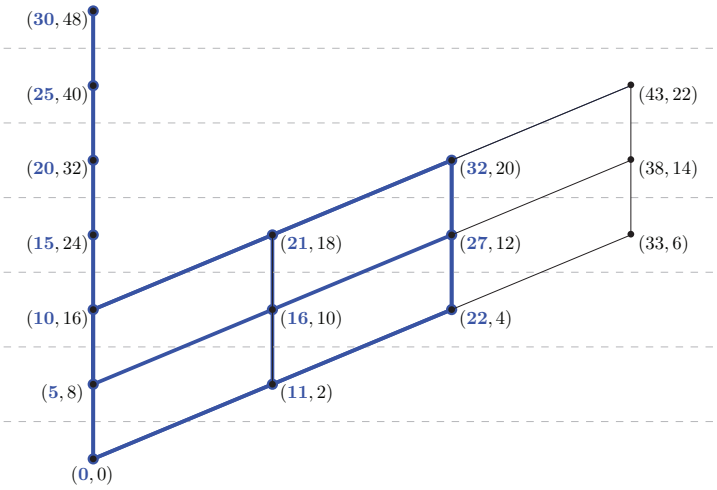


Figure 1: The posets $(\mathrm{Ap}_1, \leq_1)$ (in blue) and $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ (in black) for $\mathcal{S}_1 = \langle 5, 11, 13 \rangle$.

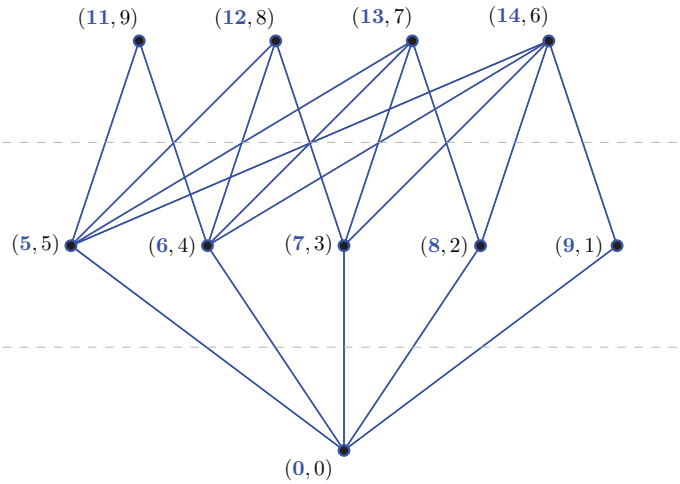


Figure 2: The posets $(\mathrm{Ap}_1, \leq_1)$ (in blue) and $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ (in black) for $\mathcal{S}_1 = \langle 5, 6, 7, 8, 9, 10 \rangle$.

Example 10. *For the sequence* $5 < 6 < 7 < 8 < 9 < 10$*, one has that* $a_1 = 5 > 4 = n - 2$*. Therefore, the Apery sets* $(\mathrm{Ap}_1, \leq_1)$ *and* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ *are isomorphic. Hence, by Theorem 5, the Betti sequences of* $k[\mathcal{C}_1]$ *and* $k[\mathcal{C}]$ *coincide. One can check that both are* $(1, 11, 30, 35, 19, 4)$*. The posets* $(\mathrm{Ap}_1, \leq_1)$ *and* $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ *in this example are shown in Figure 2.*

In [1, Sect. 6], the authors studied the canonical projections of the projective monomial curve $\mathcal{C}$ defined by an arithmetic sequence $a_1 < \cdots < a_n$ of relatively prime integers, i.e., the curve $\pi_r(\mathcal{C})$ obtained as the Zariski closure of the image of $\mathcal{C}$ under the r-th canonical projection $\pi_r : \mathbb{P}_k^n \dashrightarrow \mathbb{P}_k^{n-1}$, $(p_0 : \cdots : p_n) \dashmapsto (p_0 : \cdots : p_{r-1} : p_{r+1} : \cdots : p_n)$. We know that $\pi_r(\mathcal{C})$ is the projective monomial curve associated to the sequence $a_1 < \cdots < a_{r-1} < a_{r+1} < \cdots < a_n$.

In Proposition 11, for any $r \in \{2, \ldots, n-1\}$, we consider $\mathcal{A}_1 = \{a_1, \ldots, a_n\} \setminus \{a_r\}$, the numerical semigroup $\mathcal{S}_1 = \mathcal{S}_{\mathcal{A}_1}$, and its homogenization $\mathcal{S}$, and we characterize when the posets $(\mathrm{Ap}_1, \leq_1)$ and $(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}})$ are isomorphic.

Proposition 11. *Consider* $a_1 < \ldots < a_n$ *an arithmetic sequence of relatively prime integers with* $n \geq 4$*, and take* $r \in \{2, \ldots, n-1\}$*. Set* $\mathcal{A}_1 := \{a_1, \ldots, a_n\} \setminus \{a_r\}$*, and let* $\mathcal{S}_1$ *be the numerical semigroup generated by* $\mathcal{A}_1$*, and* $\mathcal{S}$ *its homogenization. Then,*

$$(\mathrm{AP}_{\mathcal{S}}, \leq_{\mathcal{S}}) \simeq (\mathrm{Ap}_1, \leq_1) \iff \begin{cases} a_1 > n-2 \text{ and } a_1 \neq n, & \text{if } r = 2, \\ a_1 \geq n \text{ and } r \leq a_1 - n + 1, & \text{if } 3 \leq r \leq n-2, \\ a_1 \geq n-2, & \text{if } r = n-1. \end{cases}$$

Consequently, if the previous condition holds, then $\beta_i(k[\mathcal{C}_1]) = \beta_i(k[\mathcal{C}])$*, for all* i*.*

Example 12. *For the sequence* $9 < 10 < 11 < 12 < 13$*, the Betti sequences of* $k[\mathcal{C}_1]$ *and* $k[\mathcal{C}]$ *coincide by Proposition 9. Indeed, it is* $(1, 10, 20, 15, 4)$ *for both curves. The parameters of this arithmetic sequence are* $a_1 = 9$*,* $e = 1$ *and* $n = 5$*. Hence, the Betti sequences of* $k[\pi_r(\mathcal{C}_1)]$ *and* $k[\pi_r(\mathcal{C})]$ *coincide for* $r = 2, 3, 4$ *by Proposition 11. One can check that the Betti sequence of* $k[\pi_2(\mathcal{C})]$ *and* $k[\pi_4(\mathcal{C})]$ *is* $(1, 5, 6, 2)$*, and the Betti sequence of* $k[\pi_3(\mathcal{C})]$ *is* $(1, 8, 12, 5)$*.*

References

[1] I. Bermejo, E. García-Llorente, I. García-Marco, and M. Morales. Noether resolutions in dimension 2. *J. Algebra* **482** (2017), 398-426.

[2] A. Campillo and P. Gimenez, Syzygies of affine toric varieties, *J. Algebra* **225** (2000), 142–161.

[3] M. P. Cavaliere and G. Niesi, On monomial curves and Cohen-Macaulay type, *Manuscripta Math.* **42** (1983), 147–159.

[4] W. Decker, G.-M. Greuel, G. Pfister, and H. Schönemann. Singular 4-3-0 — A computer algebra system for polynomial computations. http://www.singular.uni-kl.de, 2022.

[5] P. Gimenez and M. González Sánchez. Castelnuovo-Mumford regularity of projective monomial curves via sumsets, *Mediterr. J. Math.* **20**, 287 (2023), 24 pp.

[6] S. Goto, N. Suzuki, and K. Watanabe. On affine semigroup rings, *Jap. J. Math.* **2** (1976), 1–12.

[7] J. Saha, I. Sengupta, and P. Srivastava. Betti sequence of the projective closure of affine monomial curves. *J. Symb. Comput.* **119** (2023), 101-111.

The flexibility among 3-decompositions

Irene Heinrich [*1] and Lena Volk [†1]

[1]Dept. of Mathematics, Technische Universität Darmstadt, 64289 Darmstadt, Germany

Abstract

The 3-decomposition conjecture, postulated by Hoffmann-Ostenhof in 2011, is a major open question about the structure of cubic graphs: *Can the edge set of every cubic graph be decomposed into a spanning tree, a disjoint union of cycles, and a matching?* To date, the conjecture remains wide open. Towards a deeper structural understanding of 3-decompositions, we investigate the set of all 3-decompositions of a graph as a whole. On the one side, we provide a graph class that displays extremal behaviour: up to isomorphism, only one 3-decomposition exists. On the other side, we show that in general, 3-decompositions are more flexible. This contrasts the existing approaches which focus on the construction of precisely one decomposition of the considered graph. We exploit these insights towards a verification of the 3-decomposition conjecture on Bilu-Linial expanders.

1 Introduction

All graphs in this paper are simple and finite. A *3-decomposition* of a cubic graph G is a triple (T, C, M) of subgraphs of G where T is a spanning tree of G, C is 2-regular, and M is a matching such that $\{E(T), E(C), E(M)\}$ is a partition of $E(G)$. (See Figure 1 for examples of 3-decompositions.) The 3-decomposition conjecture, postulated by Hoffmann-Ostenhof [6], is a central open question about the structure of cubic graphs.

3-Decomposition Conjecture. *Every connected cubic graph has a 3-decomposition.*

The 3-decomposition conjecture has received great interest, and numerous results verify the conjecture on subclasses (e.g., planar [7], treewidth-3 [5], pathwidth-4 [2], and claw-free graphs [1]). Li and Cui [10] proved that the following weaker variant of the 3-decomposition conjecture is true: Every connected cubic graph can be decomposed into a spanning tree, a disjoint union of cycles, and a disjoint union of paths of length at most 2. There is ample literature on 3-decompositions when the considered

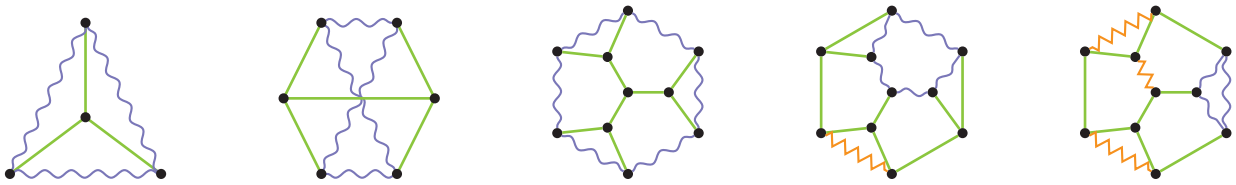


Figure 1: 3-decompositions of K_4, $K_{3,3}$, and three distinct 3-decompositions of the tricorn graph G_T. Spanning tree edges are straight and green, cycle edges are wavy blue, and matching-edges are zigzag-shaped and orange. It holds $\min_{\text{MATCH}}(G_T) = 0$ and $\max_{\text{MATCH}}(G_T) = 3$.

*Email: heinrich@mathematik.tu-darmstadt.de. The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (EngageS: grant agreement No. 820148).

†Email: volk@mathematik.tu-darmstadt.de.

graph G admits one of the following two extremes: a Hamiltonian path (a tree maximizing the number of degree-2 vertices) or a HIST [6] (a spanning tree which is *homeomorphically irreducible*, i.e., free of degree-2 vertices). However, little is known about the structure of the set of *all* 3-decompositions of a cubic graph. Towards a deeper structural understanding of cubic graphs we analyze the set of all 3-decompositions of a graph (class). We focus on the following three questions.

Q 1. *Are there graphs with a unique 3-decomposition?*

Consider the two graph invariants

$$\min_{\text{MATCH}}(G) \coloneqq \min\{\|M\| \colon (T, C, M) \text{ is a 3-decomposition of } G\}, \text{ and}$$
$$\max_{\text{MATCH}}(G) \coloneqq \max\{\|M\| \colon (T, C, M) \text{ is a 3-decomposition of } G\},$$

where $\|\cdot\|$ denotes the size (i.e., the number of edges) of a graph.

Q 2. *Which graphs (or graph classes) are extremal with respect to* $\min_{\text{MATCH}}$ *and* $\max_{\text{MATCH}}$*, respectively? How flexible is the set of all 3-decompositions of a graph with respect to the number of matching edges it contains?*

Q 3. *How can we exploit the observed flexibility among 3-decompositions towards proving the 3-decomposition conjecture?*

Our contribution. We positively answer Question 1 by providing an infinite class of graphs with the property that each graph in the class has a unique 3-decomposition up to isomorphism (Theorem 4). It is noteworthy that all graphs in this class have a HIST and it is known that a HIST of a cubic graph naturally corresponds to a 3-decomposition [6]. Hence, we further investigate for which graphs there exist HIST-free 3-decompositions. Assuming the 3-decomposition conjecture to hold, we prove that every graph of connectivity 2 has a HIST-free 3-decomposition (Theorem 5). We used the computer to verify that apart from K_4 and $K_{3,3}$ every 3-connected cubic graph of order at most 20 has a 3-decomposition without a HIST (Theorem 6).

Concerning Question 2, we prove that there exists a family of graphs $(H_n)_{n\in\mathbb{N}}$ with $\min_{\text{MATCH}}(H_n) = 0$ for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} \max_{\text{MATCH}}(H_n) = \infty$ (Proposition 8). We complement this result by proving the existence of two other graph families $(G_n)_{n\in\mathbb{N}}$ and $(G'_n)_{n\in\mathbb{N}}$ that satisfy $\lim_{n\to\infty} \max_{\text{MATCH}}(G_n) = 0$ (Theorem 4) and $\lim_{n\to\infty} \min_{\text{MATCH}}(G'_n) = \infty$ (Proposition 7).

We give a partial answer to Question 3 in Section 5 where we highlight that the flexibility of 3-decompositions can be exploited in order to provide a tighter analysis of 3-decompositions of Bilu-Linial expanders. This broadens our understanding of 3-decompositions since expander graphs show completely different behavior compared to the previously studied classes.

Due to space restrictions, some of the proofs are omitted or shortened to proof sketches.

Further related work. The recent results on the 3-decomposition conjecture are surveyed in the introduction of [2]. Hoffmann-Ostenhof, Noguchi, and Ozeki studied the existence of HISTs in cubic graphs [8]. Deciding whether a graph allows for a HIST is in general an intractable problem, which remains intractable even if the input is restricted to the class of cubic graphs [4].

2 Preliminaries

For two integers a and b we set $[a, b] \coloneqq \{a, a+1, \ldots, b\}$. We denote the path of order n by P_n, the complete graph of order n by K_n, and the complete bipartite graph with one part on n vertices and the other part on m vertices by $K_{n,m}$. The $\dot{K}_4$ is a graph obtained from K_4 by subdividing precisely one edge. Analogously, the $\dot{K}_{3,3}$ is obtained by subdividing an edge of the $K_{3,3}$. If u and v are vertices of a tree T, then uTv denotes the unique u-v-path in T. For a graph G and an edge subset $E' \subseteq E(G)$

we set $G[E']$ to be the graph with edge set E' and vertex set $\{v \in V(G) \colon \exists u \in V(G) \colon uv \in E'\}$. A non-empty graph G is *k-connected* (*k-edge connected*) if for each two distinct vertices u and v of G there are at least k internally vertex-disjoint (edge-disjoint) u-v-paths in G. The maximum number $k \in \mathbb{N}$ such that G is k-connected (k-edge connected) is the *connectivity* (*edge connectivity*) of G. In contrast to general graphs, the connectivity and the edge-connectivity of a cubic graph are equal. If $E' \subseteq E(G)$ such that $G[E(G) \setminus E']$ has more components than G, then E' is an *$|E'|$-edge separator*, otherwise we call $G[E']$ *non-separating*. Let $E' \subseteq E(G)$. If there exists a bipartition $\{U, W\}$ of $V(G)$ such that $E' = \{uw \in E(G) \colon u \in U, w \in W\}$, then E' is a *cut set* of G. A *bridge* is a 1-edge separator. If G is cubic and has a 3-decomposition (T, C, M), then G is *3-decomposable*.

Lemma 1. *Let G be a cubic graph with a 3-decomposition (T, C, M).*

1. *$\|G\| = 3/2|V(G)|$ and $\|C\| + \|M\| = \|G\| - \|T\| = |V(G)|/2 + 1$.*
2. *C and M are non-separating subgraphs of G.*
3. *Each vertex $v \in V(G)$ is either a degree-3 vertex of T, or a degree-2 vertex of T and contained in M, or a degree-1 vertex of T and contained in C. In particular, $\|C\| \geq 3$*

Observation 2 (Reformulation of [8], Theorem 2)**.** *If G is a cubic graph, then $\min_{\text{MATCH}}(G) = 0$ if and only if there exists a HIST T of G, which is the case precisely if $(T, G[E(G) \setminus E(T), \emptyset])$ is a 3-decomposition of G.*

Lemma 3. *If G is a 3-decomposable graph and $\ell \coloneqq \min\{\|C\| \colon C$ is a non-separating cycle in $G\}$, then*

$$0 \leq \min_{\text{MATCH}}(G) \leq \max_{\text{MATCH}}(G) \leq 1/2|V(G)| + 1 - \ell \leq 1/2|V(G)| - 2.$$

3 Graphs with unique 3-decompositions

In this section, we tackle Question 1. In fact, already among the smallest cubic graphs there are two examples of graphs with a unique 3-decomposition up to isomorphism: K_4 and $K_{3,3}$. The K_4 decomposes into a $K_{1,3}$, a 3-cycle, and an empty matching; the $K_{3,3}$ decomposes into a tree known as the *H-graph*, a 4-cycle, and an empty matching (see Figure 1). Observe that, in accordance with Observation 2, each of the two trees is a HIST. We argue that the decompositions are unique: By Lemma 1.3 each of the decompositions contains a cycle. Observe that a shortest cycle in K_4 is a 3-cycle and each two 3-cycles of K_4 can be mapped to each other by an automorphism of K_4. The remaining edges form a $K_{1,3}$. In particular, no larger cycle can be part of a 3-decomposition of K_4. The uniqueness of the decomposition of $K_{3,3}$ can be proven in analogy to this. In fact, there are infinitely many graphs with this property:

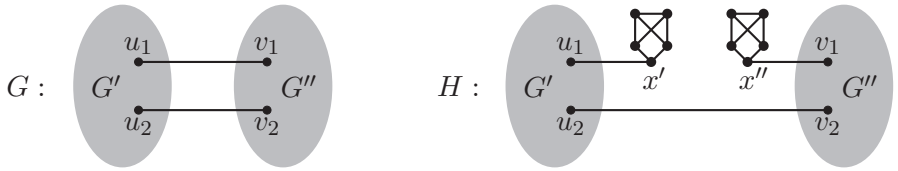
Theorem 4. *There exists an infinite family $\mathcal{G}$ of graphs which have precisely one 3-decomposition up to isomorphism. Further, $\min_{\text{MATCH}}(G) = \max_{\text{MATCH}}(G) = 0$ for all $G \in \mathcal{G}$.*

Proof sketch. The class $\mathcal{T}$ of homeomorphically irreducible subcubic trees contains infinitely many non-isomorphic trees. Let $\mathcal{G}$ be the family of cubic graphs obtained by the following construction: Let $T \in \mathcal{T}$. For each leaf ℓ of T let K^ℓ be either a copy of $\dot{K}_4$ or $\dot{K}_{3,3}$. Take the disjoint union of T and the graphs in $\{K^\ell \colon \deg_T(\ell) = 1\}$ and identify the degree-2 vertex of K^ℓ with ℓ for each leaf ℓ of T. The ingredients for the uniqueness proof are as follows: For a graph $G \in \mathcal{G}$ all edges of the corresponding tree $T \in \mathcal{T}$ are bridges of the construction, further each appended $\dot{K}_4$ or $\dot{K}_{3,3}$ has (up to isomorphism) precisely one non-separating cycle. The resulting decomposition is free of matching-edges. □

The situation observed at the beginning of this section (the only option of obtaining a 3-decomposition corresponds to a HIST) never occurs in the setting of connectivity-2 graphs if the 3-decomposition conjecture holds:

Theorem 5. *If the 3-decomposition conjecture holds, then every cubic graph with connectivity 2 has a 3-decomposition with a non-empty matching.*

Proof. We show the following stronger claim: If the 3-decomposition conjecture holds, then every connected cubic graph which has a 2-edge separator of non-incident edges has a 3-decomposition with a non-empty matching. The theorem follows immediately from this claim since if two incident edges e_1, e_2 form a 2-edge separator of a cubic graph, then the unique edge incident to e_1 and e_2 is a bridge (and, hence, the connectivity is at most 1). Assume that the 3-decomposition conjecture holds and let G be a cubic graph with a HIST T and a 2-edge separator of non-incident edges $\{u_1v_1, u_2v_2\}$. The graph $G \setminus \{u_1v_1, u_2v_2\}$ has precisely two components G' and G''.



Set $C := G[E(G) \setminus E(T)]$ and consider the 3-decomposition $(T, C, \emptyset)$ of G. Since $\{u_1v_1, u_2v_2\}$ is a separator we obtain $\{u_1v_1, u_2v_2\} \cap E(C) = \emptyset$ and, hence $\{u_1v_1, u_2v_2\} \subseteq E(T)$. Precisely one of the following situations occurs: $u_1Tu_2 \subseteq G'$ or $v_1Tv_2 \subseteq G''$. We may assume that $u_1Tu_2 \subseteq G'$.

We construct a graph H as follows: add two new vertices x' and x'' to G, remove u_1v_1, and add the edges u_1x' and $x''v_1$. Take the disjoint union of this graph with two copies K' and K'' of $\dot{K}_4$ and identify x' (resp. x'') with the degree-2 vertex of K' (resp. K''). The resulting graph H has a 3-decomposition (T_H, C_H, M_H) by assumption. Since u_1v_1 and u_2v_2 are not incident $\|u_1Tu_2\| \geq 1$ and we may choose an edge $e \in E(u_1Tu_2)$. Now, merge the 3-decomposition induced by the HIST and the one of H together to one for G. Take the decomposition from H in G'' and in G' a slight modification of the decomposition induced by the original HIST: Remove the edge e from the spanning tree part of T in G' to disconnect u_1 and u_2 in the spanning forest in G' and instead connect them via the spanning tree T_H in G'', which connects v_1 and v_2. We may add e to the matching since the decomposition used on G' so far had an empty matching. More formally $((T \cap G') \cup (T_H \cap G'') \cup (u_1, v_1) \cup (u_2, v_2) \setminus \{e\}, (C \cap G') \cup (C_H \cap G''), \{e\} \cup (M_H \cap G''))$ is a 3-decomposition with a non-empty matching for G. □

Theorem 6. *Apart from K_4 and $K_{3,3}$, every 3-connected cubic graph of order at most 20 has a 3-decomposition with a non-empty matching*[1].

4 Flexibility among 3-decompositions

Proposition 7. *For every $n \in \mathbb{N}$ there exists a 2-connected cubic graph G'_n with $\min_{\text{MATCH}}(G'_n) = n$.*

Proof sketch. We refrain from giving a technical description of G'_n and refer to Figure 2 for the construction and a 3-decomposition of G'_n with precisely n matching-edges. In particular, the graph G'_n is 3-decomposable and $\min_{\text{MATCH}}(G'_n) \leq n$. Assume that (T_n, C_n, M_n) is a 3-decomposition of G'_n. Observe that the only non-separating cycles in G'_n are the four triangles (in Figure 2: two triangles on the left and two triangles on the right of the drawing). At most one of the two left triangles and at most one of the two right triangles can be contained in C_n since C_n is a disjoint union of separating cycles by Lemma 3. Further, since C_n is non-empty we obtain $\|C_n\| \in \{3, 6\}$. From Lemma 1 follows $\|C_n\| + \|M_n\| = n + 6$ and with this, $\|M_n\| \geq n$. □

Theorem 8. *For every odd number $n \in \mathbb{N}$ there exists a cubic graph H_n with $\min_{\text{MATCH}}(H_n) = 0$ and $\max_{\text{MATCH}}(H_n) = n$. Further, there exists a 3-decomposition of H_n with $n - 3l$ edges in the matching for every $l \in [0, {}^{(n+1)}/_4]$.*

[1] https://gitlab.rlp.net/obachtle/reductions-for-the-3-decomposition-conjecture, March 2024.

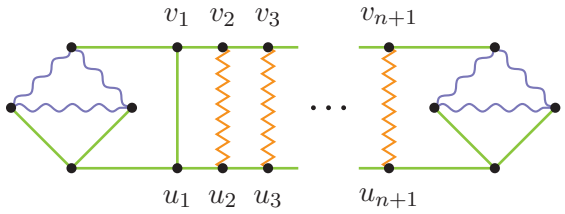


Figure 2: The graph G'_n is a 2-connected cubic graph with $\min_{\text{MATCH}}(G'_n) = n$.

Proof sketch. We only discuss the two extreme cases in this sketch. Fix an odd number $n \in \mathbb{N}$ and set $k := {}^{(n+3)}/_2$. Let $P = v_1 v_2 \dots v_k$ be the k-vertex path. Let Q^1 and Q^k be two copies of P_3 and let $K^2, K^3, \dots, K^{k-1}$ be $k-2$ copies of $K_{1,3}$. Take the disjoint union of P, Q^1, Q^k, and all K^i for $i \in [2, k-1]$. Now, identify the degree-2 vertex of Q_1 (resp. Q_k) with v_1 (resp. v_k). Further, for each $i \in [2, k-1]$ identify v_i with a degree-1 vertex of K^i. The resulting tree T has $2k-2$ degree-3 and $2k$ degree-1 vertices. Choose a planar embedding of T and connect the leaves of T by the outer facial cycle. Then, the resulting graph H_n is cubic and has the HIST T. Thus, $\min_{\text{MATCH}}(H_n) = 0$. Further, $\max_{\text{MATCH}}(H_n) = n$ since the shortest non-separating cycle is of length 3 and a 3-decomposition with n matching-edges can be obtained as follows: Let C_n be the triangle induced by the vertices of Q^1 in H_n. The following edges form the matching M_n: for $i \in [2, k-1]$ the edge of the outer face joining two vertices of K^i and the edge joining the degree-3 vertex of the K^i to P, and the edge of the outer face joining two vertices of Q^k. Let $T_n = H_n[E(H_n) - E(C_n) - E(M_n)]$. The triple (T_n, C_n, M_n) is a 3-decomposition of H_n with $\|M_n\| = n$. For $n = 3$ the 3-decompositions are depicted in the third and the fifth graph in Figure 1. □

5 3-Decompositions of Bilu-Linial Expanders

Bilu and Linial [3] give a concrete construction for a family of expander graphs by a series of lifting operations associated to random signings. See [9] for a survey on expanders. In the following, we investigate how 3-decompositions can be lifted. This illustrates how exploiting the flexibility of 3-decompositions, yields a fruitful approach to verify the 3-decomposition conjecture for more classes of graphs. The *2-lift* of a graph G equipped with a *signing* $s\colon E(G) \to \{-1, 1\}$ is the graph $\text{lift}(G, s)$ with

$$V(\text{lift}(G,s)) = \{v_0 \colon v \in V(G)\} \cup \{v_1 \colon v \in V(G)\},$$
$$E(\text{lift}(G,s)) = \bigcup_{uv \in s^{-1}(1)} \{u_0v_0, u_1v_1\} \cup \bigcup_{uv \in s^{-1}(-1)} \{u_0v_1, u_1v_0\}.$$

The vertices v_0 and v_1 are *fibers* of v and $\deg_{\text{lift}(G,s)}(v_0) = \deg_{\text{lift}(G,s)}(v_1) = \deg_G(v)$. For a subgraph H of G, we set $\text{lift}(H, s) := \text{lift}(H, s|_{E(H)})$. Observe that $\text{lift}(H, s)$ is a subgraph of $\text{lift}(G, s)$. The *signing* of a path $P \subseteq G$ is $s(P) := \prod_{e \in E(P)} s(e)$. In general, the existence of HISTs is not preserved by 2-lifts: If G is a cubic graph with a HIST T and $s \equiv -1$, then $\text{lift}(G, s)$ is bipartite and $|V(\text{lift}(G, s))|$ is a multiple of 4. It follows with [8, Corollary 3] that $\text{lift}(G, s)$ does not have a HIST. In contrast to this, 3-decompositions can be lifted under certain preconditions on the signing and the matching-edges.

Note that the lift of a connected graph is not necessarily connected again. E.g., if G is connected and $s \equiv 1$, then $\text{lift}(G, s)$ is isomorphic to the disjoint union of two copies of G. The assumptions of Theorem 10 ensure that the considered lift is connected. In the following, we characterize signings which yield a disconnected lift in order to show that 3-decompositions can be lifted in this case.

Lemma 9. *Let G be a connected graph with a signing s. The following are equivalent:*

1. $\text{lift}(G, s)$ *is disconnected.*
2. $s^{-1}(-1)$ *is empty or a cut set of G.*
3. $\text{lift}(G, s)$ *is isomorphic to two disjoint copies of G.*

In particular, if G is 3-decomposable and $\mathrm{lift}(G, s)$ *is disconnected, then each of the two components of* $\mathrm{lift}(G, s)$ *is 3-decomposable.*

Theorem 10. *Let (T, C, M) be a 3-decomposition of a cubic graph G with a signing s. If there exists $xy \in E(M)$ such that $s(xy) = -1$ and $s(xTy) = 1$, then the following is a 3-decomposition of* $\mathrm{lift}(G, s)$*:*

$$(\mathrm{lift}(T, s) + x_0y_1, \mathrm{lift}(C, s), \mathrm{lift}(M, s) - x_0y_1).$$

A random variable $s\colon E(G) \to \{-1, 1\}$ is a *random signing* of G if the sign of each edge is chosen uniformly at random.

Lemma 11. *Let G be a graph with a 3-decomposition (T, C, M). If s is a random signing of G, then*

$$\mathbb{P}\left[\exists xy \in E(M)\colon s(xy) = -1 \wedge s(xTy) = 1\right] = 1 - (3/4)^{\|M\|}.$$

Corollary 12. *If G is a cubic graph and s is the random signing on G, then the probability that the construction of Theorem 10 yields a 3-decomposition of* $\mathrm{lift}(G, s)$ *is maximized if the considered 3-decomposition (T, C, M) of G satisfies $\|M\| = \max_{\mathrm{MATCH}}(G)$.*

When iteratively applying the lifting operation, the number of edges in the matching m_k of the k-th lift G_k is $2^k(m_0 - 1) + 1$. Thus, the probability that iteratively applying Theorem 10 yields a 3-decomposition of G_k is at least $\prod_{l=0}^{k-1}\left(1 - (3/4)^{m_l}\right)$.
One can significantly improve this bound using that each lift yields at least two valid 3-decompositions (use x_1y_0 instead of x_0y_1 in Theorem 10) and, hence, two distinct possible edges in the matching.

6 Further research

The most pressing question is whether the flexibility of 3-decompositions can be exploited in order to prove the 3-decomposition conjecture on expander graphs or on symmetric graphs. Further, it is desirable to classify all graphs which have a unique 3-decomposition up to isomorphism.

References

[1] E. Aboomahigir, M. Ahanjideh, and S. Akbari. Decomposing claw-free subcubic graphs and 4-chordal subcubic graphs. *Discret. Appl. Math.*, 296:52–55, 2021.

[2] O. Bachtler and I. Heinrich. Reductions for the 3-decomposition conjecture. In *LAGOS*, volume 223 of *Procedia Computer Science*, pages 96–103. Elsevier, 2023.

[3] Y. Bilu and N. Linial. Constructing expander graphs by 2-lifts and discrepancy vs. spectral gap. In *FOCS*, pages 404–412. IEEE Computer Society, 2004.

[4] R. J. Douglas. NP-completeness and degree restricted spanning trees. *Discret. Math.*, 105(1-3):41–47, 1992.

[5] I. Heinrich. *On Graph Decomposition: Hajos' Conjecture, the Clustering Coefficient, and Dominating Sets.* PhD thesis, TU Kaiserslautern, 2019.

[6] A. Hoffmann-Ostenhof. *Nowhere-zero flows and structures in cubic graphs.* PhD thesis, University Vienna, 2011.

[7] A. Hoffmann-Ostenhof, T. Kaiser, and K. Ozeki. Decomposing planar cubic graphs. *J. Graph Theory*, 88(4):631–640, 2018.

[8] A. Hoffmann-Ostenhof, K. Noguchi, and K. Ozeki. On homeomorphically irreducible spanning trees in cubic graphs. *J. Graph Theory*, 89(2):93–100, 2018.

[9] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(4):439–561, 2006.

[10] R. Li and Q. Cui. Spanning trees in subcubic graphs. *Ars Comb.*, 117:411–415, 2014.

Computing edge-colored ultrahomogeneous graphs *

Irene Heinrich, Eda Kaja, and Pascal Schweitzer

TU Darmstadt, Darmstadt, Germany

Abstract

We develop a practical algorithm to enumerate all ultrahomogeneous edge-colored graphs up to a specified order. As input, the algorithm can take either a list of all coherent configurations or all transitive permutation groups. Efficiency is achieved by pruning lexicographic products quickly. We provide numerical data on the number of objects up to isomorphism for orders up to 34.

1 Introduction

Ultrahomogeneous structures are classical objects in model theory with applications in algebra and combinatorics. A structure $\mathcal{R}$ is *ultrahomogeneous* if every isomorphism between two induced substructures of $\mathcal{R}$ extends to an automorphism of $\mathcal{R}$. We are interested in algorithms for generating and handling finite ultrahomogeneous structures. In this paper we develop algorithms for the base case of structures with irreflexive binary relations. These structures can always be translated into loopless edge-colored graphs. In our search for ultrahomogeneous structures in the base case, it suffices to consider vertex-monochromatic coherent configurations. These are binary relational structures that satisfy certain regularity conditions implied by ultrahomogeneity. There is a well-known Galois correspondence between coherent configurations and permutation groups. There are thus two starting points that we can take for our search: coherent configurations or transitive groups. Since we are only interested in ultrahomogeneous structures we can limit ourselves to so-called 2-closed permutation groups on the permutation group side.

For various subclasses of finite ultrahomogeneous structures, explicit classifications are known. This is the case for simple graphs [3, 12], directed [10], 3-edge-colored [11], vertex-colored [9], and vertex-colored oriented [8] graphs. None of these classifications allow arbitrarily many edge colors. Also the primitive permutation groups of finite binary ultrahomogeneous structures have recently been classified [4].

As the class of graphs considered becomes larger, the classification of ultrahomogeneous objects becomes ever more complicated. A computer assisted approach seems to be in order.

Results. We develop a practical algorithm to enumerate all ultrahomogeneous edge-colored graphs up to a specified order. The algorithm takes as input a list of all coherent configurations or all transitive permutation groups of at most the given order.

Techniques. First, we provide a fast practical algorithm that checks whether a given object is ultrahomogeneous. Our algorithm can take either coherent configurations or permutation groups as input and checks whether the graph induced by the input is ultrahomogeneous (Subsection 3.1).

The lexicographic product operation of graphs preserves ultrahomogeneity and it turns out that many ultrahomogeneous objects are in fact lexicographic products of smaller ultrahomogeneous objects.

*The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (EngageS: Grant Agreement No. 820148) and from the German Research Foundation DFG (SFB-TRR 195 "Symbolic Tools in Mathematics and their Application"). Emails: lastname@mathematik.tu-darmstadt.de.

Second, to avoid a combinatorial explosion, we develop a linear time algorithm to prune such products. The algorithm takes as input a coherent configuration and determines that it is "not a lexicographic product" or renders the input as "a product or not ultrahomogeneous" (Subsection 3.2). The crux here is that in the latter case, we do not need to process the input since both non-ultrahomogeneous objects and products can be discarded. This allows our algorithm to avoid checking regularity of the input and thus run in time $O(kn)$, where k is the rank of the configuration (number of colors/binary relations) and n the number of vertices. The computational results are aggregated in Section 4.

2 Preliminaries

For $k \in \mathbb{N}$ we set $[k] \coloneqq \{1, 2, \dots, k\}$ and for a k-tuple $t = (v_1, v_2, \dots, v_k)$ we set $\pi_i(t) \coloneqq v_i$. The *restriction* of a map ψ to a set U is $\psi|_U$. If φ is a restriction of ψ, then ψ is an *extension* of φ.

A *binary relational structure* is a tuple $\mathcal{R} = (V, R_1, R_2, \dots, R_k)$ where V is a set of *vertices* and $R_i \subseteq V^2$ for each $i \in [k]$. We set $V(\mathcal{R}) \coloneqq V$. Throughout this paper we assume, for our purposes w.l.o.g., that $\{R_i \colon i \in [k]\}$ is a partition of V^2, there exists a $d \in [k]$ such that $R_d = \{(v, v) \colon v \in V\}$ is the *diagonal* of $\mathcal{R}$, and for each $i \in [k]$ either R_i is symmetric or there exists $j \in [k]$ such that $R_j = \{(v, u) \colon (u, v) \in R_i\}$. All relational structures in this paper are binary and finite. Two relational structures $\mathcal{R} = (V, R_1, R_2, \dots, R_k)$ and $\mathcal{S} = (W, S_1, S_2, \dots, S_k)$ are *isomorphic* if there exists a bijection $\varphi \colon V \to W$ such that for all $i \in [k]$ it holds that $(u, v) \in R_i$ if and only if $(\varphi(u), \varphi(v)) \in S_i$. In this case φ is an *isomorphism* from $\mathcal{R}$ to $\mathcal{S}$. If, additionally, $\mathcal{R} = \mathcal{S}$, then φ is an *automorphism* of $\mathcal{R}$. The automorphisms of $\mathcal{R}$ form a group, denoted by $\mathrm{Aut}(\mathcal{R})$. For a subset U of V the *induced substructure* of $\mathcal{R}$ on U is $\mathcal{R}[U] \coloneqq (U, R_1 \cap U^2, R_2 \cap U^2, \dots, R_k \cap U^2)$. A *partial isomorphism* from $\mathcal{R}$ to itself is an isomorphism between two induced substructures of $\mathcal{R}$. If every partial isomorphism from $\mathcal{R}$ to itself extends to an automorphism of $\mathcal{R}$, then $\mathcal{R}$ is *ultrahomogeneous*. The structures $\mathcal{R}$ and $\mathcal{S}$ are *equivalent* if there exists a permutation σ of $[k]$ such that $\mathcal{R}$ is isomorphic to $(W, S_{\sigma(1)}, S_{\sigma(2)}, \dots, S_{\sigma(k)})$.

A (vertex) *coloring* of $\mathcal{R}$ is a map $\chi \colon V(\mathcal{R}) \to C$ where C is some set of *colors*. Then $(\mathcal{R}, \chi)$ is a *colored relational structure*. An inclusion-wise maximal set $U \subseteq V(\mathcal{R})$ with $|\chi(U)| = 1$ is a color class of $(\mathcal{R}, \chi)$. The definitions of isomorphisms, automorphisms, and ultrahomogeneity directly transfer to the context of colored structures, where it is important that isomorphisms preserve vertex colors. Note that there is a 1:1-correspondence between colored relational structures we consider and complete edge-colored digraphs, where the edge colors of the digraph correspond to the indices of the relations.

Coherent configurations. A relational structure $\mathcal{R} = (V, R_1, R_2, \dots, R_k)$ is a *coherent configuration*[1] if for every choice of three indices $a, b, c \in [k]$ and $(u, w) \in R_a$ the number of elements $v \in V$ such that $(u, v) \in R_b$ and $(v, w) \in R_c$ is a constant $\lambda^a_{b,c}$ which is independent of the choice of u and w. A coherent configuration $\mathcal{R}$ is *symmetric* if all the relations are symmetric.

Lexicographic products. Let $\mathcal{R} = (V, R_1, R_2, \dots, R_k)$ and $\mathcal{S} = (W, S_1, S_2, \dots, S_\ell)$ be two relational structures such that R_k is the diagonal of $\mathcal{R}$ and S_ℓ is the diagonal of S. The *lexicographic product* of $\mathcal{R}$ and $\mathcal{S}$ is the relational structure $\mathcal{R} \cdot \mathcal{S} = (V \times W, \dot{R}_1, \dot{R}_2, \dots, \dot{R}_{k-1}, \dot{S}_1, \dot{S}_2, \dots, \dot{S}_\ell)$ where $\dot{R}_i = \{((v, w), (v'w')) \in (V \times W)^2 \colon (v, v') \in R_i\}$ and $\dot{S}_j = \{((v, w), (v, w')) \colon v \in V, (w, w') \in S_j\}$. We emphasize that the index i of $\dot{R}_i$ is at most $k - 1$ (otherwise the diagonal $\dot{S}_\ell$ would have a non-empty intersection with $\dot{R}_k$). Observe that the relations of $\mathcal{R} \cdot \mathcal{S}$ indeed form a partition $V \times W$. We say that $\mathcal{R}$ or $\mathcal{S}$ is a *trivial factor* of $\mathcal{R} \cdot \mathcal{S}$ if $|V| = 1$ or $|W| = 1$, respectively.

Lemma 1. *If* $|W| \geq 2$ *and* $\min_{i \in [k]}\{|R_i|\} \geq |V|$, *then* $\max_{j \in [\ell]} |\dot{S}_j| < \min_{i \in [k-1]} |\dot{R}_i|$. *In particular, if* $\mathcal{R}$ *is a coherent configuration and* $\mathcal{S}$ *is a non-trivial factor of* $\mathcal{R} \cdot \mathcal{S}$, *then* $\max_{j \in [\ell]} |\dot{S}_j| < \min_{i \in [k-1]} |\dot{R}_i|$.

[1] Technically these are colored coherent configurations and usually the underlying uncolored object is considered, i.e., the ordering of the relations is ignored. Certain coherent configurations are sometimes called association schemes.

In general, we say that a structure *is a lexicographic product* whenever it is equivalent to a lexicographic product. We say that $\mathcal{R}$ is *prime* if $|V(\mathcal{R})| \geq 2$ and every structure $\mathcal{R}'$ equivalent to $\mathcal{R}$ satisfies: $\mathcal{R}' = \mathcal{S}_1 \cdot \mathcal{S}_2$ implies $\min\{|V(\mathcal{S}_1)|, |V(\mathcal{S}_2)|\} = 1$.

Groups. The *symmetric group* of a non-empty set V is $\mathrm{Sym}(V)$. A *permutation group* Γ on V is a subgroup of $\mathrm{Sym}(V)$. For $v \in V$ and $\gamma \in \Gamma$ we set $v^\gamma := \gamma(v)$. An *action* of Γ on V is a homomorphism ϕ from Γ to $\mathrm{Sym}(V)$. The image of Γ under ϕ is a subgroup of $\mathrm{Sym}(V)$ called the permutation group *induced* by Γ on V, denoted Γ^V. The *orbit* of $x \in V$ is $x^\Gamma := \{x^\gamma \colon \gamma \in \Gamma\}$. We say Γ is *transitive* on V if $x^\Gamma = V$ for all $x \in V$. The *stabilizer* of $x \in V$ is $\mathrm{Stab}_\Gamma(x) := \{\gamma \in \Gamma \colon x^\gamma = x\}$. The *pointwise stabilizer of* $X \subseteq V$ is $\mathrm{pwStab}_\Gamma(X) := \bigcap_{x \in X} \mathrm{Stab}_\Gamma(x)$. If Γ is a transitive permutation group on V, then the partiton of $V \times V$ into orbits of Γ is a coherent configuration $\mathcal{R}(\Gamma)$. Note that $\Gamma \leq \mathrm{Aut}(\mathcal{R}(\Gamma))$. If equality holds, then Γ is 2-*closed*, that is, Γ equals its 2-*closure*, which is the largest subgroup of $\mathrm{Sym}(V)$ which preserves the orbits of Γ on $V \times V$. A coherent configuration $\mathcal{R}$ such that $\mathcal{R} = \mathcal{R}(\Gamma)$ for some transitive permutation group Γ is called *Schurian*.

Block systems. Let $\Gamma \leq \mathrm{Sym}(V)$ be transitive. A *block* is a set $B \subseteq V$ such that $B^\gamma = B$ or $B^\gamma \cap B = \emptyset$ for all $\gamma \in \Gamma$. If $|B| \in \{1, |V|\}$, then B is *trivial*. If B is a block, then $\mathcal{B} := \{B^\gamma \colon \gamma \in \Gamma\}$ is a *block system* of V. Note that $\mathcal{B}$ is a partition of V which is invariant under the action of Γ. A Schurian coherent configuration $\mathcal{R}$ is *imprimitive* if $\mathrm{Aut}(\mathcal{R})$ is imprimitive, that is $\mathrm{Aut}(\mathcal{R})$ admits a non-trivial block system.

Lemma 2. *If $\mathcal{R}$ is a coherent configuration, then up to equivalence there is a unique factorization of $\mathcal{R}$ into prime factors with respect to the lexicographic product.*

Proof sketch. We first observe that, up to equivalence the lexicographic product is associative. Next, it can be shown that if $\mathcal{R} \cdot \mathcal{S} = \mathcal{R}' \cdot \mathcal{S}'$ then $\mathcal{S} \leq \mathcal{S}'$ or $\mathcal{S}' \geq \mathcal{S}'$, meaning $\mathcal{S}$ is an induced substructure of $\mathcal{S}'$ or vice versa. Finally we observe that if $\mathcal{S} < \mathcal{S}'$ then $\mathcal{R}' \cdot \mathcal{S}'$ is equivalent to $\mathcal{R}' \cdot \mathcal{T} \cdot \mathcal{S}$ for some structure $\mathcal{T}$. □

As for graphs [11], for coherent configurations lexicographic products also preserve ultrahomogeneity.

Lemma 3. *If $\mathcal{R}$ and $\mathcal{S}$ are relational structures, then $\mathcal{R} \cdot \mathcal{S}$ is ultrahomogeneous if and only if both structures $\mathcal{R}$ and $\mathcal{S}$ are ultrahomogeneous.*

3 Algorithms

Let $\mathcal{R} = (V, R_1, R_2, \ldots, R_k)$ be a coherent configuration. For our practical computations and their analysis we assume that $V = [|V|]$ and that we are given $\mathcal{R}$ as a $|V| \times |V|$-*adjacency matrix* $A(\mathcal{R})$ with $A_{i,j} = s$ precisely if $(i, j) \in R_s$.

Definition 4. *For a subset $W \subseteq V$, the* neighborhood partition $\mathcal{P}_\mathcal{R}(W)$ *is the partition of $V \setminus W$ where two elements $i, j \in V \setminus W$ are in the same part if and only if $A(\mathcal{R})_{w,i} = A(\mathcal{R})_{w,j}$ for every $w \in W$.*

3.1 Checking ultrahomogeneity

Lemma 5. *If $(\mathcal{R}, \chi)$ is a colored binary relational structure, then $(\mathcal{R}, \chi)$ is ultrahomogeneous if and only if the following conditions hold for every color class C of $(\mathcal{R}, \chi)$:*

1. *C is an orbit of $\mathrm{Aut}((\mathcal{R}, \chi))$.*
2. *For some (and thus by Part 1 every) $v_c \in C$ the structure $(\mathcal{R}[V(\mathcal{R}) \setminus \{v_c\}], \chi^{v_c})$ is ultrahomogeneous, where χ^{v_c} is a coloring whose color classes form the meet of the neighborhood partition $\mathcal{P}_\mathcal{R}(\{v_c\})$ and the color classes of χ (i.e., it is the coarsest partition which is finer than both of them).*

function `is_ultrahomogeneous`(A, W);
Input : an adjacency matrix A of a coherent configuration $\mathcal{R}$ and a subset $W \subseteq V(\mathcal{R})$
Output: **true** if $\mathcal{R}$ is ultrahomogeneous with respect to W and **false** otherwise
if *there is a part in* $\mathcal{P}_\mathcal{R}(W)$ *on which* pwStab(W) *does not act transitively* **then return false**;
$L \coloneqq$ a list containing precisely one vertex of every part in $\mathcal{P}_\mathcal{R}(W)$;
for $v \in L$ **do**
  if `is_ultrahomogeneous`$(A, W \cup \{v\}) ==$ **false then return false**;
end
return true;

Algorithm 1: Checking if a coherent configuration $\mathcal{R}$ is ultrahomogeneous. The input is the adjacency matrix A of $\mathcal{R}$ and a set of vertices W (default: empty).

Theorem 6. *A coherent configuration $\mathcal{R}$ is ultrahomogeneous if and only if Algorithm 1 returns "*true*" when called on the input (A, W) with $A = A(\mathcal{R})$ and $W = \emptyset$.*

Proof. For $\{w_1, w_2, \ldots, w_s\} \subseteq V(\mathcal{R})$ observe that the neighborhood partition $\mathcal{P}_\mathcal{R}(\{w_1, w_2, \ldots, w_i\})$ is precisely the partition of $\mathcal{R}[V(\mathcal{R}) \setminus \{w_1, w_2, \ldots, w_i\}]$ into the color classes with respect to $(((\chi^{w_1})^{w_2})^{\cdots})^{w_i}$ (for the definition of this coloring, see Lemma 5). Recursively applying Lemma 5 yields the theorem. □

Ignoring the running time of basic group theoretic algorithms (i.e., using the Schreier-Sims algorithm to compute point-wise stabilizers), the running time of Algorithm 1 can be bounded using the number of irredundant bases up to equivalence under the group action. However, using some heuristics in particular to deal with permutations of the sequences of chosen points, one can significantly reduce this running time requirement.

3.2 Checking lexicographic products

Lemmas 2 and 3 imply that once we have, up to some order, the number of ultrahomogeneous relational structures that are not a lexicographic product, we can compute the number of all such structures, including the products. We therefore develop a fast algorithm that can discard lexicographic products.

Input : the adjacency matrix $A(\mathcal{R})$ of a coherent configuration $\mathcal{R} = (V, R_1, R_2, \ldots, R_k)$ where R_1 is the diagonal of $\mathcal{R}$ and $|R_i| \leq |R_j|$ whenever $i \leq j$
Output: either "not a lexicographic product" or "lexicographic product or not ultrahomogeneous"
if $k \leq 2$ **then return** "not a lexicographic product";
min_$j \coloneqq 2$;
choose $v_0 \in V(\mathcal{R})$;
for i **from** *1* **to** $k - 1$ **do**
  choose $v \in V(\mathcal{R})$ with $(v_0, v) \in R_i$;
  for $w \in V(\mathcal{R}) \setminus \{v_0, v\}$ **do**
    if $A_{v_0 w} \neq A_{vw}$ **then** min_$j \coloneqq \max(\text{min_}j, A_{v_0,w} + 1, A_{v,w} + 1)$;
  end
  if $i + 1 ==$ min_j **then return** "lexicographic product or not ultrahomogeneous";
end
return "not a lexicographic product" ;

Algorithm 2: Check if a coherent configuration is a non-trivial lexicographic product.

Theorem 7. *The output of Algorithm 2 is correct.*

Assuming the relations are already ordered by size, the running time of Algorithm 2 is $O(kn)$ where k is the number of relations (rank) and n the number of vertices. Note that the fact that not even the entire input has to be checked is achieved by leveraging the assumed ultrahomogeneity.

4 Computations

Ultrahomogeneity implies coherence, and there is a complete database of coherent configurations of order at most 34 [6] (see also the paper series of Hanaki and Myamoto [5, 7]). Complete data for 38 is also available.

Our approach for the generation of ultrahomogeneous binary relational structures is to run our ultrahomogeneity test (Algorithm 1) on the configurations.

Thin coherent configurations are omitted in the data base of coherent configurations [6]. Since every transitive thin coherent configuration is ultrahomogeneous and the thin coherent configurations correspond exactly to the transitive permutation groups, they exactly account for the difference.

We used SageMath [13] for our computations. The coherent configurations are given via adjacency matrices. To filter out some of them we have the following observation.

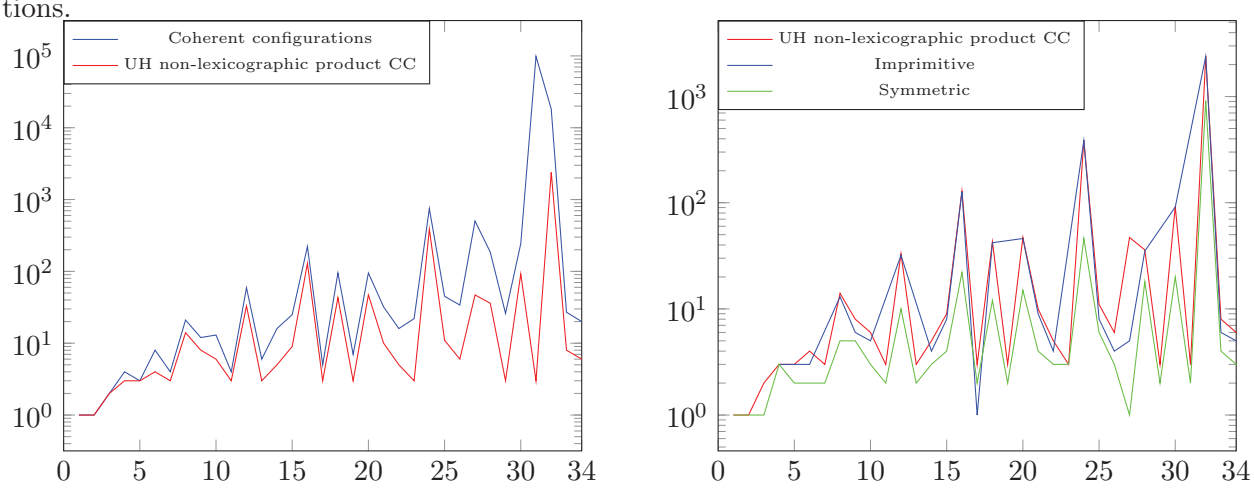
Lemma 8. *If a coherent configuration is ultrahomogeneous, then it is Schurian.*

Proof. Suppose $\mathcal{R} = (V, R_1, R_2, \ldots, R_k)$ is an ultrahomogeneous coherent configuration. We argue that $\mathcal{R}' \coloneqq \mathcal{R}(\mathrm{Aut}(\mathcal{R}))$ is equivalent to $\mathcal{R}$. It is clear that $\mathcal{R}'$ is at least as fine as $\mathcal{R}$, that is, if (v, w) and (v', w') are in the same relation of $\mathcal{R}'$ then they are in the same relation of $\mathcal{R}$. For the other direction, if (v, w) and (v', w') are in the same relation of $\mathcal{R}$ then by ultrahomogeneity there is an automorphism mapping (v, w) to (v', w'), and thus the two pairs are in the same relation of $\mathcal{R}'$. □

Thus we may restrict our attention to Schurian coherent configurations. Using Algorithm 2 we filter out the lexicographic products and then apply Algorithm 1 to the remaining configurations.

The fact that we can limit ourselves to Schurian coherent configurations is crucial since this gives us an alternative for order 31. Rather than considering the 98307 coherent configurations of order 31, we make use of GAP [2] and the *AssociationSchemes* [1] package. By the Galois correspondence, Schurian coherent configurations are in 1:1-relation with 2-closed groups. Hence we first compute the list consisting of the 2-closures of the 12 transitive groups of degree 31. There are 8 resulting Schurian coherent configurations coming from these groups (including one thin coherent configuration), whose adjacency matrices can be obtained using the *AssociationSchemes* package, and then we apply Algorithm 1 to check for ultrahomogeneity. This approach is equivalent to working with the matrices, and in this particular case it reduced the workload significantly. Indeed, it turns out that there are orders for which there are significantly fewer transitive groups, while there are other orders for which there are significantly fewer coherent configurations. The total computation was less than one day on a personal computer (Intel i7 at 2.8 GHz).

Figure 1: Numerical data surrounding ultrahomogeneity of vertex-monochromatic coherent configurations.



In Figure 1 we depict the number of ultrahomogeneous relational structures of order up to 34. On the left side, we present the total number of ultrahomogeneous coherent configurations which are not lexicographic products, compared to the total number of homogeneous coherent configurations. On the right side, we show the number of imprimitive coherent configurations and symmetric coherent configurations within the overall count of ultrahomogeneous coherent configurations.

5 Future work

We generated all ultrahomogeneous edge-colored graphs up to order 34. In particular by the pruning of lexicographic products, our algorithms are comfortably efficient enough to compute the number of ultrahomogeneous graphs in the order ranges in which the coherent configurations are available. However, there is ample room for speeding up the algorithms using additional pruning. Algorithm 1 can be sped up by considering only canonical sequences of points v chosen recursively. More pressing is an analysis of the ultrahomogeneous graphs that are not lexicographic products. Certain other ultrahomogeneity preserving general constructions are known, but the question of whether we can use product structures to provide a concise classification, preferably admitting efficient algorithms, remains.

References

[1] J. Bamberg, A. Hanaki, and J. Lansdown. AssociationSchemes: A GAP package for working with association schemes and homogeneous coherent configurations, Version 3.0.0. 2023.

[2] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.13.0*, 2024.

[3] A. Gardiner. Homogeneous graphs. *J. Combinatorial Theory Ser. B*, 20(1):94–102, 1976.

[4] N. Gill, M. W. Liebeck, and P. Spiga. *Cherlin's conjecture for finite primitive binary permutation groups*, volume 2302 of *Lecture Notes in Mathematics*. Springer, Cham, 2022.

[5] A. Hanaki, H. Kharaghani, A. Mohammadian, and B. Tayfeh-Rezaie. Classification of skew-Hadamard matrices of order 32 and association schemes of order 31. *J. Combin. Des.*, 28(6):421–427, 2020.

[6] A. Hanaki and I. Miyamoto. Classification of association schemes with small vertices. `http://math.shinshu-u.ac.jp/~hanaki/as/`.

[7] A. Hanaki and I. Miyamoto. Classification of association schemes of small order. *Discrete Math.*, 264(1-3):75–80, 2003.

[8] I. Heinrich, E. Kaja, and P. Schweitzer. Finite vertex-colored ultrahomogeneous oriented graphs. In *Graph-Theoretic Concepts in Computer Science - 50th International Workshop, WG 2024, Gozd Martuljek, Slovenia, June 19-21, 2024*, Lecture Notes in Computer Science, 2024. To appear.

[9] I. Heinrich, T. Schneider, and P. Schweitzer. Classification of finite highly regular vertex-coloured graphs. *arXiv preprint arXiv:2012.01058*, 2020.

[10] A. H. Lachlan. Finite homogeneous simple digraphs. In *Proceedings of the Herbrand Symposium Logic Colloquium '81*, pages 189–208. North-Holland Publishing Company, 1982.

[11] A. H. Lachlan. Binary homogeneous structures. I, II. *Proc. London Math. Soc. (3)*, 52(3):385–411, 412–426, 1986.

[12] J. Sheehan. Smoothly embeddable subgraphs. *J. London Math. Soc. (2)*, 9:212–218, 1974/75.

[13] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.3)*, 2024. `https://www.sagemath.org`.

Extended abstract of Regular polytopes, sphere packings and Apollonian sections*

Iván Rasskin†[1]

[1]Laboratoire d'Informatique et des Systèmes, Aix-Marseille Université, Campus de Luminy, France

Abstract

In this paper, we explore the geometry and the arithmetic of a family of polytopal sphere packings induced by regular polytopes in any dimension. We prove that every integral polytope is crystallographic and we show that there are 11 crystallographic regular polytopes in any dimension. After introducing the notion of Apollonian section, we determine which Platonic crystallographic packings emerge as cross sections of the Apollonian arrangements of the regular 4-polytopes. Additionally, we compute the Möbius spectrum of every regular polytope.

1 Introduction

Apollonian circle packings and their generalizations are currently active areas of research in geometric number theory [9, 10, 2]. In dimension 2, some variants of integral Apollonian packings have been explored by substituting the building block with a different circle packing modeled on a polyhedron [8, 22, 23, 3, 5, 14]. While every polyhedron can be employed to construct a packing, not all of them admit an integral structure like the Apollonian one. A fundamental question regarding the determination of which polyhedra are *integral* in this sense is still wide open [12, 5].

Similarly, in dimension 3, a family of crystallographic/Apollonian-like sphere packings arise by iteratively reflecting an initial sphere packing modeled on a 4-polytope as in Figure 1. Integral crystallographic packings modeled on the 4-simplex [20, 11] and the 4-cross polytope [13, 7, 19, 16] have been extensively studied. Unlike polyhedra, not every 4-polytope is *crystallographic*, in the sense that it serves as a suitable model for a crystallographic packing. In this paper, we delve into the crystallography and the integrality of regular polytopes in any dimension.

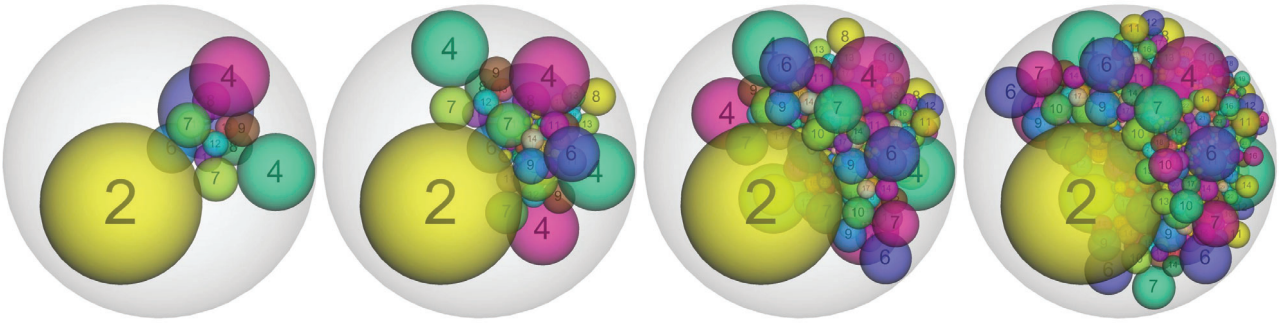


Figure 1: An integral hypercubic crystallographic packing after 0, 1, 2 and 3 iterations. The labels are the *bends* (reciprocal of the radii) of the spheres.

*The full version of this work can be found in [18], which is a recent update of a previous preprint including some partial results of this version. This work is currently under review and it is partially contained in the PhD thesis of the author [17]. This research is supported by the CNRS and the Austrian Science Fund FWF projects F-5503 and P-34763

†Email: ivan.rasskin@lis-lab.fr. Research of I. R. supported by the CNRS

2 Preliminaries on sphere packings and edge-scribable polytopes

An *oriented hypersphere*, or simply *sphere*, of $\widehat{\mathbb{R}^d} := \mathbb{R}^d \cup \{\infty\}$, is the image of a spherical cap of $\mathbb{S}^d$ under the stereographic projection. Every sphere S is uniquely defined by its center $c \in \widehat{\mathbb{R}^d}$ and its bend $b \in \mathbb{R}$ (the recripocal of the *oriented* radius), or if S is a half-space, by its normal vector $\widehat{n} \in \mathbb{S}^{d-1}$ pointing to the interior and the signed distance $\delta \in \mathbb{R}$ between its boundary and the origin. The *inversive coordinates* of S are represented by the $(d+2)$-dimensional real vector

$$\mathbf{i}(S) = \begin{cases} \left(bc, \dfrac{\overline{b}-b}{2}, \dfrac{\overline{b}+b}{2}\right)^T & \text{if } b \neq 0, \\ \\ (\widehat{n}, \delta, \delta)^T & \text{otherwise} \end{cases} \tag{1}$$

where $\overline{b} = b\|c\|^2 - \frac{1}{b}$ is the *co-bend* of S. The co-bend is the bend of S after inversion through the unit sphere. The *inversive product* of two spheres S, S' of $\widehat{\mathbb{R}^d}$ is the real value

$$\langle S, S' \rangle = \mathbf{i}(S)^T \mathbf{Q}_{d+2} \mathbf{i}(S) \tag{2}$$

where $\mathbf{Q}_{d+2}$ is the diagonal matrix $\mathrm{diag}(1, \ldots, 1, -1)$ of size $d+2$. The inversive product encodes the relative position of two spheres S and S' according to the following criteria:

$$\langle S, S' \rangle \begin{cases} < -1 & \text{if } S \cap S' = \emptyset, \\ = -1 & \text{if } \partial S \text{ and } \partial S' \text{ are tangent and } \mathrm{int}(S) \cap \mathrm{int}(S') = \emptyset, \\ = 1 & \text{if } \partial S \text{ and } \partial S' \text{ are tangent and } S \subseteq S' \text{ or } S' \subseteq S, \\ > 1 & \text{if } \partial S \cap \partial S' = \emptyset \text{ and } S \subset S' \text{ or } S' \subset S. \end{cases} \tag{3}$$

An arrangement of spheres $\mathcal{S}$ in $\widehat{\mathbb{R}^d}$, possible infinite, is a *packing* if their interiors are mutually disjoint. The group of Möbius transformations of $\widehat{\mathbb{R}^d}$ preserves the inversive product and acts linearly on the inversive coordinates as an orthogonal subgroup of $\mathrm{SL}_{d+2}(\mathbb{R})$ with respect to $\mathbf{Q}_{d+2}$.

For every $d \geq 1$, we denote the *polar* of a subset $X \subset \mathbb{R}^d$ by $X^* = \{u \in \mathbb{R}^d \mid \langle u, v \rangle \leq 1 \text{ for all } v \in X\}$. The *stereographic sphere* of a point $v \in \mathbb{R}^d$ *outside* $\mathbb{S}^{d-1}$ (i.e. with $\|v\| > 1$) is the sphere S_v of $\widehat{\mathbb{R}^{d-1}}$ obtained by the stereographic projection of the spherical cap $\{-v\}^* \cap \mathbb{S}^{d-1}$. For any d-polytope $\mathcal{P}$ with vertices outside the unit sphere, the *(sphere) arrangement projection* of $\mathcal{P}$ is defined as the arrangement $\mathcal{S}_\mathcal{P}$ of the stereographic spheres of the vertices of $\mathcal{P}$.

A d-polytope is termed *edge-scribed* if its edges are tangent to the unit sphere [6]. If, in addition, the barycenter of the contact points is the origin, it is referred to as *canonical* [24]. A d-polytope is considered *edge-scribable* if it admits an edge-scribed realization [6]. In dimension $d \geq 3$, all the edge-scribed realizations of an edge-scribable d-polytope $\mathcal{P}$ are equivalent up to Möbius transformations to a unique canonical realization $\mathcal{P}_0$ (see [21, 14] for more details).

The arrangement projection of an edge-scribed polytope is a packing. Reciprocally, we say that a sphere packing $\mathcal{S}_\mathcal{P}$ in $\widehat{\mathbb{R}^d}$ with $d \geq 2$, is *polytopal* if there is an edge-scribable $(d+1)$-polytope $\mathcal{P}$ and a Möbius transformation μ such that $\mathcal{S}_\mathcal{P} = \mu \cdot \mathcal{S}_{\mathcal{P}_0}$. The combinatorial structure of $\mathcal{S}_\mathcal{P}$ is encoded by the corresponding edge-scribable polytope $\mathcal{P}$. The vertices and the edges of $\mathcal{P}$ are in bijection to the spheres and the tangency relations of $\mathcal{S}_\mathcal{P}$. The facets of $\mathcal{P}$ correspond to the *dual spheres* of $\mathcal{S}_\mathcal{P}$ which are the spheres forming the *dual arrangement* $\mathcal{S}^*_\mathcal{P} := \mu \cdot \mathcal{S}_{\mathcal{P}^*_0}$. The *Apollonian arrangement* of $\mathcal{S}_\mathcal{P}$ is defined as the orbit space $\mathscr{P}(\mathcal{S}_\mathcal{P}) := \langle \mathcal{S}^*_\mathcal{P} \rangle \cdot \mathcal{S}_\mathcal{P}$ where $\langle \mathcal{S}^*_\mathcal{P} \rangle$ denotes the group generated by inversions through the dual spheres. We denote by $\mathscr{P}_{\{p_1,\ldots,p_d\}}$ the Apollonian arrangement of a regular polytope with Schläfli symbol $\{p_1, \ldots, p_d\}$.

2.1 Crystallographic polytopes

In dimension 2, the Apollonian arrangements of 3-polytopes are packings, but this is not true in general [14]. In higher dimensions, Apollonian arrangements which are packings belong to the family of *crystallographic sphere packings* introduced by Kontorovich and Nakamura in [12]. These are dense infinite sphere packings obtained as the orbit space $\mathscr{P} = \langle \widetilde{\mathcal{S}} \rangle \cdot \mathcal{S}$, where $\mathcal{S}$ is a finite sphere packing called the *cluster*, $\langle \widetilde{\mathcal{S}} \rangle$ is a geometrically finite subgroup of the group of Möbius transformations generated by the inversions through a finite arrangement of spheres $\widetilde{\mathcal{S}}$, called the *co-cluster*, satisfying that every sphere of $\mathcal{S}$ is disjoint, tangent or orthogonal to every sphere of $\widetilde{\mathcal{S}}$.

Definition 1. *For every $d \geq 3$, an edge-scribable d-polytope $\mathcal{P}$ is crystallographic if any Apollonian arrangement $\mathscr{P}(\mathcal{S}_\mathcal{P}) = \langle \mathcal{S}_\mathcal{P}^* \rangle \cdot \mathcal{S}_\mathcal{P}$ is a sphere packing in dimension $d-1$.*

Crystallographic polytopes exist only in dimension $3 \leq d \leq 19$ [2]. From a Boyd's remark in [4], we have that an edge-scribable polytope $\mathcal{P}$ is crystallographic when the dihedral angles of $\mathcal{P}$, viewed as an hyperideal hyperbolic polytope, satisfy the *crystallographic restriction*. This restriction dictates that the periode of every rotation obtained as the product of two reflections through the facets is either $2, 3, 4, 6, \infty$, imposing a condition on the dihedral angles. On the other hand, the dihedral angle α of two adjacents facets f and f' of $\mathcal{P}$ is equal to the *intersection angle* of the corresponding dual spheres of $S_f, S_{f'} \in \mathcal{S}_\mathcal{P}^*$, as defined in [15]. This angle can be computed from their inversive product by $\langle S_f, S_{f'} \rangle = \cos(\alpha)$. Therefore, the crystallographic restriction can be reformuled in terms of the inversive product of the dual spheres, as described in Lemma 2.

Lemma 2. *For any $d \geq 3$, an edge-scribable d-polytope $\mathcal{P}$ is crystallographic if and only if for any two dual spheres S_f, $S_{f'}$ of a polytopal sphere packing $\mathcal{S}_\mathcal{P}$, we have $|\langle S_f, S_{f'} \rangle| \in \{0, \frac{1}{2}, \frac{\sqrt{2}}{2}, \frac{\sqrt{3}}{2}\} \cup [1, \infty)$.*

2.2 Integral polytopes

In [12], Kontorovich and Nakamura defined a 3-polytope $\mathcal{P}$ to be *integral*[1] if there is a crystallographic circle packing modeled on $\mathcal{P}$ where the bends of the spheres are all integers. The fundamental question regarding the determination of which 3-polytopes are integral is still wide open. In [5], Chait-Roth, Cui, and Stier studied the integral 3-polytopes with few vertices. Based on previous works of Nakamura and Kontorovich, they gave the following enumeration of the integral uniform 3-polytopes.

Theorem 3 (Th. 26 [5])**.** *There are only 8 integral uniform 3-polytopes: the tetrahedron, the octahedron, the cube, the cuboctahedron, the truncated tetrahedron, the truncated octahedron, the 3-prism and the 6-prism.*

In higher dimensions, the previous definition of integral 3-polytope can be naturally extended for any edge-scribable polytope.

Definition 4. *For any $d \geq 3$, an edge-scribable d-polytope $\mathcal{P}$ is integral if it admits an Apollonian arrangement $\mathscr{P}(\mathcal{S}_\mathcal{P})$ where the bends of the spheres are in $\mathbb{Z}$.*

A priori, an edge-scribable polytope might be integral and non-crystallographic, meaning that it could admit an Apollonian arrangement where the bends of the spheres are integers and the spheres overlap. Indeed, this is the case if we adapt the definition of integral polytope for number rings other than $\mathbb{Z}$. For instance, the 600-cell is integral in $\mathbb{Z}[\varphi]$, but is not crystallographic (see Figure 2).

[1] This definition of *integral polytope* differs from the one commonly employed in combinatorics, which involves polytopes with integer vertex coordinates, also known as *lattice polytopes*.

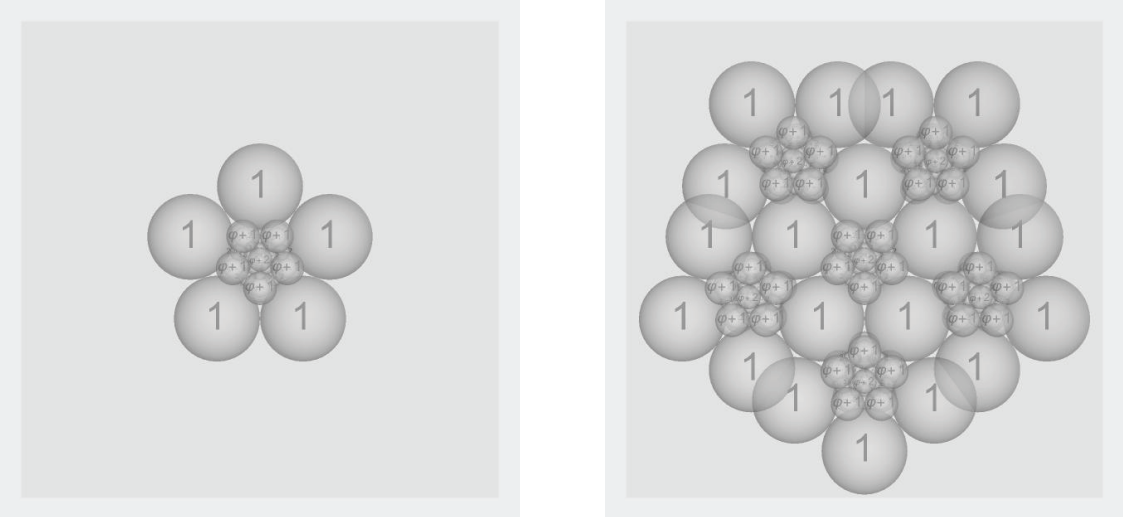


Figure 2: (Left) A polytopal sphere packing modeled on the 600-cell labelled with the bends; (right) the first reflections of its Apollonian arrangement which is integral in $\mathbb{Z}[\varphi]$ and is not a packing.

3 Main results

3.1 The relation between crystallography and integrality

In this paper, we prove the following condition for determining the integrality of edge-scribable polytopes.

Lemma 5. *For any $d \geq 3$, if an edge-scribable d-polytope $\mathcal{P}$ is integral, then for any two dual spheres S_f, $S_{f'}$ of any polytopal sphere packing $\mathcal{S}_\mathcal{P}$, we have $|\langle S_f, S_{f'}\rangle| \in \{\frac{\sqrt{n}}{2} \,|\, n \in \mathbb{N}\}$.*

With this lemma we can easily identify a mistake in the list of the integral uniform 3-polytopes of Chait-Roth, Cuit and Stier (Th. 3): the 6-prism is not integral, since it contains two dual spheres whose inversive product is $-5/3 \notin \{\pm\frac{\sqrt{n}}{2}\}_{n\in\mathbb{N}}$. Another straightforward consequence follows from Lemmas 2 and 5, and gives us the relation between crystallographic and integral polytopes in higher dimensions.

Theorem 6. *Every integral polytope is crystallographic.*

In the case of regular polytopes, we have the following.

Theorem 7. *For every $d \geq 3$, the only crystallographic regular d-polytopes are:*

($d = 3$) *the five Platonic solids,*

($d = 4$) *all the regular 4-polytopes except the 600-cell,*

($d = 6$) *the 6-cross polytope.*

Moreover, all these are integral except the icosahedron, the dodecahedron and the 120-cell which are integral in $\mathbb{Z}[\varphi]$.

3.2 Apollonian sections

The study of cross-sections is a classic approach for extracting patterns of crystallographic sphere packings [4, 1]. In this paper, we introduce an algebraic tool called *Apollonian section* which proves to be useful for identifing which Platonic crystallographic circle packings emerge as cross-sections of the Apollonian arrangements of the regular 4-polytopes.

Theorem 8. *There are the following relations between the Apollonian arrangements of the regular d-polytopes for $d = 3, 4$:*

$$\begin{aligned} \mathscr{P}_{\{3,3\}} &\subset \mathscr{P}_{\{3,3,3\}}, \\ \mathscr{P}_{\{3,3\}}, \mathscr{P}_{\{3,4\}}, \mathscr{P}_{\{4,3\}} &\subset \mathscr{P}_{\{3,3,4\}}, \\ \mathscr{P}_{\{4,3\}} &\subset \mathscr{P}_{\{4,3,3\}}, \\ \mathscr{P}_{\{3,4\}}, \mathscr{P}_{\{4,3\}} &\subset \mathscr{P}_{\{3,4,3\}}, \\ \mathscr{P}_{\{3,3\}}, \mathscr{P}_{\{3,5\}} &\subset \mathscr{P}_{\{3,3,5\}}, \\ \mathscr{P}_{\{5,3\}} &\subset \mathscr{P}_{\{5,3,3\}}, \end{aligned} \tag{4}$$

where "$\mathscr{P}_{\{p,q\}} \subset \mathscr{P}_{\{r,s,t\}}$" means that $\mathscr{P}_{\{p,q\}}$ can be obtained as a cross-section of $\mathscr{P}_{\{r,s,t\}}$.

Some of these cross-sections have been used as a geometric framework for obtaining results in geometric knot theory, as discussed in [16]. Another important feature of this approach is that it enable us to determine whether a cross-section preserves integrality.

Corollary 9. *Every integral Platonic crystallographic circle packing can be obtained as a cross-section of an integral regular crystallographic sphere packing.*

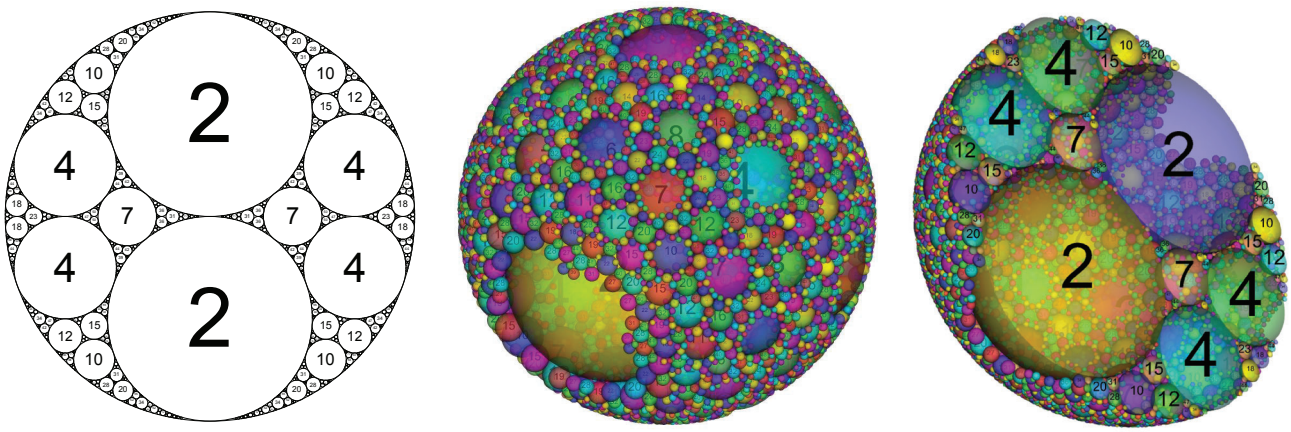


Figure 3: (Left) An integral octahedral crystallographic circle packing $\mathscr{P}_{\{3,4\}}$ obtained as a cross-section (right) of an integral orthoplicial crystallographic sphere packing $\mathscr{P}_{\{3,3,4\}}$ (center).

3.3 The Möbius spectrum of the regular polytopes

In [14], Ramírez Alfonsín and the author introduced a spectral invariant of every edge-scribable d-polytope $\mathcal{P}$ with $d \geq 3$ called the *Möbius spectrum* $\mathfrak{M}(\mathcal{P})$. This is defined as the multiset of the eigenvalues of the Gramian of any polytopal sphere packing $\mathcal{S}_{\mathcal{P}}$. Due to the Möbius uniqueness of edge-scribable polytopes, $\mathfrak{M}(\mathcal{P})$ does not depend on the packing. It is currently unknown whether there exist two combinatorially different edge-scribable polytopes with the same Möbius spectrum. In this paper, we compute the Möbius spectrum of every regular polytope $\mathcal{P}$ in terms of the number of vertices and another geometric invariant called the *canonical length* $\ell_{\mathcal{P}}$, defined as half the edge-length of a canonical realization of $\mathcal{P}$.

Theorem 10. *For any $d \geq 3$, the Möbius spectrum of every regular d-polytope $\mathcal{P}$ with n vertices is*

$$\mathfrak{M}(\mathcal{P}) = (-n\ell_{\mathcal{P}}^{-2}, \frac{n}{d}(1 + \ell_{\mathcal{P}}^{-2})^{(d)}, 0^{(n-d-1)}). \tag{5}$$

References

[1] A. Baragar. Higher dimensional Apollonian packings, revisited. *Geometriae Dedicata* **195** (2018), 137–161.

[2] N. Bogachev, A. Kolpakov, and A. Kontorovich. Kleinian sphere packings, reflection groups, and arithmeticity. *Mathematics of Computation* **93** (2024), 505–521.

[3] A. Bolt, S. Butler, and E. Hovland. Apollonian ring packings. *Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham* **283** (2018).

[4] D. W. Boyd. A new class of infinite sphere packings. *Pacific Journal of Mathematics* **50** (1974), 383–398.

[5] D. Chait-Roth, A. Cui, and Z. Stier. A taxonomy of crystallographic sphere packings. *Journal of Number Theory* **207** (2020), 196–246.

[6] H. Chen and A. Padrol. Scribability problems for polytopes. *European Journal of Combinatorics* **64** (2017), 1–26.

[7] D. Dias. The Local-Global Principle for Integral Generalized Apollonian Sphere Packings. *arXiv: Number Theory* (2014).

[8] G. Guettler and C. L. Mallows. A generalization of Apollonian packing of circles. *Journal of Combinatorics* **1** (2008).

[9] M. Kapovich and A. Kontorovich. On superintegral kleinian sphere packings, bugs, and arithmetic groups. *Journal für die Reine und Angewandte Mathematik (Crelles Journal)* (2023).

[10] C. Kertzer, S. Haag, K. E. Stange, and J. Rickards. "The local-global conjecture for Apollonian circle packings is false". *2024 Joint Mathematics Meetings (JMM 2024)*. AMS. 2024.

[11] A. Kontorovich. The local-global principle for integral Soddy sphere packings. *Journal of Modern Dynamics* **15** (2019), 209–236.

[12] A. Kontorovich and K. Nakamura. Geometry and arithmetic of crystallographic sphere packings. *Proceedings of the National Academy of Sciences* **116** (2019), 436–441.

[13] K. Nakamura. *The local-global principle for integral bends in orthoplicial Apollonian sphere packings*. 2014. arXiv: `1401.2980 [math.NT]`.

[14] J. L. Ramírez Alfonsín and I. Rasskin. A polytopal generalization of Apollonian packings and Descartes' theorem (2021). arXiv: `2107.09432 [math.CO]`.

[15] J. L. Ramírez Alfonsín and I. Rasskin. Ball packings for links. *European Journal of Combinatorics* **96** (2021), 103351.

[16] J. L. Ramírez Alfonsín and I. Rasskin. Links in orthoplicial Apollonian packings (2023). arXiv: `2301.03089 [math.GT]`.

[17] I. Rasskin. "A polytopal approach to Apollonian packings and discrete knotted structures". Theses. Université de Montpellier, 2021.

[18] I. Rasskin. Regular polytopes, sphere packings and Apollonian sections (2023). arXiv: `2109.00655v2`.

[19] A. Sheydvasser. Quaternion orders and sphere packings. *Journal of Number Theory* **204** (2019), 41–98.

[20] F. Soddy. The Kiss Precise. *Nature* **137** (1936), 1021–1021.

[21] B. A. Springborn. A unique representation of polyhedral types. Centering via Möbius transformations. *Mathematische Zeitschrift* **249** (2005), 513–517.

[22] K. E. Stange. The Apollonian structure of Bianchi groups. *Transactions of the American Mathematical Society* **370** (2015).

[23] X. Zhang. On the local-global principle for integral Apollonian 3-circle packings: *Journal für die Reine und Angewandte Mathematik (Crelles Journal)* **737** (2018), 71–110.

[24] G. M. Ziegler. *Lectures on polytopes*. Vol. 152. Springer Science & Business Media, 1994.

Disconnected Common Graphs via Supersaturation*

Jae-baek Lee† and Jonathan A. Noel‡

Department of Mathematics and Statistics, University of Victoria, Canada

Abstract

A graph H is said to be *common* if the number of monochromatic labelled copies of H in a 2-colouring of the edges of a large complete graph is asymptotically minimized by a random colouring. It is well known that the disjoint union of two common graphs may be uncommon; e.g., K_2 and K_3 are common, but their disjoint union is not. We investigate the commonality of disjoint unions of multiple copies of K_3 and K_2. As a consequence of our results, we obtain an example of a pair of uncommon graphs whose disjoint union is common. Our approach is to reduce the problem of showing that certain disconnected graphs are common to a constrained optimization problem in which the constraints are derived from supersaturation bounds related to Razborov's Triangle Density Theorem. We also improve bounds on the Ramsey multiplicity constant of a triangle with a pendant edge and the disjoint union of K_3 and K_2.

1 Introduction

In one of the first applications of the probabilistic method, Erdős [6] showed that a random colouring of the edges of a clique on $(1-o(1))2^{-1/2}e^{-1}k2^{k/2}$ vertices with red and blue contains no monochromatic complete graph on k vertices with positive probability; this implies a lower bound on the *Ramsey number* of the complete graph K_k, i.e. the smallest N for which every 2-colouring of the edges of K_N contains a monochromatic K_k. To this day, Erdős' bound has been improved only slightly by Spencer [24]. One of the core themes in Ramsey theory is that random colourings tend to perform well in avoiding certain monochromatic substructures.

This intuition extends to the closely related area of "Ramsey multiplicity" in which the goal is to minimize the number of monochromatic labelled copies of a given graph H in a red/blue colouring of the edges of K_N asymptotically as N tends to infinity. A graph H is said to be *common* if this asymptotic minimum is achieved by a sequence of random colourings. A famous result of Goodman [10] implies that K_3 is common (see Theorem 7). Inspired by this, Erdős [5] conjectured that K_k is common for all k and, nearly two decades later, Burr and Rosta [4] conjectured that every graph H is common. Sidorenko [22] observed that the *paw graph* P consisting of a triangle with a pendant edge is uncommon. Around the same time, Thomason [25] showed that K_k is uncommon for all $k \geq 4$; thus, the aforementioned conjectures are both false. Later, Jagger, Sťovíček and Thomason [14] proved that every graph H containing a K_4 is uncommon. In particular, almost every graph is uncommon. In recent years, there has been a steady flow of results proving that the members of certain families of graphs are common or uncommon [16, 17, 11, 1, 2, 12, 15]. In spite of this, the task of classifying common graphs seems hopelessly difficult.

The main goal of this paper is to provide a new approach for bounding the number of monochromatic copies of certain disconnected graphs in a colouring of K_N and to use it to obtain several new families

*The full version of this work can be found in [18] and will be published elsewhere.

†Email: dlwoqor0923@uvic.ca

‡Email: noelj@uvic.ca Research of J. A. Noel supported by NSERC and a university start-up grant.

of common graphs. Given graphs H_1 and H_2, let $H_1 \sqcup H_2$ denote their disjoint union; also, for a graph F and $\ell \geq 1$, let $\ell \cdot F$ be the disjoint union of ℓ copies of F. The argument of Sidorenko [22] that the paw graph is uncommon also shows that $K_3 \sqcup K_2$ is uncommon (with the same proof). Most of our results focus on the commonality of unions of several copies of K_3 and K_2. Our first result is as follows.

Theorem 1. *For $0 \leq \ell \leq 2$, the graph $(2 \cdot K_3) \sqcup (\ell \cdot K_2)$ is common.*

We also show that this is best possible in the sense that $(2 \cdot K_3) \sqcup (3 \cdot K_2)$ is uncommon; see Proposition 10. Since K_3 and K_2 are both common, Sidorenko's result [22] that $K_3 \sqcup K_2$ is uncommon tells us that the disjoint union of two common graphs can be uncommon. Using Theorem 1, we find that the opposite phenomenon is also possible; the disjoint union of two uncommon graphs can be common. In fact, the disjoint union of two copies of a single uncommon graph can be common.

Corollary 2. *There exists an uncommon graph H such that $H \sqcup H$ is common.*

Proof. Consider $H = K_3 \sqcup K_2$. The fact that H is uncommon was shown by Sidorenko [22], and the fact that $H \sqcup H$ is common follows from Theorem 1 with $\ell = 2$. □

We remark that our results also allow us to obtain new examples of graphs H_1 and H_2 such that H_1 is common, H_2 is uncommon and $H_1 \sqcup H_2$ is common. However, the existence of such a pair of graphs was already known; see [17, Subsection 1.1]. We also prove a general result on disjoint unions of triangles and edges, provided that the number of triangles is at least three.

Theorem 3. *For $k \geq 3$ and $0 \leq \ell \leq 5k/3 (\approx 1.666k)$, the graph $(k \cdot K_3) \sqcup (\ell \cdot K_2)$ is common.*

Theorem 4. *For $k \geq 1$ and $\ell = \lceil 1.9665k \rceil$, the graph $(k \cdot K_3) \sqcup (\ell \cdot K_2)$ is uncommon.*

2 Preliminary

Several of the results in this paper are best understood in the context of graph limits. A *kernel* is a bounded measurable function $U : [0,1]^2 \to \mathbb{R}$ such that $U(x,y) = U(y,x)$ for all $x, y \in [0,1]$. A *graphon* is a kernel such that $0 \leq W(x,y) \leq 1$ for all $x, y \in [0,1]$. The set of all graphons is denoted $\mathcal{W}_0$. Given a graph G, let $v(G) := |V(G)|$ and $e(G) := |E(G)|$. A graph G is said to be *empty* if $e(G) = 0$. Each graph G can be associated to a graphon W_G by dividing $[0,1]$ into $v(G)$ intervals $I_1, \ldots, I_{v(G)}$ of equal measure corresponding to the vertices of G and setting W_G equal to 1 on $I_i \times I_j$ if the ith and jth vertices are adjacent and 0 otherwise. The *homomorphism density* of a graph H in a kernel U is defined by

$$t(H,U) := \int_{[0,1]^{V(H)}} \prod_{uv \in E(H)} W(x_u, x_v) dx_{V(H)}$$

where $x_{V(H)} = (x_v : v \in V(H))$. We refer the reader to [19] for more background on graph limits. The *Ramsey multiplicity constant* of a graph H is defined to be

$$c(H) := \inf_{W \in \mathcal{W}_0} (t(H,W) + t(H, 1-W)).$$

In this language, a graph H is *common* if and only if

$$c(H) = 2(1/2)^{e(H)}. \tag{1}$$

As stated above, $K_3 \sqcup K_2$ and the paw graph P are uncommon. We obtain, to our knowledge, the tightest known upper bounds on the Ramsey multiplicity constants of these two graphs; for the former graph, we also obtain a reasonably tight lower bound which is proven without the assistance of the flag algebra method.

Theorem 5. $0.121423 < c(K_3 \sqcup K_2) < 0.121450$.

Theorem 6. *The paw graph P satisfies $c(P) < 0.121415$.*

Note that, for every graph H such that $c(H)$ is currently known, either H is common or $c(H)$ is achieved by a "Turán graphon" W_{K_k} for some $k \geq 3$ [8, 13]. To our knowledge, Theorem 5 is the closest that any result has come to determining $c(H)$ for a graph H which does not fit into either of these two categories. The lower bound in Theorem 5 can be improved by at least 0.022% using the flag algebra method; however, such a proof would most likely be verifiable only with heavy computer assistance, and is thus unlikely to provide much in terms of valuable insights. Several of the known results on common graphs actually establish stronger inequalities than (1). Following [2], a non-empty graph H is said to be *strongly common* if

$$t(H, W) + t(H, 1 - W) \geq t(K_2, W)^{e(H)} + t(K_2, 1 - W)^{e(H)} \tag{2}$$

for every graphon W. A simple application of Jensen's Inequality tells us that every strongly common graph is common. A classical example of a strongly common graph is K_3; see Theorem 7. A non-empty graph H is said to be *Sidorenko* if

$$t(H, W) \geq t(K_2, W)^{e(H)} \tag{3}$$

for every graphon W. Clearly, every Sidorenko graph is strongly common which, in turn, implies that every such graph is common. By taking $W = W_{K_2}$, one can see that every Sidorenko graph must be bipartite. Sidorenko's Conjecture [23] famously states that every bipartite graph is Sidorenko. Currently, every bipartite graph H which is known to be common is also known to be Sidorenko. Also, the only known examples of strongly common graphs which are not Sidorenko are the odd cycles [2, 10, 15].

Our strategy for obtaining new examples of common graphs relies on strong correlation inequalities, such as (2) and (3). Given this, it is natural to wonder whether all common graphs are strongly common; this question was raised in [2]. As it turns out, this is far from true. For example, $K_3 \sqcup K_3$ is common but not strongly common, and there are many other examples as well.

3 Key Ideas

Our approach is to reduce the problem of showing that certain disconnected graphs are common to a constrained optimization problem, in which the constraints are derived from supersaturation bounds related to Razborov's Triangle Density Theorem. For the purposes of proving the lower bound of Theorem 5, it will be enough to use the following theorem which was first announced by Fisher [7]; as mentioned in [21], the proof contained a hole that can be patched using a later result of [9]. A new proof was found by Razborov [20] prior to proving the general Triangle Density Theorem in [21].

Theorem 7 (Goodman's Theorem [10])**.** *K_3 is strongly common.*

Theorem 8 (Fisher [7] and Goldwurm and Santini [9]; see also Razborov [20])**.** *Every graphon W with $t(K_2, W) \leq 2/3$ satisfies*

$$t(K_3, W) \geq \frac{1}{9}\left(-2\left(2 + \sqrt{4 - 6t(K_2, W)}\right) + 3t(K_2, W)\left(3 + \sqrt{4 - 6t(K_2, W)}\right)\right)$$

Theorem 9 (Bollobás [3])**.** *Every graphon W satisfies*

$$t(K_3, W) \geq \frac{4}{3}t(K_2, W) - \frac{2}{3}.$$

To prove Thoerem 4, the upper bound of Theorem 5 and Theorem 6, the graphons that we will use are all of the same general form. For $n \geq 1$, let $\triangle_n$ be the set of all vectors $\vec{z}$ of length n with non-negative entries that sum to one. Given $\vec{z} \in \triangle_n$ and an $n \times n$ symmetric matrix A with entries in

$[0,1]$, let $W_{\vec{z},A}$ be defined as follows. First, divide $[0,1]$ into n intervals $I_1,\dots,I_n$ such that the measure of I_i is equal to $\vec{z}_i$. Next, for each $1 \le i,j \le n$, define $W_{\vec{z},A}$ to be equal to $A_{i,j}$ for all $(x,y) \in I_i \times I_j$. It is easily observed that, for any graph H,

$$t(H, W_{\vec{z},A}) = \sum_{f:V(H)\to[n]} \prod_{v\in V(H)} \vec{z}_{f(v)} \prod_{uv\in E(H)} A_{f(u),f(v)}. \tag{4}$$

Using this construction, we could prove Theorem 4. Let $k \ge 1$ and $\ell = \lceil 1.9665k \rceil$. We show that $H = (k \cdot K_3) \sqcup (\ell \cdot K_2)$ is uncommon. Define $\alpha = \ell/k$ and note that $1.9665 \le \alpha \le 2$. Let

$$p := 1 - 2^{-1/(3+\alpha)}.$$

We let W be the graphon $W_{\vec{z},A}$ where $\vec{z} = (1/2, 1/2)$ and A is a 2×2 matrix whose diagonal entries are p and off-diagonal entries are 1.

Proposition 10. *The graph $(2 \cdot K_3) \sqcup (3 \cdot K_2)$ is uncommon*

Proof. We prove that the graph $H = (2 \cdot K_3) \sqcup (3 \cdot K_2)$ is uncommon. For $z \in [0, 1/2]$ and $y \in [0,1]$, we define $W_{z,y} := W_{\vec{z},A}$ where $\vec{z} = (1-2z, z, z) \in \triangle_3$ and A is the symmetric 3×3 matrix in which $A(1,2) = A(1,3) = 1$, $A(2,3) = y$ and $A(i,i) = 0$ for $1 \le i \le 3$. Setting $z = 0.28$ and $y = 0.42$ yields $h(z,y) = 0.00390226 < 2 \cdot \left(\frac{1}{2}\right)^9$, which completes the proof. □

Proposition 11. *$(3 \cdot P) \sqcup (2 \cdot K_2)$ is uncommon.*

Proof. Let $H = (3 \cdot P) \sqcup (2 \cdot K_2)$. Once again, we use the graphon $W_{z,y}$ from the previous three proofs. This time, we set $z = 0.429919$ and $y = 0.43222$. Thus, $t(H, W_{z,y}) + t(H, 1 - W_{z,y}) < 0.000121856 < 2(1/2)^{14}$ and the result follows.

□

Using the same construction above with different values of y and z, we could get the upper bound of Theorem 5 and Theorem 6

References

[1] N. Behague, N. Morrison, and J. A. Noel, *Common pairs of graphs*, E-print arXiv:2208.02045v3, 2023.

[2] ———, *Off-diagonal commonality of graphs via entropy*, E-print arXiv:2307.03788v1, 2023.

[3] B. Bollobás, *Relations between sets of complete subgraphs*, Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen, 1975), Congressus Numerantium, No. XV, Utilitas Math., Winnipeg, Man., 1976, pp. 79–84.

[4] S. A. Burr and V. Rosta, *On the Ramsey multiplicities of graphs—problems and recent results*, J. Graph Theory **4** (1980), no. 4, 347–361.

[5] P. Erdős, *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), 459–464.

[6] P. Erdős, *Some remarks on the theory of graphs*, Bull. Amer. Math. Soc. **53** (1947), 292–294.

[7] D. C. Fisher, *Lower bounds on the number of triangles in a graph*, J. Graph Theory **13** (1989), no. 4, 505–512.

[8] J. Fox and Y. Wigderson, *Ramsey multiplicity and the Turán coloring*, Adv. Comb. (2023), Paper No. 2, 39.

[9] M. Goldwurm and M. Santini, *Clique polynomials have a unique root of smallest modulus*, Inform. Process. Lett. **75** (2000), no. 3, 127–132.

[10] A. W. Goodman, *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), 778–783.

[11] A. Grzesik, J. Lee, B. Lidický, and J. Volec, *On tripartite common graphs*, Combin. Probab. Comput. **31** (2022), no. 5, 907–923.

[12] H. Hatami, J. Hladký, D. Kráľ, S. Norine, and A. Razborov, *Non-three-colourable common graphs exist*, Combin. Probab. Comput. **21** (2012), no. 5, 734–742.

[13] J. Hyde, J.-B. Lee, and J. A. Noel, *Turán colourings in off-diagonal ramsey multiplicity*, 2024, E-print arXiv:2309.06959.

[14] C. Jagger, P. Šťovíček, and A. Thomason, *Multiplicities of subgraphs*, Combinatorica **16** (1996), no. 1, 123–141.

[15] J. S. Kim and J. Lee, *Extended commonality of paths and cycles via Schur convexity*, J. Combin. Theory Ser. B **166** (2024), 109–122.

[16] D. Kráľ, J. A. Noel, S. Norin, J. Volec, and F. Wei, *Non-bipartite k-common graphs*, Combinatorica **42** (2022), no. 1, 87–114.

[17] D. Kráľ, J. Volec, and F. Wei, *Common graphs with arbitrary chromatic number*, E-print arXiv:2206.05800v1, 2022.

[18] J.-B. Lee and J. A. Noel, *Disconnected common graphs via supersaturation*, 2023, E-print arXiv:2303.09296.

[19] L. Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.

[20] A. A. Razborov, *Flag algebras*, J. Symbolic Logic **72** (2007), no. 4, 1239–1282.

[21] ———, *On the minimal density of triangles in graphs*, Combin. Probab. Comput. **17** (2008), no. 4, 603–618.

[22] A. F. Sidorenko, *Cycles in graphs and functional inequalities*, Mat. Zametki **46** (1989), no. 5, 72–79, 104.

[23] ———, *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), no. 2, 201–204.

[24] J. Spencer, *Ramsey's theorem—a new lower bound*, J. Combinatorial Theory Ser. A **18** (1975), 108–115.

[25] A. Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, J. Lond. Math. Soc. (2) **39** (1989), no. 2, 246–255.

Bicolored point sets admitting non-crossing alternating Hamiltonian paths*

Jan Soukup†[1]

[1]Department of Applied Mathematics, Charles University, Faculty of Mathematics and Physics, Malostranské nám. 25, 118 00 Praha 1, Czech Republic.

Abstract

Consider a bicolored point set P in general position in the plane consisting of red and blue points such that the number of blue points differs from the number of red points by at most one. We show that if a subset of the red points forms the vertices of a convex polygon separating the blue points, lying inside the polygon, from the remaining red points, lying outside the polygon, then the points of P can be connected by non-crossing straight-line segments so that the resulting graph is a properly colored Hamiltonian path.

1 Introduction

In geometric graph theory it is a common problem to decide whether a given graph can be drawn in the plane on a given point set so that the edges are represented by non-crossing straight-line segments. For example, deciding whether a given general planar graph has a non-crossing straight-line drawing on a given point set is NP-complete [7].

There are many interesting unanswered questions when considering bicolored point sets instead (see the comprehensive survey by Kano and Urrutia [11]). We restrict ourselves to drawings of bipartite graphs on bicolored point sets where edges are drawn as non-crossing straight-line segments between points of different colors. This question remains interesting even for paths. Let B and R denote a set of blue points and a set of red points in the plane, respectively, such that $R \cup B$ is in general position, i.e., no three points are collinear. We call a non-intersecting path on $R \cup B$ whose edges are straight-line segments and every segment connects two points of $R \cup B$ of distinct colors, an *alternating path*. If such an alternating path connects all points of $R \cup B$, we call it an *alternating Hamiltonian path*. If such an alternating Hamiltonian path shares the first and last vertex (but otherwise is still non-intersecting), we call it an *alternating Hamiltonian cycle*.

If $||R| - |B|| \leq 1$ and R can be separated from B by a line, then Abellanas et al. [1] showed that there always exists an alternating Hamiltonian path on $R \cup B$. This fact together with the well-known Ham sandwich theorem implies that if $|R| = |B|$, then there always exists an alternating path on $R \cup B$ connecting at least half of the points. This trivial lower bound on the length of an alternating path that always exists is the best known according to our knowledge. This bound was improved by a small linear factor by Mulzer and Valtr [12] for point sets in *convex* position, i.e., when the points form the vertices of a convex polygon. On the other hand, if we do not assume that R and B are separated by a line, then there are examples where $|R| = |B| \geq 8$ and no alternating Hamiltonian path on $R \cup B$ exists, even if $R \cup B$ is in convex position. Moreover, for $R \cup B$ in convex position with $|R| = |B| = n$, Csóka et al. [9] showed that there are configurations where the longest alternating path on $R \cup B$ has size at most $(4 - 2\sqrt{2})n + o(n)$.

*The full version of this work can be found in [13] and will be published elsewhere. Supported by project 23-04949X of the Czech Science Foundation (GAČR) and by the grant SVV–2023–260699.

†Email: soukup@kam.mff.cuni.cz

As we have seen above, an alternating Hamiltonian path does not exist on every point set but it exists if $||R| - |B|| \leq 1$ and R can be separated from B by a line. Another sufficient condition was found by Cibulka et al. [8]. They looked more closely at configurations where R and B form a so-called double chain and showed that if $||R| - |B|| \leq 1$ and each chain of the double-chain contains at least one-fifth of all points, then there exists an alternating Hamiltonian path on $R \cup B$.

The final sufficient condition that we know of was found by Abellanas et al. [1]. They showed that if $||R| - |B|| \leq 1$, the points of R are vertices of a convex polygon, and all points of B are inside this polygon, then there exists an alternating Hamiltonian path on $R \cup B$.

In this paper, we generalize this last result, and by doing so, we extend the known family of configurations of points for which there exists an alternating Hamiltonian path on $R \cup B$. Specifically, we prove the following theorem.

Theorem 1. *Let R be a set of red points and B be a set of blue points such that $R \cup B$ is in general position. Let P be a convex polygon whose vertices are formed by a subset of R. Assume that the remaining points of R lie outside of P, points of B lie in the interior of P, and $||R| - |B|| \leq 1$. Then there exists an alternating Hamiltonian path on $R \cup B$.*

When $|R| = |B|$ we even find an alternating Hamiltonian cycle.

2 Preliminaries and an outline of the proof

By a *polygonal region* we understand a closed, possibly unbounded, region in the plane whose boundary (possibly empty) consists of finitely many non-crossing straight-line segments or half-lines connected into a polygonal chain. A bounded polygonal region is a polygon. A polygon can be defined by an ordered set of its vertices; in that case, we assume that the vertices lie on the boundary of the polygon in the clockwise direction, and we use index arithmetic modulo the number of vertices. A *diagonal* of a convex polygon (or a polygonal region) is any segment connecting two points on the boundary of the polygon. For an edge e of a convex polygon (or polygonal region), the closed half-plane to the side of e that is disjoint with the polygon's interior is denoted by $\mathrm{out}(e)$. For two points a, b in the plane, we denote by ab the segment connecting them. The *convex hull* of a set of points X, denoted by $\mathrm{conv}(X)$, is the smallest convex set that contains X.

Recall that B and R always denote the set of blue points and the set of red points, respectively. Moreover, B and R are always disjoint, and $R \cup B$ is always in general position. For a region T of the plane, $\|T\|_R$ and $\|T\|_B$ denotes the number of red points inside T and the number of blue points inside T, respectively. For the points on the boundaries of regions, we specify if they belong to the region or not (we will need to assign every point to exactly one part of a partition of the plane into polygonal regions).

Our primary result, Theorem 1, is a generalization of the following theorem proved by Abellanas et al. [1].

Theorem 2 ([1])**.** *Let R be a set of red points and B be a set of blue points such that $R \cup B$ is in general position. Let R form the vertices of the polygon $\mathrm{conv}(R \cup B)$, the points of B lie in the interior of $\mathrm{conv}(R \cup B)$, and $||R| - |B|| \leq 1$. Then there exists an alternating Hamiltonian path on $R \cup B$.*

Our improvement lies in the fact that the polygon P can be formed by a subset of R (instead of the whole R), whereas the remaining points of R remain outside of P. The approach in the proof of Theorem 2 in the case when $|R| = |B|$ is to split the polygon formed by R into convex polygons, each containing exactly one edge of the polygon and one blue point from inside the polygon, and then connect by straight-line segments each of the blue points to the vertices of the edge that is inside the same part. In this way, alternating paths of length two are formed inside each part of the partition. Moreover, they share their end vertices, and so, together, they form an alternating Hamiltonian cycle (this cycle is non-crossing since each of the small paths lies in its own part of the partition).

We proceed similarly with only one significant distinction. Namely, we partition the whole plane into convex parts so that every edge of the polygon is a diagonal of one part, and each part contains one more blue point than it contains red points (not counting the vertices of the polygon). Inside each of these parts, we find an alternating Hamiltonian path. And these paths together form an alternating Hamiltonian cycle as before.

In section 3 we outline how we split the plane and in section 4 how to find the alternating Hamiltonian path.

3 Partitioning theorem

For the partitioning of the plane, we prove the following theorem.

Theorem 3. *Let $P = (p_1, \dots, p_s)$ be a convex polygon, B be a set of blue points in the interior of P, and R be a set of red points outside of P such that $s = |B| - |R|$ and $R \cup B \cup \{p_1, \dots p_s\}$ is in general position. Then there exists a partition of the plane into convex polygonal regions $Q_1, \dots, Q_s$ such that each $p_i p_{i+1}$ is a diagonal of Q_i and for every i, we have $\|Q_i\|_B - \|Q_i\|_R = 1$. Moreover, every point of $R \cup B$ is counted in exactly one Q_i. That is, if a point of $R \cup B$ lies on the common boundary of more Q_i's it is assigned to only one of them.*

For the case of $s = 3$, i.e., when P is a triangle, we prove the following stronger lemma.

Lemma 4. *Let Q be a convex polygonal region, $P = (p_1, p_2, p_3)$ be a triangle inside Q, B be a set of blue points in the interior of P, and R be a set of red points outside P but inside Q such that $R \cup B \cup \{p_1, p_2, p_3\}$ is in general position. Additionally let $n_1, \dots, n_3$ be integers satisfying the following conditions.*

1. *$|B| - |R| = n_1 + n_2 + n_3$.*
2. *For every nonempty subset I of $\{1, \dots, 3\}$,*

$$\sum_{i \in I} n_i \geq - \left\| Q \cap \bigcup_{i \in I} \operatorname{out}(p_i p_{i+1}) \right\|_R . \tag{1}$$

Then there exists a point y in P different from p_1, p_2 and p_3 such that the half-lines yp_1, yp_2 and yp_3 split Q into three parts Q_1, Q_2 and Q_3. Moreover, there exists an assignment of points of $B \cup R$ that lie on the boundaries of Q_1, Q_2 and Q_3 into adjacent parts so that $\|Q_i\|_B - \|Q_i\|_R = n_i$.

For an example partition according to Lemma 4, see Figure 1.

Note that the conditions established by Inequation (1) are necessary: When I contains only one index i, the part Q_i is split by $p_i p_{i+1}$ into two regions, one inside P and one outside of P. The part outside of P is inside $\operatorname{out}(p_i p_{i+1})$, and so $\|Q_i\|_R \leq \|Q \cap \operatorname{out}(p_i p_{i+1})\|_R$. Together with a trivial condition $\|Q_i\|_B \geq 0$ we get $\|Q_i\|_B - \|Q_i\|_R \geq - \|Q \cap \operatorname{out}(p_i p_{i+1})\|_R$, which is exactly one of the conditions. It can be analogously observed for larger cardinalities of I's.

Furthermore, note that in the case when Q is the plane and all n_i's are equal to 1, the conditions established by Inequation (1) always hold, and Lemma 4 implies Theorem 3 when P is a triangle.

In the proof of Lemma 4, we employ a standard technique (see Akiyama and Alon [2]) and substitute points with disks of the same area and work with the area of the disks instead of the number of points. This is helpful because the boundaries of polygonal regions have an area of size zero, and so the area of all disks will be precisely distributed between the interiors of the polygonal regions of the partition. We find the point y using a well-known result in fixed point theory: Knaster–Kuratowski–Mazurkiewicz lemma (see [6, Theorem 5.1] for a simple proof). At the end of the proof, we return from disks back to points and we have to solve the problem where to assign points whose disks intersect the boundaries of the partition. We present the details in the full version.

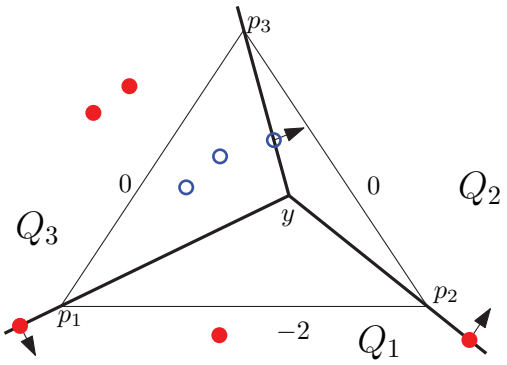


Figure 1: Illustration of a partition from Lemma 4. Region Q_1 contains two fewer blue points than it contains red points. Region Q_2 contains the same number of blue points as red points. And the same holds for Q_3. The arrows indicate to which regions belong the points on the boundaries.

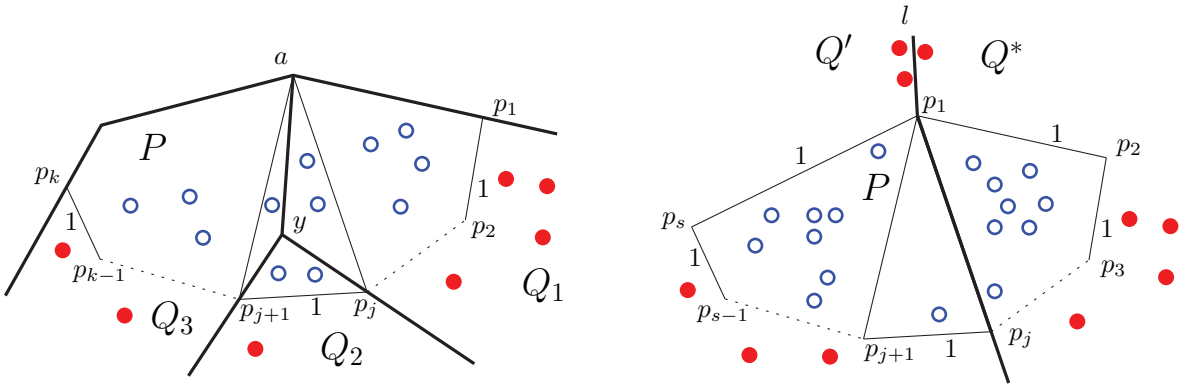


Figure 2: Left: Inside the polygon P we find a triangle ap_jp_{j+1} and apply Lemma 4 to split the polygonal region into three parts Q_1, Q_2, Q_3. The induction hypothesis can then be applied to Q_1 and Q_3 to obtain a complete partition of the polygonal region.
Right: In the first step of the partitioning of the convex polygon P we sometimes use a half-line l and partition regions Q^* and Q' by induction.

To prove Theorem 3, we use induction on the number of vertices of the polygon P. The main idea is to find a suitable triangle formed by vertices of P, and apply Lemma 4 to this triangle. We set the numbers n_1, n_2, n_3 so that we obtain three polygonal regions Q_1, Q_2 and Q_3 each containing $\|Q_i\|_B - \|Q_i\|_R$ edges of P. Then we partition Q_1, Q_2 and Q_3 by induction. Note that except for the very first step, we are partitioning bounded polygonal regions instead of the plane, and the polygon P is already partially split but that is only easier. See Figure 2, left.

Unfortunately, finding the very first triangle is not always possible. However, if that is the case, then we can find a diagonal p_1p_j of P and a half-line l shooting from p_1 such that l and the half-line p_1p_j split the plane into two polygonal regions Q^* and Q' that can be partitioned by the normal induction process. See Figure 2, right. We present the details in the full version.

4 Conclusion of the proof

To finish the proof of Theorem 1, we use a result proved by Abellanas et al. [1] about point sets with color classes separated by a line. We use a slightly modified version that easily follows from the proof of the original version.

Theorem 5 ([1])**.** *Let R be a set of red points and B be a set of blue points such that $R \cup B$ is in*

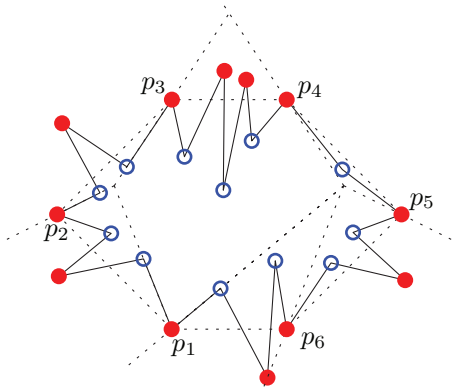


Figure 3: An alternating Hamiltonian cycle in a case when 6 red points form a polygon separating the remaining 6 red points from 12 blue points lying inside the polygon.

general position. Assume that $|R| - |B| = 1$ *and that there are two points* $r_1, r_2 \in R$ *such that the line* r_1r_2 *separates* R *from* B *and that* r_1, r_2 *are vertices of the convex hull* $\mathrm{conv}(R \cup B)$*. Then there exists an alternating Hamiltonian path on* $R \cup B$ *with end vertices* r_1, r_2*.*

We apply this theorem several times to the partition obtained by Theorem 3 to finish the proof of Theorem 1.

(Idea of the) proof of Theorem 1. We may assume that $|R| = |B|$ and prove that there exists an alternating Hamiltonian cycle, otherwise, we could add one point and remove it at the end. Let $R' = R \setminus P$. That is, R' contains exactly the points of R that are not vertices of P. Therefore, $s = |B| - |R'|$.

By Theorem 3 applied on the polygon P, the set of blue points B and the set of red points R', there exists a partition of the plane into convex polygonal regions $Q_1, \ldots, Q_s$ such that for every i, the edge p_ip_{i+1} is a diagonal of Q_i, and for every i, the region Q_i contains exactly one more blue point than red points of R'.

By Theorem 5 applied to each Q_i separately, we obtain an alternating Hamiltonian path in each Q_i with ends in p_i and p_{i+1} covering all red and blue points inside Q_i. These paths are connected together in the end vertices p_i. Therefore, together they form an alternating Hamiltonian cycle. See Figure 3 for an illustration.

□

5 Conclusion and open questions

The main technical part of our proof is Theorem 3. We believe that the following stronger version that also generalizes Lemma 4 holds.

Conjecture 6. *Let* Q *be a convex polygonal region,* $P = (p_1, \ldots, p_s)$ *be a convex polygon inside* Q*,* B *be a set of blue points in the interior of* P*, and* R *be a set of red points outside* P *but inside* Q *such that* $R \cup B \cup \{p_1, \ldots p_s\}$ *is in general position. Additionally let* $n_1, \ldots, n_s$ *be integers satisfying the following conditions.*

1. $|B| - |R| = n_1 + \cdots + n_s$.
2. *For every nonempty cyclic interval of indices* I *from* $\{1, \ldots, s\}$,

$$\sum_{i \in I} n_i \geq -\left\| Q \cap \bigcup_{i \in I} \mathrm{out}(p_ip_{i+1}) \right\|_R . \tag{2}$$

Then there exists a partition of Q into convex polygonal regions $Q_1, \ldots, Q_s$ such that for every i, the segment p_ip_{i+1} is a diagonal of Q_i and $\|Q_i\|_B - \|Q_i\|_R = n_i$. Moreover, every point of $R \cup B$ is counted in exactly one Q_i.

Similarly, as for Lemma 4 we observe that the conditions established by Inequation (2) are necessary.

The case when there are no red points outside of P and every n_i is a positive integer was already proved by García and Tejel [10] and later by Aurenhammer [3]. The case with points outside of P seems to be more difficult (for example, even some negative n_i's can satisfy the conditions established by Inequation (2) in that case).

Similar problems of finding partitions of colored point sets into subsets with disjoint convex hulls such that the sets of points of all color classes are partitioned as evenly as possible is well studied, see [4, 5]. However, we were not able to apply the results directly because we have the additional restriction that p_ip_{i+1}'s have to be diagonals of the convex hulls in the partition. We managed to prove Conjecture 6 only in the case when $s = 3$ (Lemma 4) and that proved crucial in proving Theorem 3.

References

[1] Abellanas, M., García, J., Hernández, G., Noy, M., Ramos, P.: Bipartite embeddings of trees in the plane. Discrete Appl. Math. **93**(2-3), 141–148 (1999), doi:10.1016/S0166-218X(99)00042-6

[2] Akiyama, J., Alon, N.: Disjoint simplices and geometric hypergraphs. Annals of the New York Academy of Sciences **555**(1), 1–3 (1989), doi:10.1111/j.1749-6632.1989.tb22429.x

[3] Aurenhammer, F.: Weighted skeletons and fixed-share decomposition. Comput. Geom. **40**(2), 93–101 (2008), ISSN 0925-7721, doi:10.1016/j.comgeo.2007.08.002

[4] Bespamyatnikh, S., Kirkpatrick, D., Snoeyink, J.: Generalizing ham sandwich cuts to equitable subdivisions. vol. 24, pp. 605–622 (Jan 2000), ISSN 1432-0444, doi:10.1007/s004540010065

[5] Blagojević, P.V., Rote, G., Steinmeyer, J.K., Ziegler, G.M.: Convex equipartitions of colored point sets. Discrete & Computational Geometry **61**, 355–363 (2019), doi:10.1007/s00454-017-9959-7

[6] Border, K.C.: Fixed point theorems with applications to economics and game theory. Cambridge University Press, Cambridge (1989), ISBN 0-521-38808-2

[7] Cabello, S.: Planar embeddability of the vertices of a graph using a fixed point set is NP-hard. J. Graph Algorithms Appl. **10**(2), 353–363 (2006), doi:10.7155/jgaa.00132

[8] Cibulka, J., Kynčl, J., Mészáros, V., Stolař, R., Valtr, P.: Universal sets for straight-line embeddings of bicolored graphs. In: Thirty essays on geometric graph theory, pp. 101–119, Springer, New York (2013), doi:10.1007/978-1-4614-0110-0_8

[9] Csóka, E., Blázsik, Z.L., Király, Z., Lenger, D.: Upper bounds for the necklace folding problems. Journal of Combinatorial Theory, Series B **157**, 123–143 (2022), doi:https://doi.org/10.1016/j.jctb.2022.05.012

[10] García, A., Tejel, J.: Dividiendo una nube de puntos en regiones convexas. In: Actas VI Encuentros de Geometría Computacional, pp. 169–174, Barcelona (1995)

[11] Kano, M., Urrutia, J.: Discrete geometry on colored point sets in the plane—a survey. Graphs Combin. **37**(1), 1–53 (2021), doi:10.1007/s00373-020-02210-8

[12] Mulzer, W., Valtr, P.: Long Alternating Paths Exist. In: 36th International Symposium on Computational Geometry (SoCG 2020), Leibniz International Proceedings in Informatics (LIPIcs), vol. 164, pp. 57:1–57:16, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany (2020), doi:10.4230/LIPIcs.SoCG.2020.57

[13] Soukup, J.: Bicolored point sets admitting non-crossing alternating hamiltonian paths. arXiv preprint (2024), doi:10.48550/arXiv.2404.06105

On homogeneous matroid ports*

Jaume Martí-Farré†[1] and Anna de Mier‡[1,2]

[1]Dept. of Mathematics, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain
[2]Institut de Matemàtiques (IMTech) UPC

Abstract

A matroid port is a clutter (antichain of sets) determined by the collection of circuits of a matroid that contain a fixed point. We study the problem of determining matroid ports all whose elements have the same size h. This problem has been studied from the cryptographic perspective and it is related to an analogous problem in matroid theory. We give some general results and then focus on the binary case. We recast some known results and find all 4-homogeneous binary matroid ports.

1 Introduction

A *clutter* is a collection of sets that are mutually incomparable with respect to inclusion; the *support* of a clutter Δ is $\cup_{A\in\Delta}A$. For a matroid $\mathcal{M}$ on the ground set E, the collection of its circuits $\mathcal{C}(\mathcal{M})$ forms a clutter (as do the collections of bases or hyperplanes). We are interested in another clutter derived from the circuits of a matroid. For an element $p \in E$, the *matroid port of* $\mathcal{M}$ *at* p is the clutter $\mathcal{M}_p$ defined as[1]

$$\mathcal{M}_p = \{C - p : C \in \mathcal{C}(\mathcal{M}), p \in C\}.$$

A clutter Δ with support Ω is said to be a *matroid port* if Δ is the port of some matroid $\mathcal{M}_\Delta$ with ground set $E = \Omega \cup p$ with $p \notin \Omega$ (we refer to Section 2 for more details on how $\mathcal{M}_\Delta$ is constructed). Matroid ports were introduced by Lehman [2] in connection with game theory, and they are a key structure in the theory of secret sharing [6], both for the characterization of ideal access structures and to obtain bounds on the optimal information rate of the scheme.

In this paper we focus on the problem of determining the matroid ports all whose elements have the same size. We say that a clutter Δ is h-homogeneous if $|A| = h$ for all $A \in \Delta$, and we call h the *rank* of the clutter[2]. Our motivation for studying this problem is twofold, as we next explain.

A particular family of h-homogeneous matroid ports are those that arise as ports of matroids all whose circuits have size $h + 1$. Unfortunately, determining such matroids is a hard problem. In [8], Murty determined the binary matroids all whose circuits have a given size. As we will see throughout this work, there are usually many more h-homogeneous binary matroid ports than matroids all whose circuits have size $h+1$, since being an h-homogeneous matroid port only gives information on the sizes of the circuits that contain p. In particular, for even h it is proved in [8] that there is only one binary matroid with circuits of size $h + 1$ (a single circuit), but we will see in Section 4 that 4-homogeneous

*The full version of this work will be published elsewhere.

†Email: jaume.marti@upc.edu. Research of J. M.-F. is supported by AGAUR, Generalitat de Catalunya, under project SGR-Cat 2021 00595 and by Universitat Politècnica de Catalunya under funds AGRUP-UPC.

‡Email: anna.de.mier@upc.edu. Research of A. dM. is supported by the Grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

[1]We use the common convention in matroid theory to write $A - b, A \cup b$ instead of $A - \{b\}, A \cup \{b\}$.

[2]In the few places we use the word "rank" to refer to the rank of a matroid it will be said clearly.

binary matroid ports are much richer. The question of determining matroids whose circuit-sizes belong to a small, fixed set, has also been studied [3], but the results do not seem applicable in our situation.

In addition to the relationship with the purely matroid theoretic question studied by Murty, the problem of determining h-homogeneous matroid ports is of interest in the context of secret sharing schemes in cryptography. Specifically, from the results in [6] it follows that to provide a complete description of h-homogeneous matroid ports is equivalent to characterizing the access structures of the ideal secret sharing schemes whose minimal qualified subsets have h participants. As far as we know, the only results in this direction have been obtained in [1, 4, 5] for the cases $h = 2$ and $h = 3$. The tools used in these works are mostly of a cryptographic nature. It is also one of our goals to present these results in a unified, more combinatorial way.

This work has two main contributions. First, we develop some reductions and general results that are applicable to any h and to all kind of ports, whether binary or not. Then, we treat the concrete case of $h = 4$ in the binary case, providing a complete description of 4-homogeneous binary matroid ports. The proofs will be included in a full version of this extended abstract.

We conclude this introduction by mentioning that we are not aware of results for the case $h \geqslant 5$ in general, or for the non-binary case for $h = 3, 4$. We feel that many of our techniques can be applied in the binary case for $h \geqslant 5$, but we are far from even conjecturing a list of 5-homogeneous binary matroid ports.

2 Preliminaries and reductions

We review in this section several results and characterizations about matroid ports and binary matroid ports. We refer to Oxley's book [9] for all terms and results in matroid theory.

A connected matroid $\mathcal{M}$ is determined by any one of its ports $\mathcal{M}_p$. We state next the description of the circuits of $\mathcal{M}$ in terms of the sets in the port. We use the notation $A_1 \ominus A_2$ to denote the symmetric difference of A_1 and A_2. Also, by $\min(\{B_1, \ldots, B_m\})$ we denote the inclusion minimal elements of $\{B_1, \ldots, B_m\}$.

Theorem 1. *([2], [9, Thm. 4.3.3]) Let Δ be a matroid port with support Ω. Then there is a unique connected matroid $\mathcal{M}_\Delta$ on $\Omega \cup p$. Moreover, the circuits of $\mathcal{M}_\Delta$ are*

$$\mathcal{C}(\mathcal{M}_\Delta) = \big\{A \cup p \,:\, A \in \Delta\big\} \cup \min(\big\{A_1 \ominus_\Delta A_2 \,:\, A_1, A_2 \in \Delta, A_1 \neq A_2\big\}),$$

where

$$A_1 \ominus_\Delta A_2 = (A_1 \cup A_2) \setminus \bigcap_{\substack{A \subseteq A_1 \cup A_2 \\ A \in \Delta}} A.$$

A clutter Δ on a finite set E is said to be a $\mathbb{K}$-*representable matroid port* if it is the port of a matroid $\mathcal{M}$ representable over the field $\mathbb{K}$. In particular, $\mathbb{F}_2$-representable matroid ports will be called *binary*.

Before stating characterizations for matroid ports and binary matroid ports, we need one more definition. The *blocker* of a clutter Δ on E is the clutter $b(\Delta) = \min\left(\{B \subseteq E : B \cap A \neq \emptyset \text{ for all } A \in \Delta\}\right)$. It is well-known that $b(b(\Delta)) = \Delta$. For matroid ports, it is not difficult to check that $b(\mathcal{M}_p) = (\mathcal{M}^*)_p$, where $\mathcal{M}^*$ is the dual of $\mathcal{M}$. Thus, the blocker of a matroid port is again a matroid port. As the dual of a $\mathbb{K}$-representable matroid is also $\mathbb{K}$-representable, the blocker of a binary matroid port is also a binary matroid port. Note though that the blocker of an h-homogeneous clutter need not be homogeneous.

Theorem 2. *1. The clutter Δ is a matroid port if and only if whenever $A_1, A_2, A_3 \in \Delta$ and $x \in (A_2 \cap A_3) \setminus A_1$ there is $A \in \Delta$ with $A \subseteq \big((A_1 \ominus_\Delta A_2) \cup A_3\big) \setminus \{x\}$.*

2. The clutter Δ is a binary matroid port if and only if either of the following two statements holds:

(a) For all $A \in \Delta$ and for $B \in b(\Delta)$ the intersection $A \cap B$ has odd cardinality.

(b) If $A_1, A_2, A_3 \in \Delta$, then there exists $A \in \Delta$ with $A \subseteq A_1 \ominus A_2 \ominus A_3$.

For a proof of this theorem, we refer to [10] and [11]. We mention that there are several other characterizations of matroid ports that range from excluded minors [10, 11] to independent sequences and bounds on the optimal information rates in secret sharing schemes [6, Theorem 4.4]. Although we do not use these characterizations, we use the notion of minors of matroid ports in some constructions, so we review them here.

Let Δ be a clutter with support Ω and let $Z \subset \Omega$. The *deletion of* Z is the clutter $\Delta \setminus Z$ given by $\Delta \setminus Z = \{A \subseteq \Omega \setminus Z : A \in \Delta\}$; we refer to the clutter $\Delta \setminus (\Omega \setminus Z)$ as the *restriction* of Δ to Z, denoted by $\Delta|Z$. The *contraction of* Z is the clutter Δ/Z given by $\Delta/Z = \min(\{A \subseteq \Omega \setminus Z : A \setminus Z \in \Delta\})$. When Δ is the clutter of circuits of a matroid, these definitions give the usual notions of deletion and contraction in matroids (and we write $\mathcal{M}/Z$ instead of $\mathcal{C}(\mathcal{M})/Z$, and so on). Every clutter that can be obtained from Δ by repeatedly applying the operations $\setminus$ and $/$ is called a *minor* of Δ. Minors of matroid ports are matroid ports: indeed, we have $\mathcal{M}_p \setminus Z = (\mathcal{M} \setminus Z)_p$ and $\mathcal{M}_p/Z = (\mathcal{M}/Z)_p$.

We say that the clutter Δ is *path-connected* if for all $x, y \in \Omega$ there is a sequence $A_1, \ldots, A_k$ with $A_i \in \Delta$, $x \in A_1$, $y \in A_k$ and $A_i \cap A_j \neq \emptyset$ (if the sets of the clutter are thought of as the hyperedges of a hypergraph, path-connectivity becomes connectivity in the hypergraph). We say that x and y are at distance k in Δ if k is the smallest integer for which such a sequence exists. For any clutter Δ there exists a unique partition $\Omega = \Omega_1 \cup \cdots \cup \Omega_s$ such that the restrictions $\Delta|\Omega_1, \ldots, \Delta|\Omega_s$ are path-connected and $\Delta = \Delta|\Omega_1 \cup \cdots \cup \Delta|\Omega_s$. In this situation we say that $\Delta|\Omega_1, \ldots, \Delta|\Omega_s$ are the *path-connected components* of Δ. It is clear that a clutter is h-homogeneous if and only if each of its path-connected components is h-homogeneous.

In the particular case that the clutter is a matroid port $\mathcal{M}_p$, being path-connected is equivalent to the matroid $\mathcal{M}$ not being a parallel connection of two smaller matroids (see [9, Sec. 7.1] for the definition and properties of parallel connection). From this one obtains the following characterization.

Lemma 3. *Let Δ be a clutter whose path-connected components are $\Delta_1, \ldots, \Delta_m$. Then Δ is a matroid port if and only if $\Delta_1, \ldots, \Delta_m$ are matroid ports. Moreover, for any field $\mathbb{K}$, Δ is a $\mathbb{K}$-representable matroid port if and only if $\Delta_1, \ldots, \Delta_m$ are $\mathbb{K}$-representable matroid ports.*

It was shown in [7] that the diameter of a path-connected matroid port is at most two (by the diameter we mean, as usual, the maximum of the distances). The following lemma gives a condition for adding sets to an h-homogeneous binary matroid port of diameter 2 so that the result is an h-homogeneous binary matroid port of diameter 1. The proof is based on checking condition 2.(a) in Theorem 2.

Lemma 4. *Let Δ be a path-connected, h-homogeneous binary matroid port with support Ω such that its blocker $b(\Delta)$ contains the pairs $\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_s, y_s\}$ (with all the x_i, y_i different among them). Then $s \leqslant h$. For $s = h$, let $z_3, \ldots, z_h \notin \Omega$ and define $\Delta' = \Delta \cup \{\{x_i, y_i, z_3, \ldots, z_h\} : 1 \leqslant i \leqslant h\}$. If $b(\Delta)$ contains no set of the form $\{v_1, \ldots, v_t\}$ with $v_i \in \{x_i, y_i\}$ and $t < h$ then the clutter Δ' is an h-homogeneous binary matroid port.*

In addition to restricting to path-connected clutters, we introduce some other reductions. Two elements x, y in a clutter Δ are *equivalent* if for any $A \in \Delta$ we have $|A \cap \{x, y\}| \leqslant 1$, and if $|A \cap \{x, y\}| = 1$ then $A \ominus \{x, y\} \in \Delta$. It is easy to check that equivalent elements in a matroid port $\mathcal{M}_p$ correspond to parallel elements in $\mathcal{M}$. An element q in a clutter is called *universal* if $q \in \bigcap_{A \in \Delta} A$. If a matroid port $\mathcal{M}_p$ has a universal element q, then $\{p, q\}$ are a series pair in $\mathcal{M}$ (a parallel pair in $\mathcal{M}^*$). The *reduction* of a clutter Δ is the clutter Δ^{red} obtained by removing all but one copy of each equivalence class and removing all universal elements.

Lemma 5. *A clutter Δ is a matroid port (resp. is a $\mathbb{K}$-representable matroid port) if and only if its reduced clutter Δ^{red} is so. Moreover, if Δ has exactly t universal elements, then Δ is h-homogeneous if and only if Δ^{red} is $(h - t)$-homogeneous.*

We say that a clutter is *reduced* if it is path-connected and $\Delta^{\text{red}} = \Delta$. By Lemmas 3 and 5, we can restrict to h-homogeneous reduced matroid ports (although in some of the results in the following section we drop this restriction if the result is sufficiently simple to state it in general).

3 Homogeneous matroid ports of ranks 1, 2, 3, $n-1$ and n

Let Δ be an h-homogeneous clutter with support Ω, where $|\Omega| = n$. For $h = 1$ and $h = n$, the characterizations of Theorem 2 imply that $\Delta = \big\{\{x_1\}, \ldots, \{x_n\}\big\}$ and $\Delta = \big\{\{x_1, \ldots, x_n\}\big\}$ are binary matroid ports. The following proposition deals with the case $h = n-1$. The proof consists in checking that the collection $\mathcal{C}(\mathcal{M}_\Delta)$ as in Theorem 1 is indeed the collection of circuits of a matroid.

Proposition 6. *Let Δ be an $(n-1)$-homogeneous clutter with support Ω with $n = |\Omega|$. Then, Δ is a matroid port. Furthermore, Δ is a binary matroid port if and only if $|\Delta| = 2$.*

Now let us consider the case $h = 2$. Note that a 2-homogeneous clutter with support Ω can be thought of as the edges of a graph with vertex set Ω (and no isolated vertices). For the complete multipartite graph $G = K_{n_1,\ldots,n_\ell}$, for $\ell \geqslant 2$ and $n_i \geqslant 1$, the clutter $E(G)$ is a matroid port. Indeed, the reduced clutter $E(K_{n_1,\ldots,n_\ell})^{\text{red}}$ is the set of edges of a complete graph K_ℓ, and $E(K_\ell)$ is the port of a uniform matroid $U_{2,\ell+1}$ at any of its points. By adding equivalent elements we obtain $E(K_{n_1,\ldots,n_\ell})$. As $U_{2,4}$ is the excluded minor for binary matroids, the only one of these ports that is binary is $E(K_{n_1,n_2})$.

The following result states that these are the only 2-homogeneous matroid ports. It can be proved directly by using the characterizations of matroid ports in the previous section, or by combining results about secret sharing schemes from [1] and [6].

Theorem 7. *Let Δ be a path connected 2-homogeneous clutter. Then Δ is a matroid port if and only if Δ is isomorphic to $E(K_{n_1,\ldots,n_\ell})$, and it is is a binary matroid port if and only if Δ is isomorphic to $E(K_{n_1,n_2})$.*

Finally, in Theorem 8 we present the description of 3-homogeneous reduced binary matroid ports. The clutters involved in this theorem are defined as follows. The clutter $\Delta_{3,0}$ is the clutter whose elements are the 7 lines of the Fano plane; that is, $\Delta_{3,0} = \{\{a_1, a_2, a_3\}, \{a_1, a_4, a_7\}, \{a_1, a_5, a_6\}, \{a_2, a_4, a_6\}, \{a_2, a_5, a_7\}, \{a_3, a_4, a_5\}, \{a_3, a_6, a_7\}\}$ (as noted in [5], $\Delta_{3,0}$ is the port of the binary affine cube $\mathrm{AG}(3,2)$ at any of its points). The clutter $\Delta_{3,1}$ is $\Delta_{3,0} \setminus a_7$ (which can also be thought of as the set of 3-cycles of K_4). We remark that $\Delta_{3,0}$ can be constructed from $\Delta_{3,1}$ by using Lemma 4 (and checking first that $\Delta_{3,1}$ is binary, for instance by using Theorem 2).

Theorem 8. *Let Δ be a reduced 3-homogeneous clutter. Then, Δ is a binary matroid port if and only if Δ is isomorphic either $\Delta_{3,0}$ or $\Delta_{3,1}$.*

There are several ways of proving this theorem; one of them is by carefully analysing the circuits of $\mathcal{M}_\Delta$ as given by Theorem 1 and showing that if Δ is a 3-homogeneous binary matroid port, then $\mathcal{M}_\Delta$ is a matroid all whose circuits are of size 4. Thus, one can then apply the results by Murty [8].

Alternatively, Theorem 8 follows from the results in [5] and [6] by relating matroid ports to ideal secret sharing schemes. Moreover, from these papers one can show that if Δ a 3-homogeneous non-binary matroid port then Δ is the port of a matroid of rank three. An example is the clutter defined by the non-Pappus configuration, obtained by taking all 3-element subsets of a set of size 9 except those that are lines in the non-Pappus matroid (see [5, Example 3.5]). As far as we know, the complete description of non-binary 3-homogeneous matroid ports is an open problem.

4 Homogeneous binary matroid ports of rank 4

For 4-homogeneous matroid ports, there are no known results from the secret sharing community, and the results from [8] do not shed much light, as there is only one binary matroid all whose circuits have

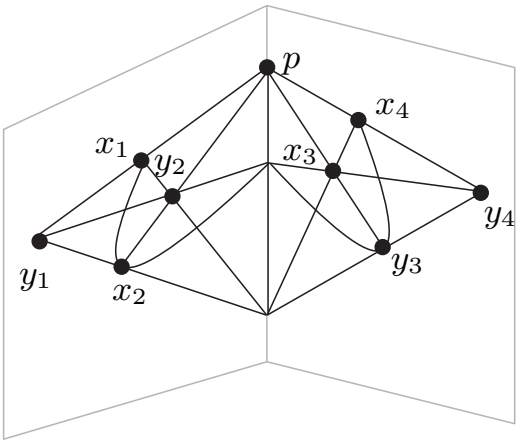


Figure 1: A geometric representation of the binary 4-spike. The hyperplanes avoiding p are in bijection with the sets in the port $\Delta_{4,1}$.

size 5 (namely, a single 5-circuit). Note that the corresponding matroid port has just one set, a case already commented at the beginning of Section 3.

Before stating our main result, we need to introduce yet one more reduction, which can be seen as an extension of equivalent elements.

Let Δ be a 4-homogeneous binary matroid port. We say that the pairs $\{a_1, a_2\}$ and $\{b_1, b_2\}$ are *clones* if the following three conditions hold

- $|A \cap \{a_1, a_2, b_1, b_2\}| \neq 1$ for all $A \in \Delta$;
- $\{a_i, b_j\} \not\subseteq A$ for any $A \in \Delta$ and $1 \leqslant i, j \leqslant 2$;
- $\{a_1, a_2, x, y\} \in \Delta$ if and only if $\{b_1, b_2, x, y\} \in \Delta$ for all $x, y \in E$.

Since the property of being binary and 4-homogeneous is preserved under deletion, the removal of one of the pairs $\{a_1, a_2\}$ or $\{b_1, b_2\}$ still leaves a binary 4-homogeneous matroid port. Next we consider the reverse process, that is, when it is possible to add a pair of clones. The proof of the following lemma consists in checking condition 2.(b) in Theorem 2.

Lemma 9. *Let Δ be a binary 4-homogeneous matroid port such that there are two elements $\{a_1, a_2\}$ with the property that $|A \cap \{a_1, a_2\}| \in \{0, 2\}$ for all $A \in \Delta$. Let b_1, b_2 be two elements not in the support of Δ. Then the clutter $\Delta' = \Delta \cup \{A \setminus \{a_1, a_2\} \cup \{b_1, b_2\} \mid \{a_1, a_2\} \subseteq A \in \Delta\}$ is binary.*

Thus, we can assume that Δ is clone-free. Theorem 10 states that there are only four 4-homogeneous binary matroid ports that are reduced and clone-free. We next define these four clutters, and show that they are indeed binary.

The clutter $\Delta_{4,1}$ is the following

$$\begin{aligned}\Delta_{4,1} = \{&\{x_1, x_2, y_3, x_4\}, \{x_1, x_3, y_2, x_4\}, \{x_2, x_3, y_1, x_4\}, \{y_1, y_2, y_3, x_4\},\\ &\{x_1, y_2, y_3, y_4\}, \{x_2, y_1, y_3, y_4\}, \{x_3, y_1, y_2, y_4\}, \{x_1, x_2, x_3, y_4\}\}.\end{aligned}$$

It is a binary matroid port since $b(\Delta_{4,1})$ is the port of the binary 4-spike at its tip (Figure 1 shows a geometric representation of the binary 4-spike; the elements of $\Delta_{4,1}$ are given by the complements of the planes that do not contain p, with p removed).

If we apply Lemma 4 to $\Delta_{4,1}$, we obtain the binary matroid port $\Delta_{4,0}$:

$$\begin{aligned}\Delta_{4,0} = \{&\{x_1, x_2, y_3, x_4\}, \{x_1, x_3, y_2, x_4\}, \{x_2, x_3, y_1, x_4\}, \{y_1, y_2, y_3, x_4\},\\ &\{x_1, y_2, y_3, y_4\}, \{x_2, y_1, y_3, y_4\}, \{x_3, y_1, y_2, y_4\}, \{x_1, x_2, x_3, y_4\},\\ &\{x_1, y_1, z_1, z_2\}, \{x_2, y_2, z_1, z_2\}, \{x_3, y_3, z_1, z_2\}, \{x_4, y_4, z_1, z_2\}\}.\end{aligned}$$

Finally, $\Delta_{4,2} = \Delta_{4,0} \setminus x_4$ and $\Delta_{4,3} = \Delta_{4,0} \setminus \{x_4, y_1\}$. All other deletions of $\Delta_{4,0}$ give ports isomorphic to $\Delta_{4,1}$, $\Delta_{4,2}$ or $\Delta_{4,3}$, or are not reduced. Note that $\Delta_{4,1} = \Delta_{4,0} \setminus z_1$, but checking that $\Delta_{4,0}$ is binary directly is more involved than constructing it from $\Delta_{4,1}$.

Theorem 10. *Let Δ be a reduced 4-homogeneous clutter without clones. Then, Δ is a binary matroid port if and only if Δ is isomorphic to either $\Delta_{4,0}$, or $\Delta_{4,1}$, or $\Delta_{4,2}$, or $\Delta_{4,3}$.*

The proof of Theorem 10 is long and entirely new, in the sense that it does not rely on previous results from either the matroid or the cryptographic communities. We give a very short sketch of the main ideas for the interested reader.

We start with a reduced and clone-free binary matroid port Δ; being reduced, the clutter Δ must contain at least three sets. We exploit the fact that the matroid $\mathcal{M}_\Delta$ is binary and the characterizations of Theorem 2 to find bounds on the sizes of the intersections of the sets in Δ. From this a careful case analysis follows, leading to the four clutters $\Delta_{4,0}, \Delta_{4,1}, \Delta_{4,2}$ and $\Delta_{4,3}$.

The final part of the proof consists in showing that the clutters $\Delta_{4,i}$, for $i \in \{0, 1, 2, 3\}$, are terminal, in the following sense: there is no reduced and clone-free binary matroid port that strictly contains $\Delta_{4,0}$; any reduced and clone-free binary matroid port that strictly contains $\Delta_{4,3}$ also contains a clutter isomorphic to $\Delta_{4,2}$; and any reduced and clone-free binary matroid port that strictly contains $\Delta_{4,1}$ or $\Delta_{4,2}$ is isomorphic to $\Delta_{4,0}$.

References

[1] E. F. Brickell and D. M. Davenport. On the classification of ideal secret sharing schemes (extended abstract). In *Advances in cryptology—CRYPTO '89 (Santa Barbara, CA, 1989)*, volume 435 of *Lecture Notes in Comput. Sci.*, pages 278–285. Springer, New York, 1990.

[2] A. Lehman. A solution of the Shannon switching game. *J. Soc. Indust. Appl. Math.*, 12:687–725, 1964.

[3] M. Lemos, T. J. Reid, and H. Wu. On the circuit-spectrum of binary matroids. *European J. Combin.*, 32(6):861–869, 2011.

[4] J. Martí-Farré and C. Padró. Secret sharing schemes on sparse homogeneous access structures with rank three. *Electron. J. Combin.*, 11(1):Research Paper 72, 16, 2004.

[5] J. Martí-Farré and C. Padró. Ideal secret sharing schemes whose minimal qualified subsets have at most three participants. *Des. Codes Cryptogr.*, 52(1):1–14, 2009.

[6] J. Martí-Farré and C. Padró. On secret sharing schemes, matroids and polymatroids. *J. Math. Cryptol.*, 4(2):95–120, 2010.

[7] J. Martí-Farré, C. Padró, and L. Vázquez. On the diameter of matroid ports. *Electron. J. Combin.*, 15(1):Note 27, 9, 2008.

[8] U. S. R. Murty. Equicardinal matroids. *J. Combinatorial Theory Ser. B*, 11:120–126, 1971.

[9] J. Oxley. *Matroid theory*, volume 21 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, second edition, 2011.

[10] P. D. Seymour. The forbidden minors of binary clutters. *J. London Math. Soc. (2)*, 12(3):356–360, 1975/76.

[11] P. D. Seymour. A forbidden minor characterization of matroid ports. *Quart. J. Math. Oxford Ser. (2)*, 27(108):407–413, 1976.

Enumeration of unlabelled chordal graphs with bounded tree-width Discrete Mathematics Days 2024

Jordi Castellví*[1] and Clément Requilé†[2]

[1]Centre de Recerca Matemàtica, Barcelona, Spain
[2]Departament de Matemàtiques and Institut de Matemàtiques de la Universitat Politècnica de Catalunya, Barcelona, Spain

1 Introduction

For $k \geq 1$, a k-tree is defined recursively as either a complete graph on $k+1$ vertices or a graph obtained by adding a new vertex incident to a k-clique of a smaller k-tree. The class of k-trees plays an important rôle in graph theory as it allows for an alternative definition of the *tree-width* of a graph g as the minimum k such that g is a subgraph of a k-tree. In particular, k-trees are the maximal graphs with tree-width at most k. Tree-width is stable under taking minors, thus from [11] the number of graphs with n vertices and bounded tree-width grows like ρ^n, for some $\rho > 1$, up to some lower order terms and a factor $n!$ in the case of labelled graphs. However, determining the value of ρ is a notorious open problem.

An approach for the enumeration of constrained classes of graphs admitting some recursive decomposition is to derive a functional equation satisfied by the associated generating function, then compute asymptotic estimates of its coefficients using methods from complex analysis [8]. A natural operation to decompose graphs with bounded tree-width is known as the *clique-sum* of two graphs and consists in distinguishing a clique of the same size in each of the graphs and identifying together the vertices of the two cliques to obtain one unique graph, then removing any subset of the edges of the new clique. This latter step makes the clique-sum operation intractable in the setting of [8]. Note, however, that this new clique becomes a separator of the resulting graph.

A graph is chordal if every cycle of length greater than three contains at least one chord. Alternatively, Dirac proved in [6] that a graph is chordal if and only if every minimal separator is a clique. This characterisation makes the clique-sum operation amenable to recursive methods by restricting it to chordal graphs. Wormald first used it in [14] to obtain the generating function of labelled chordal graphs from a recursive system of equations; based on the fact that k-connected chordal graphs can be uniquely decomposed into their $(k+1)$-connected components, and by rooting the k-connected ones at k-cliques it is possible to derive an equation defining the generating function of k-connected graphs in terms of that of the $(k+1)$-connected ones. If we now consider chordal graphs with tree-width bounded by some $t \geq 1$, then one obtains a finite system of equations from the connected to the t-connected level, which is in fact composed of the class of t-trees. Thus one can see chordal graphs with tree-width at most t as a natural generalisation of t-trees. This work was continued in [3] to obtain an estimate for the number labelled chordal graphs with tree-width at most $t \geq 1$ and n vertices of the form

$$c\, n^{-5/2}\, \gamma^n \qquad \text{as } n \to \infty. \tag{1}$$

*Email: jcastellvi@crm.cat. Research of J. C. supported by the Spanish State Research Agency, through the Severo Ochoa and María de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M).

†Email: clement.requile@upc.edu. Research of C. R. supported by the Spanish State Research Agency through projects MTM2017-82166-P and PID2020-113082GB-I00, and the grant Beatriu de Pinós BP2019, funded by the H2020 COFUND project No 801370 and AGAUR (the Catalan agency for management of university and research grants).

for some $c > 0$ and $\gamma > 1$ depending on t and computable numerically up to any precision.

In his seminal work [13], Pólya developed a theory to encode symmetries of labelled combinatorial structures, and thus opened the way to enumerate their unlabelled counterparts. Using Pólya's theory, Otter [12] was able to enumerate unlabelled trees from the rooted ones. His method was generalised in what is known as the *dissymmetry theorem for tree-decomposable structures* [1]. Building on that theory, Gainer-Dewar managed to derive a system of equations from which the ordinary generating function of unlabelled k-trees can be computed [9]. An alternative derivation was later designed in [10], and was instrumental in obtaining in [7] an asymptotic estimate for the number of unlabelled k-trees with n vertices in the form of (1).

In this work, we generalise [9, 10] and derive a system of equations defining the ordinary generating function of unlabelled chordal graphs with bounded tree-width. Our method is based on the decomposition into $(k+1)$-connected components of chordal graphs rooted at k-cliques, similarly to [14] and [3], and requires a non-trivial extension of Pólya theory to rooted structures started in [2]. This decomposition is tree-like in the sense of [1], and we can apply a dissymmetry theorem to "unroot" our graphs analogous to the tools developed in [2].

Theorem 1. *Let $t \geq 1$ and $k \in [t]$. Then the class of unlabelled k-connected chordal graphs with tree-width at most t can be derived from a grammar. Furthermore, this grammar can be translated into a finite system of equations that completely defines the associated ordinary generating function.*

The derivation from Theorem1 implies an efficient algorithm to compute the number of unlabelled k-connected chordal graphs with tree-width at most t and n vertices. In a long version of this work we intend to prove an asymptotic estimate in the form of (1), following [7] and [3]. Using our grammar, one can also derive structural results on large random graphs in the class, similarly to what was done in [3] and [5], as well as design a Boltzmann sampler to generate large uniform random graphs.

2 An extension of Pólya theory

2.1 Extended cycle index sums

Let $\mathcal{A}$ be a class of labelled graphs. A *symmetry* of $\mathcal{A}$ is a tuple (a, σ) where $\mathsf{a} \in \mathcal{A}$ and σ is an automorphism of a. The set of all symmetries of $\mathcal{A}$ is denoted by $\mathcal{S}(\mathcal{A})$. For $(\mathsf{a}, \sigma) \in \mathcal{S}(\mathcal{A})$ and a k-clique $K = \{v_1, \dots, v_i\}$ of a, with $k \geq 1$, note that there exists a smallest positive integer j such that $\sigma^j(K) = K$. We then say that the k-cliques $K, \sigma(K), \dots, \sigma^{j-1}(K)$ form a cycle of length j. Note that σ^j restricted to the vertices of K may be different from the identity. And we say that the cycle of k-cliques has type μ, if $\mu \vdash k$ is the cycle structure of the permutation σ^j restricted to the vertices of K. If we now let $c_{\mu,j}(\mathsf{a}, \sigma)$ be the number of cycles of k-cliques of a of length j and type $\mu \vdash k$, then the *extended weight-monomial* of a symmetry (a, σ) of size n is defined as

$$w_{(\mathsf{a},\sigma)} := \frac{1}{n!} \prod_{k=1}^{n} \prod_{\mu \vdash k} \prod_{j \geq 1} s_{\mu,j}^{c_{\mu,j}(\mathsf{a},\sigma)}.$$

From there, we define the *extended cycle index sum* of $\mathcal{A}$ as the sum of extended weight-monomials of symmetries of $\mathcal{A}$

$$X_{\mathcal{A}} := \sum_{(\mathsf{a},\sigma) \in \mathcal{S}(\mathcal{A})} w_{(A,\sigma)}.$$

It is a formal power series and a refinement of the (classical) cycle index sum, as the latter can be recovered setting $s_{\lambda,j} = 1$, for all $\lambda \vdash k > 1$, and $s_{(1),j} = s_j$. In order to recover the (ordinary) generating function of an unlabelled class from its cycle index sum, we recall Pólya's classical result.

Proposition 2 (Pólya [13])**.** *Let $\mathcal{A}$ be a class of labelled graphs and $\mathcal{U}$ be the class obtained by unlabelling the graphs in $\mathcal{A}$. Then, if we denote by $Z_{\mathcal{A}}(s_1, s_2, s_3, \dots)$ the cycle index sum of $\mathcal{A}$ and by $U(x)$ the ordinary generating function of $\mathcal{U}$, we have $U(x) = Z_{\mathcal{A}}(s_i \to x^i)_{i \geq 1} = Z_{\mathcal{A}}(x, x^2, x^3, \dots)$.*

2.2 Symmetries of graphs rooted at cliques

For $k \geq 1$, a graph in $\mathcal{A}$ is said to be *rooted at a k-clique* if one of its k-cliques K is distinguished, and the vertices of K are ordered instead of labelled. We denote by $\mathcal{A}^{(k)}$ the class of graphs in $\mathcal{A}$ that are rooted at a k-clique, and for $\mathsf{a} \in \mathcal{A}^{(k)}$ we let $r(\mathsf{a})$ be its root clique. Then an automorphism σ of $\mathsf{a} \in \mathcal{A}^{(k)}$ is also required to map $r(\mathsf{a})$ to itself, maybe permuting its vertices.

We consider permutations with cycle type $\lambda = (\lambda_1^{n_1}, \dots, \lambda_k^{n_k}) \vdash k$ with a canonical ordering of their cycles, and recall that in that case there are $\alpha(\lambda) := k!/(\lambda_1^{n_1} \dots \lambda_k^{n_k} n_1! \dots n_k!)$ many permutations with cycle type λ. A symmetry $(\mathsf{a}, \sigma) \in \mathcal{S}(\mathcal{A}^{(k)})$ is said to be *λ-rooted* if $\sigma_{|r(\mathsf{a})}$ has cycle type λ and the order of the vertices of $r(\mathsf{a})$ respects the canonical ordering of λ. We denote by $\mathcal{S}_\lambda(\mathcal{A}^{(k)})$ the set of all λ-rooted symmetries of $\mathcal{A}^{(k)}$. And for $(\mathsf{a}, \sigma) \in \mathcal{S}_\lambda(\mathcal{A}^{(k)})$ and $i \in [n]$, we let $c^*_{\mu,j}(\mathsf{a}, \sigma)$ be the number of cycles of i-cliques of a under the action of σ with length j and type $\mu \vdash i$, but this time each one of those i-cliques is not entirely contained in $r(\mathsf{a})$.

Then the *λ-rooted cycle index sum* of $\mathcal{A}^{(k)}$ can be similarly defined as

$$X^\lambda_{\mathcal{A}^{(k)}} := \sum_{(\mathsf{a},\sigma)\in\mathcal{S}_\lambda(\mathcal{A}^{(k)})} \frac{1}{n!} \prod_{i=1}^{n} \prod_{\mu\vdash i} \prod_{j\geq 1} s_{\mu,j}^{c^*_{\mu,j}(\mathsf{a},\sigma)}.$$

In practice, the λ-rooted cycle index sum of $\mathcal{A}^{(k)}$ can be computed from that of $\mathcal{A}$ via a formal derivative:

$$X^\lambda_{\mathcal{A}^{(k)}} = \frac{k!}{\alpha(\lambda)\kappa(\lambda)} \frac{\partial}{\partial s_{\lambda,1}} X_{\mathcal{A}}, \qquad \text{with} \quad \kappa(\lambda) := \prod_{i=1}^{k-1} \prod_{\mu\vdash i} \prod_{j\geq 1} s_{\mu,j}^{c_{\mu,j}(K_k,\sigma)}, \tag{2}$$

where K_k is the complete graph on k vertices and σ is one of its automorphisms.

However, in order to reverse this operation, that is computing the extended cycle index sum of some class $\mathcal{A}$ of unrooted graphs from the λ-rooted cycle index sums of $\mathcal{A}^{(k)}$, λ-rooted symmetries do not carry enough information and one requires to point symmetries at cycles (see [2]).

2.3 Symmetries of cycle-pointed graphs

For a class of labelled graphs $\mathcal{A}$, rooted or not, a symmetry $(\mathsf{a}, \sigma) \in \mathcal{S}(\mathcal{A})$ is said to be *cycle-pointed* when one of the cycles C of cliques of σ is distinguished. If C is a cycle of k-cliques ($k \geq 1$), then the graph a is said to be *k-cycle-pointed*, or *k-pointed* for short. The symmetries of a k-pointed graph a are then the symmetries (a, σ) for which σ is pointed at a cycle of k-cliques of a, however if a is rooted at a k-clique then none of its symmetries can be pointed at a cycle of cliques totally contained in $r(\mathsf{a})$.

We denote by $\mathcal{A}^{\bullet_k}$ the class of k-pointed graphs in $\mathcal{A}$ and by $\mathcal{S}_p(\mathcal{A}^{\bullet_k})$ the class of its cycle-pointed symmetries. The extended cycle index sum of $\mathcal{A}^{\bullet_k}$ is defined as

$$X_{\mathcal{A}^{\bullet_k}} := \sum_{(\mathsf{a},\sigma)\in\mathcal{S}_p(\mathcal{A}^{\bullet_k})} \frac{\ell}{|\mathsf{a}|!} t_{\lambda,\ell} \prod_{i=1}^{|\mathsf{a}|} \prod_{\mu\vdash i} \prod_{j\geq 1} s_{\mu,j}^{c^\bullet_{\mu,j}(\mathsf{a},\sigma)},$$

where λ is the type of the pointed cycle of (a, σ), ℓ its length, and $c^\bullet_{\mu,j}(\mathsf{a}, \sigma)$ is the number of unpointed cycles of i-cliques of a with length j and type $\mu \vdash i$. In practice, and following [2], one can derive the cycle index sum of $\mathcal{A}^{\bullet_k}$ from that of $\mathcal{A}$

$$X_{\mathcal{A}^{\bullet_k}} = \sum_{\lambda\vdash k} \sum_{j\geq 1} j t_{\lambda,j} \frac{\partial}{\partial s_{\lambda,j}} X_{\mathcal{A}}. \tag{3}$$

Thus, if every graph in $\mathcal{A}$ has at least one k-clique then $X_{\mathcal{A}}$ is completely determined by $X_{\mathcal{A}^{\bullet_k}}$. Denote by Ψ the operator such that $X_{\mathcal{A}} = \Psi(X_{\mathcal{A}^{\bullet_k}})$. Then we finally have

$$X_{\mathcal{A}} = \Psi\left(X_{\mathcal{A}^{\bullet_k}}\right). \tag{4}$$

Finally, reproducing the proof methods developed in [13] and [1], combinatorial construction rules can be readily translated into extended cycle index sums. For instance, if we let $\mathcal{A}$ and $\mathcal{B}$ be classes of labelled graphs, rooted or not, and $k \geq 1$ then, using the language from [8], we have

$$X_{\mathcal{A}+\mathcal{B}} = X_{\mathcal{A}} + X_{\mathcal{B}} \qquad \text{and} \qquad X_{\mathcal{A}\times\mathcal{B}} = X_{\mathcal{A}} \cdot X_{\mathcal{B}}, \tag{5}$$

$$X_{\mathcal{A}^{\bullet_k}+\mathcal{B}^{\bullet_k}} = X_{\mathcal{A}^{\bullet_k}} + X_{\mathcal{B}^{\bullet_k}} \qquad \text{and} \qquad X_{\mathcal{A}^{\bullet_k}\times\mathcal{B}} = X_{\mathcal{B}\times\mathcal{A}^{\bullet_k}} = X_{\mathcal{B}} \cdot X_{\mathcal{A}^{\bullet_k}}, \tag{6}$$

$$(\mathcal{A}+\mathcal{B})^{\bullet_k} = \mathcal{A}^{\bullet_k} + \mathcal{B}^{\bullet_k} \qquad \text{and} \qquad (\mathcal{A}\times\mathcal{B})^{\bullet_k} = \mathcal{A}\times\mathcal{B}^{\bullet_k} + \mathcal{A}^{\bullet_k}\times\mathcal{B}. \tag{7}$$

Note that the same considerations apply to rooted or non-rooted graphs with a distinguished subgraph. Precisely, automorphisms have to preserve the root and/or the distinguished subgraph, though maybe permuting its vertices, and the cycles of cliques entirely contained in the root of a cycle-pointed symmetry are not taken into account in its extended weight-monommial. Furthemore, extended cycles index sum, λ-rooted symmetries and λ-cycle index sums are defined for rooted and/or k-pointed graphs in the same way.

3 Chordal clique-sums and symmetries

3.1 Substitutions of cliques

Let $k \geq 1$, $\mathcal{A}$ be a class of (possibly rooted) labelled graphs, and $\mathcal{B}$ be a class of graphs rooted at a k-clique. The *clique substitution* $\mathcal{A} \circ_k \mathcal{B}$ is the class obtained by identifying each k-clique of graphs $\mathsf{a} \in \mathcal{A}$ with the root of graphs in $\mathcal{B}$. This results in a graph for which the base graph a is now a distinguished subgraph.

If $C = (c_1, \ldots, c_\ell)$ is a cycle of cliques and $C_1, \ldots, C_k$ is a sequence of k copies of C, then the *composed cycle* of $C_1, \ldots, C_k$ is the cycle of cliques of length ℓk such that for $i \in [k-1]$ and $j \in [\ell]$, the clique coming after c_j in C_i is c_j in C_{i+1}, and the clique coming after c_j in C_k is $c_{j+1 \mod \ell}$ in C_1. And if we denote by $(X_{\mathcal{A}})^{[j]}$ the cycle index sum resulting from multiplying the second subindex of all variables by j, that is, $s_{\lambda,i} \to s_{\lambda,ij}$, then the extended cycle index sum of the class $\mathcal{A} \circ_k \mathcal{B}$ is defined using the classical composition of cycle index sums

$$X_{\mathcal{A}\circ_k\mathcal{B}} = X_{\mathcal{A}}\left(s_{\lambda,j} \to s_{\lambda,j} \cdot \left(X_{\mathcal{B}}^{\lambda}\right)^{[j]}\right)_{\lambda\vdash k, j\geq 1}. \tag{8}$$

If additionally the graphs in both $\mathcal{A}$ and $\mathcal{B}$ are unpointed, then we also define the *k-pointed substitution* $\mathcal{A}^{\bullet_k} \odot_k \mathcal{B}$ as the class of all k-pointed graphs obtained by the following procedure: let first $(\mathsf{a}, \sigma) \in \mathcal{S}_p(\mathcal{A}^{\bullet_k})$ be a symmetry pointed at some cycle $C = (c_1, \ldots, c_k)$ with type λ, take k copies of a graph $\mathsf{p} \in \mathcal{B}^{\bullet_k}$ admiting a λ-rooted symmetry, and identify each clique in C with the root of one of the copies of p following the canonial order of λ. Second, for every unpointed cycle $D \neq C$ of (a, σ) with type μ, choose a graph $\mathsf{b} \in \mathcal{B}$ admitting a μ-rooted symmetry, take $|D|$ copies of b and identify each clique in D with the root of one of the copies of b. The base graph a is now a distinguished subgraph of the resulting graph and the vertices (except possibly the ones in the root) are assigned unique labels such that the relative order is preserved.

Furthermore, if $C_1, \ldots, C_k$ are the pointed cycles of the copies of p pasted respectively at $c_1, \ldots, c_k$ then the pointed cycle of the resulting graph is the composed cycle of $C_1, \ldots, C_k$. Note that this construction also works if p is a j-pointed graph with $j \neq k$. In that case the resulting graph is j-pointed. Adapting [2], it can then be shown that the k-pointed substitution $\mathcal{A}^{\bullet_k} \odot_k \mathcal{B}$ is a class of k-pointed graphs, that is, every graph admits a symmetry pointed at the cycle of k-cliques. The extended cycle index sum of $\mathcal{A}^{\bullet_k} \odot_k (\mathcal{B}, \mathcal{P})$ is then obtained via the *k-pointed plethystic composition* of their extended cycle index sums, whose definition is an extension of [2]

$$X_{\mathcal{A}^{\bullet_k}\odot_k\mathcal{B}} = X_{\mathcal{A}^{\bullet_k}} \odot_k X_{\mathcal{B}} := X_{\mathcal{A}^{\bullet_k}}\left(s_{\lambda,i} \to \left(X_{\mathcal{B}}^{\lambda}\right)^{[i]}, t_{\mu,j} \to \left(X_{\mathcal{B}^{\bullet_k}}^{\mu}\right)^{[j]}\right), \tag{9}$$

and where $i, j \geq 1$, and λ and μ both range over the partitions of k.

3.2 Starlike chordal clique-sum

Let $\mathcal{A}$ be a class of labelled graphs. The *starlike chordal clique-sum* of $\mathcal{A}$ is the class of rooted graphs $\mathsf{star}(\mathcal{A})$ obtained by taking a multiset of graphs from $\mathcal{A}^{(k)}$, identifying their rooted k-cliques together, and relabelling all the other vertices respecting their previous relative order. Then the λ-rooted extended cycle index sum of $\mathsf{star}(\mathcal{A})$ is

$$X^{\lambda}_{\mathsf{star}(\mathcal{A})} = \exp\left(\sum_{j\geq 1}\frac{1}{j}\left(X^{\lambda^j}_{\mathcal{A}^{(k)}}\right)^{[j]}\right), \tag{10}$$

where λ^j denotes the cycle type of σ^j if σ has cycle type λ.

Similarly, for $\ell \geq 1$ we define the *ℓ-pointed starlike chordal clique-sum* of $\mathcal{A}$ as the class $\mathsf{star}_\ell(\mathcal{A})$ of all the rooted and k-pointed graphs obtained by choosing some multiset S of elements of $\mathcal{A}^{(k)}$, considering $j \geq \ell$ disjoint copies $\mathsf{p}_1, \ldots, \mathsf{p}_j$ of some $\mathsf{p} \in \mathcal{A}^{\bullet_k}$, and then identifying together the rooted cliques of all the graphs in S and the rooted cliques of $\mathsf{p}_1, \ldots, \mathsf{p}_j$. Again, we relabel the non-root vertices to preserve the relative order. Furthermore, if $C_1, \ldots, C_j$ are the pointed cycles of $\mathsf{p}_1, \ldots, \mathsf{p}_j$, respectively, then the pointed cycle of the resulting graph is the composed cycle of $C_1, \ldots, C_j$. One can then show that every graph in $\mathsf{star}_\ell(\mathcal{A})$ admits a symmetry containing the pointed cycle of k-cliques as one of its cycles of cliques. In fact, the converse is also true and we have $\mathsf{star}(\mathcal{A})^{\bullet_k} = \mathsf{star}_1(\mathcal{A})$.

If, for $\ell \geq 1$, we now let

$$Z_{\mathsf{set}_\ell}(s_1, t_1, s_2, t_2, \ldots) = \sum_{j\geq \ell} j\, t_j \frac{\partial}{\partial s_j} \exp\left(\sum_{i\geq 1}\frac{s_i}{i}\right),$$

then the λ-cycle index sum of $\mathsf{star}_\ell(\mathcal{A})$ is given by

$$X^{\lambda}_{\mathsf{star}_\ell(\mathcal{A})} = Z_{\mathsf{set}_\ell}\left(s_i \to \left(X^{\lambda^i}_{\mathcal{A}^{(k)}}\right)^{[i]}, t_j \to \left(X^{\lambda^j}_{\mathcal{A}^{\bullet_k}}\right)^{[j]}\right)_{i\geq 1, j\geq \ell}. \tag{11}$$

4 Counting chordal graphs with bounded tree-width

Fix $t \geq 1$, and for any $k \in [t+1]$ let $\mathcal{G}_k$ (resp. $\mathcal{G}_k^{\bullet_k}$) be the class of labelled k-connected chordal graphs with tree-width at most t (resp. and that are k-pointed). Note that $\mathcal{G}_{t+1}$ is reduced to the $(t+1)$-clique, with cycle index sum

$$X_{\mathcal{G}_{t+1}} = \frac{1}{(t+1)!}\sum_{\lambda\vdash t+1}\alpha(\lambda)\kappa(\lambda). \tag{12}$$

From there, the relation (2) gives us the class $\mathcal{G}^{(t)}_{t+1}$. Fix now some $k \in [t]$ and some $\lambda \vdash k$. Then, adapting the scheme from [3] and provided we know $\mathcal{G}^{(k)}_{k+1}$, we can obtain by iteration the class $\mathcal{G}^{(k)}_k$ as a solution of the recursive equation

$$\mathcal{G}^{(k)}_k = \mathsf{star}\left(\mathcal{G}^{(k)}_{k+1} \circ_k \mathcal{G}^{(k)}_k\right). \tag{13}$$

To now obtain $\mathcal{G}_k$ from $\mathcal{G}^{(k)}_k$, we proceed following [2], by using the dissymmetry theorem for tree-decomposable classes [1] on $\mathcal{G}^{\bullet_k}_k$, which can be derived by adapting [3]. This gives

$$\mathcal{G}^{\bullet_k}_k \simeq \mathcal{G}^{(k)}_k + (\mathcal{G}_{k+1})^{\bullet_k}_{\geq 2} \odot_k \mathcal{G}^{(k)}_k + \mathsf{star}_2\left(\mathcal{G}^{(k)}_{k+1} \circ_k \mathcal{G}^{(k)}_k\right), \tag{14}$$

where the extended cycle index sum of $(\mathcal{G}_{k+1})^{\bullet_k}_{\geq 2}$ is defined as in (3) but without the terms with $j = 1$. Equations (13) and (14) relating combinatorial classes can then be translated into relations between extended cycle index sums using the various identities derived in Sections 2 and 3.

Finally, starting from (12) and by successive iterations of the recursive step (13), the cycle pointing step (3) and the unrooting step, composed of (14) together with (2) and (4), we can obtain $\mathcal{G}_k$ for any $k \in [t]$. In practice, we are able to compute its extended cycle index sum and the associated generating function, using Proposition 2, whose n-th coefficient is the number of unlabelled graphs with n vertices. We provide an effective implementation of the algorithm computing any term of the generating function on this repository, and as an example display next the first numbers of unlabelled chordal graphs with tree-width at most t, connectivity k and up to ten vertices.

	$t=1$	$t=2$	$t=3$	$t=4$
$k=1$	1 1 1 2 3 6 11 23 47 106	1 1 2 4 11 35 124 500 2224 10640	1 1 2 5 14 53 234 1265 8015 58490	1 1 2 5 15 57 266 1556 11187 97859
$k=2$	-	0 1 1 1 2 5 12 39 136 529	0 1 1 2 4 14 55 293 1842 13491	0 1 1 2 5 17 75 455 3486 32907
$k=3$	-	-	0 0 1 1 1 2 5 15 58 275	0 0 1 1 2 4 14 62 391 3182
$k=4$	-	-	-	0 0 0 1 1 1 2 5 15 64

The last non-empty line of column t corresponds to unlabelled t-trees, while the line $k=1$ of the second column corresponds to connected chordal series-parallel graphs with OEIS sequence A243788. To the extent of our knowledges, the other sequences are new. Note that an algorithm was designed to compute, among others, the first numbers of unlabelled chordal planar graphs (OEIS sequence A243787). The first discrepancy between A243787 and the line $k=1$ of the third column is given by the unique non-planar connected chordal graph with tree-width three and six vertices: it is the starlike chordal sum of three K_4's at a common triangle. By adapting [4] to the context of Pólya theory, we believe that a similar program could be developed in order to obtain additional terms of the ordinary generating function of unlabelled chordal planar gaphs with n vertices as well as an asymptotic estimate in the form of (1).

References

[1] F. Bergeron, G. Labelle, P. Leroux, *Combinatorial Species and Tree-like Structures*, translated by M. Readdy, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1997.

[2] M. Bodirsky, É. Fusy, M. Kang, S. Vigerske, Boltzmann samplers, Pólya theory, and cycle pointing, *SIAM Journal on Computing* **40:3** (2011), 721–769.

[3] J. Castellví, M. Drmota, M. Noy, C. Requilé, Chordal graphs with bounded tree-width, *Advances in Applied Mathematics* **157** (2024), 102700.

[4] J. Castellví, M. Noy, C. Requilé, Enumeration of chordal planar graphs and maps, *Discrete Mathematics* **346:1** (2023), 113163.

[5] J. Castellví, B. Stufler, Limits of chordal graphs with bounded tree-width, preprint, 2023, `arxiv:2310.20423`.

[6] G. A. Dirac, On rigid circuit graphs, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* **25** (1961), 71–76.

[7] M. Drmota, E. Y. Jin, An asymptotic analysis of labeled and unlabeled k-trees, *Algorithmica* **75** (2016), 579–605.

[8] P. Flajolet, R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.

[9] A. Gainer-Dewar, Γ-Species and the enumeration of k-trees, *The Electronic Journal of Combinatorics* **19:4** (2012), P45.

[10] A. Gainer-Dewar, I. M. Gessel, Counting unlabeled k-trees, *Journal of Combinatorial Theory, Series A* **126** (2014), 177–193.

[11] S. Norine, P. Seymour, R. Thomas, P. Wollan, Proper minor-closed families are small, *Journal of Combinatorial Theory. Series B* **96:5** (2006), 754–757.

[12] R. Otter, The number of trees, *Annals of Mathematics* **49** (1948), 583—599.

[13] G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen, *Acta Mathematica*, **68** (1937), 145–254.

[14] N. C. Wormald, Counting labelled chordal graphs, *Graphs and Combinatorics*, **1:2** (1985), 193–200.

The Borsuk number of a graph*

José Cáceres†[1], Delia Garijo‡[2], Alberto Márquez§[2], and Rodrigo I. Silveira¶[3]

[1]Dpto. de Matemáticas, Universidad de Almería, Spain
[2]Dpto. de Matemática Aplicada I, Universidad de Sevilla, Spain
[3]Dept. de Matemàtiques, Universitat Politècnica de Catalunya, Spain

Abstract

The *Borsuk problem* asks for the smallest number such that any bounded set in n-dimensional space can be cut into that many subsets with smaller diameter. It is a classical problem in combinatorial geometry that has been subject of much attention over the years, and research on variants of the problem continues nowadays in a plethora of directions. In this work, we propose a formulation of the problem in the context of graphs. Depending on how the graph is partitioned, we consider two different settings dealing either with the usual notion of diameter in abstract graphs, or with the so-called *continuous diameter* for the locus of plane geometric graphs. We present a complexity result, exact computations and upper bounds on the parameters associated to the problem.

1 Introduction

In 1933, Borsuk posed the question of whether every bounded set X in $\mathbb{R}^d$ could be partitioned into $d+1$ closed (sub)sets each with diameter smaller than that of X [1]. In this context, the diameter is defined as the maximum of the distances between two points in the set, under the Euclidean metric. This leads to the concept of *Borsuk number*. For a set $X \subset \mathbb{R}^d$, the Borsuk number $b(X)$ is the smallest number such that X can be partitioned into $b(X)$ subsets, each with diameter smaller than X. Borsuk's question can be thus stated as whether $b(X) \leq d+1$, for any bounded $X \subset \mathbb{R}^d$. The answer to this question was shown to be positive for $d = 2, 3$ [4, 10], and for general d for centrally symmetric convex bodies [11] and smooth convex bodies [6]. The general answer turned out to be negative, as shown in 1993 by Kahn and Kalai [8]. Since then, researchers have been trying to figure out the smallest dimension for which the partition does not exist, being $d = 64$ the currently best [7]. Many variants of the Borsuk problem have also been studied, see [12] for a recent survey.

We present a formulation of the problem in the context of graphs. Conceptually, we define the *Borsuk number* of a graph as the smallest number $b(G)$ such that G can be partitioned into $b(G)$ subgraphs, each with smaller diameter than the original graph. However, we need to define carefully how a graph can be partitioned. We propose two natural ways to do this, which lead to two variants of the problem: the *discrete* and the *continuous* Borsuk number of a graph. We define these formally in Section 2. Sections 3–5 contain our study on both parameters, encompassing a complexity result, exact computations and upper bounds. Proofs are omitted due to the page limit, although we very briefly explain the key ideas to prove our main results; they are based on an accurate analysis of how shortest paths and distances can change when modifying a graph.

*This research is supported by Grant PID2019-104129GB-I00 funded by MICIU/AEI/10.13039/501100011033.
†Email: jcaceres@ual.es.
‡Email: dgarijo@us.es.
§Email: almar@us.es.
¶Email: rodrigo.silveira@upc.edu.

1.1 Preliminaries

The *distance* between two vertices in an abstract graph G is the length of a shortest path connecting them. The *diameter* of G, denoted by $\mathrm{diam}_d(G)$, is the maximum distance between any two vertices of G. A *plane geometric graph* is an undirected graph $G = (V(G), E(G))$ whose vertices are points in $\mathbb{R}^2$, and whose edges are straight-line segments, connecting pairs of points, that intersect only at their endpoints. Each edge e has a length, $|e|$, equal to the Euclidean distance between its endpoints. The *locus* $\mathcal{G}$ of a plane geometric graph G is the set of all points of the Euclidean plane that are on (the edges of) G. In contrast to (abstract) graphs, in $\mathcal{G}$, there can be an infinite number of pairs of points whose distance is equal to the diameter. Here, the distance between two points is again the length of a shortest path between the points, but now such a path will contain up to two fragments of edges if the points are not vertices. The *diameter* of $\mathcal{G}$ or *continuous diameter* of G, $\mathrm{diam}_c(\mathcal{G})$, is the maximum distance between any two points in $\mathcal{G}$. Two points whose distance attains this value are called *diametral points*, and the shortest paths connecting diametral points are *diametral paths*. Problems dealing with the continuous diameter of a graph, also called *generalized* diameter [3], have received considerable attention recently, see [2, 5]. In the continuous case, we treat G and $\mathcal{G}$, interchangeably, as a closed point set, and assume that the distance between the endpoints of edge e is $|e|$.

2 Definitions of Borsuk number

2.1 Continuous Borsuk number

We consider a plane geometric graph G and partition its locus $\mathcal{G}$ by a sequence of cuts with straight lines. A line ℓ naturally partitions $\mathcal{G}$ into two geometric subgraphs (possibly, one empty). Moreover, to guarantee that the partition by ℓ does not produce a disconnected subgraph, we add to both subgraphs the longest segment in ℓ that has its endpoints in $\mathcal{G} \cup \ell$; this maximal segment is denoted by s. So, actually, the partition gives two subgraphs of $\mathcal{G} \cup \ell$, which are:

$$\mathcal{G}_1 = (\ell^+ \cap \mathcal{G}) \cup s \quad \text{and} \quad \mathcal{G}_2 = (\ell^- \cap \mathcal{G}) \cup s,$$

where ℓ^+ and ℓ^- are, respectively, the open half-planes above and below ℓ (right–left for vertical lines.)

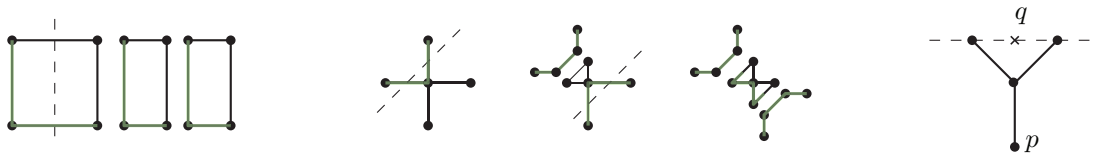
We define the *continuous Borsuk number of* G or *Borsuk number of* $\mathcal{G}$, and denote it by $b_c(\mathcal{G})$, as the minimum cardinality of a partition of $\mathcal{G}$ by lines $\ell_1, \ldots, \ell_k$ into subgraphs $\mathcal{G}_1, \ldots, \mathcal{G}_{k+1}$ such that $\max\{\mathrm{diam}_c(\mathcal{G}_1), \ldots, \mathrm{diam}_c(\mathcal{G}_{k+1})\} < \mathrm{diam}_c(\mathcal{G})$. In order to guarantee that the intersection with a line creates at most two subgraphs, each line ℓ_i is inserted only into one of the existing subgraphs.

Figure 1(a) illustrates this definition for a square. After partitioning the square with a vertical line ℓ (dashed) through its center point, we obtain two subgraphs: all points of $\mathcal{G}$ on each halfplane induced by ℓ, union the maximal segment in ℓ intersecting $\mathcal{G}$. Since this partitions the graph into two subgraphs (of $\mathcal{G} \cup \ell$), each with smaller diameter than that of $\mathcal{G}$, its continuous Borsuk number is two (best possible). However, sometimes more subgraphs are needed. The example in Figure 1(b) shows a 4-star graph, requiring at least two lines, giving continuous Borsuk number three. Note that the continuous diameter can increase when inserting a line, due to distances between points on the graph and new points on the line, see Figure 1(c).

One of the main open questions in this continuous setting is whether $b_c(\mathcal{G})$ can be upper-bounded by a constant. The following proposition gives a linear upper bound on the number of vertices of G.

Proposition 1. *Let $\mathcal{G}$ be the locus of a plane geometric graph with n vertices. Then, $b_c(\mathcal{G}) \leq 2n - 1$.*

Proof. (Brief sketch.) Consider a direction not parallel to any of the edges of $\mathcal{G}$; assume for simplicity that this is the vertical direction. For $\varepsilon > 0$, we split $\mathcal{G}$ by $2n$ vertical lines into $2n-1$ subgraphs; there are two lines associated to each vertex, one to the left and the other to the right, both at distance ε from the vertex. Thus, there are $2n - 1$ vertical strips, each containing either only portions of edges of $\mathcal{G}$, or only one vertex and portions of edges. Each resulting graph, $\mathcal{G}_1, \ldots, \mathcal{G}_{2n-1}$, is in one of these



(a) (b) (c)

Figure 1: (a) A square with side length 1 and diameter 2 (given by green paths), and a partition with a line; (b) a 4-star partitioned into three subgraphs by inserting two lines; (c) the continuous diameter increases when inserting the dashed line into the tree (p, q is a diametral pair.)

strips. Analyzing the different types of diametral pairs of points that may have been generated in these graphs $\mathcal{G}_i$, we can prove that their diameter is smaller than $\mathrm{diam}_c(\mathcal{G})$. It is worth noting that this construction does not work using only n lines, since the width of the strips containing a vertex of $\mathcal{G}$ must tend to zero, in order to avoid diametral pairs of points located on the inserted lines whose distance is larger than the original diameter. □

2.2 Discrete Borsuk number

We now consider partitions of an abstract graph G by simply deleting edges; here all edges have the same length, equal to 1. The *discrete Borsuk number of* G, denoted by $b_d(G)$, is the minimum cardinality of a partition of G by deleting edges into subgraphs $G_1, \ldots, G_k$ (of G) such that $\max\{\mathrm{diam}_d(G_1), \ldots, \mathrm{diam}_d(G_k)\} < \mathrm{diam}_d(G)$. The following observation gives some simple examples.

Observation 2. *(i) If G is a path or a cycle of even length, $b_d(G) = 2$.*

(ii) If G is a cycle of odd length, $b_d(G) = 3$.

(iii) If G is a star graph on $k+1$ vertices, $b_d(G) = k$.

In Section 5, we study the Borsuk number of an arbitrary tree T, in both, the discrete and the continuous setting. We show that while $b_c(\mathcal{T})$ is bounded by a constant, $b_d(T)$ can be linear with the number of vertices (as happens for the star). This linearity of the discrete Borsuk number also occurs in other families of graphs, such as unicycle graphs and maximal outerplanar graphs that are not trees.

3 Computational complexity

The problem of deciding whether the discrete Borsuk number of a graph G is below a given threshold is related to the *minimum clique cover problem*. A *clique cover* of a graph G is a partition of its vertex set into cliques. The *clique cover number* of G is the minimum size of a clique cover. The *minimum clique cover problem* seeks for a minimum clique cover.

Lemma 3. *The clique cover number of a non-complete graph G is an upper bound of $b_d(G)$, and both numbers coincide when* $\mathrm{diam}_d(G) = 2$.

Theorem 4. *Let G be a graph, and let k be a positive integer number. The problem of deciding whether $b_d(G) < k$ is NP–complete.*

Proof. Let G be a graph such that $\mathrm{diam}_d(G) > 1$ (otherwise, G is a complete graph, and $b_d(G)$ is simply the number of vertices of G, since all edges need to be removed to have each connected component with diameter zero.) The *cone* C_G of G is the graph obtained from G by adding a new vertex adjacent to all the vertices in G. The graph G has a clique cover of size k if and only if C_G has a clique cover of size k. Since $\mathrm{diam}_d(C_G) = 2$, by Lemma 3, the clique cover number of C_G is precisely $b_d(C_G)$. Thus, the result follows from the fact that deciding whether the clique cover number of an arbitrary graph is below a given threshold is an NP–complete problem [9]. □

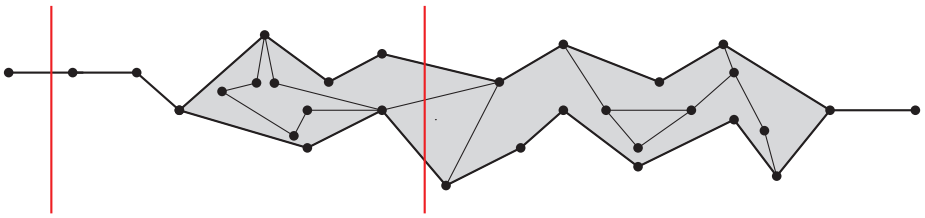


Figure 2: A graph that is monotone with respect to the x-axis; $G \cup \mathcal{F}_I$ consists of the graph (in black) and the gray region. Red vertical lines either intersect $G \cup \mathcal{F}_I$ at a single point or at a segment.

We conjecture that the problem in the continuous setting is also NP-hard, but, at the moment, a proof remains as future work.

4 Continuous Borsuk number of monotone graphs

Let G be a (plane geometric) graph, and let $G \cup \mathcal{F}_I$ be the part of the plane formed by the graph itself and all its interior faces. The graph G is said to be *ℓ-monotone* if the intersection of any line perpendicular to ℓ with $G \cup \mathcal{F}_I$ is either a single point or a segment; see Figure 2. We extend naturally this concept to the locus $\mathcal{G}$. For an ℓ-monotone graph $\mathcal{G}$, and a line ℓ' perpendicular to ℓ that is moving from left to right (parameterized by $\ell \cap \ell'$), we define the functions $d^+(\ell') = \mathrm{diam}_c((\ell'^+ \cap \mathcal{G}) \cup s')$ and $d^-(\ell') = \mathrm{diam}_c((\ell'^- \cap \mathcal{G}) \cup s')$, where s' is the maximal segment of ℓ' intersecting $\mathcal{G}$.

Lemma 5. *The functions $d^+(\ell')$ and $d^-(\ell')$ are monotone, respectively, decreasing and increasing.*

The continuous diameter can increase when partitioning a graph (see Figure 1b) but, as a straightforward consequence of the preceding lemma, we obtain that this is not true for monotone graphs.

Corollary 6. *The functions $d^+(\ell')$ and $d^-(\ell')$ associated to an ℓ-monotone graph $\mathcal{G}$ are upper-bounded by* $\mathrm{diam}_c(\mathcal{G})$.

In order to bound the continuous Borsuk number of a monotone graph, we introduce the concept of *diametral set*. The diametral set $D(p, q) \subseteq \mathcal{G}$ of a diametral pair p, q is defined as the union of all the shortest paths connecting p and q. Note that, for example, a cycle has an infinite number of diametral pairs of points, but only one distinct diametral set, which is the whole cycle (the union of the two diametral paths for each diametral pair is the same, the cycle). Thus, while a graph can have an infinite number of diametral pairs of points, we next state that this is not the case for diametral sets, which is key to prove Theorem 8 below.

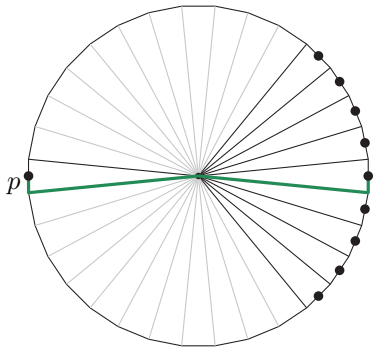
Lemma 7. *Let $\mathcal{G}$ be the locus of any plane geometric graph with n vertices. The number of distinct diametral sets of $\mathcal{G}$ is in $O(n^2)$.*

Theorem 8. *Let $\mathcal{G}$ be an ℓ-monotone graph such that there are no $k+1$ disjoint diametral sets. Then, $b_c(\mathcal{G}) \leq k + 2$.*

Proof. (Brief sketch.) By Corollary 6, in order to reduce the original diameter when cutting by lines, it suffices to intersect the $O(n^2)$ diametral sets of Lemma 7, with lines perpendicular to ℓ, since the new points on the cutting lines cannot cause an increase of $\mathrm{diam}_c(\mathcal{G})$. If we shorten one of the shortest paths connecting two diametral points, their distance decreases, and so each of the $O(n^2)$ diametral set only needs to be intersected once. We can prove that all the sets can be intersected using $k + 1$ lines. The idea is to project each diametral set onto the line ℓ so that each set determines an interval on the line. Then, for each interval, we define a line that intersects it, and also crosses all the intervals overlapping with that one. This produces a sequence of subsets of diametral sets $\mathcal{D}_0 \supset \mathcal{D}_1 \supset \mathcal{D}_2 \ldots$, where $\mathcal{D}_0$ is the set of all diametral sets of $\mathcal{G}$, satisfying that the diametral sets in $\mathcal{D}_{i-2} \backslash \mathcal{D}_{i-1}$ do not

intersect those in $\mathcal{D}_{i-1}\backslash\mathcal{D}_i$. Hence, we can find at most $k+1$ of the lines defined above, otherwise we would have $k+1$ disjoint diametral sets. □

We note that the previous bound can be attained, at least for $k = 1$. Consider, for example, the wheel graph on 33 vertices, $\mathcal{W}_{33}$, embedded in the plane such that its outer boundary is a regular 32-sided polygon, and the distance from the wheel center to each polygon vertex is one. This implies that each side has length $s = 2\sin(\pi/32) \approx 0.19$. Any two vertices of the polygon are connected by a path of length two, through the wheel center. This path is shorter than going along the boundary as soon as the other vertex is more than ten vertices away along the boundary (since $11s > 2$). It follows that the diametral pairs of this graph are given by pairs of midpoints of polygon sides that are at distance $2+s \approx 2.19$. In fact, each midpoint has nine points at exactly that distance, corresponding to the midpoint exactly opposite, plus those of the first four sides neighboring the opposite side, in each direction. See side figure for the nine diametral pairs involving p.



Next we argue that subdividing by one line is not enough to decrease the diameter of $\mathcal{W}_{33}$. Any line intersecting the wheel will leave at least 15 complete triangles of the wheel on one side. These triangles are contiguous, and form a fan. The diameter of any such a fan with 13 or more triangles remains the same as the original one, $2+s$. It follows that two lines are necessary. Moreover, they are also sufficient, since two parallel lines at a very small distance that enclose the center will result in three subgraphs with smaller diameter. Therefore, $b_c(\mathcal{W}_{33}) = 3 = k+2$.

5 Borsuk number of trees

In this section, we first compute $b_d(T)$ for an arbitrary tree T, and then we move to the continuous version of the problem, which behaves differently.

Proposition 9. *The discrete Borsuk number of any tree T with n vertices can be computed in $O(n)$ time. Furthermore,*

(i) If the center of T is not a unique vertex, then $b_d(T) = 2$.

(ii) If the center of T is a vertex v, then $b_d(T) = b_d(T') = \delta_{T'}(v)$, where T' is the subtree of T induced by the vertices of all diametral paths, and $\delta_{T'}(v)$ is the degree of v in T'.

While $b_d(T)$ depends on the center of T, we next show that the continuous Borsuk number is upper-bounded by a constant. We apply the following lemma that states that lines intersecting a tree at its center cannot cause an increase of the diameter of the tree.

Lemma 10. *Let $\mathcal{T}$ be the locus of a tree with center point $\mathcal{C}$, and let ℓ be a line that passes through $\mathcal{C}$. Then, $\max\{\mathrm{diam}_c((\ell^+ \cap \mathcal{T}) \cup s), \mathrm{diam}_c((\ell^- \cap \mathcal{T}) \cup s)\} \leq \mathrm{diam}_c(\mathcal{T})$, where s is the longest segment in ℓ that has its endpoints in $\mathcal{T} \cup \ell$.*

Lemma 10 also holds for lines that intersect the tree, not at the center, but infinitely close to it. Further, with lines that go exactly through the center $\mathcal{C}$, we cannot guarantee that the diameters obtained after cutting are strictly smaller than $\mathrm{diam}_c(\mathcal{T})$ (for example, the star graph with three edges of the same length and not contained in the same half-plane through the center). However, Proposition 11 below states that when a tree has Borsuk number 2, we can always find a line intersecting $\mathcal{T}$ at a point infinitely close to the center giving a correct partition (that is, the diameter decreases with respect to the original). This is an important step in order to design an algorithm for deciding whether the continuous Borsuk number of a tree is 2 or 3 which, by Theorem 12, are its possible values.

Proposition 11. *Let $\mathcal{T}$ be the locus of a tree with center point $\mathcal{C}$. If $b_c(\mathcal{T}) = 2$ then there exists a sequence of lines $\{\ell_i\}_{i\geq 0}$ satisfying that:*

(i) $\{d_T(t_i, \mathcal{C})\}_{i\geq 0}$ *approaches zero, where t_i is the closest point in $\mathcal{T} \cap \ell_i$ to $\mathcal{C}$.*

(ii) there exists $j \geq 0$ such that for every $i \geq j$, $\max\{\mathrm{diam}_c((\ell_i^+ \cap \mathcal{T}) \cup s_i), \mathrm{diam}_c((\ell_i^- \cap \mathcal{T}) \cup s_i)\} < \mathrm{diam}_c(\mathcal{T})$, *where s_i is the longest segment in ℓ_i that has its endpoints in $T \cup \ell_i$.*

Theorem 12. *Let $\mathcal{T}$ be the locus of a tree. Then, $b_c(\mathcal{T}) \leq 3$.*

Proof. (Brief sketch.) Consider a line ℓ that passes through the center point $\mathcal{C}$ of $\mathcal{T}$, and splits the tree into two graphs $\mathcal{T}_1$ and $\mathcal{T}_2$. We can assume that ℓ does not contain any edge incident or containing $\mathcal{C}$ as an interior point. By Lemma 10, the diameters of $\mathcal{T}_1$ and $\mathcal{T}_2$ are at most $\mathrm{diam}_c(\mathcal{T})$. A case analysis of how distances change after inserting the line, lets us conclude that $d_{T_1}(p, \mathcal{C}) \leq \mathrm{diam}_c(\mathcal{T})/2$ for every point p on $\mathcal{T}_1$ (analogous for $\mathcal{T}_2$). This fact is the key tool to prove that the diameters of $\mathcal{T}_1$ and $\mathcal{T}_2$ are strictly smaller than $\mathrm{diam}_c(\mathcal{T})$ when $\mathcal{C}$ is not a vertex of T, which leads to $b_c(\mathcal{T}) = 2$.

If $\mathcal{C}$ is a vertex of T, a diametral pair of $\mathcal{T}_1$ or $\mathcal{T}_2$ may consists of two leaves at distance $\mathrm{diam}_c(\mathcal{T})$. Then, we may need two lines to decrease the diameter; for example, this is the case in the star graph with all edges of the same length and such that no half-plane through the center contains all of them. It suffices to take two parallel lines to ℓ, one slightly above and the other below. This gives $b_c(\mathcal{T}) \leq 3$. □

6 Conclusions

We have introduced the concept of Borsuk number of a graph in a discrete and a continuous setting. Let us mention that this is ongoing research. In the continuous setting, we are currently focusing on proving the NP-hardness of computing $b_c(\mathcal{G})$, and whether there is a polynomial time algorithm to decide whether $b_c(G) = 2$. In addition, we are trying to answer the question of whether $b_c(\mathcal{G})$ can be upper-bounded by a constant, and designing an algorithm for trees as mentioned above. We are also delving deeper into the discrete version, currently studying the Borsuk number of unicycle graphs to better understand the behavior of this parameter.

References

[1] K. Borsuk. Three theorems on the n-dimensional sphere. *Fund. Math*, 20:177–190, 1933.

[2] J. L. De Carufel, C. Grimm, S. Schirra, and M. Smid. Minimizing the continuous diameter when augmenting a tree with a shortcut. In *Algorithms and Data Structures. WADS 2017. LNCS 10389*, pages 301– 312, 2017.

[3] C. E. Chen and R. S. Garfinkel. The generalized diameter of a graph. *Networks*, 12(3):335–340, 1982.

[4] H. G. Eggleston. Covering a three-dimensional set with sets of smaller diameter. *J. London Math.*, 30:11—24, 1955.

[5] D. Garijo, A. Márquez, N. Rodríguez, and R. I. Silveira. Computing optimal shortcuts for networks. *Eur. J. Oper. Res.*, 279(1):26–37, 2019.

[6] H Hadwiger. Mitteilung betreffend meine Note: Überdeckung einer Menge durch Mengen kleineren Durchmessers. *Commentarii Mathematici Helvetici*, 19(1):72–73, 1946.

[7] T. Jenrich and A. E. Brouwer. A 64-dimensional counterexample to Borsuk's conjecture. *Electron. J. Comb.*, 21(4):4, 2014.

[8] J. Kahn and G. Kalai. A counterexample to Borsuk's conjecture. *B. Am. Math. Soc.*, 29(1):60–62, 1993.

[9] R. M. Karp. Reducibility among combinatorial problems. In R.E. Miller, J.W. Thatcher, and J.D. Bohlinger, editors, *Complexity of Computer Computations*, Plenum Press, New York, pages 85–103. 1972.

[10] J. Perkal. Sur la subdivision des ensembles en parties de diamètre inférieur. *Colloq. Math.*, 1:45, 1947.

[11] A. S. Rissling. Das Borsuksche Problem in dreidimensionalen Räumen konstanter Krümmung. *Ukr. Geom. Sb.*, 11:78–83, 1971.

[12] C. Zong. Borsuk's partition conjecture. *Jpn. J. Math.*, 16:185–201, 2021.

A canonical van der Waerden theorem in random sets

José D. Alvarado*[1], Yoshiharu Kohayakawa†[1], Patrick Morris‡[2], Guilherme O. Mota§[1], and Miquel Ortega¶[2]

[1]Instituto de Matemática e Estatística , Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, Brazil
[2]Departament de Matemàtiques, Universitat Politècnica de Catalunya (UPC), Carrer de Pau Gargallo 14, 08028 Barcelona, Spain

Abstract

The canonical van der Waerden theorem states that, for large enough n, any colouring of $[n]$ gives rise to monochromatic or rainbow k-APs. In this work, we are interested in sparse random versions of this result. More concretely, we determine the threshold at which the binomial random set $[n]_p$ inherits the canonical van der Waerden properties of $[n]$.

1 Introduction

Arithmetic Ramsey theory is a branch of combinatorics that studies what sort of arithmetic structure must appear in all possible – although frequently restricted to finite – colourings of the integers. One of the first and most celebrated results in the field is van der Waerden's theorem [23], which states that any finite colouring of the integers contains monochromatic arithmetic progressions of arbitrary length.

There are several paths to generalizing van der Waerden's theorem, which also provide a better understanding of the underlying phenomenon responsible for the appearance of such arithmetic structure. A first option consists in studying when can one guarantee the existence of other objects besides arithmetic progressions. An example of such a generalization is the work of Rado [15], who studied the case of general linear structures. In fact, he was able to characterize precisely those homogeneous linear systems of equations that admit monochromatic solutions in any finite colouring of the integers (for more details see, for example, [9]).

A second possibility for generalizing van der Waerden's theorem consists in dropping the restriction on the number of colours, and studying what kind of structure one may still find for all possible colourings, using a finite numbers of colours or not. It turns out that, in the case of arithmetic progressions, the needed piece of structure to complete the puzzle is that of *rainbow* progressions, namely, arithmetic progressions where every element has a different colour. This gives rise to the canonical van der Waerden theorem, first proved by Erdős and Graham [6], where an arithmetic progression which is monochromatic or rainbow is called *canonical*.

Theorem 1 (Canonical van der Waerden)**.** *For any integer $k \geq 1$, there exists large enough n such that the following holds. Any colouring of $[n] = \{1, \ldots, n\}$ contains a canonical arithmetic progression of length k.*

*Email: josealvarado.mat17@ime.usp.br. Partially supported by FAPESP (2020/10796-0)
†Email: yoshi@ime.usp.br. Partially supported by CNPq (406248/2021-4, 407970/2023-1, 315258/2023-3).
‡Email: pmorrismaths@gmail.com. Partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Walter Benjamin program - project number 504502205.
§Email: mota@ime.usp.br. Partially supported by (315916/2023-0, 406248/2021-4) and FAPESP (2018/04876-1).
¶Email: miquel.ortega.sanchez-colomer@upc.edu. Partially supported by FPI PRE2021-099120

Note that Theorem 1 clearly implies that any colouring of the integers (possibly with an infinite numbers of colours) contains arbitrarily long canonical arithmetic progressions. A well-known compactness argument, with a nice additional idea, can be applied to prove that this "infinite" version in fact does imply the "finite" version in the theorem above (see, e.g., [7, p. 29]).

A final path to generalizing van der Waerden's theorem is changing the ambient set where the theorem holds. Instead of colouring the integers, one might consider looking for monochromatic arithmetic progressions in colourings of some particular subset of the integers. For example, a common way to do this is proving an analogue of the theorem in random sets (see, for example, [19]). The theorem also holds when the ambient set is substituted by a set that is pseudorandom enough, meaning, in very informal terms, that it has statistical properties similar to those of a random set (see [5] and the references therein for an in-depth discussion).

In fact, the focus of this work will be a sparse random version of Theorem 1. Consider the random set $[n]_p$, where every element of $[n]$ is sampled independently with probability p. We study how large must p be for $[n]_p$ to satisfy an analogue of the canonical van der Waerden theorem. More formally, a threshold for a monotone property $\mathcal{P}$ is a function $p^* = p^*(n)$ such that

$$\lim_{n\to\infty} P([n]_p \in \mathcal{P}) = \begin{cases} 0 \text{ if } p \leq cp^* \\ 1 \text{ if } p \geq Cp^* \end{cases}$$

for constants $C, c > 0$. The case of the statement that guarantees $[n]_p \notin \mathcal{P}$ with high probability is referred to as the 0-*statement* and the other one as the 1-*statement*.

For example, Rödl and Ruciński [18,19] established such a threshold for the van der Waerden property (and, more generally, a Rado type theorem). To state it precisely, let us say a set is (r, k)-*van der Waerden* if any r-colouring contains a monochromatic arithmetic progression of length k. Their result, for the case of arithmetic progressions, reads as follows.

Theorem 2 (Sparse van der Waerden)**.** *Given integers $k \geq 3$ and $r \geq 2$, there exist constants $C, c > 0$ such that:*

- *For $p > Cn^{-1/(k-1)}$, the random set $[n]_p$ is a.a.s. (r, k)-van der Waerden.*
- *For $p < cn^{-1/(k-1)}$, the random set $[n]_p$ is a.a.s. not (r, k)-van der Waerden.*

See also [8] for further discussion of Theorem 2, where sharpness of the threshold is proved.

The main contribution of this work consists in proving the same kind of statement for the canonical van der Waerden theorem. Indeed, let us say a set is *canonically k-van der Waerden* if every colouring contains a canonical arithmetic progression of length k. We prove the following.

Theorem 3 (Sparse canonical van der Waerden)**.** *Given a natural number $k \geq 3$, there exist constants $C, c > 0$ such that:*

- *For $p > Cn^{-1/(k-1)}$, the random set $[n]_p$ is a.a.s. canonically k-van der Waerden.*
- *For $p < cn^{-1/(k-1)}$, the random set $[n]_p$ is a.a.s. not canonically k-van der Waerden.*

Note that the 0-statement follows from the corresponding 0-statement in Theorem 2 with $r = 2$, since a rainbow arithmetic progression of length k cannot be formed with only two colours.

Results such as this one, where a known theorem over discrete ambient sets is translated to sparser settings, and particularly to sparse random subsets, have become a common theme in modern combinatorics in the last decades. In the case of Ramsey's theorem, this was first carried out in a seminal series of papers by Rödl and Ruciński [16–18], and both the 0-statement and the 1-statement had delicate and involved proofs. The 1-statement was later reproved with a short and elegant argument by Nenadov and Steger [14], using the method of hypergraph containers [3]. The ideas of Nenadov and Steger

have also been used in arithmetic Ramsey theory, for example, to give short proofs of the 1-statement in sparse random versions of Rado's theorem (see [10] or [22]). Our proof uses the ideas of [14] and the method of hypergraph containers, although a straightforward application of their methods is not possible because their argument relies on the boundedness of the number of colours.

Sparse random versions of canonical Ramsey theorems are much more recent. In these, one must somehow overcome the difficulty of having a possibly unbounded number of colours in a given colouring. In a breakthrough result, Kamčev and Schacht [11] have proven a sparse analogue of the canonical Ramsey theorem for cliques, using the transference principle of Conlon and Gowers [4]. In independent work, a subset of the authors [1] prove a canonical Ramsey theorem when the colourings are constrained by some prefixed lists of colours. This is then one of of the ingredients to establish a canonical Ramsey theorem for even cycles [2]. These results use a combination of ideas from the method of containers and the work of Rödl and Ruciński [19].

The current work is inspired in the previous work of [1, 2] and [11], and aims to prove an analogous result in the arithmetic setting. It turns out that, when looking for arithmetic progressions, the situation for canonical theorems is different from theirs, since we only look for monochromatic or rainbow copies of arithmetic progressions, whereas when dealing with graphs, one might also find *lexicographic* copies of graphs (see [1] or [11] for more details). In the course of proving our result, we use a new set of ideas which allow for a streamlined proof in the arithmetic setting.

2 Sketch of the proof

In this section we give a rough sketch of the proof of Theorem 3 and some of the ideas involved in it. As we noted before, we concentrate only on the proof of the 1-statement.

The proof of the 1-statement starts out with a basic dichotomy over a given colouring of $[n]_p$. If such a colouring has a colour with positive density, we are able to apply the sparse version of Szemerédi's theorem (see Theorem 5 below) to find a monochromatic arithmetic progression of length k or k-AP for short). If, on the other hand, all colours are sparse, it turns out that we may find a rainbow k-AP. In order to split according to this criterion, we introduce bounded colourings.

Definition 4. *An r-colouring $\chi\colon A \to [r]$ of a set $A \subset \mathbb{N}$ is α-bounded for $\alpha > 0$ if $|\chi^{-1}(i)| \leq \alpha|A|$ for all $i \in [r]$, that is, all colours have density at most α in A.*

2.1 The dense colour case

We say a set A is (δ, k)-*Szemerédi* if every subset of A of size greater than $\delta|A|$ contains a k-AP. The sparse random version of Szemerédi's theorem, proven by Conlon and Gowers [4] and Schacht [21] independently, establishes the threshold where $[n]_p$ satisfies this condition.

Theorem 5 (Szemerédi's theorem for sparse random sets)**.** *Given $\delta > 0$ and a natural number $k \geq 3$, there exists a constant C such that, for $p > Cn^{-1/(k-1)}$, the random set $[n]_p$ is a.a.s. (δ, k)-Szemerédi.*

For our purposes, this can be rephrased in terms of bounded colourings.

Corollary 6. *Given $\alpha > 0$ and a natural number $k \geq 3$, there exists $C = C(\alpha, k)$ such that the set $[n]_p$ a.a.s. satisfies the following property for $p > Cn^{-1/(k-1)}$. Every colouring of $[n]_p$ that is not α-bounded contains a monochromatic k-AP.*

2.2 Searching for rainbows

On account of the previous observation, it suffices to establish the following to prove Theorem 3.

Proposition 7. *Given $k > 0$, there exist $C = C(k)$ and $\alpha = \alpha(k)$ such that the set $[n]_p$ a.a.s. satisfies the following property for $p > Cn^{-1/(k-1)}$. Every α-bounded colouring of $[n]_p$ contains a rainbow k-AP.*

Once picking a suitable constant α, successively merging the smallest colours, we may reduce the proof of Proposition 7 to the case of α-bounded colourings with at most $r = \lceil 2/\alpha \rceil$ colours. This leaves us in a better position, since now the number of colours is bounded in terms of α and we may use container type arguments.

2.2.1 A container theorem

In very loose terms, the hypergraph container theorem [3, 20] is a way to cluster independent sets of a sufficiently regular hypergraph. This allows one to control the probability of lying in one of these clusters when a simple union bound over all possible independent sets would be too large. Still in somewhat vague terms, it gives, for every sufficiently regular hypergraph $\mathcal{H}$, a collection of *containers* $\mathcal{C} \subset \mathcal{P}(V(\mathcal{H}))$ that satisfy the following properties:

- The containers are almost independent. They contain few edges of $\mathcal{H}$, usually in the sense that $e(\mathcal{H}[C]) < \varepsilon e(\mathcal{H})$ for $\varepsilon > 0$ as small as needed and every $C \in \mathcal{C}$.
- Every independent set in $\mathcal{H}$ is contained in one of the containers.
- There are few containers. More precisely, every independent set has a small subset (its *fingerprint*) that is uniquely associated to a container. The number of containers may be bounded by the total amount of small subsets.

In order to prove Proposition 7 we just look for rainbow k-APs, so we apply the container theorem to the rainbow copy hypergraph $\mathcal{H} = \mathcal{H}(n, k, r)$ with vertex set consisting of r copies of $[n]$, one for every possible colour, and edge set formed by all possible rainbow k-APs. More formally, $\mathcal{H}$ is the k-uniform hypergraph with vertex set $V(\mathcal{H}) = [n] \times [r]$, and edge set

$$E(\mathcal{H}) = \left\{ \{(n_1, c_1), \dots, (n_k, c_k)\} \in \binom{V(\mathcal{H})}{k} : (n_1, \dots, n_k) \text{ forms a } k\text{-AP and } c_i \neq c_j \, \forall i \neq j \right\}.$$

It is useful to identify subsets of $V(\mathcal{H})$ with the following notion of an $[r]$-coloured set.

Definition 8. *An $[r]$-colouring of a set $A \subset [n]$ is a function $\chi \colon A \to \mathcal{P}([r])$. Such a pair (A, χ) forms an $[r]$-coloured set. A subcolouring (A', χ') of (A, χ) is a colouring such that $A' \subset A$ and $\chi'(i) \subset \chi(i)$ for all $i \in A'$.*

Indeed, an $[r]$-coloured subset (A, χ) with $A \subset [n]$ can also be thought of as a subset of $[n] \times [r] = V(\mathcal{H})$ in a natural way, and we write $A_\chi \subset V(\mathcal{H})$ for such a set. An application of the hypergraph container theorem then gives the following (for a very similar application of the container method, see [12, 13]).

Theorem 9. *For any $k \in \mathbb{N}$ and $\varepsilon > 0$, there exists a constant $C = C(k, \varepsilon) > 0$, a collection $\mathcal{A}$ of $[r]$-coloured sets, and a function $f \colon \mathcal{P}([n])^r \to \mathcal{A}$ satisfying:*

- *Every $A_\chi \in \mathcal{A}$ satisfies $e(\mathcal{H}[A_\chi]) < \varepsilon n^2$, that is, there are less than εn^2 rainbow k-APs compatible with the $[r]$-coloured set A_χ.*
- *For every $[r]$-coloured set I_χ with no rainbow k-AP, there exists a subcolouring $S_\psi \subset I_\chi$ such that*

$$|S| \leq C n^{1-1/(k-1)} \quad \text{and} \quad I_\chi \subset f(S_\psi).$$

2.2.2 A supersaturation result

In order to use Theorem 9, we must study what we can deduce about our containers from the fact that they admit few rainbow k-APs. In most applications of the hypergraph container theorem this is stated as a supersaturation result, which gives conditions that guarantee the existence of many solutions. Here, we state it in the contrapositive form, which turns out to be slightly more comfortable for this application. We obtain the following.

Proposition 10. *Given k, there exists $M = M(k)$ such that the following holds. For every r, there is an $\varepsilon = \varepsilon(k, r) > 0$ such that any $[r]$-coloured subset (χ, A) satisfying $|A| \geq n/2$ and $e(\mathcal{H}[A_\chi]) < \varepsilon n^2$ admits a subset of colours $L \subset [r]$ with $|L| \leq M$ and a subset $B \subset A$ with*

$$|B| \geq n/8 \quad \text{and} \quad \chi(B) \subset L, \tag{1}$$

that is, one can find a large number of values in A which only use colours in L, and the size of L is bounded independently of the number of colours r.

We remark that the crucial part of Proposition 10 is that the bound M on $|L|$ does not depend on the total number of colours r, but only on k. The proof of this result starts out by noticing that the set

$$C = \{x \in A \colon |\chi(x)| \geq k\}$$

is small, say $|C| < n/4$. Otherwise, by a supersaturated version of Szemerédi's theorem, we find many k-APs in C, and, since every element of C has at least k different colours to choose from, every k-AP in C gives rise to at least one rainbow k-AP.

Let $A' = A \setminus C$. By somewhat more delicate counting arguments, one can prove that, for most values of $x \in A'$, the list $\chi(x)$ is made up uniquely of colours that appear at least in βn other values of A', for some $\beta = \beta(k) > 0$. These colours make up our list L, and the bound on its size follows from the fact that there are at most $|A'|k$ pairs of value and colour, so that

$$|L| \leq \frac{|A'|k}{\beta n} \leq \frac{k}{\beta}.$$

2.2.3 Putting it together

Proposition 10 implies, for ε small enough, that each container obtained in Theorem 9 is either small ($|A| \leq n/2$), or there exists $B \subset A$ satisfying (1). From Theorem 9 it follows that, if the set $[n]_p$ can be coloured with an α-bounded colouring with no rainbow k-AP, then $[n]_p \subset A$ for some $A_\chi \in \mathcal{A}$. If $|A| \leq n/2$, it is exponentially unlikely that $[n]_p \subset A$, so we focus on the latter case. In fact, we expect $[n]_p$ to have size close to np and intersect the corresponding B in about $|B|p \geq np/8$ positions. Assuming such estimates would leave us in a very good position, since any colouring of $[n]_p$ compatible with B would have a colour of density at least

$$\frac{|B|p}{npM} \geq \frac{1}{8M},$$

which, setting $\alpha \ll 1/8M$, would contradict the α-boundedness of the colouring. Thus, the conclusion of Proposition 7 fails only when $[n]_p$ or its intersection with B have large deviations from their expected value. Using Chernoff bounds to obtain exponential concentration and some standard estimates involving a union bound over all possible containers gives Proposition 7.

References

[1] J. D. Alvarado, Y. Kohayakawa, P. Morris, and G. O. Mota, *A canonical Ramsey theorem with list constraints in random graphs*, Procedia Computer Science **223** (2023), 13–19, XII Latin-American Algorithms, Graphs and Optimization Symposium (LAGOS 2023).

[2] ———, *A canonical Ramsey theorem for even cycles in random graphs*, manuscript (2024).

[3] J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709. MR 3327533

[4] D. Conlon and W. T. Gowers, *Combinatorial theorems in sparse random sets*, Ann. of Math. (2) **184** (2016), no. 2, 367–454. MR 3548529

[5] D. Conlon, J. Fox, and Y. Zhao, *Extremal results in sparse pseudorandom graphs*, Adv. Math. **256** (2014), 206–290. MR 3177293

[6] P. Erdős and R. L. Graham, *Old and new problems and results in combinatorial number theory*, Monographies de L'Enseignement Mathématique [Monographs of L'Enseignement Mathématique], vol. 28, Université de Genève, L'Enseignement Mathématique, Geneva, 1980. MR 592420

[7] S. Farhangi, *On refinements of van der Waerden's theorem*, Master's thesis, Virginia Polytechnic Institute and State University, 2016.

[8] E. Friedgut, H. Hàn, Y. Person, and M. Schacht, *A sharp threshold for van der Waerden's theorem in random subsets*, Discrete Anal. (2016), Paper No. 7, 20pp. MR 3533306

[9] R. L. Graham, B. L. Rothschild, and J. H. Spencer, *Ramsey theory*, Wiley-Interscience Series in Discrete Mathematics, John Wiley & Sons, Inc., New York, 1980, A Wiley-Interscience Publication. MR 591457

[10] R. Hancock, K. Staden, and A. Treglown, *Independent sets in hypergraphs and Ramsey properties of graphs and the integers*, SIAM J. Discrete Math. **33** (2019), no. 1, 153–188. MR 3899160

[11] N. Kamčev and M. Schacht, *Canonical colourings in random graphs*, 2023. Available as arXiv:2303.11206.

[12] X. Li, H. Broersma, and L. Wang, *Integer colorings with no rainbow* 3*-term arithmetic progression*, Electron. J. Combin. **29** (2022), no. 2, Paper No. 2.28, 15pp. MR 4417188

[13] H. Lin, G. Wang, and W. Zhou, *Integer colorings with no rainbow* k*-term arithmetic progression*, European J. Combin. **104** (2022), Paper No. 103547, 12pp. MR 4414806

[14] R. Nenadov and A. Steger, *A short proof of the random Ramsey theorem*, Combin. Probab. Comput. **25** (2016), no. 1, 130–144. MR 3438289

[15] R. Rado, *Studien zur Kombinatorik*, Math. Z. **36** (1934), 424–480 (German).

[16] V. Rödl and A. Ruciński, *Lower bounds on probability thresholds for Ramsey properties*, Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993, pp. 317–346. MR 1249720

[17] ———, *Random graphs with monochromatic triangles in every edge coloring*, Random Structures Algorithms **5** (1994), no. 2, 253–270. MR 1262978

[18] ———, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942. MR 1276825

[19] ———, *Rado partition theorem for random subsets of integers*, Proc. London Math. Soc. (3) **74** (1997), no. 3, 481–502. MR 1434440

[20] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992. MR 3385638

[21] M. Schacht, *Extremal results for random discrete structures*, Ann. of Math. (2) **184** (2016), no. 2, 333–365. MR 3548528

[22] C. Spiegel, *A note on sparse supersaturation and extremal results for linear homogeneous systems*, Electron. J. Combin. **24** (2017), no. 3, Paper No. 3.38, 19pp. MR 3691555

[23] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wiskd., II. Ser. **15** (1927), 212–216 (German).

Multi-Objective Linear Integer Programming Based in Test Sets

M. I. Hartillo-Hermoso*[1], H. Jiménez-Tafur†[2], and J.M. Ucha-Enríquez ‡[1]

[1]Dpto. Matemática Aplicada I, Universidad de Sevilla, Spain
[2]Dpto. de Matemáticas. Universidad Pedagógica Nacional, Bogotá, Colombia

Abstract

We introduce a new exact algorithm for Multi-objective Linear Integer problems based on the classical ϵ-constraint method and algebraic test sets computed with Gröbner bases. Our method takes advantage of test sets 1) to identify which IPs have to be solved in an ϵ-constraint framework and 2) using reduction with test-sets instead of solving with an optimizer. We show that the computational results are promising in some families of examples.

1 Introduction

Problems in the real world involve multiple objectives. Due to conflict among these objectives, finding a feasible solution that simultaneously optimizes all objectives is often impossible. As decision makers usually need a complete knowledge of the best decisions they can take from those different points of view, generating the set of *efficient solutions* (i.e., solutions for which it is impossible to improve the value of one objective without worsening the value of at least one other objective) is a primary goal in multi-objective optimization. Multi-objective Integer Programming (MOIP) is the branch that deals with this kind of problem in the case of integer variables, and the linear case (MOILP) is the one in which we will concentrate.

Generation methods compute the whole space of Pareto optimal solutions. Among these type of methods, we have the *weighted sum* of objectives approach and the *ϵ-constraint technique*, that generates a grid in the objective space with ranges between the costs of *ideal* and *nadir* points. In ϵ-constraint methods, for each point in the upper bound set (cf. [9, 5]) a single-objective problem is solved, avoiding incremental movements through the grid.

In [14, 10, 15, 19, 13] different approaches to apply this ϵ-constrained setting in MOLIP can be found. Two additional algebraic approaches to MOIP have been presented: the one proposed in [3], that introduces the so called *partial Gröbner bases*, and [4] that generalized for several cost functions the ideas presented in [1] for single-objective problems. Unfortunately these two algebraic proposals can not manage big examples, to the best of our knowledge.

Our approach is based on the so-called *test sets* associated to single-objective Linear Integer Programming problems (LIP), taking advantage of their special characteristics. A test set is a set of directions that guides the movement from any feasible point until the optimum of the LIP is reached. So LIPs are solved by *reduction* with these test sets, instead of passing them to an optimizer. It is proved in [18] that *Gröbner bases* provide the minimal test set for a fixed total ordering compatible with the linear cost function of the considered program. These test sets do not depend on the right hand sides (RHS) of the constraints. Interested readers can consult the references [17, 2].

*Email: hartillo@us.es
†Email: hjimenezt@pedagogica.edu.co
‡Email: ucha@us.es

We will show how our method takes advantage of the features of test sets to manage the ϵ-constraint setting efficiently: most of the typical redundant computations are circumvented and we only provide new efficient solutions. Although the computation of Gröbner bases can be a hard task, very sensitive to the number of variables (cf. [16]) in our experiments the algorithm is fairly competitive in the *unbounded knapsack problem*.

This paper is a generalization of a previous work of the authors ([12]) for the biobjective case.

2 Preliminaries

A multi-objective linear integer optimization problem (MOLIP) in standard form can be stated as

$$\begin{array}{ll} \min & c_1(\mathbf{x}), \ldots, c_p(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in \mathbb{Z}^n_{\geq 0} \end{array} \tag{1}$$

for $A \in \mathbb{Z}^{m\times n}$, $\operatorname{rank}(A) = m$, $\mathbf{b} \in \mathbb{Z}^m$ and $c_1, \ldots, c_p$ with $p \geq 2$ linear functions with integer coefficients. In general there is no feasible point that minimizes all the cost functions, so we are interested in obtaining the *efficient points*, that is those feasible points $\mathbf{x}^\star$ such that there is no feasible $\mathbf{x}$ with $c_k(\mathbf{x}) \leq c_k(\mathbf{x}^\star)$ with at least one strict inequality for $k = 1, \ldots, p$. If $\mathbf{x}^\star$ is an efficient point, $(c_1(\mathbf{x}^\star), \ldots, c_p(\mathbf{x}^\star))$ is a *non-dominated (or Pareto) point* in the decision space. If we replace the condition $c_k(\mathbf{x}) \leq c_k(\mathbf{x}^\star)$ for $c_k(\mathbf{x}) < c_k(\mathbf{x}^\star)$ we obtain *weakly efficient* points. We will denote $\mathcal{X}$ the set of efficient points and $\mathcal{N}$ the set of non-dominated points, the *Pareto frontier*.

We will assume that the feasible region for problem (1) is finite, so the Pareto frontier $\mathcal{N}$ is finite as well. In this paper we present an algorithm to obtain a set $\mathcal{X}^\star \subset \mathcal{X}$ that is a *minimal complete set of efficient points* (that is, if $\mathbf{x}^a, \mathbf{x}^b \in \mathcal{X}^\star$ then $(c_1(\mathbf{x}^a), \ldots, c_p(\mathbf{x}^a)) \neq (c_1(\mathbf{x}^b), \ldots, c_p(\mathbf{x}^b))$ and $|\mathcal{X}^\star| = |\mathcal{N}|$, as in [8])

The *ϵ-constraint technique*, (see [11]), one of the best known techniques to address problem (1), manages many problems of the form

$$\begin{array}{ll} \min & c_k(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & c_j(\mathbf{x}) \leq \epsilon_j, \ j = 1, \ldots, p \ (j \neq k) \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0} \end{array} \tag{2}$$

for fixed $k = 1, \ldots, p$ and suitable values of ϵ_j in order to solve Problem (1). Optimal points of Problem 2 are always weakly efficient. Furthermore we can identify the efficient solutions, as the following theorem of [8] states:

Theorem 1. *A feasible solution $\mathbf{x}^\star$ of a linear MOIP is efficient if and only if there exists a $(\epsilon_1, \ldots, \epsilon_p) \in \mathbb{R}^p$ such that $\mathbf{x}^\star$ is an optimal solution of the corresponding problems (2) for $k = 1, \ldots, p$.*

Thus we have families of IPs for which only the right hand side (RHS) varies, so it is natural to consider at this point one algebraic tool called the *test set* of a given LIP. Given the family of LIPs in standard form (no inequalities)

$$\begin{array}{ll} \min & c(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0} \end{array} \tag{3}$$

for $A \in \mathbb{Z}^{m\times n}$, $\operatorname{rank}(A) = m$, $\mathbf{b} \in \mathbb{Z}^m$ and c a linear function with coefficients in $\mathbb{Z}^n$, in general there is not only one optimal point but several ones with the same cost. We can refine the cost function considering a total order $\prec_c$ that first compares two points by the cost c and breaks ties according to a chosen *term order* $\prec$ (see [6]). If we consider problem (3) replacing the cost function by $\prec_c$, it does not affect the optimal value but, as it is a total order, it insures a unique optimum.

Definition 2. *A* test set with respect to $\prec_c$ *of the family of problems (3) for fixed A is a set $\mathcal{T} \subset \{\mathbf{t} \in \mathbb{Z}^n : A\mathbf{t} = 0\}$ valid for any RHS, with the following properties:*

1. *For any feasible, non-optimal solution $\mathbf{x}$ of (3) for some $\mathbf{b}$, there exists $\mathbf{t} \in \mathcal{T}$ such that $\mathbf{x} - \mathbf{t}$ is feasible and $\mathbf{x} - \mathbf{t} \prec_c \mathbf{x}$.*

2. *Given the optimal solution $\mathbf{x}^\star$ of (3) for some $\mathbf{b}$, we have that $\mathbf{x}^\star - \mathbf{t}$ is not feasible for any $\mathbf{t} \in \mathcal{T}$.*

There exists a test set for any given LIP that can be computed with Gröbner bases with respect to $\prec_c$ ([18]). The existence of test sets for an LIP implies a straightforward algorithm to find its optimum: we start from any feasible point and subtract elements of the testset as long as we obtain feasible points. We will refer to this process as *reduction* of a feasible point with the test set.

So given the family of problems (2) for a fixed k, using test sets to solve them requires only 1) the computation of *one* test set for *all* the problems and 2) the reduction of a feasible point of each problem with the test set. It is very important to underline that, if the test set is available, the reduction process is very often faster than passing the IP to an optimizer. In addition, we will see that test sets guide us during the task of choosing which values of ϵ_j produce new efficient solutions, avoiding many redundant LIPs to be solved. At last, in contrast with several methods that compute first weakly efficient solutions and filter them in a second step, we will see that using a suitable total order we obtain efficient points directly.

3 Characterization of efficient points using test sets

To solve the problem (1) we will adopt a recursive scheme. We will obtain a minimal set of efficient solution of the problems

$$\begin{array}{ll} \min & c_1(\mathbf{x}), \ldots, c_i(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \in \mathbb{Z}^n_{\geq 0} \end{array} \tag{4}$$

for $i = 2, \ldots, p$ and for this purpose we will use the ϵ-constraint method and manage the problems $P_i(\epsilon_1, \ldots, \epsilon_{i-1})$ (in standard form)

$$\begin{array}{ll} \min & c_i(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & c_1(\mathbf{x}) + r_1 = \epsilon_1, \\ & \vdots \\ & c_{i-1}(\mathbf{x}) + r_i = \epsilon_{i-1}, \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0}, \end{array} \tag{5}$$

for $i = 1, \ldots, p$ and $(\epsilon_1, \ldots, \epsilon_{i-1}) \in \mathbb{R}^{i-1}$.

For a given $i, 2 \leq i \leq p$, let us note $\prec_{\hat{c}_i}$ the total order that first compares two feasible points with respect to c_i and to break ties uses successively $c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_p$ and finally a chosen term order $\prec$ if it were necessary. We will denote $\mathcal{T}_{\hat{i}} \subset \mathbb{Z}^{n+(i-1)}$ the test-set for problem (5) with respect to the total order $\prec_{\hat{c}_i}$ (the elements of this test set have $i-1$ additional variables because of the slack variables added to the problem to put it in standard form).

The following result provides a characterization of the efficient points in this context.

Theorem 3. *A feasible point $(\mathbf{x}^*, \mathbf{0}) \in \mathbb{Z}^{n+(p-1)}$ is the optimal solution of $P_p(c_1(\mathbf{x}^\star), \ldots, c_{p-1}(\mathbf{x}^\star))$ with respect to the total order $\prec_{\hat{c}_p}$ if and only if $\mathbf{x}^\star$ is an efficient solution of (4) and among the ones with costs $(c_1(\mathbf{x}^\star), \ldots, c_{p-1}(\mathbf{x}^\star))$ is the smallest one with respect to $\prec_{\hat{c}_p}$*

Corollary 4. *If $(\mathbf{x}^*, \mathbf{t}) \in \mathbb{Z}^{n+(p-1)}$ con $\mathbf{t} \geq 0$ is the optimal solution of $P_p(\epsilon)$ for some $\epsilon \in \mathbb{Z}^{n+(p-1)}$ with respect to $\prec_{\hat{c}_p}$ then $\mathbf{x}^\star$ is an efficient solution of (1).*

Theorem 3 provides in particular a way to obtain the first point of our set of representatives of the non-dominated set of points of problem (1), the one with minimum c_1:

Corollary 5. *[8, Lemma 5.2.] Let $\mathbf{x}_1^\star$ be the optimal solution of*

$$\min\{c_1(\mathbf{x}) : A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbb{Z}^n_{\geq 0}\}$$

with respect to the ordering $\prec_{\hat{c}_1}$. Then $\mathbf{x}_1^\star$ is an efficient solution of (1) with minimum cost c_1.

4 Recursive construction of a minimal set of efficient solutions

Given the set of efficient solutions $\mathcal{X}$, let us denote $\mathcal{X}^\star \subset \mathcal{X}$ the minimal complete set of efficient points whose elements have the property of being the smallest ones with respect to $\prec_{\hat{c}_p}$ among the points that have the same costs. So by definition there is one efficient solution corresponding to each element in the Pareto frontier $\mathcal{N}$. The next result show how the elements $\mathcal{X}^\star$ can be obtained:

Theorem 6. *Let $\mathbf{x} \in \mathcal{X}^\star$. Then one of the following statements is true:*

1. *$c'(\mathbf{x}^*) = (c_1(\mathbf{x}^*), \ldots, c_{p-1}(\mathbf{x}^\star))$ belongs to the Pareto frontier of problem (4) for $i = p-1$*
2. *There exists a solution $\mathbf{x}'$ of $P_p(\epsilon')$ for some $\epsilon' \in \mathbb{R}^{n+(p-1)}$ such that $c_i(\mathbf{x}') \leq c_i(\mathbf{x}^\star)$ for $1 \leq i \leq p-1$ with at least an strict inequality and there exists $(\mathbf{t}, \mathbf{r}) \in \mathcal{T}_p$ such that $\mathbf{t} \leq \mathbf{x}^\star$ and $\mathbf{r} \geq \mathbf{0}, \mathbf{r} \neq \mathbf{0}$ (componentwise) and $\mathbf{r} = c'(\mathbf{x}^\star) - c'(\mathbf{x}')$.*

The theorem above assures, by induction, that the elements of $\mathcal{X}^\star$ come from solving problem (4) for some $i = 1, \ldots, p-1$ (that is, belong to the solution of the problem taking into account only the first i cost functions) or from reducing elements of the form $(\mathbf{x}^\star, \mathbf{r})$ for some $\mathbf{x}^\star$ efficient solution of problem (4) for some $i = 1, \ldots, p-1$ and some $\mathbf{r}$ that produce an element $(\mathbf{x}^\star, \mathbf{r})$ that is reducible and whose reduction with respect to $\mathcal{T}_p$. Its reduction produces a new element in $\mathcal{X}^\star$.

Theorem 7. *Let $\mathbf{x}^*$ be an efficient solution of*

$$\begin{array}{ll} \min & c_1(\mathbf{x}), \ldots, c_i(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \quad \mathbf{x} \in \mathcal{Z}^n_{\geq 0} \end{array} \tag{6}$$

with respect to $\preceq_{\hat{c}_i}$ for some $i, 1 \leq i \leq p-1$ then $\mathbf{x}^\star$ is efficient for

$$\begin{array}{ll} \min & c_1(\mathbf{x}), \ldots, c_i(\mathbf{x}), c_{i+1}(\mathbf{x}) \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \quad \mathbf{x} \in \mathbb{Z}^n_{\geq 0} \end{array} \tag{7}$$

with respect to $\preceq_{\hat{c}_{i+1}}$.

Algorithm 1 Algorithm to obtain a minimal set of efficient solutions of a MOILP with p objetives ($p \geq 2$)

Require: vector of cost functions $(c_1, c_2, \ldots, c_p)$, A and $\mathbf{b}$ of problem (1)
  Compute $\mathcal{T}_1$
  $\mathbf{e}_1 \leftarrow$ the solution of $\min\{c_1(\mathbf{x})$ s.t. $A\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathbf{Z}^n_{\geq 0}\}$ with respect to $\prec_{\hat{c}_1}$
  $\mathcal{X}' \leftarrow \{\mathbf{e}_1\}$
  $P \leftarrow \{\mathbf{e}_1\}$
  for $i = 2, \ldots, p$ **do**
    Compute $\mathcal{T}_i$
    for all $\mathbf{x} \in P$ **do**
      $P := P \setminus \{\mathbf{x}\}$
      Compute $\tilde{G}_{\mathbf{x}}$
      if $\tilde{G}_{\mathbf{x}} \neq \emptyset$ **then**
        $G_{\text{jumps}} \leftarrow \{(\mathbf{x}, \mathbf{r})$ such that there exists $(\mathbf{t}, \mathbf{r}) \in \tilde{G}_{\mathbf{x}}\}$
        for all $(\mathbf{x}, \mathbf{r}) \in G_{\text{jumps}}$ **do**
          $(\mathbf{y}, \mathbf{t}) \leftarrow$ optimal solution of $P_i(\epsilon_1, \ldots, \epsilon_{i-1})$ with respect to $\prec_{\hat{c}_i}$ and initial feasible solution $(\mathbf{x}, \mathbf{r})$.
          if $\mathbf{y} \notin \mathcal{X}'$ **then**
            $\mathcal{X}' \leftarrow \mathcal{X}' \cup \{\mathbf{y}\}$
            $P \leftarrow P \cup \{\mathbf{y}\}$
          end if
        end for
      end if
    end for
  end for
  OUPUT: A minimal set of efficient solutions with respect to $\prec_{\hat{c}_p}$

Algorithm 1 takes into account our previous results and produce a minimal set of efficient points for a given problem (1). For a given $\mathbf{x}$ and a given test-set $\mathcal{T}_i$ we will denote $G_{\mathbf{x}} = \{(\mathbf{t}, \mathbf{r}) \in \mathcal{T} : \mathbf{t} \leq \mathbf{x}, \mathbf{r} \geq \mathbf{0}, \mathbf{t} \neq \mathbf{0}\}$ and $\tilde{G}_{\mathbf{x}}$ the subset of elements $(\mathbf{t}, \mathbf{r})$ of $G_{\mathbf{x}}$ with their last $i-1$ components non comparable.

5 Conclusions

We have introduced a new exact algorithm to obtain a minimal set of efficient points for MOLIPs. It is based on the classical ϵ-constraint method and test sets for a family of IPs computed via Gröbner bases with respect to an order that, properly chosen, guides us in the process of obtaining only efficient solutions and avoiding most of unnecessary computations.

Computational experiments are promising for unbounded knapsack problems (that could be hard to treat with the usual techniques of the binary case). We have have been able to solve problems up to 100 variables for 3 objectives and 75 variables for 4 and 5 objectives (as far as we know the biggest examples proposed in the literature). We have treated too some examples of multi-objective redundancy allocation problems (as in [7]) with excellent results.

References

[1] D Bertsimas, G Perakis, and S Tayur. A new algebraic geometry algorithm for integer programming. *Management Science*, 46(7):999–1008, 2000.

[2] D. Bertsimas and R. Weismantel. *Optimization over integers*. Dynamic ideas, 2005.

[3] V. Blanco and J. Puerto. Partial Gröbner bases for multiobjective integer linear optimization. *SIAM J. Discrete Math.*, 23(2):571–595, 2009.

[4] V. Blanco and J. Puerto. Some algebraic methods for solving multiobjective polynomial integer programs. *J. Symbolic Comput.*, 46(5):511–533, 2011.

[5] K. Bringmann, T. Friedrich, C. Igel, and T. Voß. Speeding up many-objective optimization by Monte Carlo approximations. *Artificial Intelligence*, 204:22–29, 2013.

[6] D. A. Cox, J. Little, and D. O'Shea. *Using algebraic geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2005.

[7] A. Murat D. Cao and R.B. Chinnam. Efficient exact optimization of multi-objective redundancy allocation problems in series-parallel systems. *Reliability Engineering & System Safety*, 111:154–163, 2013.

[8] M. Ehrgott. *Multicriteria Optimization*. Springer, Berlin, second edition, 2005.

[9] M. Ehrgott and X. Gandibleux. Bound sets for biobjective combinatorial optimization problems. *Comput. Oper. Res.*, 34(9):2674–2694, 2007.

[10] M. Ehrgott and S. Ruzika. Improved ϵ-constraint method for multiobjective programming. *J. Optim. Theory Appl.*, 138(3):375–396, 2008.

[11] Y. Haimes, L.S. Lasdon, and D.A. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics*, 1(3):296–297, 1971.

[12] M. I. Hartillo-Hermoso, Haydee H. Jiménez-Tafur, and J. M. Ucha-Enríquez. An exact algebraic ϵ-constraint method for bi-objective linear integer programming based on test sets. *European J. Oper. Res.*, 282(2):453–463, 2020.

[13] G. Kirlik and S. Sayın. A new algorithm for generating all nondominated solutions of multiobjective discrete optimization problems. *European Journal of Operational Research*, 232:479–488, 2014.

[14] M. Laumanns, L. Thiele, and E. Zitzler. An efficient, adaptative parameter variation scheme for metaheuristics based on the epsilon-constraint method. *European Journal of Operational Research*, 169:932–942, 2006.

[15] G. Mavrotas and K. Florios. An improved version of the augmented ϵ-constraint method (AUGMECON2) for finding the exact pareto set in multi-objective integer programming problems. *Applied Mathematics and Computation*, 219:9652–9669, 2013.

[16] E. W. Mayr and Albert R. A. R: Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. in Math.*, 46(3):305–329, 1982.

[17] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. American Mathematical Society, Providence, RI, 1996.

[18] R.R. Thomas. A geometric Buchberger algorithm for integer programming. *Mathematics of Operations Research*, 20(4):864–?884, 1995.

[19] W. Zhang and M. Reimann. A simple augmented ϵ-constraint method for multi-objective mathematical integer programming problems. *European Journal of Operational Research*, 234:15–24, 2014.

Rainbow connectivity of multilayered random geometric graphs

Josep Díaz*[1], Öznur Yaşar Diner†[2], Maria Serna‡[1], and Oriol Serra§[1]

[1]Universitat Politècnica de Catalunya, Barcelona
[2]Kadir Has University, Istanbul

Abstract

An edge-colored multigraph G is rainbow connected if every pair of vertices is joined by at least one rainbow path, i.e., a path where no two edges are of the same color. In the context of multilayered networks, we introduce the notion of multilayered random geometric graphs, from $h \geq 2$ independent random geometric graphs $G(n, r)$ on the unit square. We define an edge-coloring by coloring the edges according to the copy of $G(n, r)$ they belong to and study the rainbow connectivity of the resulting edge-colored multigraph. We show that $r(n) = \left(\frac{\ln n}{n^{h-1}}\right)^{1/2h}$, is a threshold of the radius for the property of being rainbow connected. This complements the known analogous results for the multilayered graphs defined on the Erdős–Rényi random model.

1 Introduction

Complex networks are used to simulate large-scale real-world systems, which may consist of various interconnected sub-networks or topologies. For instance, this could involve different transportation systems and coordinating schedules between them, modeling interactions across different topologies of the network. Barrat et al. [1] proposed a new network model to represent the emerging large network systems, which include coexisting interacting different topologies. Those network models are known as *layered complex networks, multiplex networks* or as *multilayered networks*. In a multilayered network, each type of interaction of the agents gets its own layer, like a social network having a different layer for each relationship, such as friendship or professional connections [6]. Recently, there's been a lot of interest in adapting tools used in the analysis for single-layer networks to the study of multilayered ones, both in deterministic and random models [2]. In the present work, we explore thresholds for the *rainbow connectivity of the multilayered random geometric graphs*.

A *random geometric graph* (RGG), $G(n, r)$, where $r = r(n)$ on the unit square $I = [0, 1]^2$ is defined as follows: Given n vertices and a radii $r(n) \in [0, \sqrt{2}]$, n vertices are sprinkled independently and uniformly at random (u.a.r.) in the unit square $I = [0, 1]^2$. Two vertices are adjacent if and only if their Euclidean distance is less than or equal to $r(n)$.

Random geometric graphs provide a natural framework for the design and analysis of wireless networks. For further information on random geometric graphs, one may refer to Penrose [10] or to the more recent survey by Walters [12]. Random geometric graphs exhibit a sharp threshold behavior with respect to connectivity [7]: As the value of r increases, there is a critical threshold value r_c such that

*Email: diaz@cs.upc.edu Research of J. D. is supported by the Spanish Agencia Estatal de Investigación [PID-2020-112581GB-C21, MOTION]

†Email: oznur.yasar@khas.edu.tr Research of Ö. Y. is supported by the Spanish Agencia Estatal de Investigación [PID-2020-112581GB-C21, MOTION]

‡Email: mjserna@cs.upc.edu Research of M. S. is supported by the Spanish Agencia Estatal de Investigación [PID-2020-112581GB-C21, MOTION]

§Email: oriol.serra@upc.edu Research of O. S. is supported by the Spanish Agencia Estatal de Investigación [PID2020-113082GB-I00, CONTREWA]

when $r < r_c$, the graph is typically disconnected, while for $r > r_c$, the graph is typically connected. The threshold for connectivity of $G(n,r)$ is $r_c \sim \sqrt{\frac{\ln n}{\pi n}}$. Notice r_c is also a threshold for the disappearance of isolated vertices in $G(n,r)$.

For any random geometric graph, $G(n,r)$, the expected degree $|N_{G(n,r)}(v)|$ is w.h.p.[1] $n\pi r^2, \forall v \in V(G)$. Equivalently the expected degree is concentrated around its mean. Regarding the diameter, $diam(G)$, of a random geometric graph $G(n,r)$, Díaz et al. [5] showed that if $r = \Omega(r_c)$ then $diam(G) = (1+o(1))\frac{\sqrt{2}}{r}$.

We now introduce a general definition for the random model of edge colored multigraphs obtained by the superposition of a collection of random geometric graphs on the same set of vertices. Formally, a *multilayered geometric graph* $G(n,r,h,b)$ is defined by three parameters, n the number of nodes, r the radii of connectivity, and h the number of layers, together with a position assignment $b : [n] \to \underbrace{[0,1]^2 \times \cdots \times [0,1]^2}_{h}$. For $i \in [n]$, we denote $b(i) = (b_1^i, \ldots, b_h^i)$, where $b_k^i \in [0,1]^2$. The multigraph $G(n,r,h,b)$ has vertex set $[n]$ and an edge (i,j) with color k, $1 \le k \le h$, if the Euclidean distance between b_k^i and b_k^j is at most r. Note that, for $k \in [h]$, r and the positions $(b_k^i)_n$, a geometric graph $G_k(n,r)$ is defined by the edges with color k. Thus, $G(n,r,h,b)$ can be seen as the colored union of h geometric graphs, all with the same vertex set and radius. Observe that $G(n,r,h,b)$ is defined on I^{2h}. We refer to $G_k(n,r)$ as the k-th layer of $G(n,r,h,b)$.

A *multilayered random geometric graph* $G(n,r,h)$ is obtained when the position assignment b of the vertices is selected independently, for each vertex and layer, uniformly at random in $[0,1]^2$. Thus, the k-th layer is an RGG. This definition is given for dimension two and it can be extended to points in a multidimensional space by redefining the scope of the position function.

Given an edge–colored graph G, we say G is *rainbow connected* if, between any pair of vertices $u, v \in V(G)$, there is a path with edges of pairwise distinct colors. Chartrand et al [4] introduced the study of the rainbow connectivity of graphs as a strong property to secure strong connectivity in graphs and networks. Since then, variants of rainbow connectivity have been applied to different deterministic models of graphs, see for ex. the survey of Li et al. [8] for further details on the extension of rainbow connectivity to other graph models.

The study of rainbow connectivity has been addressed in the context of multilayered binomial random graphs by Bradshaw and Mohar [3]. The authors give sharp concentration results on three values on the number h of layers needed to ensure rainbow connectivity of the resulting multilayered binomial random graph $G(n,p)$ with appropriate values of p. The results have been extended by Shang [11] to ensure rainbow connectivity k in the same model, namely, the existence of k internally disjoint rainbow paths joining every pair of vertices in the multilayered graph.

In this paper, we are interested in studying the rainbow connectivity of a multilayered random geometric graph $G(n,r,h)$. In particular, for every fixed h, we are interested in the minimum value of r (as a function of n) such that w.h.p. the multilayered random geometric graph $G(n,r,h)$ is rainbow connected. Dually, for fixed values of r we want to determine the minimum number of layers h such that $G(n,r,h)$ is rainbow connected. The latter parameter can be defined as the rainbow connectivity of the multilayered random geometric graph.

Main results: Our main results are lower and upper bounds of the value of r, to asymptotically assure that w.h.p. $G(n,r,h)$, do have or do not have the property of being rainbow connected.

Theorem 1. *Let $h \ge 2$ be an integer and let $G = G(n,h,r)$ be an h–layered random geometric graph. Then, if*

$$r(n) \ge \left(\frac{\ln n}{n^{h-1}}\right)^{1/2h},$$

[1] w.h.p. means with high probability, i.e. with probability tending to 1 as $n \to \infty$.

then w.h.p. G is rainbow connected.

Moreover, there is a constant $0 < c \leq 1$ such that, if

$$r(n) < c\left(\frac{\ln n}{n^{h-1}}\right)^{1/2h},$$

then w.h.p. G is not rainbow connected.

Notice that Theorem 1 can be re-stated as a threshold of h for the rainbow connectivity of multilayered geometric random graph G.

Corollary 2. *Let $r = r(n)$ with $r(n) = o(1)$. Set*

$$h_0 = \left\lceil \frac{\log n + \log\log n}{\log nr^2} \right\rceil.$$

The multilayered random geometric graph $G(n, r, h)$ is w.h.p. rainbow connected if $h \leq h_0$, while if $h > h_0$ it is w.h.p. not rainbow connected.

2 Rainbow Connectivity of Two-layered Random Geometric Graphs

The proof of Theorem 1 requires a special argument for the case $h = 2$. We give below the proof of this case which also illustrates the techniques for general $h > 2$.

Proposition 3. *Let $G(n, r, 2)$ be a two–layered random geometric graph. If*

$$r(n) \geq \left(\frac{\ln n}{n}\right)^{1/4},$$

then G is w.h.p. rainbow connected.

Moreover, there is a positive constant $c > 0$ such that, if

$$r(n) \leq c\left(\frac{\ln n}{n}\right)^{1/4},$$

then w.h.p. G is not rainbow connected.

Proof. Denote by $G_1(n, r)$ and $G_2(n, r)$ the two layers of G, with the value of $r = r(n)$ given in the statement of the proposition. For each pair $v_i, v_j \in V$, let X_{v_i,v_j} denote the indicator random variable

$$X_{v_i,v_j} = \begin{cases} 1 & \text{if there is not a rainbow path between } v_i \text{ and } v_j \text{ in } G\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

Let v_k be different from v_i and v_j. Let A_{v_k} be the event that v_k is joined to v_i in $G_1(n, r)$ and to v_j in $G_2(n, r)$ or vice versa, namely,

$$A_{v_k} = \{\{v_i \in \mathcal{B}_1(v_k)\} \cap \{v_j \in \mathcal{B}_2(v_k)\}\} \cup \{\{v_j \in \mathcal{B}_1(v_k)\} \cap \{v_i \in \mathcal{B}_2(v_k)\}\}\,,$$

where $\mathcal{B}_i(v)$ denotes the set of neighbours of v in G_i, $i = 1, 2$. By taking into account the boundary effects on the unit square, we have $\Pr(v_i \in \mathcal{B}(v_j)) = \pi r^2 + o(r^2)$. We have,

$$(\pi r^2 + o(r^2))^2 \leq \Pr(A_{v_k}) \leq 2(\pi r^2 + o(r^2))^2\,.$$

Let $A_{v_iv_j}$ denote the event that v_i and v_j are joined by an edge either in $G_1(n, r)$ or in $G_2(n, r)$, that is

$$A_{v_i,v_j} = \{v_i \in \mathcal{B}_1(v_j)\} \cup \{v_i \in \mathcal{B}_2(v_j)\},$$

so that

$$\Pr(A_{v_i,v_j}) = 2\pi r^2 + o(r^2).$$

For given v_i and v_j, the event that they are joined by a rainbow path in G is $(\cup_{k\neq i,j} A_{v_k}) \cup A_{v_i,v_j}$. Therefore, since A_{v_k} and A_{v_i,v_j} are independent, for every sufficient large n we have

$$\begin{aligned}\mathbb{E}(X_{v_i,v_j}) &= \Pr(\overline{(\cup_{k\neq i,j} A_{v_k}) \cup (A_{v_i,v_j})}) = \Pr((\cap_{k\neq i,j} \overline{A_{v_k}}) \cap \overline{(A_{v_i,v_j})} \\ &\leq (1-(\pi r^2)^2 + o(r^2))^{n-2} \cdot (1 - 2\pi r^2 + o(r^2)) \\ &\leq (1-(\pi r^2)^2 + o(r^2))^{n}\,.\end{aligned}$$

Let X be a random variable counting the number of pairs $\{v_i, v_j\}$ that are not joined by a rainbow path in G. Then $X = \sum_{i<j} X_{v_i,v_j}$ and, by plugging in the inequality for $r(n)$,

$$\begin{aligned}\mathbb{E}(X) = \sum_{i<j} \mathbb{E}(X_{v_i,v_j}) &\leq \binom{n}{2}(1-(\pi r^2)^2 + o(r^2))^n \\ &\leq e^{2\log n}\left(1 - \pi^2 \frac{\log n}{n} + o\left(\frac{\log n}{n}\right)\right)^n \leq e^{(2-\pi^2)\log n + o(\log n)}\end{aligned}$$

By Markov's inequality, it follows that $\Pr(X \geq 1) \leq \mathbb{E}(X) \to 0$, as $n \to \infty$. It follows that w.h.p. G is rainbow connected, which proves the first part of the statement.

For the second part, let $r(n) \leq c(\log n/n)^{1/4}$ for some positive small constant c to be specified later. By using the upper bounds on the probabilities of the events A_{v_k} and A_{v_i,v_j},

$$\mathbb{E}(X_{v_i,v_j}) \geq (1 - 2(\pi r^2 + o(r^2))^2)^{n-2}(1 - 2\pi r^2 + o(r^2)) \geq \left(1 - 2c^4\pi^2 \frac{\ln n}{n}\right)^{n-1}.$$

$$\mathbb{E}(X_{v_i,v_j}) \geq (1 - 2(\pi r^2 + o(r^2))^2)^{n-2} \geq \left(1 - 2c^4\pi^2 \frac{\ln n}{n} + o\left(\frac{\log n}{n}\right)\right)^{n-2}$$

Let $X_{v_i} = \sum_{j\neq i} X_{v_i,v_j}$ denote the number of vertices v_j not joined with v_i by a rainbow path in G. We have, with $c' = 2c^4\pi^2$,

$$\mathbb{E}(X_{v_i}) \geq (n-2)\left(1 - c'\frac{\ln(n-1)}{n-1} + o\left(\frac{\log n}{n}\right)\right)^{n-2} \sim e^{(1-c')\ln n} = n^{1-c'}.$$

By choosing $c < (2/\pi^2)^{1/4}$ we have $c' < 1$, so that $\mathbb{E}(X_i) \to \infty$ with $n \to \infty$. Since X_{v_i} is a sum of independent random variables, by Chernoff inequality we have $\Pr(X_{v_i} = 0) \leq e^{-n^{1-c''}/2}$ for each $1 > c'' > c'$. It follows that G is w.h.p. not rainbow connected. □

3 Proof of Theorem 1

The proof of Theorem 1, for $h > 2$, is sketched below.

A key property of multilayered random geometric graphs is their local expanding properties.

Lemma 4. *Let $h > 2$ be fixed and let $G = G(n, r, h)$ be a multilayered random geometric graph. Let $u \in V(G)$ a fixed vertex and denote by $N_j(u)$ the set of vertices reached from u by rainbow paths of length j starting at u, the i–th edge along the path colored i. Let $M = nr^2$. Then, for $1 \leq j \leq h-1$ we have that w.h.p.*

$$|N_j(u)| = \Theta(M^j).$$

The proof of Lemma 4 uses the fact that the probability that the size of the image of a random map $g : [m] \to [k]$ deviates from m more than a constant $a > 0$ is at most $2\exp(-2(a - m^2/2k)^2/m)$. This fact in turn follows by a direct application of the McDiarmid concentration inequalities [9].

Lemma 4 provides the existence of rainbow paths from a given vertex to all vertices in the graph.

Proposition 5. *Let $h > 2$ be fixed and let $G = G(n, h, r)$ be an h–multilayered random geometric graph. Let $u \in V(G)$. If*

$$r \geq \left(\frac{\ln n}{n^{h-1}}\right)^{1/2h},$$

then w.h.p. there is a rainbow path from u to every other vertex in G.

Proof. Let us consider first the case that $h \geq 3$ is odd, i.e., $h = 2k + 1$, for some $k > 1$. Denote by $G_i = G_i(n, r)$ the i-thlayer of G. For $I \subseteq [h]$, we denote by $G_I(n, r)$ the layered graph formed by the layers included in I. For a pair i, j of distinct vertices in $V(G)$ and a permutation σ of $\{1, 2, 3, \ldots, h\}$, let $P(i, j; \sigma)$ denote the set of rainbow paths of length h joining i and j with the first edge in $G_{\sigma(1)}$ and the last one in $G_{\sigma(h)}$. For a permutation σ, let $I_1(\sigma) = \{\sigma(1), \ldots, \sigma(k)\}$

Let $A = N_{k,\sigma}(i)$ be the set of vertices reached from i by rainbow paths of length k starting at j following the color order determined by σ. Let $B = N_{k,\sigma}(j)$ be the set of vertices reached from j by rainbow paths of length k starting at j following the color order determined by following σ in reversed order with the k-th edge along the path colored $k+2$. From Lemma 4, $|A|, |B| = \Theta((nr^2)^k) = \Theta(n^k r^{2k})$

Let $X_{i,j}$ denote the number of rainbow paths of length h joining i and j with the first edge in $G_{\sigma(1)}$, the second edge in $G_{\sigma(2)}$ and so on. For a pair $(k, k') \in A \times B$ with $k \neq k'$, let $Y_{k,k'}$ be the indicator function that k and k' are neighbours in $G_{\sigma(k+1)}$. We have $\mathbb{E}(Y_{k,k'}) = \pi r^2$, the probability that the vertices k' and k are adjacent in $G_{\sigma(k+1)}$. Then,

$$X_{ij} = \sum_{k,k'} Y_{k,k'},$$

where the sum runs through all pairs $(k, k') \in A \times B$ with $k \neq k'$. We observe that the variables $Y_{k,k'}$ are independent. When the pairs $(k, k'), (l, l')$ are disjoint it is clear that $Y_{k,k'}, Y_{l,l'}$ are independent. When $k = l$, say, then $\Pr(Y_{k,k'} = 1, Y_{k,l'} = 1)$ is the probability that k' and l' are both adjacent to k, which is the product $\Pr(Y_{k,k'} = 1) \Pr(Y_{k,l'} = 1)$.

Let us fix $r(n) \geq \left(\frac{\ln n}{n^{h-1}}\right)^{1/2h}$. Note that $N_{h-1}(u) \ll n$, so each $(h-1)$–layered subgrah of G is not w.h.p. rainbow connected. Then it follows that w.h.p. the sets A and B, for $i \neq j$ not connected by a rainbow path of length $h - 1$ are disjoint. In this case, the events $Y_{k,k'}$ are independent, therefore

$$\begin{aligned}\Pr(X_{i,j} = 0) = \Pr(\cap_{k,k'}\{Y_{k,k'} = 0\}) &= \prod_{k,k'} \Pr(Y_{k,k'} = 0) \\ &= (1 - \pi r^2)^{(n^k r^{2k})^2} \leq e^{-\pi n^{2k} r^{4k+2}}.\end{aligned}$$

By using the union bound on all pairs i, j and the lower bound on r,

$$\Pr(\cap_{i,j}\{X_{ij} \geq 1\}) = 1 - \Pr(\cup_{i,j} X_{i,j} = 0) \geq 1 - n^2 e^{\pi n^{2k} r^{4k+2}},$$

As $k = (h-1)/2$, by the lower bound on r,

$$n^{2k} r^{4k+2} = n^{h-1} r^{2h} \geq (\log n),$$

Therefore, the last term in the bound on $\Pr(\cap_{i,j}\{X_{ij} \geq 1\})$ is $o(1)$ as $n \to \infty$. Hence w.h.p. all pairs i, j are connected by a rainbow path of length h.

For even h, the result is obtained by an extension of the argument used for $h = 2$ in Proposition 3. □

For the lower bound on $r(n)$, an application of the second moment method as the one given in Proposition 3 for the case $h = 2$ can be extended to $h > 2$.

4 Conclusions

The main purpose of this paper is to identify the threshold for the radius to get a rainbow-connected multilayered random geometric graph, as obtained in Theorem 1. As mentioned in the Introduction, the analogous problem of determining the threshold for h so that the multilayered binomial random graph is rainbow connected was addressed by Bradshaw and Mohar [3].

We believe that the model of multilayered random geometric graphs is very appealing and leads to a host of interesting problems. One may think of a dynamic setting where n individuals perform random walks within the cube and communicate with the close neighbors at discrete times $t_1 < t_2 < \cdots < t_h$. The rainbow connectivity in this setting measures the number of instants needed so that every individual can communicate with each of the other ones. A natural immediate extension is to address the threshold to get rainbow connectivity k, as achieved in the case of multilayered binomial random graphs by Shang [11].

There is a vast literature addressing rainbow problems in random graph models, and this paper is meant to open the path to these problems in the context of multilayered random geometric graphs. It would also be interesting to find asymptotic estimates on r such that h copies produce a rainbow clique of size $\sqrt{h}$.

We observe that, for large h, the threshold of r for rainbow connectivity approaches the connectivity threshold of random geometric graphs. The arguments in the proof, however, apply only for constant h. For h growing with n, the correlation between distinct edges in our model decreases and the model gets closer to the random binomial graph, where the results are expected to behave differently and the geometric aspects of the model become irrelevant.

References

[1] A. Barrat, M. Barthélemy, R. Pastor-Satorras and A. Vespignani, The architecture of complex weighted networks, *Proc. National Academy of Sciences* **101-11** (2004), 3747–3752.

[2] G. Bianconi, *Multilayer Networks: Structure and Function*. Oxford University Press, Oxford, 2018.

[3] P. Bradshaw, B. Mohar, A Rainbow Connectivity Threshold for Random Graph Families, in: *Nesetril, J., Perarnau, G., Rué, J., Serra, O. (eds) Extended Abstracts EuroComb 2021. Trends in Mathematics*, vol 14. Birkhäuser, Cham, 2021, 842–847.

[4] G. Chartraud, G. L. Johns, K. A. Mckeon and P. Zang The rainbow connectivity of a graph. *Networks* **54-2** (2009), 75–81.

[5] J. Díaz, D. Mitsche, G. Perarnau, and X. Pérez, On the relation between graph distance and Euclidean distance in random geometric graphs, *Advances in Applied Probability* **48-3** (2016) 848-864.

[6] M. E. Dickison, M. Magnani M, L. Rossi, *Multilayer Social Networks*. Cambridge University Press, 2016.

[7] A. Goel and S. Rai and B. Krishnamachari, Sharp thresholds for monotone properties in random geometric graphs, *Annals of Applied Probability* **15** (2005), 364–370.

[8] X. Li, Y. Shi, and Y. Sun, Rainbow connections of graphs: a survey *Graphs Comb.* **29** (2013), 1–38.

[9] M.D. McDiarmid. On the methods of bounded differences, in: *Surveys in Combinatorics*, Cambridge University Press, 1989, 148–188.

[10] M. Penrose, *Random Geometric Graphs*, Oxford Studies in Probability. Oxford U.P., 2003.

[11] Y. Shang, Concentration of rainbow k-connectivity of a multiplex random graph, *Theoretical Computer Science* **951** (2023), 113771.

[12] M. Walters, Random Geometric Graphs, in: *Surveys in Combinatorics*, Cambridge U.P., 2011, 365—401.

[13] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*, Cambridge U.P., 1994.

Polytope neural networks*

Juan L. Valerdi†

1 Introduction

A major challenge in the theory of neural networks is to precisely characterize the functions they can represent [2]. This topic differs from universal approximation theorems [1], which aim to guarantee the existence of neural networks that approximate functions well. Although it is well known that feedforward neural networks with ReLU activation are continuous piecewise linear (CPWL) functions [2, 3], the minimum number of layers required to represent any CPWL function remains an open question.

A potential way to solve this problem is through the concept of depth of a polytope given by neural networks.

Definition 1. *The collection of polytope neural networks with depth m is defined as*

$$\Delta(m)=\left\{\sum_{i=1}^{p} conv\{P_i,Q_i\} \;\middle|\; P_i,Q_i\in\Delta(m-1)\right\},$$

where the sum corresponds to Minkowski sum and $conv\{P_i,Q_i\}$ means the convex hull of $P_i\cup Q_i$. The base set $\Delta(0)$ represents the polytopes consisting of a single point.

Definition 2. *A polytope P is said to have (minimal) depth m, denoted as $d(P)=m$, if $P\in\Delta(m)$ and $P\notin\Delta(m-1)$.*

Neural networks are traditionally named after their building object or operation. For example, ReLU neural networks use ReLU activation, and convolutional neural networks [3] are based on convolution kernels. In a similar manner, the naming of polytope neural networks is derived from their underlying object.

The connection between ReLU and polytope neural networks can be found through tropical geometry [10]. Any ReLU network can be decomposed into the difference of two convex CPWL functions, which can be mapped to polytopes via Newton polytopes. In particular, understanding the functions representable by ReLU neural networks of a given depth is equivalent to studying which polytopes can be constructed at that depth, as defined in Definition 1.

The open question for ReLU networks reduces to whether the function $\max\{x_1,x_2,\dots,x_n,0\}$ can be represented with minimal depth $\lceil\log_2(n+1)\rceil$. This question can be rephrased in the language of polytopes as follows.

Conjecture 3 (Hertrich et al. [6])**.** *Let S be an n-simplex, then $d(S+P)=\lceil\log_2(n+1)\rceil$, for any polytope P with $d(P)<\lceil\log_2(n+1)\rceil$.*

*The full version of this work can be found in [8] and will be published elsewhere.

†Email: j.valerdi11@gmail.com

Our understanding of the sets $\Delta(m)$, beyond the case of $m = 1$, which corresponds to the set of zonotopes, remains limited. The conjecture is known to be true for $n = 2$ and $n = 3$ [2, 7]. However, to this date, the only contribution addressing Conjecture 3 for any n has been made by Haase et al. [5], who have proven it for lattice polytopes. Their approach involved relating depth with subdivision and volume properties of Minkowski sums and convex hulls.

The goal of this work is to advance our knowledge of polytope neural networks relevant to Conjecture 3. We show basic depth properties from Minkowski sums, convex hulls, number of vertices, faces, affine transformations, and indecomposable polytopes. More significantly, key findings include depth characterization of polygons; identification of polytopes with an increasing number of vertices, exhibiting small depth and others with arbitrary large depth; and most importantly, depth computation for simplices.

Acknowledgements. I extend my gratitude to Ansgar Freyer for providing the proof of Theorem 12 for $n = 4$, which was expanded to the general case with minor adjustments. I also thank Francisco Santos for his hospitality during my visits to the University of Cantabria, and for valuable discussions on this work, including presentation enhancements and the ideation and proof of Theorem 14.

2 Basic properties

To develop the main results in Section 3, it is necessary to establish some basic depth properties for polytopes. We assume $\mathbb{R}^n$ as the ambient space throughout.

We begin by computing depth bounds for Minkowski sums and convex hulls, which are the fundamental operations in Definition 1.

Proposition 4. *Let P_1, P_2 be polytopes with $d(P_i) \leq m_i$. Then, $d(P_1 + P_2) \leq \max\{m_1, m_2\}$ and $d(\mathit{conv}\{P_1, P_2\}) \leq \max\{m_1, m_2\} + 1$.*

Proof. If $d(P_i) \leq m_i$, then $P_i \in \Delta(\max\{m_1, m_2\})$. This implies $\operatorname{conv}\{P_1, P_2\} \in \Delta(\max\{m_1, m_2\} + 1)$ by definition, and therefore $d(\operatorname{conv}\{P_1, P_2\}) \leq \max\{m_1, m_2\} + 1$.

Also by definition, consider the decomposition

$$P_i = \sum_{j=1}^{q_i} \operatorname{conv}\{Q_{j,i}, R_{j,i}\},$$

where $Q_{j,i}, R_{j,i} \in \Delta(\max\{m_1, m_2\} - 1)$ for all $i = 1, 2$ and $j = 1, \ldots, q_i$. Consequently, $d(P_1 + P_2) \leq \max\{m_1, m_2\}$ as

$$P_1 + P_2 = \sum_{j=1}^{q_1} \operatorname{conv}\{Q_{j,1}, R_{j,1}\} + \sum_{j=1}^{q_2} \operatorname{conv}\{Q_{j,2}, R_{j,2}\} \in \Delta(\max\{m_1, m_2\}). \qquad \square$$

Now, using Proposition 4 we can bound the depth of a polytope by its vertices.

Proposition 5. *If a polytope P is given by its vertices $P = \mathit{conv}\{x_1, \ldots, x_p\}$, then $d(P) \leq \lceil \log_2 p \rceil$.*

Proof. By definition, $d(\{x_1\}) = 0$ and $d(\operatorname{conv}\{x_1, x_2\}) = 1$. Supposing the statement is true up to $p - 1$, consider a polytope $P = \operatorname{conv}\{x_1, \ldots, x_p\}$ and decompose it as

$$P = \operatorname{conv}\{\operatorname{conv}\{x_1, \ldots, x_k\}, \operatorname{conv}\{x_{k+1}, \ldots, x_p\}\},$$

where k is the largest integer power of 2 such that $k < p$. Using the induction hypothesis, we obtain that $d(\operatorname{conv}\{x_1, \ldots, x_k\}) \leq \log_2 k$ and $d(\operatorname{conv}\{x_{k+1}, \ldots, x_p\}) \leq \lceil \log_2(p-k) \rceil$. Therefore, by Proposition 4, we conclude $d(P) \leq \log_2 k + 1 = \lceil \log_2 p \rceil$. $\square$

Other basic properties concerns the depth of a polytope in relation to its faces and affine transformations.

Proposition 6. *Any face $F \neq \emptyset$ of a polytope P satisfies $d(F) \leq d(P)$.*

Proof. For $d(P) = 0$, there is nothing to prove. If $d(P) = 1$, then P is a zonotope, and any face F is also a zonotope; therefore, $d(F) \leq 1$. For the sake of induction, suppose the statement is true up to depth $m-1$ and consider $d(P) = m$. By definition,

$$P = \sum_{i=1}^{q} \text{conv}\{P_i, Q_i\}, \quad P_i, Q_i \in \Delta(m-1).$$

A face F of P is then expressed as

$$F = \sum_{i=1}^{q} \text{conv}\{F_i, G_i\},$$

where F_i, G_i are faces of P_i, Q_i respectively. By the induction hypothesis, $d(F_i) \leq m-1$ and $d(G_i) \leq m-1$, and consequently $F_i, G_i \in \Delta(m-1)$ for all i. Therefore, $F \in \Delta(m)$ and $d(F) \leq d(P)$. □

Proposition 7. *Let P be a polytope in $\mathbb{R}^n$ and $\varphi : \mathbb{R}^n \to A$ be an affine transformation, where A is an affine subspace of $\mathbb{R}^d$. Then, $d(\varphi(P)) \leq d(P)$, with equality holding if φ is invertible.*

Proof. Let $\varphi(x) = Mx + c$, where $M \in \mathbb{R}^{d\times n}$ and $c \in \mathbb{R}^d$. For the case $d(P) = 0$ consider $P = \{a\}$, then $\varphi(P) = \{Ma + c\}$, which implies $d(\varphi(P)) = 0$. For the purpose of induction, assume that the statement, in the general case, is true up to $m-1$. Let $d(P) = m$ and express it as

$$P = \sum_{i=1}^{p} \text{conv}\{P_i, Q_i\}, \quad P_i, Q_i \in \Delta(m-1).$$

Then,

$$\varphi(P) = \varphi\left(\sum_{i=1}^{p} \text{conv}\{P_i, Q_i\}\right) = M\sum_{i=1}^{p} \text{conv}\{P_i, Q_i\} + c = \sum_{i=1}^{p} \text{conv}\{MP_i, MQ_i\} + \{c\}.$$

Utilizing the induction hypothesis and Proposition 4, we deduce that $d(\varphi(P)) \leq m$. In the case of φ being invertible, we get

$$d(P) = d(\varphi^{-1}(\varphi(P))) \leq d(\varphi(P)) \leq d(P). \qquad \square$$

A class of polytopes in which computing their depth may be easier is that of indecomposable polytopes. Two polytopes, P and Q, are said to be *positively homothetic*, if $P = \lambda Q + w$ for some $\lambda > 0$ and $w \in \mathbb{R}^n$. A polytope P is said to be *indecomposable* if any decomposition $P = \sum_{i=1}^{k} P_i$ is only possible when P_i is positively homothetic to P for all $i = 1, \ldots, k$.

Proposition 8. *If P is an indecomposable polytope, then there exist polytopes P_1, P_2 such that $P = conv\{P_1, P_2\}$ and $d(P) = \max\{d(P_1), d(P_2)\} + 1$.*

Proof. By definition, there exist $P_i, Q_i \in \Delta(d(P)-1), i = 1, \ldots, k$, such that

$$P = \sum_{i=1}^{k} \text{conv}\{P_i, Q_i\},$$

where an index j necessarily satisfies $\max\{d(P_j), d(Q_j)\} + 1 = d(P)$. As P is indecomposable, there exist $\lambda_j > 0$ and $w_j \in \mathbb{R}^n$ such that $P = \lambda_j \text{conv}\{P_j, Q_j\} + w_j = \text{conv}\{\lambda_j P_j + w_j, \lambda_j Q_j + w_j\}$, and by Proposition 7,

$$d(P) = \max\{d(P_j), d(Q_j)\} + 1 = \max\{d(\lambda_j P_j + w_j), d(\lambda_j Q_j + w_j)\} + 1. \qquad \square$$

3 Main results

We first present a full depth characterization for polygons.

Theorem 9. *Any polygon P satisfies $d(P) \leq 2$.*

Proof. Let P be a polygon. If P is a zonotope, then $d(P) = 1$; whereas, if P is a triangle, then $d(P) = 2$ due to Proposition 5 and the fact that P is not a zonotope. Suppose that P is neither a zonotope nor a triangle; then, it can be decomposed as $P = \sum_{i=1}^{k} P_i$, where P_i is a zonotope or a triangle for all $i = 1, \ldots, k$ [4]. Therefore, $d(P) \leq 2$ by Proposition 4. □

From Theorem 9, we deduce that a polygon can have depth 0 if it consists of a single point, depth 1 if it is a zonotope, or depth 2 otherwise.

We continue with zonotopes and (bi)pyramids, as example of polytopes which can have large number of vertices and small depth.

Proposition 10. *Any n-(bi)pyramid, $n \geq 3$, with a zonotope base has depth 2.*

Proof. A 3-(bi)pyramid P includes triangular facets, therefore it is not a zonotope, and thus $d(P) \geq 2$. Assuming that up to $n-1$, (bi)pyramids has depth greater than or equal to 2, let's consider a facet F of an n-(bi)pyramid P containing an/the apex. Since F is a pyramid of dimension $n-1$, then $d(F) \geq 2$ based on the induction hypothesis. Consequently, $d(P) \geq d(F) \geq 2$ by Proposition 6.

Now, consider P an arbitrary n-(bi)pyramid with a zonotope base Z and apex (or apices) A. Then, $2 \leq d(P) = d(\text{conv}\{Z, \text{conv } A\}) \leq 2$ according to Proposition 4. □

Theorem 11. *Let $v_p = 2\sum_{i=0}^{n-1} \binom{p-1}{i}$ for $p \geq n$. For each p satisfying this condition, there exist polytopes with v_p vertices and depth 1, and also with $v_p + 1$ vertices and depth 2.*

Proof. Let $g_i = [\mathbf{0}, b_i]$, where $i = 1, \ldots, p$, represent line segments with $b_1, \ldots, b_p$ denoting points in $\mathbb{R}^n$ in general position. The zonotope $Z = \sum_{i=1}^{p} g_i$ has depth 1 and has v_p vertices given the generators are in general position [9]. Lifting Z to $\mathbb{R}^{n+1}$ by adding 0 to the new coordinate allows the construction of a pyramid P with Z as its base. Therefore, $d(P) = 2$ by Proposition 10. □

In Theorem 11, we constructed two families of polytopes, zonotopes and pyramids, which exhibit an increasing number of vertices and possess depths of 1 and 2, respectively. This indicates that depth bounds from Proposition 5 may be far from the true depth of a polytope. However, this bound based on vertices cannot be further refined, as it is tight for simplices.

We next present two approaches for calculating the depth of simplices. The first approach leverages the face structure and indecomposability of simplices, while the second approach results from a more general finding regarding polytopes containing complete subgraphs.

Theorem 12. *Any n-simplex has minimal depth $\lceil \log_2(n+1) \rceil$.*

Proof. We know that 2-simplices have depth 2. Let's make the assumption that, for $k = 3, \ldots, n-1$, k-simplices have depth $\lceil \log_2(k+1) \rceil$ and consider an n-simplex P. Given that P is indecomposable [4], we can employ Proposition 8 to get a pair of polytopes P_1, P_2 such that $P = \text{conv}\{P_1, P_2\}$ and $\max\{d(P_1), d(P_2)\} = d(P) - 1$.

Without loss of generality, one of the P_i, let's say P_1, contains at least $q = \lceil \frac{n+1}{2} \rceil$ points that are vertices of P. Consider $F = \text{conv}\{x_1, \ldots, x_q\}$, where $x_i, i = 1, \ldots, q$ are vertices of P contained in P_1. Then, F is a $(q-1)$-simplex and a face of P. Let H be a supporting hyperplane of P associated with F. From

$$F = H \cap F \subset H \cap P_1 \subset H \cap P = F,$$

we deduce that F is also a face of P_1. By the induction hypothesis,

$$d(F) = \left\lceil \log_2 \left\lceil \frac{n+1}{2} \right\rceil \right\rceil = \lceil \log_2(n+1) \rceil - 1$$

Referring to Proposition 5, Proposition 6, and Proposition 8, we derive that

$$\lceil \log_2(n+1) \rceil - 1 \leq d(P_1) \leq \max\{d(P_1), d(P_2)\} = d(P) - 1 \leq \lceil \log_2(n+1) \rceil - 1,$$

thus concluding that $d(P) = \lceil \log_2(n+1) \rceil$. □

For the second approach we will compute the depth of 2-neighbourly polytopes, for which we need the following result.

Lemma 13. *If the graph of a polytope $G(P)$ contains a complete subgraph with $p \geq 3$ vertices, and P can be decomposed as $P = \sum_{i=1}^{k} P_i$, then at least one of $G(P_j)$ also contains a complete subgraph with p vertices.*

Proof. Consider that u, v, w are vertices of P in the complete subgraph of $G(P)$ with $p \geq 3$ vertices. Given that any vertex of P can be uniquely represented as the sum of vertices of $P_i, i = 1, \ldots, k$, let u_i, v_i, w_i be those vertices for P_i that represent u, v, w respectively. Therefore, we can express the edges $[u, v], [u, w], [v, w]$ as

$$[u, v] = \sum_{i=1}^{k} [u_i, v_i], \quad [u, w] = \sum_{i=1}^{k} [u_i, w_i], \quad [v, w] = \sum_{i=1}^{k} [v_i, w_i].$$

The edges $[u_i, v_i], [u_i, w_i], [v_i, w_i]$ are parallel to $[u, v], [u, w], [v, w]$ respectively, and because u, v, w form a triangle in $G(P)$, it follows that their ratios of edge lengths satisfies

$$\frac{|u_i - v_i|}{|u - v|} = \frac{|u_i - w_i|}{|u - w|} = \frac{|v_i - w_i|}{|v - w|}.$$

This implies there exists an index j for which these ratios are nonzero, implying that vertices u_j, v_j, w_j form a triangle in $G(P_j)$. Extending this reasoning to any other vertex z in the complete subgraph, by applying the same logic with vertices u, v, z, it is deduced that u_j, v_j, z_j also form a triangle in $G(P_j)$, and this pattern continues with other vertices. □

Theorem 14. *If the graph of a polytope $G(P)$ contains a complete subgraph with $p \geq 3$ vertices, then $d(P) \geq \lceil \log_2 p \rceil$.*

Proof. Suppose a subgraph of $G(P)$ is complete and contains $p = 3$ or $p = 4$ vertices. If we assume $d(P) = 1$, then $P = \sum_{i=1}^{k} P_i$, where each P_i is a segment. This contradicts Lemma 13, which implies that at least one P_i must include p vertices. Therefore, we conclude $d(P) \geq 2$.

For the sake of induction, let's assume that the result holds for all cases up to $p - 1$. Now, consider that $G(P)$ includes a complete subgraph consisting of p vertices. By definition, we can express P as

$$P = \sum_{i=1}^{k} \operatorname{conv}\{P_i, Q_i\}, \quad \text{where } d(P_i), d(Q_i) \leq d(P) - 1.$$

According to Lemma 13, there exists an index j for which $G(\operatorname{conv}\{P_j, Q_j\})$ also contains a complete subgraph K with p vertices. Without loss of generality, we can assume that P_j contains at least $\lceil \frac{p}{2} \rceil$ vertices of K, and consequently the complete subgraph induced by those vertices. Using the induction hypothesis we obtain

$$d(P) - 1 \geq d(P_j) \geq \left\lceil \log_2 \left\lceil \frac{p}{2} \right\rceil \right\rceil = \lceil \log_2 p \rceil - 1,$$

from which it follows $d(P) \geq \lceil \log_2 p \rceil$. □

Corollary 15. *Any 2-neighbourly polytope P with p vertices satisfies $d(P) = \lceil \log_2 p \rceil$.*

Proof. It is a direct consequence of Theorem 14 and Proposition 5. □

Corollary 16. *Any n-simplex has depth $\lceil \log_2(n+1) \rceil$.*

Another important consequence of Theorem 14 is that allows to find a family of polytopes with the same dimension and increasingly large depth.

Corollary 17. *For every $p > n \geq 4$ the cyclic n-polytope with p vertices has depth $\lceil \log_2(p+1) \rceil$.*

4 Concluding remarks

Knowing that n-simplices has depth $\lceil \log_2(n+1) \rceil$ reveals one part of Conjecture 3, and together with Proposition 4, we have obtained an upper depth bound for the conjecture. However, a tight lower bound is still needed to prove it.

In ReLU neural networks, from which Conjecture 3 originated, it has been proven that, for CPWL functions f and g, if their depth satisfy $d(f) < d(g)$, then $d(f+g) = d(g)$ [8]. If this result also holds true in polytope neural networks, it could solve the conjecture. However, the existing proof for CPWL functions is inapplicable to polytopes, as it requires the inverse for the sum operation.

Another interesting contrast between polytope and ReLU networks is found in Corollary 17, where cyclic n-polytopes, for $n \geq 4$, have arbitrary large depth. Instead, for a fixed domain $\mathbb{R}^n$, all CPWL functions can be computed by ReLU neural networks with a depth of $\lceil \log_2(n+1) \rceil$. For polytopes, this contrast is also seen with Theorem 9, where polygons are shown to have a maximum depth of 2.

References

[1] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems*, 2(4):303–314, 1989.

[2] Ronald DeVore, Boris Hanin, and Guergana Petrova. Neural network approximation. *Acta Numerica*, 30:327–444, 2021.

[3] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[4] Branko Grünbaum. *Convex Polytopes*. Springer, 2003.

[5] Christian Alexander Haase, Christoph Hertrich, and Georg Loho. Lower bounds on the depth of integral relu neural networks via lattice polytopes. In *International Conference on Learning Representations*, 2023.

[6] Christoph Hertrich, Amitabh Basu, Marco Di Summa, and Martin Skutella. Towards lower bounds on the depth of relu neural networks. *SIAM Journal on Discrete Mathematics*, 37(2):997–1029, 2023.

[7] Anirbit Mukherjee and Amitabh Basu. Lower bounds over boolean inputs for deep neural networks with relu gates. *arXiv preprint arXiv:1711.03073*, 2017.

[8] Juan L. Valerdi. On minimal depth in neural networks. *arXiv preprint arXiv:2402.15315*, 2024.

[9] Thomas Zaslavsky. *Facing up to arrangements: Face-count formulas for partitions of space by hyperplanes: Face-count formulas for partitions of space by hyperplanes*, volume 154. American Mathematical Soc., 1975.

[10] Liwen Zhang, Gregory Naitzat, and Lek-Heng Lim. Tropical geometry of deep neural networks. In *International Conference on Machine Learning*, pages 5824–5832. PMLR, 2018.

Expressing the coefficients of the chromatic polynomial in terms of induced subgraphs: a systematic approach*

Kerri Morgan[†1] and Lluís Vena[‡2]

[1]School of Science (Mathematical Sciences), RMIT University
[2]Department of Mathematics, Universitat Politècnica de Catalunya - BarcelonaTech (UPC)

Abstract

We follow works of Whitney, Farrell, and Morgan and Delbourgo, to express the coefficients of the chromatic polynomial $P(G;\lambda)$ of a graph G in the variable λ in terms of the number of (induced) subgraphs of G: the coefficient of $\lambda^{|G|-p}$ is given as a polynomial on variables $\binom{x_i}{k}$ with integer coefficients, and where the x_i are the number of induced copies of a 2–connected graphs with $\leq p+1$ vertices that are not formed by gluing two 2–connected graphs through a common clique. Our main contribution is that the finding of these expressions can be systematised, and that they do not depend on the 2–connected graphs with $\leq p+1$ vertices that are formed by gluing two 2–connected graphs through a common clique. As an application, we give an alternative proof of the chromatic uniqueness of the wheels with an odd number of vertices.

1 Introduction

The chromatic polynomial of a graph G, $P(G;\lambda)$, gives, as its evaluations on the positive integers n, the number of proper colourings of a graph using n colours. In particular, the chromatic polynomial has $0, 1, \ldots, \chi(G)-1$ as roots. In general, it can be defined as the polynomial that is λ on a graph on a single vertex, 0 if the graph has any loops, multiplicative over connected components, and such that $P(G;\lambda) = P(G-e;\lambda) - P(G/e;\lambda)$ when e is a non-loop edge. In general, using [6], the chromatic polynomial can be given as:

$$P(G;\lambda) = \sum_{A\subseteq E(G)} (-1)^{|A|}\lambda^{k(A)} \tag{1}$$

where $k(A)$ is the number of components of the graph $G = (V(G), A)$. In particular, the coefficient of $\lambda^{|V(G)|-p}$ is given, up to a sign, by the number of subsets of edges spanning a subgraph of G with $|V(G)|-p$ components. Whitney [6] gave the following expression for chromatic polynomial of a graph G of order n as $P(G;\lambda) = \sum_{i,j}(-1)^{i+j}m_{ij}\lambda^{n-i}$ where m_{ij} is the number of 2-connected subgraphs of G of rank i and nullity j. He [7] showed that this could be expressed as $P(G;\lambda) = \sum_i m_i\lambda^{n-i}$ where $m_i = \sum_j(-1)^{i+j}m_{ij}$ and $(-1)^i m_i$ is the number of subgraphs of G with i edges and containing no broken circuits. Building on this work, Farrell [4] showed that the coefficients of the chromatic polynomial could be expressed as

$$P(G;\lambda) = \sum_i c_{n-i}\lambda^{n-i} \tag{2}$$

where the c_{n-i} is an expression in the counts of 2-connected induced subgraphs of G (we count subgraphs and induced subgraphs in terms of edge sets of G, see (3)), however, the arguments for the general case would be quite involved.

*This work started in the MATRIX Workshop on Uniqueness and Discernment in Graph Polynomials. MATRIX is Australia's international research institute for the mathematical sciences.

[†]Email: kerri.morgan@rmit.edu.au

[‡]Email: lluis.vena@upc.edu. Supported by the Grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

In the main result of this work, Theorem 2, we give a more precise description of how c_{n-i} can be written as an expression in the counts of 2-connected induced subgraphs, and also allows to easily implement an algorithm that finds such expression: the complexity of the algorithm for $n-i$ depends on the cube of the number of connected graphs on $i+1$ vertices (see Section 3 below).

Given A, a set of edges of G, the graph $(V(A), A)$ is the graph with A as its set of edges, and where $V(A) = \{v \in V(G) \mid \exists e \in A, v \text{ adjacent to } e\}$ is its set of edges. Given a graph H, we let

$$\mathrm{sube}(H,G) = \sum_{A \subseteq E(G)} \mathbf{1}_{(V(A),A) \text{ isomorphic to } H} \quad \Big\| \quad \mathrm{inde}(H,G) = \sum_{A \subseteq E(G)} \mathbf{1}_{G \text{ restricted to } V(A) \text{ isomorphic to } H} \tag{3}$$

be, respectively, the *number of subgraphs* isomorphic to H in G and the *number of induced subgraphs* isomorphic to H in G.[1] Regarding the previous work on some specific coefficients, the following are found in [4]:

Theorem 1. *[4, Thm 1, 2] The coefficients c_{n-3}, c_{n-4} from (2) in $P(G;\lambda)$ equals,[2], respectively*

$$-\binom{m}{3} + (m-2)t + C_4 - 2K_4 := -\binom{inde(K_2,G)}{3} + (inde(K_2,G)-2)inde(K_3,G) + inde(C_4,G) - 2inde(K_4,G) \ ,$$

$$\binom{m}{4} - \binom{m-2}{2}t + \binom{t}{2} - (m-3)\cdot C_4 + (2m-9)\cdot K_4$$
$$- 6\cdot inde(K_5,G) - inde(C_5,G) + inde(\theta_{2,2,2},G) + 3inde(W_5,G) + 2inde(W_5 \setminus \{spoke\},G) \ .$$

At this point it is worth observing that Whitney's [6] main interest was to give a general account on the expressions that appear in the general coefficient of λ^{n-i} in terms of 2-connected subgraphs, while in [3, 4] the primary focus was to give an expression in terms of induced subgraphs and with the minimum number of terms as possible; the price to pay was that only the first terms could be computed (with reasonable effort) exactly. In the proof of Theorem 2, we follow the arguments of both [6, 3, 4] with the aim of giving a general account of the coefficients (in the style of [6]), but in terms of induced subgraphs (as in [3, 4]).

Remark. The fact that coefficients c_{n-i}, and, more generally, the whole chromatic polynomial of G, depends on the counts of its finite subgraphs has been extensively used in the literature, see, for instance [1, 2, 5].

2 Our result

Let $\mathcal{B}$ denote the set of 2–connected graphs. Consider the multiset of elements of $\mathcal{B}$, $\mathcal{T} = \{T_1, ..., T_1, ..., T_r, ..., T_r\}$, with t_i copies of T_i. Then $\Gamma(\mathcal{T}) = (T_1, \ldots, T_r)$ provides a sequence of elements of $\mathcal{T}$ without repetition, $n(\mathcal{T}) = (t_1, \ldots, t_r)$ and $v(\mathcal{T}) = (|V(T_1)|, \ldots, |V(T_r)|)$ give, respectively, the sequence of the number of copies that each T_i has in $\mathcal{T}$ and the number of vertices of each graph in $\mathcal{T}$ (these two sequences have an ordering consistent with $\Gamma(\mathcal{T})$). Note that a multiset of graphs, such as $\mathcal{T}$ can be viewed as a graph, denoted as $G(\mathcal{T})$, with vertex set $\sqcup_{T\in\mathcal{T}} V(T)$ and edge set $\sqcup_{T\in\mathcal{T}} E(T)$, thus having $t_1 + \ldots + t_r$ connected components.

A 2-connected graph $G = (V, E)$ is said to be *clique-separable* if there is a partition of V into three non-empty vertex sets $V = V_1 \sqcup V_2 \sqcup V_3$ such that there are no edges between V_1 and V_3, V_2 is a complete graph on $|V_2| \geq 2$ vertices, $V_1 \sqcup V_2$ and $V_3 \sqcup V_2$ induce two 2–connected graphs with $\geq |V_2| + 1$ vertices each.

[1] The number of subgraphs (induced subgraphs) usually refers to the number of injective graph homomorphisms (injective and also preserving non-edges). We are considering the subgraphs as subsets of edges; thus the number of subgraphs of C_4 in C_4 is 1, while the number of subgraphs of C_4 in C_4 with the usual understanding is 8.

[2] Note there is a typographical error in [4, Thm 2], namely "$-\binom{t}{2}$" should actually be "$+\binom{t}{2}$".

Theorem 2. *The chromatic polynomial $P(G;\lambda)$ can be computed as*

$$P(G;\lambda)=\sum_{p=0}^{|G|}\left[\sum_{\substack{\mathcal{T}\ multiset\ of\ \mathcal{B}\\(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})\leq p}} c_p(\mathcal{T})\prod_{i\in[dim(n(\mathcal{T}))]}\binom{inde(\Gamma(\mathcal{T})_i,G)}{n(\mathcal{T})_i}\right]\lambda^{|G|-p} \tag{4}$$

$$P(G;\lambda)=\sum_{p=0}^{|G|}\left[\sum_{\substack{\mathcal{T}\ multiset\ of\ \mathcal{B}\\(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})\leq p}} s_p(\mathcal{T})\prod_{i\in[dim(n(\mathcal{T}))]}\binom{sube(\Gamma(\mathcal{T})_i,G)}{n(\mathcal{T})_i}\right]\lambda^{|G|-p} \tag{5}$$

where: $(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})$ is the usual scalar product of two vectors, $inde(\cdot,G)$ and $sube(\cdot,G)$ are given by (3), both $c_p(\mathcal{T})$ and $s_p(\mathcal{T})$ are integers depending solely on $\mathcal{T}$ and p (not on G), and $|G|:=|V(G)|$. Furthermore:

(i) $c_p(\mathcal{T})=0$ if $(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})>p$

(ii) $c_p(\mathcal{T})=0$ if a $T\in\mathcal{T}$ is clique-separable

(iii) if $\mathcal{T}=\{T\}$ and $|T|=p+1$,

$$c_p(\mathcal{T})=\sum_{A\subseteq E(T),\ (V(T),A)\ 2-connected}(-1)^{|A|}$$

(iv) if $\mathcal{T}=\{T\}$ and $|T|=p+1$ and $i\geq 1$,

$$c_{p+i}(\mathcal{T})=-\sum_{\substack{\mathcal{T}'\ multiset\ of\ \mathcal{B},\ \mathcal{T}'\neq\mathcal{T}\\ \mathcal{T}'\ containing\ subgraphs\ of\ T}} c_{p+i}(\mathcal{T}')\prod_{j\in[dim(n(\mathcal{T}'))]}\binom{inde(\Gamma(\mathcal{T}')_j,T)}{n(\mathcal{T}')_j}$$

(v) When $|\mathcal{T}|=t\geq 2$ and for each $i\geq 0$ we have:

$$c_{(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})+i}(\mathcal{T})=\sum_{k_1+\ldots+k_t=i,\ k_s\geq 0}\ \prod_{T_t\in\mathcal{T}} c_{|T_t|-1+k_t}(\{T_t\})\ .$$

(vi) For each $p\geq 0$ and $\mathcal{T}$ multiset of $\mathcal{B}$, $c_p(\mathcal{T})$ are determined by (i),(ii),(iii), (iv), (v).

Before proceeding to the proof, we highlight that, in the proof and in the statement of Theorem 2, the H in (3) that are used are 2–connected.

Sketch of the proof. First we show (5) by translating the summation over edges as a sum of "independent" combinations of 2–connected blocks which would form the subgraph in question. Since we are considering these 2–connected blocks as being combined independently, we should substract the instances where these independent 2–connected blocks are combined into larger 2–connected blocks, such as when 3 edges are combined to form a triangle. Once (5) is obtained, we show (4) using that the number of instances of a subgraph T can be counted using induced graphs that are supergraphs of T on the same vertex set. We complete the argument using Vandermonde's involution formula together with Pólya and Ostrowski result from 1920 which implies that the polynomials $\binom{mx}{k}$ with positive integers m and k and variable x can be written in terms of $\binom{x}{i}$, $1\leq i\leq k$ using integer coefficients. Part (iii) follows by examining the contribution to c_{n-i} in (2) by only one 2–connected block with $i+1$ vertices. Parts (ii), (v), and (iv) follow by the multiplicative properties of the chromatic polynomial over 2–connected components, and its behaviour over clique-join graphs. In particular, given (4) and the $c_p(\mathcal{T})$ as unknowns we consider certain chromatic polynomials which, when closely examined, gives the equations and relations described in (ii), (v), and (iv). These constructions are described below.

Proof of (ii). Let A be a 2-connected graph which is a clique-join of two other graphs (so it is clique-separable). The proof goes by induction on the number of edges and vertices of A. Let A_1, A_2 be the two 2-connected components that are joined by a clique. Then we consider G_1 the graph formed by 5 vertex-disjoint copies of A and G_2 the graph obtained by the disjoint union of: 2 vertex-disjoint copies of A_1, 2 vertex-disjoint copies of A_2, and a graph formed by 3 copies of A_1 and 3 copies of A_2 on the same clique (and in such a way that the number of induced copies of A in the resulting graph is 9 by choosing one of the copies of A_1 and one of the copies of A_2, independently). The chromatic polynomial of G_1 and G_2 is the same in both cases:

$$\left(\frac{P(A_1;\lambda)P(A_2;\lambda)}{\lambda(\lambda-1)\cdots(\lambda-q)}\right)^5 = \frac{\left(\frac{P(A_1;\lambda)P(A_2;\lambda)}{\lambda(\lambda-1)\cdots(\lambda-q)}\right)\left(\frac{P(A_1;\lambda)P(A_2;\lambda)}{\lambda(\lambda-1)\cdots(\lambda-q)}\right)\left(\frac{P(A_1;\lambda)P(A_2;\lambda)}{\lambda(\lambda-1)\cdots(\lambda-q)}\right)}{\lambda(\lambda-1)\cdots(\lambda-q)} P(A_1;\lambda)^2P(A_2;\lambda)^2$$

where $q+1$ is the size of the clique by which they are joined in A. Moreover:

- If a 2-connected graph T is an induced graph of one of the parts (either a subgraph of A_1 or a subgraph of A_2), then the number of induced copies of T in G_1 and in G_2 (counted as in (3)) is the same.
- There are strictly more induced copies of A in G_2 than in the former (5 in G_1 versus 9 in G_2 if we assume A_1 and A_2 are different, otherwise is 5 versus at least 9 or at most $\binom{6}{2} = 15$ depending on the join interaction of A_1 and A_2 with respect to the common clique).
- Any induced copy of a 2–connected graph that contains a part in A_1 and a part of A_2 (and a strict induced subgraph of A) would consists on two graphs joined through a clique, and thus the corresponding coefficient in any chromatic polynomial and for any monomial is zero by induction.

By the previous argument for any induced graph completely contained in A_1 or A_2 the induced graph accounts are the same, and any graph that contains a part in A_1 and a part in A_2 is a clique-join and thus, by an inductive argument, does not appear in the summation making up the coefficients. Therefore, the fact that there are strictly more induced copies of A in G_2 than in G_1 implies that the corresponding coefficient of A for λ in $P(A;\lambda)$ should be zero. By adding some isolated vertices, we can conclude the same for all the coefficients involving the number of induced copies of A, when $\mathcal{T} = \{A, \dots, A\}$, are zero. For a general multiset $\mathcal{T}$ containing A, it follows from (v) and the fact that all the coefficients are zero when $\mathcal{T} = \{A\}$ as we have just shown.

Proof of (iv). Consider the chromatic polynomial of T, a 2–connected graph and $i \geq 1$ isolated vertices; $T \sqcup \{v_j\}_{j\in[i]}$. Then, $P(T \sqcup \{v_j\}_{j\in[i]};\lambda) = P(T;\lambda)\lambda^i$. From (4), we can determine that there are no 2–connected components with larger number of vertices (or edges) than T, so all the terms $\binom{\mathrm{inde}(\Gamma(\mathcal{T})_i, T\sqcup\{v_j\}_{j\in[i]})}{n(\mathcal{T})_i}$ are zero unless $\Gamma(\mathcal{T})_i$ is an induced subgraph of T. In particular, the only coefficients $c_p(\mathcal{T})$ that are multiplying non-zero terms of the type $\prod_{j\in[\dim(n(\mathcal{T}))]}\binom{\mathrm{inde}(\Gamma(\mathcal{T})_j, G)}{n(\mathcal{T})_j}$ are those where all the graphs in $\mathcal{T}$ are induced subgraphs of T. This implies that, if $p = |T| - 1 + i$ we have

$$0 = \left[\sum_{\substack{\mathcal{T} \text{ multiset of } \mathcal{B},\ (v(\mathcal{T})-(1,\dots,1))\cdot n(\mathcal{T}) \leq |T|-1+i \\ \mathcal{T} \text{ containing only induced subgraphs of } T}} c_{|T|-1+i}(\mathcal{T}) \prod_{j\in[\dim(n(\mathcal{T}))]} \binom{\mathrm{inde}(\Gamma(\mathcal{T})_j, G)}{n(\mathcal{T})_j}\right] \tag{6}$$

where the zero comes from the fact that all the coefficients multiplying monomials from $P(T \sqcup \{v_j\}_{j\in[i]};\lambda)$ of degree $< i+1$ are zero, which implies that when $p = |T| - 1 + i$ the coefficient of $\lambda^{|T|+i-|T|+1-i} = 0$. Isolating the term $c_{|T|-1+i}(T)$ in (6) that is multiplying $\binom{\mathrm{inde}(T,\, G)}{1} = 1$ gives (iv).

Proof of (v). We have to show that, for each $i \geq 0$, $c_{(v(\mathcal{T})-(1,\dots,1))\cdot n(\mathcal{T})+i}(\mathcal{T}) = \sum_{k_1+\dots+k_t=i, k_s\geq 0, \text{ with } |\mathcal{T}|=t} \prod_{T_j\in\mathcal{T}} c_{|T_j|-1+k_j}(\{T_j\})$. Given $\mathcal{T} = \{T_1, \dots, T_1, T_2, \dots, T_2, \dots, T_r, \dots, T_r\}$ with t_i copies of each T_i, consider the chromatic polynomial of G, the graph obtained from the disjoint union of t_j copies of the graph T_j, for each $j \in [r]$, and i isolated vertices, so: $P(G;\lambda) = P(\sqcup_{j\in[r], s\in[t_i]} T_j \sqcup \{v\}_{j\in[i]};\lambda) =$

$\lambda^i \prod_{j\in[r]} P(T_j;\lambda)^{t_j}$. By the disjoint unionness of G (in terms of the graphs of $\mathcal{T}$), for each $T_j \in \mathcal{B}$, $\mathrm{inde}(T_j, G) = \sum_{T\in\mathcal{T}} \mathrm{inde}(T_j, T)$. Using Vandermonde's involution formula to the latter, the monomial involving the coefficient $c_{(v(\mathcal{T})-(1,\ldots,1))\cdot n(\mathcal{T})+i}(\mathcal{T})$ which multiplies the term $\prod_{j\in[\dim(n(\mathcal{T}))]} \binom{\mathrm{inde}(\Gamma(\mathcal{T})_j,G)}{n(\mathcal{T})_j} = \prod_{j\in[\dim(n(\mathcal{T}))]} \binom{\mathrm{inde}(T_j,G)}{t_j}$ depends on the multiplication of the terms where $\binom{\mathrm{inde}(T_i,T_i)}{1}$ from the polynomial $P(T_i;\lambda)$,[3] while making that the powers of the λ coincide at the end. Thus the formula (v) follows.

3 Comments and consequences of our main result

On the $c_p(\mathcal{T})$ of chromatically equivalent graphs. Observing Theorem 2, it is natural to ask the relationship between pairs of graphs G and H, their chromatic polynomials $P(G;\lambda)$ and $P(H;\lambda)$, and the relationship between their corresponding sequences $\{c_p(G)\}_p$ and $\{c_p(H)\}_p$. Even though there is obviously a relation, it is non-trivial. In particular, there are pairs of graphs with the same sequences $\{c_p(G)\}_p = \{c_p(H)\}_p$, yet $P(G;\lambda) \neq P(H;\lambda)$, for instance, any pair of graphs that are clique-joins, yet they have different chromatic polynomials (even perhaps they have a different number of vertices). On the other hand, the following two graphs, shown as Figure 1 have the same chromatic polynomial:

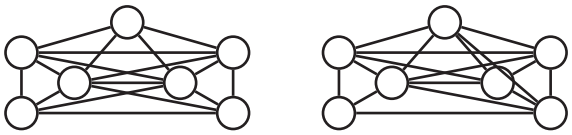


Figure 1: Graph G on the left. Graph H on the right

$P(G;\lambda) = P(H;\lambda) = \lambda^7 - 17\lambda^6 + 118\lambda^5 - 425\lambda^4 + 829\lambda^3 - 818\lambda^2 + 312\lambda$, yet their sequences $\{c_p(G)\}_p$ differ. Indeed, G from Figure 1 is clique-separable (vertices $V_1 = \{\text{top vertex}\}, V_2 = \{\text{middle four}\}, V_3 = \{\text{lower two}\}$, so all its coefficients are 0, however, H is not clique-separable, and in particular, it has a non-zero coefficient $c_6(H) = 30$.

Algorithmic questions. Expression (4) in Theorem 2 can be used in order to find $c_{n-p}(\mathcal{T})$ as follows. One can set a linear system with one equation for each connected graph G on $p+1$ vertices using the value of the coefficient of λ^{n-p} in the expression for $P(G;\lambda)$, with the unknown coefficients $c_{n-p}(\mathcal{T})$ in the expression (4) as variables, and with the corresponding expression of the number of induced subgraphs as coefficients of the equation (for each graph G, these numbers can be computed). The number of connected graphs on $p+1$ vertices is denoted by k_p. k_p is then an upper bound on the number of variables, and on the number of equations as well. Thus the linear system can be solved in time $O(k_p^2)$. Also, for each graph G, its linear equation can be set up in time $2^{|E(G)|}$ times checking the 2–connected isomorphism type of the subset of edges; since $2^{|E(G)|} = O(k_p)$ and checking the 2–connected isomorphism has $O(p!k_p) = O(k_p^2)$ complexity, a relatively easy algorithm on time $O(k_p^3)$ can be implemented.

On wheels. As W_{2n-1} are 3-colourable, no induced copy of K_4 can be found in a graph chromatically equivalent to them. Then Lemma 3 gives an alternative proof that W_{2n-1} are chromatically unique [8].

Lemma 3. *If G is a graph on $n \geq 5$ vertices and chromatically equivalent to the wheel W_n (that is, $P(G;\lambda) = P(W_n;\lambda)$), and $G \ncong W_n$, then G has at least 2 induced C_4 and an induced K_4.*

[3]The lower term of the binomial coefficient cannot be strictly larger, as there are no sufficient copies in T_i. If it is strictly smaller, then will be accounted by some $\binom{\mathrm{inde}(T_j,T_j)}{i}$ where i is strictly larger. Further, for $\binom{\mathrm{inde}(T_i,T_j)}{t}$ with $i \neq j$, the only way of carrying such term is when the product picks another $\binom{\mathrm{inde}(T_a,T_b)}{t}$, with $i > 0$ and where T_a is not a subgraph of T_b, and thus the term will become 0.

Proof. G and W_n should have the same number of vertices. Since the chromatic coefficients of λ^{n-1}, and λ^{n-2} are the same for G and W_n, they have the same number of edges and triangles.

Since any induced 2–connected subgraph of the wheel with n vertices (maximal induced cycle of size $n-1$) with ≥ 4 and $\leq n-2$ vertices is a clique-join, we may use Theorem 2 (ii) and (4), the expressions configuring the coefficients of λ^{n-p} for $p \in [3, n-3]$ for the wheel only depends on the number of triangles and number of edges, and thus these expressions gets balanced with those parts from G (as they have the same number of triangles and edges).

Now, since the chromatic number of the wheels is 4, there are no induced copies of K_j, $j \geq 5$ in G. By Theorem 2 or Theorem 1, the graph G will have induced K_4 if and only if it has induced C_4's; indeed, when $n \geq 6$ is due to Theorem 1 and the fact that W_n has neither induced C_4 nor K_4, and thus these numbers should balance in the coefficient λ^{n-3} as the coefficient also depends on the number of triangles and edges, but those two numbers are the same for G and for W_n (when $n = 5$, G could have a C_4 without a K_4 but then it would be the wheel as those two have the same number of triangles and triangles all should be incident with the last vertex, and thus the claim follows).

Assume for a contradiction that G has no induced $C_4, C_5, \ldots, C_k$, for some $k \geq 4$, then any induced 2–connected graph on $\leq k$ vertices has only triangles, and all the induced cycles have a chord. In particular, all of these graphs are chordal graphs. Thus, it has a perfect elimination ordering, meaning that the neighbourhood of the any removed vertex in the perfect elimination ordering is a clique. This means that, either the graph is K_i, $i \geq 4$, or it is a clique-join. Since G, $n \geq 5$ has no copies of K_i, $i \geq 5$, the only remaining case is for K_4. However, if it has a K_4, then G should also contain an induced C_4 (as claimed in the previous paragraph). Now let us focus on the coefficient c_{n-k}. Consider an induced 2-connected subgraph of G with $k+1$ vertices; if it is not a chordal graph and it is not C_{k+1}, then it contains an induced cycle on $\leq k$ vertices, which is a contradiction with the assumption. Otherwise, it is a chordal graph, and the perfect elimination ordering shows that it is either a clique on $k+1$ vertices, or a clique-join, thus not counting towards c_{n-k}. In particular, it can only be C_{k+1}, but if $n > k+2$, then W_n has no induced cycle of length C_{k+1} and the only contributions towards c_{n-k} are from edges and triangles, which is the same as for G. Since C_{k+1} has a non-zero coefficient by Theorem 2 (iii), then G cannot have an induced copy of C_{k+1} for otherwise, under the assumption of having no $C_4, C_5, \ldots, C_k$, then it would not have the same chromatic polynomial as W_n. This process can be run until c_2 for which a single copy of C_{n-1} appears in W_n, thus by the previous argument forcing a single copy of C_{n-1} in G as well. Since both have the same number of triangles ($n-1$ induced triangles and $2(n-1)$ edges is the wheel on n vertices), they end up being the same graph, and thus the assumption $G \neq W_n$ does not hold. □

Acknowledgements. The authors thank the referees for their helpful comments and references.

References

[1] M. Abért and T. Hubai. Benjamini-schramm convergence and the distribution of chromatic roots for sparse graphs. *Combinatorica*, 35:127–151, 2015.

[2] P. Csikvári and P. E. Frenkel. Benjamini–schramm continuity of root moments of graph polynomials. *European Journal of Combinatorics*, 52:302–320, 2016.

[3] D. Delbourgo and K. Morgan. On the fifth chromatic coefficient. *Australasian Journal of Combinatorics.*, 84(1):56–85, 2022.

[4] E. Farrell. On chromatic coefficient. *Discrete Mathematics*, 29:257–265, 1980.

[5] V. Patel and G. Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.

[6] H. Whitney. The coloring of graphs. *The Annals of Mathematics*, 33:688, 1932.

[7] H. Whitney. A logical expansion in mathematics. *Bulletin of the American Mathematical Society*, 8:572–579, 1932.

[8] S.-J. Xu and N.-Z. Li. The chromaticity of wheels. *Discrete Mathematics*, 51(2):207–212, 1984.

Extending the Continuum of Six-Colorings*

Konrad Mundinger[1,2], Sebastian Pokutta[1,2], Christoph Spiegel[1,2], and Max Zimmer[1,2]

[1]Technische Universität Berlin, Institute of Mathematics
[2]Zuse Institute Berlin, Department AIS2T, *lastname*@zib.de

Abstract

We present two novel six-colorings of the Euclidean plane that avoid monochromatic pairs of points at unit distance in five colors and monochromatic pairs at another specified distance d in the sixth color. Such colorings have previously been known to exist for $0.41 < \sqrt{2}-1 \leq d \leq 1/\sqrt{5} < 0.45$. Our results significantly expand that range to $0.354 \leq d \leq 0.657$, the first improvement in 30 years. The constructions underlying this notably were derived by formalizing colorings suggested by a custom machine learning approach.

1 Introduction

The Hadwiger-Nelson problem asks for the smallest number of colors needed to color the points of the Euclidean plane $\mathbb{E}^2$ without any two points a unit distance apart having the same color. Viewing the plane as an infinite graph, with an edge between any two points if and only if the distance between them is 1, motivates why this number is also referred to as the *chromatic number of the plane* and denoted by $\chi(\mathbb{E}^2)$. The problem goes back to 1950 and has since become one of the most enduring and famous open problems in combinatorial geometry and graph theory. For an extensive history of the problem and results related to it, we refer the reader to Jensen and Toft [7] as well as Soifer [9, 17].

By the de Bruijn–Erdős theorem [1], and therefore assuming the axiom of choice, the problem is equivalent to finding the largest possible chromatic number of a finite unit distance graph, that is a graph that can be embedded into the plane such that any two vertices are adjacent if and only if the corresponding points are at unit distance. The triangle is one obvious such graph, giving a lower bound of 3, and the Moser spindle [8] is the most famous example of a graph giving a lower bound of 4. There had been no improvement to that lower bound since 1950 until de Grey famously established that $\chi(\mathbb{E}^2) \geq 5$ through a graph of order 1581 in 2018 [2]. Simplifying and reducing the size of this construction has been of great interest to the extent of being the topic of a Polymath project [4, 3, 10, 11].

Regarding upper bounds, there is a large number of distinct 7-colorings of the plane that avoid monochromatic pairs at unit distance, the first of which (using a tiling of the plane with congruent regular hexagons) was already observed back in 1950 by Isbell [9, 17]. This upper bound of $\chi(\mathbb{E}^2) \leq 7$ has remained unchanged since and many variants of the original question have therefore been proposed in the hopes of shedding some light on why this problem has proven so stubborn. To state one such variant, we say that an n-coloring of the plane has *coloring type* $(d_1, \ldots, d_n)$ if color i does not realize distance d_i [14, 15]. This gives a measurement of how close this coloring is to achieving the original goal and can be seen as a defining a natural 'off-diagonal' variant of the original problem. Finding a coloring of type $(1, 1, 1, 1, 1, 1)$ would obviously improve the upper bound of $\chi(\mathbb{E}^2)$ to 6.

*This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

Stechkin found a coloring of type $(1, 1, 1, 1, 1/2, 1/2)$, which was published by Raiskii in 1970 [12], and Woodall found a coloring of type $(1, 1, 1, 1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{12})$ in 1973 [18]. The first six-coloring to feature a non-unit distance in only one color has type $(1, 1, 1, 1, 1, 1/\sqrt{5})$ and was found by Soifer in 1991 [15]. Hoffman and Soifer also found a coloring of type $(1, 1, 1, 1, 1, \sqrt{2} - 1)$ in 1993 [5, 6]. Both of these constructions are in fact part of a family that realizes $(1, 1, 1, 1, 1, d)$ for any $1/\sqrt{5} \leq d \leq \sqrt{2} - 1$ [6, 16, 17], leading Soifer [13] to pose the "still open and extremely difficult" [9] problem of determining the *continuum of six colorings* X_6, that is the set of all d for which there exists a six-coloring of the plane of type $(1, 1, 1, 1, 1, d)$. To the best of our knowledge, no improvements have been suggested in the last 30 years.

We propose two novel six-colorings of the plane, one parameterized by d and the other fixed, that together significantly expand the range of d known to be in X_6. The first is a valid coloring of type $(1, 1, 1, 1, 1, d)$ as long as $0.354 \leq d \leq 0.553$ and the second covers the range of $0.418 \leq d \leq 0.657$.

Theorem 1. *X_6 contains the closed interval* $[0.354, 0.657]$.

It should be noted that both constructions were derived by formalizing colorings that were suggested by a custom machine learning approach in which a Neural Network was trained to represent a coloring of a specified type or range of types. We will briefly touch upon this in Section 4 and otherwise go into more detail about this approach and potential other applications in a separate publication. This work is intended to give a formal justification of Theorem 1, with the first coloring being explored in Section 2 and the second in Section 3.

2 A construction for $0.354 \leq d \leq 0.553$

The first constructions is made up of four different polytopal shapes, a detailed description of which is given in the appendix. The equidiagonal pentagon and the equilateral triangle respectively described Figure 3 and Figure 4 together are be colored with the sixth color (red) in which we are avoiding points at distance d. The octagons described in Figure 5 receive three of the other five colors (orange, green, and blue) and the hexagons described in Figure 6 receive the remaining two (yellow and turquoise). All shapes are uniquely parameterized by the choice of d, with the exception of the pentagon, which has an additional degree of freedom in the form of α_1. We will later determine the range of valid α_1 depending on d numerically and see that this additional variable can be fixed by linearly interpolating between two extremal values (though other options can also be valid depending on d).

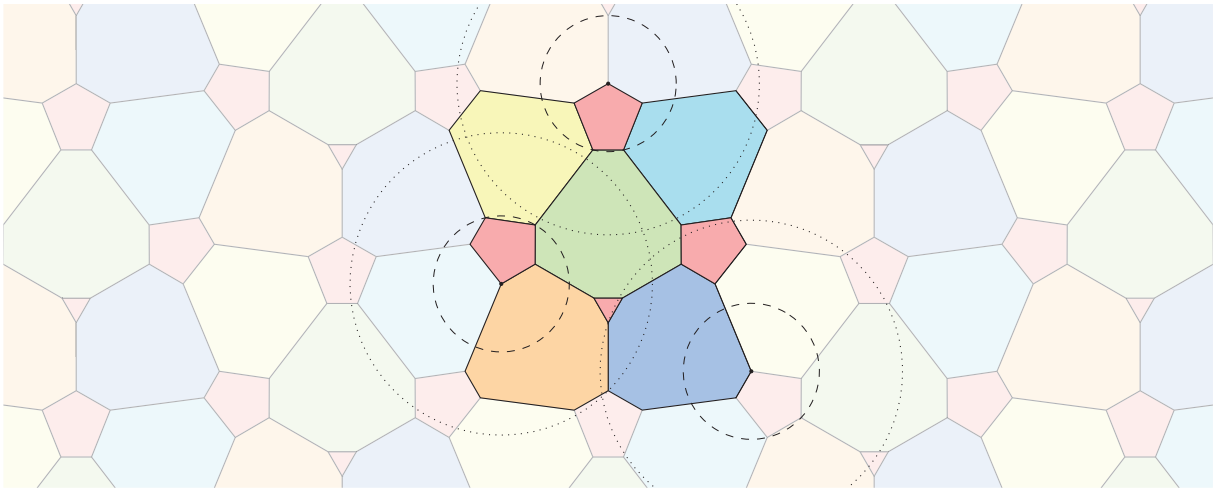


Figure 1: Illustration of the first coloring with circles at unit distance (dotted) and distance d (dashed) highlighted at three critical points.

A copy of three pentagons, one triangle, three octagons and two hexagons together form the building block of the first coloring.Note that the triangle disappears as d approaches the upper end of the valid spectrum. Looking at the overall construction in Figure 1, it is visually clear that the only conditions that are at risk making this construction invalid are given be the following set of constraints, where the variables are defined alongside the corresponding shape in the appendix:

$$s_4 \leq d \quad (1) \qquad w_2 \leq 1 \quad (4)$$
$$s_5 \geq d \quad (2) \qquad w_3 \leq 1 \quad (5)$$
$$w_1 \leq 1 \quad (3) \qquad h_1 + h_3 + d \geq 1 \quad (6)$$

Unfortunately we were unable to derive a closed form expression for the range of d for which a valid choice of α_1 can be found. However, it is easy to numerically verify that for $d \in [0.354, 0.553]$ such a choice can be made. Furthermore, by linearly interpolating between the two extreme points, that is by choosing $\alpha_1 = 113.7 + (d - 0.354)\, 14.11/0.299$, we can remove the additional degree of freedom in the definition of the pentagon. Finally, we note that there is again always an appropriate choice for the color on the boundaries between the shapes.

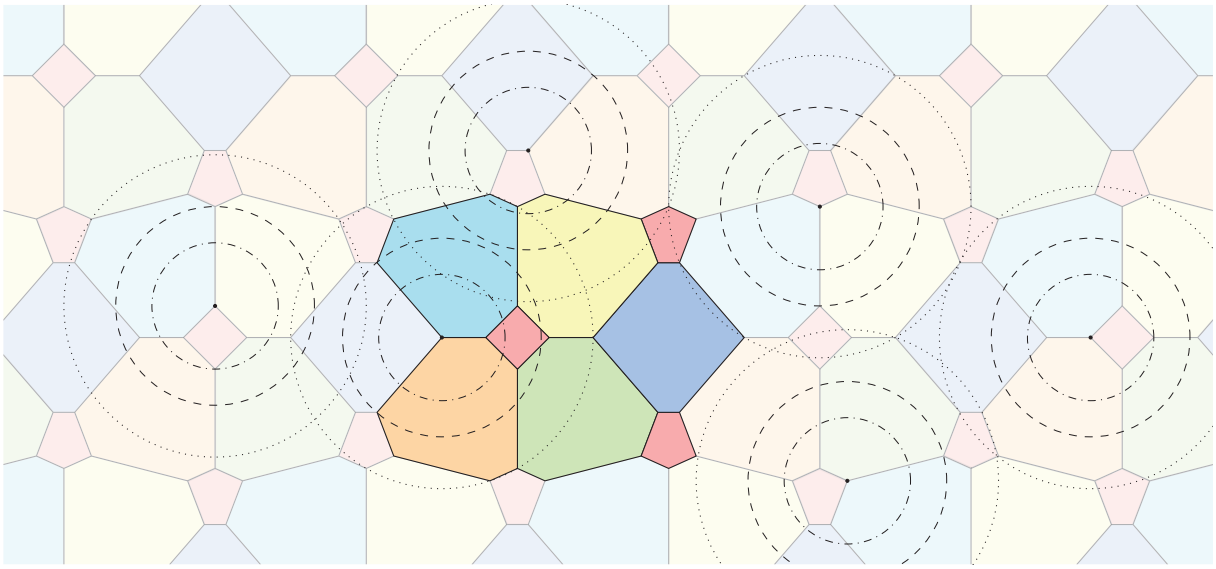


Figure 2: Illustration of the second coloring with circles at unit distance (dotted), and distance $d_{\max}$ (dashed), and distance distance $d_{\min}$ (dash-dotted) highlighted at six critical points.

3 A construction for $0.418 \leq d \leq 0.657$

Let $d_{\max}$ be the real root of $d^4 + 5\sqrt{3}d^3 + 18d^2 - 3\sqrt{3}d - 7 = 0$ closest to 0.65 and $d_{\min} = \sqrt{3} - 2\, d_{\max}$. Note that a closed form for $d_{\max}$ is given by

$$\begin{aligned} d_{\max} = &-(5\sqrt{3})/4 + 1/2\left(27/4 + 1/3\,(7290 - 15\sqrt{1821})^{1/3} + (5\,(486 + \sqrt{1821}))^{1/3}/3^{2/3}\right)^{1/2} \\ &+ 1/2\left(27/2 - 1/3\,(7290 - 15\sqrt{1821})^{1/3} - (5(486 + \sqrt{1821}))^{1/3}/3^{2/3}\right. \\ &\left. + 9/4\left(3/(27/4 + 1/3\,(7290 - 15\sqrt{1821})^{1/3} + (5(486 + \sqrt{1821}))^{1/3}/3^{2/3}))\right)^{1/2}\right)^{1/2}. \end{aligned}$$

We can easily verify numerically that $d_{\min} \leq 0.418 \leq d \leq 0.657 \leq d_{\max}$ and the second construction will in fact be valid for any $d \in [d_{\min}, d_{\max}]$. It is again made up of four different polytopal shapes, a detailed

description of which is given in the appendix. The pentagon and square described in Figure 7 together are colored with the sixth color (red) in which we are avoiding points at distance d. The heptagon described in Figure 8 receives four of the other five colors (orange, green, yellow, and turquoise) while hexagon described in Figure 8 receives the last remaining color (blue). A copy of two pentagons, one square, four heptagons and one hexagon together form the building block of the second coloring, which is illustrated in Figure 2.

4 Discussion and Outlook

We conclude by noting that there was a significant technical component to these new constructions. We developed a custom machine learning approach in which we had a Neural Network represent a (probabilistic) six-coloring of the plane. The parameters of the network were update according to a batched form of the loss given by the probabilistic likelihood that two points at unit distance (or at distance d) are monochromatic with the right color(s). The resulting output was detailed enough to inspire the above constructions, though formally describing them and verifying their correctness still required a fair amount of manual effort.

References

[1] de Bruijn, N.G., Erdős, P.: A colour problem for infinite graphs and a problem in the theory of relations. Indigationes Mathematicae **13**, 371–373 (1951)

[2] De Grey, A.D.: The chromatic number of the plane is at least 5. Geombinatorics Quarterly **XXVIII**(1), 18–31 (2018)

[3] Exoo, G., Ismailescu, D.: The chromatic number of the plane is at least 5: a new proof. Discrete & Computational Geometry **64**(1), 216–226 (2020)

[4] Heule, M.J.: Computing small unit-distance graphs with chromatic number 5. Geombinatorics Quarterly **XXVIII**(1), 32–50 (2018)

[5] Hoffman, I., Soifer, A.: Almost chromatic number of the plane. Geombinatorics **3**(2), 38–40 (1993)

[6] Hoffman, I., Soifer, A.: Another six-coloring of the plane. Discrete Mathematics **150**(1-3), 427–429 (1996)

[7] Jensen, T.R., Toft, B.: Graph coloring problems. John Wiley & Sons (2011)

[8] Moser, L., Moser, W.: Solution to problem 10. Canad. Math. Bull **4**, 187–189 (1961)

[9] Nash, J.F., Rassias, M.T.: Open problems in mathematics. Springer (2016)

[10] Parts, J.: Graph minimization, focusing on the example of 5-chromatic unit-distance graphs in the plane. arXiv preprint arXiv:2010.12665 (2020)

[11] Polymath, D.: On the chromatic number of circular disks and infinite strips in the plane. Geombinatorics Quarterly **XXX**(4), 190–201 (2021)

[12] Raiskii, D.E.: Realization of all distances in a decomposition of the space $\mathbb{R}^n$ into $n+1$ parts. Mathematical notes of the Academy of Sciences of the USSR **7**, 194–196 (1970)

[13] Soifer, A.: Six-realizable set x_6. Geombinatorics **III**(4), 140–145 (1994)

[14] Soifer, A.: Relatives of chromatic number of the plane i. Geombinatorics **1**(4), 13–17 (1992)

[15] Soifer, A.: A six-coloring of the plane. Journal of Combinatorial Theory, Series A **61**(2), 292–294 (1992)

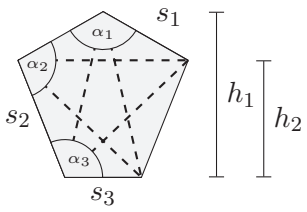
[16] Soifer, A.: An infinite class of six-colorings of the plane. Congressus Numerantium pp. 83–86 (1994)

[17] Soifer, A.: The mathematical coloring book: Mathematics of coloring and the colorful life of its creators. Springer (2009)

[18] Woodall, D.R.: Distances realized by sets covering the plane. Journal of Combinatorial Theory, Series A **14**(2), 187–200 (1973)

Appendix

A Building blocks of the first coloring

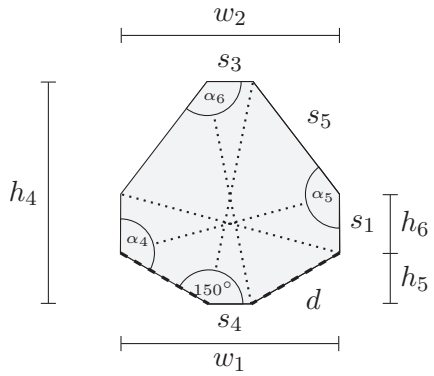


$$
\begin{aligned}
s_1 &= d/2\ \csc(\alpha_1/2)\\
t_1 &= 2\arccos\big(\csc(\alpha_1/2)/4\big) - \alpha_1\\
s_3 &= 2d\ \sin(t_1/2)\\
h_1 &= d\ \cos(t_1/2)\\
h_2 &= h_1 - (d/2)\ \cot(\alpha_1/2)\\
s_2 &= \sqrt{h_2^2 + (d-s_3)^2/4}\\
\alpha_2 &= 90^\circ - \alpha_1/2 + \arcsin(h_2/s_2)\\
\alpha_3 &= 270^\circ - \alpha_1/2 - \alpha_2
\end{aligned}
$$

Figure 3: An equidiagonal pentagon with each diagonal of length d, highlighted by dashed lines, used for the red color avoiding points at distance d in the first coloring.

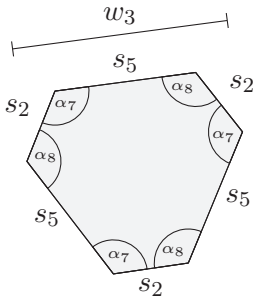
$$
\begin{aligned}
t_2 &= \big(\sqrt{1-(s_1\ \sin(30^\circ+\alpha_1/2))^2} - s_1\ \cos(30^\circ+\alpha_1/2)\big)/\sqrt{3}\\
s_4 &= \sqrt{3}\ \max(t_2-d,0)\\
h_3 &= 3/2\ \max(t_2-d,0)
\end{aligned}
$$

Figure 4: An equilateral triangle, used for the red color avoiding points at distance d in the first coloring.



$$
\begin{aligned}
w_1 &= \sqrt{3}\,t_2\\
t_3 &= 180^\circ - \arccos\big((1-w_1^2-s_1^2)/(-2w_1s_1)\big)\\
w_2 &= s_1\cos(t_3) + \sqrt{1-s_1^2\sin(t_3)^2}\\
h_4 &= \sqrt{1-(s_4+s_3)^2/4}\\
h_5 &= \sqrt{t_2^2 - w_1^2/4} - h_3 + \max(t_2-d,0)\\
h_6 &= \sqrt{s_1^2-(w_1-w_2)^2/4}\\
s_5 &= \sqrt{h_7^2+(w_2-s_3)^2/4}\\
\alpha_4 &= 180^\circ - \alpha_1/2\\
\alpha_5 &= \arctan\big(2\,h_7/(w_2-s_3)\big) + t_3\\
\alpha_6 &= 390^\circ - \alpha_4 - \alpha_5
\end{aligned}
$$

Figure 5: An axisymmetric octagon in which four of the diagonals have unit length, highlighted by dotted lines, and two of the sides have length d, highlighted by dashed lines. Used for the orange, green and blue color avoiding points at unit distance in the first coloring.



$$\alpha_7 = 360^\circ - \alpha_2 - \alpha_5$$
$$\alpha_8 = 360^\circ - \alpha_3 - \alpha_6 = 240^\circ - \alpha_7$$
$$t_4 = \sqrt{s_2^2 + s_5^2 - 2s_2 s_5 \cos(\alpha_7)}$$
$$t_5 = \arcsin(s_5 \sin(\alpha_7)/t_4)$$
$$w_3 = \sqrt{t_4^2 + s_2^2 + 2t_4 s_2 \cos(\alpha_7 + \alpha_8 + t_5)}$$

Figure 6: A hexagon with two angles and two side lengths. Used for the yellow and turqouise color avoiding points at unit distance in the first coloring. Note that it is in general *not* axisymmetric.

B Building blocks of the second coloring

Figure 7: An axisymmetric pentagon and a square together are used for the red color avoiding points at distance d in the second coloring. s_2 and s_3 are implicitly defined in Figure 8.

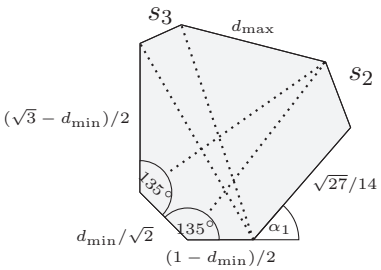
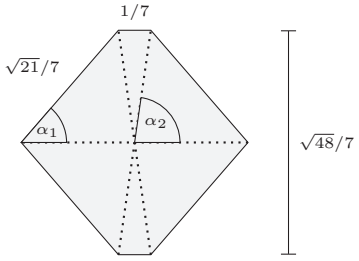


Figure 8: A heptagon in which four of the diagonals have unit length, highlighted by dotted lines. Used for the orange, green, yellow, and turquoise color avoiding points at unit distance in the second coloring. We do not give a closed form solution for s_2 and s_3 but note that they are well defined. The angle α_1 is defined in Figure 9.



$$\alpha_1 = 45^\circ + \arccos(47/49)/4$$
$$\alpha_2 = 90^\circ - \arccos(47/49)/2$$

Figure 9: A centrosymmetric hexagon in which three of the diagonals have unit length, highlighted by dotted lines. Used for the blue color avoiding points at unit distance in the second coloring.

On Ewald's and Nill's Conjectures about smooth polytopes*

Luis Crespo[1], Álvaro Pelayo[2], and Francisco Santos[1]

[1]Depto. de Matemáticas, Estadística y Computación, Univ. de Cantabria, 39005 Santander, Spain†
[2]Facultad de Ciencias Matemáticas, Universidad Complutense de Madrid, 28040 Madrid, and Real Academia de Ciencias Exactas, Físicas y Naturales de España, Spain‡

Abstract

A monotone polytope in $\mathbb{R}^n$ is a smooth reflexive polytope. These polytopes arise as the momentum polytopes of monotone symplectic toric manifolds. Ewald's well-known Conjecture from 1988 states that if P is a monotone n-polytope in $\mathbb{R}^n$ then the set $\mathbb{Z}^n \cap P \cap -P$ contains a unimodular basis of the lattice $\mathbb{Z}^n$. McDuff (2009) shows that a stronger property of a monotone polytope, which she calls *star Ewald condition*, is closely related to whether the central fiber of the corresponding monotone symplectic toric manifold is a stem. In 2009 Nill proposed a generalization of Ewald's Conjecture to smooth lattice polytopes. In this extended abstract, prepared for the Discrete Mathematics Days conference (University of Alcalá, July 3-5, 2024), we summarize the results concerning these conjectures that we have obtained in our recent article `arXiv:2310.10366`. We refer to this article for details and proofs.

1 Introduction

The goal of this extended abstract is to report on the results from our recent paper [4], which solves some broad cases of a well-known conjecture by G. Ewald from 1988 concerning monotone lattice polytopes [5], and its more recent generalization to smooth polytopes by B. Nill, from 2009 [13]. Our motivation comes partially from symplectic geometry, as we will explain, but for brevity we do not discuss our results in this direction. We refer to the original article [4] for more details, complete statements, and proofs.

Smooth polytopes in general, and smooth reflexive ones in particular, are very important in algebraic and symplectic geometry, providing a strong link between "discrete" problems in combinatorics/convex geometry and "continuous" problems concerning smooth (toric) manifolds. In fact, smooth reflexive n-dimensional polytopes are also known as *monotone n-dimensional polytopes*, as they are the images of $2n$-dimensional monotone symplectic toric manifolds under the momentum map $M \to \mathbb{R}^n$. We refer to Charton-Sabatini-Sepe [2], Godinho-Heymann-Sabatini [7] and McDuff [10], for recent works which discuss monotone polytopes from the perspective of symplectic geometry and to Batyrev [1], Cox-Little-Schenck [3, Theorem 8.3.4], Franco-Seong [6], Haase-Melnikov [8] and Nill [12] for their relation to Gorenstein Fano varieties in algebraic geometry.

Let us recall their precise definitions:

Definition 1 (Smooth polytope)**.** *An n-dimensional polytope P in $\mathbb{R}^n$ is* smooth *if it satisfies the following three properties:*

*The full version of this work can be found in [4] and will be published elsewhere.

†Email: luis.cresporuiz@unican.es, francisco.santos@unican.es. Research of L. C. and F. S. supported by PID2019-106188GB-I00 and PID2022-137283NB-C21 of MCIN/AEI/10.13039/501100011033 / FEDER, UE and by project CLaPPo (21.SI03.64658) of Universidad de Cantabria and Banco Santander.

‡Email: alvpel01@ucm.es. Research of Á. P. supported by a BBVA (Bank Bilbao Vizcaya Argentaria) Foundation Grant for Scientific Research Projects with title *From Integrability to Randomness in Symplectic and Quantum Geometry.*

- *P is* simple*: there are precisely n edges meeting a each vertex;*
- *P is* rational*: it has rational edge directions (equivalently, the normal vector to the facets are rational);*
- *the primitive edge-direction vectors at each vertex of P form a basis of the lattice $\mathbb{Z}^n$.*

Equivalently, a smooth polytope P in $\mathbb{R}^n$ is a polytope whose normal fan is simplicial, rational, and unimodular.

Definition 2 (Reflexive polytope)**.** *A* reflexive polytope *in $\mathbb{R}^n$ is a lattice polytope with the origin in its interior and whose dual polytope is also a lattice polytope. Equivalently, a lattice polytope in $\mathbb{R}^n$ is reflexive if and only if every facet-defining inequality is of the form $u_F \cdot x \leqslant 1$, where u_F is the primitive exterior normal vector to the facet.*

Definition 3 (Monotone polytope)**.** *A polytope in $\mathbb{R}^n$ is* monotone *if it is smooth and reflexive.*

There are finitely many monotone polytopes in each dimension n modulo unimodular equivalence (that is, modulo $\mathrm{GL}(n,\mathbb{Z})$ or equivalently $\mathrm{AGL}(n,\mathbb{Z})$ transformations). Up to dimension 9 they are counted in [9, 14] and, as seen in Table 1, the number of monotone polytopes increases rapidly with the dimension. Figure 1 shows the five possibilities in dimension two.

dimension	1	2	3	4	5	6	7	8	9
monotone polytopes	1	5	18	124	866	7622	72256	749892	8229721

Table 1: Number of monotone polytopes in each dimension up to 9.

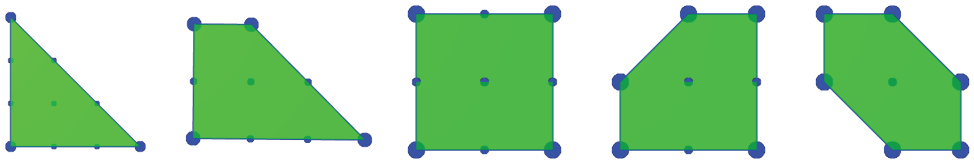


Figure 1: The five monotone polygons: monotone triangle, trapezoid, square, pentagon, and hexagon.

We are interested in understanding, both theoretically and computationally, the properties of the *Ewald set* of a monotone polytope. This set appears implicitly in the influential 1988 paper by Günter Ewald [5].

Definition 4 (Ewald set [4, Definition 1.1])**.** *The* Ewald set *of a polytope $P \subset \mathbb{R}^n$ is*

$$\mathcal{E}(P) := \mathbb{Z}^n \cap P \cap -P.$$

Its points are called Ewald points *of P.*

That is, $\mathcal{E}(P) \subset \mathbb{Z}^n$ consists of the *symmetric integral points of P*, meaning integral points $x \in \mathbb{Z}^n$ for which both $x \in P$ and $-x \in P$. Our main motivation is the following conjecture:[1]

Conjecture 5 (Ewald's Conjecture 1988 [5, Conjecture 2])**.** *Let $n \in \mathbb{N}$. If P is an n-dimensional monotone polytope in $\mathbb{R}^n$ then $\mathcal{E}(P)$ contains a unimodular basis of $\mathbb{Z}^n$.*

[1]The original formulation of Conjecture 5 refers to dual polytopes, stating that the dual of any monotone polytope P can be sent into $[-1,1]^n$, via a unimodular transformation. As pointed out by Øbro [15] this is equivalent to our formulation, used already by McDuff [11, Section 3.1] and Payne [16, Remark 4.6]. (McDuff and Payne remove the origin from $\mathcal{E}(P)$ in their definition, but for technical reasons we do not).

The conjecture has been verified computationally for $n \leqslant 7$ by Øbro [15, page 67], but little more is known about it. Both Payne and McDuff [11, 16] remark that it is not even known whether there is a monotone polytope with $\mathcal{E}(P) = \{0\}$.

Nill [13] proposed the following generalization of Conjecture 5 to smooth polytopes:

Conjecture 6 (General Ewald's Conjecture, Nill 2009 [13])**.** *Let $n \in \mathbb{N}$. If P is an n-dimensional smooth lattice polytope in $\mathbb{R}^n$ with the origin in its interior then $\mathcal{E}(P)$ contains a unimodular basis of $\mathbb{Z}^n$.*

This is clearly stronger than Conjecture 5, but it might actually be equivalent; as Nill points out, Conjecture 5 implies that $\mathcal{E}(P)$ linearly spans $\mathbb{R}^n$ for every smooth lattice polytope P with $0 \in \mathrm{Int}(P)$. (The implication is not on a dimension-by-dimension basis).

2 Three Ewald conditions and their motivation in symplectic geometry

Øbro's computational verification of Conjecture 5 for $n \leqslant 7$ shows the following strong version of it: for every facet F of P, $\mathcal{E}(P) \cap F$ contains a unimodular basis. This serves as motivation for the definition we give next. Before that let us introduce the following notation: let P be any polytope and let $\mathcal{F}$ and $\mathcal{R}$ be the sets of facets and *ridges* (that is, faces of codimension two) of P. For a face f of P we denote:

$$\mathrm{Star}(f) = \bigcup_{f \subset F \in \mathcal{F}} F; \quad \mathrm{star}(f) = \bigcup_{f \subset R \in \mathcal{R}} R; \quad \mathrm{Star}^*(f) = \mathrm{Star}(f) \setminus \mathrm{star}(f).$$

Definition 7 (Ewald conditions, McDuff [11, Definition 3.5])**.** *Let P be an n-dimensional polytope with the origin in its interior. We say that:*

1. *P satisfies the* weak Ewald condition *if $\mathcal{E}(P)$ contains a unimodular basis of $\mathbb{Z}^n$.*
2. *P satisfies the* strong Ewald condition *if, for each facet F of P, the set $\mathcal{E}(P) \cap F$ contains a unimodular basis of $\mathbb{Z}^n$.*
3. *A face f of P satisfies the* star Ewald condition *or* is star Ewald *if there exists $\lambda \in \mathcal{E}(P)$ such that $\lambda \in \mathrm{Star}^*(f)$ and $-\lambda \notin \mathrm{Star}(f)$.*
4. *P satisfies the* star Ewald condition *if every face of P satisfies it.*

The star Ewald condition is motivated by the following problem in symplectic toric geometry. It is known that every symplectic toric manifold M has a particular *central* toric orbit that is not *displaceable* by a Hamiltonian isotopy. A relevant question is whether for a given manifold this central orbit is the only non-displaceable one. If this happens then the central orbit is called a *stem*. McDuff relates displaceability of toric orbits in M to *displaceability by probes* of points in the corresponding momentum polytope (a concept that she defines). More precisely, she proves the following:

Theorem 8 (McDuff [11])**.** 1. *Let M be a symplectic toric manifold with momentum polytope P. If a point $u \in \mathrm{Int}(P)$ is displaceable by a probe then its fiber $L_u \subset M$ is displaceable by a Hamiltonian isotopy [11, Lemma 2.4].*
2. *A monotone polytope P satisfies the star Ewald condition if and only if every point of $\mathrm{Int}(P) \setminus \{0\}$ is displaceable by a probe [11, Theorem 1.2].*

It follows that if the momentum polytope of a monotone symplectic toric manifold satisfies the star Ewald condition then the central fiber is a stem.

The star Ewald condition is stronger than the weak Ewald condition by [11, Lemma 3.7]. However, there are 6-dimensional monotone polytopes for which the star Ewald condition fails [11, footnote to p. 134] (see also [4, Proposition 3.11]). Hence, the strong Ewald condition does not imply the star Ewald condition.

3 Deeply smooth polytopes satisfy the Ewald conditions

Definition 9 ([4, Definition 4.9])**.** *Let v be a vertex of a lattice smooth polytope P in $\mathbb{R}^n$, and let $u_1, \ldots, u_n$ be the primitive edge vectors at v. The parallelepiped*

$$\{v + \sum_{i=1}^{n} \lambda_i u_i \,|\, \lambda_i \in [0,1] \ \forall i\}$$

is called the corner parallelepiped *of P at v.*

We say that P is deeply smooth *if it contains the corner parallelepiped of P at v for every vertex v of P. We call P* deeply monotone *if it is deeply smooth and monotone.*

Our first main result in [4] determines a class of polytopes for which the Ewald conditions hold:

Theorem 10 ([4, Theorem 4.14])**.** *Every deeply monotone polytope satisfies the strong and star Ewald conditions (and, consequently, also the weak condition).*

As far as we know this is the first result covering a broad case of Ewald's Conjecture in arbitrary dimension. In the following table we have computed how many monotone polytopes fall within this class for $n \leqslant 6$.

dimension	monotone	deeply monotone
3	18	16
4	124	72
5	866	300
6	7622	1352

4 (Monotone) fiber bundles and neat polytopes

Now we look at the Ewald sets of bundles over polytopes.

Definition 11 (Bundle of a polytope)**.** *Let $n, k \in \mathbb{N}$. Given three polytopes $P \subset \mathbb{R}^{k+n}$, $B \subset \mathbb{R}^k$ and $Q \subset \mathbb{R}^n$, we say that P is a bundle with base B and fiber Q if the following conditions hold:*

1. *P is combinatorially equivalent to $B \times Q$.*
2. *There is a short exact sequence of linear maps*

$$0 \to \mathbb{R}^n \xrightarrow{i} \mathbb{R}^{k+n} \xrightarrow{\pi} \mathbb{R}^k \to 0$$

such that $\pi(P) = B$ and for every $x \in B$ we have that the polytope $Q_x := \pi^{-1}(x) \cap P$ is normally isomorphic to $i(Q)$ (that is, they have the same normal fan).

If P is a monotone bundle with fiber Q and base B then, with the natural identification $Q \cong \{0\} \times Q$, we have that $\mathcal{E}(Q) \subset \mathcal{E}(P)$. It is natural to ask under what conditions we have the analog property for the base: that every point in $\mathcal{E}(B)$ lifts to $\mathcal{E}(P)$. The answer is the following:

Definition 12 (Neat polytope [4, Definition 2.2])**.** *Let $m, n \in \mathbb{N}$. Let P be a smooth lattice polytope in $\mathbb{R}^n$ defined by the inequalities $Ax \leqslant c$, where $A \in \mathbb{Z}^{m \times n}$ and $c \in \mathbb{Z}^m$. For each $b \in \mathbb{Z}^m$ we define*

$$P_b := \{x \in \mathbb{R}^n : Ax \leqslant c + b\}$$

and call it the deformation of P by b. We say that P is neat *if whenever P_b and P_{-b} are normally isomorphic to (i.e., have the same normal fan as) P for a $b \in \mathbb{Z}^m$ we have that*

$$P_b \cap (-P_{-b}) \cap \mathbb{Z}^n \neq 0;$$

that is, there is an integer point $x \in P_b$ such that $-x \in P_{-b}$.

One of our results in [4] says that the condition above is precisely what is required of the fiber Q for the Ewald properties to be preserved under the fiber bundle operation:

Theorem 13 ([4, Corollary 5.10])**.** *For a lattice smooth polytope Q the following properties are equivalent:*

1. *Q is neat and satisfies the weak (resp. star) Ewald condition.*
2. *Every lattice smooth bundle P with fiber Q and base $[-1,1]$ satisfies the weak (resp. star) Ewald condition.*
3. *Every lattice smooth bundle P with fiber Q and an arbitrary base B satisfies the weak (resp. star) Ewald condition whenever B satisfies it.*

Corollary 14 ([4, Corollary 2.4])**.**

- *If Conjecture 5 holds then every monotone polytope is neat.*
- *If Conjecture 6 holds then every lattice smooth polytope is neat.*

5 The number of Ewald points

We now turn to discuss how many Ewald points a monotone polytope can have. It is easy to show that for every monotone n-polytope

$$\mathcal{E}(P) \subset \mathcal{E}([-1,1]^n) = \{-1,0,1\}^n,$$

where the first inclusion should be understood modulo unimodular equivalence. Hence, no monotone n-polytope can have more than 3^n Ewald points. Somewhat surprisingly, the number of Ewald points of the monotone cube is asymptotically attained (modulo a factor proportional to $\sqrt{n}$) by the monotone n-simplex and by any bundle with fiber the monotone simplex and base a segment.

In [4] we computed the number of Ewald points for every monotone polytope up to dimension seven. The minimum number in each dimension is as follows:

n	1	2	3	4	5	6	7
$\min \lvert\mathcal{E}(P)\rvert$	3	7	13	27	59	117	243

These numbers seem to grow exponentially, which supports the claim made in Conjecture 5. In fact, we have an explicit construction of monotone n-polytopes with $|\mathcal{E}(P)|$ growing asymptotically as $3^{2n/3}$ and which achieves *exactly* the minimal size for all $n \in [3,7]$:

Theorem 15 ([4, Corollary 6.7])**.** *For each $n \geqslant 3$ there is a monotone n-polytope P_n with*

$$|\mathcal{E}(P_n)| = \begin{cases} 13 \cdot 3^{2k-2} & \text{if } n = 3k \\ 3^{2k+1} & \text{if } n = 3k+1 \\ 59 \cdot 3^{2k-2} & \text{if } n = 3k+2 \end{cases}$$

Thus, the minimum number of Ewald points of monotone n-polytopes is of order $O(3^{2n/3})$.

6 Nill's Conjecture: a proof for $n = 2$ and partial results for higher n

In [4] we prove a strong form of Nill's Conjecture in dimension 2, in which the hypothesis is relaxed:

Theorem 16 ([4, Corollary 7.3])**.** *If P is a lattice polygon with the origin in its interior and each vertex of P is at lattice distance one from the line spanned by its two neighboring boundary lattice points, then $\mathcal{E}(P)$ contains a lattice basis.*

It seems quite challenging to make the type of arguments we use to work in dimensions 3 or higher, but in [4] we were able to prove the following two partial results.

Definition 17 ([4, Definition 7.4])**.** *Let P be a lattice polytope, F a face of it, and $x_0 \in P$. The maximum distance from x_0 to the facets containing F is called* distance from x_0 to F*. We say that x_0 is* next to F *if it is in the interior of P and at distance one from F.*

Proposition 18 ([4, Proposition 7.5])**.** *Let P be a deeply smooth n-polytope with the origin in its interior, and suppose that the origin is next to a certain vertex v. Then, $\mathcal{E}(P)$ contains the lattice basis consisting of the primitive edge vectors of P at v.*

Proposition 19 ([4, Proposition 7.7])**.** *Let P be a smooth 3-polytope with the origin in its interior, and suppose that the origin is next to a certain edge uv. Then, $\mathcal{E}(P)$ contains a lattice basis.*

References

[1] V.V. Batyrev: Dual polyhedra and mirror symmetry for Calabi-Yau hypersurfaces in toric varieties, *J. Algebraic Geom.* **3** (1994), 493–535.

[2] I. Charton, S. Sabatini, D. Sepe: Compact monotone tall complexity one T-spaces. Preprint, arXiv:2307.04198, 2023.

[3] D. Cox, J. Little, H. Schenck: *Toric Varieties*, Graduate Studies in Mathematics, **124**, American Math. Society, 2010.

[4] L. Crespo, Á. Pelayo, F. Santos, Ewald's Conjecture and integer points in algebraic and symplectic toric geometry, preprint, arXiv:2310.10366, 2023.

[5] G. Ewald: On the Classification of Toric Fano Varieties. *Discrete Comput. Geom.* **3** (1988), 49–54.

[6] S. Franco, R-K Seong: Fano 3-folds, reflexive polytopes and brane brick models. *J. High Energ. Phys.* 2022, 8 (2022). https://doi.org/10.1007/JHEP08(2022)008

[7] L. Godinho, F. von Heymann, S. Sabatini: 12, 24 and beyond, *Advances in Mathematics* **319** (2017), 472–521.

[8] C. Haase, I. V. Melnikov: The Reflexive Dimension of a Lattice Polytope, *Ann. Comb.* **10** (2006), 211–217.

[9] B. Lorenz, A. Paffenholz: Smooth Reflexive Lattice Polytopes. Data available at `http://polymake.org/polytopes/paffenholz/www/fano.html`

[10] D. McDuff: The topology of toric symplectic manifolds. *Geometry and Topology*, **15** (2011)

[11] D. McDuff: Displacing Lagrangian toric fibers via probes, In *Low-dimensional and symplectic topology.* Proceedings of the 2009 Georgia International Topology Conference held at the University of Georgia, Athens, GA, May 18–29, 2009, Proc. Sympos. Pure Math. **82**, American Mathematical Society, Providence, RI, 2011, 131–160.

[12] B. Nill: Gorenstein toric Fano varieties, *Manuscripta Math.* **116** (2005), 183–210.

[13] B. Nill, personal communication. Conjecture posed, among other places, at the workshop *Combinatorial challenges in toric varieties*, American Institute of Mathematics (AIMS), 2009.

[14] M. Øbro: *An algorithm for the classification of smooth fano polytopes*, preprint, arXiv:0704.0049, 2007.

[15] M. Øbro: Classification of smooth Fano polytopes, Ph. D. thesis, University of Aarhus 2007. https://pure.au.dk/portal/en/publications/classification-of-smooth-fano-polytopes(781f9160-c4e2-11dc-88d5-000ea68e967b).html

[16] S. Payne: Frobenius splittings in toric varieties. *Algebra and Number Theory* **3**:1 (2009), 107–118.

Integer programs with nearly totally unimodular matrices: the cographic case

Manuel Aprile*[1], Samuel Fiorini†[2], Gwenaël Joret‡[2], Stefan Kober§[2], Michał T. Seweryn¶[2], Stefan Weltge‖[3], and Yelena Yuditsky**[2]

[1]Universitá degli studi di Padova, Via Trieste 63 Padova 35121, Italy
[2]Université libre de Bruxelles, Boulevard du Triomphe, B-1050 Brussels, Belgium
[3]Technische Universität München Boltzmannstraße 3, D-85748 Garching, Germany

Abstract

It is a notorious open question whether integer programs (IPs), with an integer coefficient matrix M whose subdeterminants are all bounded by a constant Δ in absolute value, can be solved in polynomial time. We answer this question in the affirmative if we further require that, by removing a constant number of rows and columns from M, one obtains a submatrix A that is the transpose of a network matrix.

Our approach focuses on the case where A arises from M after removing k rows only, where k is a constant. We achieve our result in two main steps, the first related to the theory of IPs and the second related to graph minor theory.

First, we derive a strong proximity result for the case where A is a general totally unimodular matrix: Given an optimal solution of the linear programming relaxation, an optimal solution to the IP can be obtained by finding a constant number of augmentations by circuits of A.

Second, for the case where A is transpose of a network matrix, we reformulate the problem as a maximum constrained integer potential problem on a graph G. We observe that if G is 2-connected, then it has no rooted $K_{2,t}$-minor for $t = \Omega(k\Delta)$. We leverage this to obtain a tree-decomposition of G into highly structured graphs for which we can solve the problem locally. This allows us to solve the global problem via dynamic programming.

1 Introduction

As for most computational problems that are NP-hard, the mere input size of an integer program (IP) does not seem to capture its difficulty. Instead, several works have identified additional parameters that significantly influence the complexity of solving IPs. These include the number of integer variables (Lenstra [LJ83], see also [Kan87, Dad12, RR23]), the number of inequalities for IPs in inequality form (Lenstra [LJ83]), the number of equations for IPs in equality form (Papadimitriou [Pap81],

*Email: manuel.aprile@unipd.it. Research of M. A. supported by FNRS through research project BD-DELTA (PDR 20222190, 2021–24)

†Email: samuel.fiorini@ulb.be. Research of S. F. supported by FNRS through research project BD-DELTA (PDR 20222190, 2021–24) and *King Baudouin Foundation* through project BD-DELTA-2 (convention 2023-F2150080-233051, 2023–26)

‡Email: gwenael.joret@ulb.be. Research of G. J. supported by FNRS (PDR "Product structure of planar graphs")

§Email: stefan.kober@ulb.be. Research of S. K. supported by Deutsche Forschungsgemeinschaft (German Research Foundation) under the project 451026932

¶Email: michal.seweryn@ulb.be. Research of M. S. supported by FNRS (PDR "Product structure of planar graphs")

‖Email: weltge@tum.de. Research of S. W. supported by the Deutsche Forschungsgemeinschaft (German Research Foundation) under the project 277991500/GRK220

**Email: yelena.yuditsky@ulb.be. Research of Y. Y. supported by FNRS as a Postdoctoral Researcher

see also [EW19]), and features capturing the block structure of coefficient matrices (see for instance [CEH+21, CEP+21, EHK+22, BKK+24, CKL+24]).

Another parameter that has received particular interest is the *largest subdeterminant* of the coefficient matrix, which already appears in several works concerning the complexity of linear programs (LPs) and the geometry of their underlying polyhedra [Tar86, DF94, BDSE+14, EV17] as well as proximity results relating optimal solutions of IPs and their LP relaxations [CGST86, PWW20, CKPW22]. Consider an IP of the form

$$\max \{p^\intercal x : Mx \leq b,\, x \in \mathbb{Z}^n\}\,, \tag{IP}$$

where M is an integer matrix that is *totally Δ-modular*, i.e., the determinants of square submatrices of M are all in $\{-\Delta, \ldots, \Delta\}$. It is a basic fact that if M is totally unimodular ($\Delta = 1$), then the optimum value of (IP) is equal to the optimum value of its LP relaxation, implying that (IP) can be solved in polynomial time. In a seminal paper by Artmann, Weismantel & Zenklusen [AWZ17], it is shown that (IP) is still polynomial-time solvable if $\Delta = 2$, leading to the conjecture that this may hold for every constant Δ. Recently, Fiorini, Joret, Yuditsky & Weltge [FJWY22] answered this in the affirmative under the further restriction that M has only two nonzeros per row or column. In particular, they showed that in this setting, (IP) can be reduced to the stable set problem in graphs with bounded odd cycle packing number [BFMRV14, CFHW20, CFH+20].

We remark that the algorithm of [AWZ17] even applies to full column rank matrices $M \in \mathbb{Z}^{m\times n}$ for which only the $(n \times n)$-subdeterminants are required to be in $\{-\Delta, \ldots, \Delta\}$ for $\Delta = 2$. Further results supporting the conjecture have been recently obtained by Nägele, Santiago & Zenklusen [NSZ22] and Nägele, Nöbel, Santiago & Zenklusen [NNSZ23] who considered the special case where all size-$(n \times n)$ subdeterminants are in $\{-\Delta, 0, \Delta\}$. Interestingly, the results of [AWZ17, NSZ22, NNSZ23] are crucially centered around a reformulation of (IP) where M becomes *totally unimodular up to removing a constant number of rows*, where the additional constraints capture a constant number of congruency constraints.

In an effort to provide more evidence for the above conjecture, we initiate the study of IPs in which M is totally Δ-modular and *nearly totally unimodular*, i.e., M becomes totally unimodular after removing a constant number of rows and columns. Note that without requirements on the subdeterminants, IPs with nearly totally unimodular coefficient matrices are still NP-hard. A famous example is the densest k-subgraph problem [BCC+10, Man17], which can be seen as an IP defined by a totally unimodular matrix with two extra rows (modeling a single equality constraint). A closely related example is the partially ordered knapsack problem [KS02], which is also strongly NP-hard. Another famous example is the exact matching (or red-blue matching) problem [Maa22, MVV87], for which no deterministic polynomial-time algorithm is known (yet).

While settling the conjecture for nearly totally unimodular coefficient matrices still seems to be a difficult undertaking, we can solve it for an important case: A celebrated result by Seymour [Sey80] states that network matrices and their transposes are the main building blocks of totally unimodular matrices. To any given (weakly) connected directed graph G and spanning tree T of G, one associates the *network matrix* $A \in \{0, \pm 1\}^{E(T)\times E(G-T)}$ such that $A_{e,(v,w)}$ is equal to 1 if e is traversed in forward direction on the unique v-w-path in T, is equal to -1 if it is traversed in backward direction, and is equal to 0 otherwise. Our main result is the following.

Theorem 1. *There is a strongly polynomial-time algorithm for solving the integer program* (IP) *for the case where M is totally Δ-modular for some constant Δ and becomes the transpose of a network matrix after removing a constant number of rows and columns.*

We achieve our result in two main steps, one related to the theory of integer programming and one related to graph minor theory. For the first step, we derive a new proximity result on distances between optimal solutions of IPs and their LP relaxations. A classic result of this type was established by Cook, Gerards, Schrijver, & Tardos [CGST86] who showed that if M is totally Δ-modular, (IP) is feasible, and x^* is an optimal solution of the LP relaxation, then there exists an optimal solution z^* of (IP)

with $\|x^* - z^*\|_\infty \leq n\Delta$. It is still open whether this bound can be replaced with a function in Δ only, see [CKPW22].

A convenient consequence of this result is that, given x^*, one can efficiently enumerate the possible values of z^* for a constant number of variables. In particular, if we are given the integer program (IP) with a totally Δ-modular coefficient matrix M that becomes totally unimodular after removing k rows and ℓ columns, we may simply guess the values of the variables corresponding to the ℓ columns and solve a smaller IP for each guess.

Thus, we may assume that M (is totally Δ-modular and) is of the form $M = \left[\begin{smallmatrix} A \\ W \end{smallmatrix}\right]$, where A is totally unimodular and W is an integer matrix with only k rows. For this class of IPs, we derive a considerably strengthened proximity result: Given an optimal solution x^* of the corresponding LP relaxation, there is an optimal solution z^* of (IP) where $\|x^* - z^*\|_\infty \leq f(k, \Delta)$ for some function f that depends only on k and Δ, again provided that (IP) is feasible. In fact, by bringing (IP) into equality form, we show that x^* can be rounded to a closeby integer point from which z^* can be reached by adding a number of conformal *circuits* of $\begin{bmatrix} A & \mathbf{I} \end{bmatrix}$ that can be bounded in terms of k and Δ only. Moreover, we observe that the fact that M is totally Δ-modular implies that every circuit c satisfies $\|Wc\|_\infty \leq \Delta$.

While these findings are valid for all totally unimodular matrices A, we will see that they can be crucially exploited for the case where A is the transpose of a network matrix, which we refer to as the *cographic case*. For these instances, it is convenient to reformulate the original problem (IP) as a particular instance of a *maximum constrained integer potential problem*

$$\max \left\{ p^\intercal y : \ell(v,w) \leq y(v) - y(w) \leq u(v,w) \text{ for all } (v,w) \in E(G),\, Wy \leq d,\, y \in \mathbb{Z}^{V(G)} \right\}, \quad \text{(MCIPP)}$$

where G is a directed graph, $p \in \mathbb{Z}^{V(G)}$, $\ell, u \in \mathbb{Z}^{E(G)}$, $W \in \mathbb{Z}^{[k] \times V(G)}$ and $d \in \mathbb{Z}^k$, and moreover each row of $p^\intercal$ or W sums up to zero. Notice that the first constraints are still given by a totally unimodular matrix, and hence we may regard $Wy \leq d$ as extra (or complicating) constraints. With this reformulation, the circuits of $\begin{bmatrix} A & \mathbf{I} \end{bmatrix}$ turn into vertex subsets $S \subseteq V(G)$ with the property that both induced subgraphs $G[S]$ and $G[\overline{S}]$ are (weakly) connected, where $\overline{S} := V(G) \setminus S$. We call such sets *doubly connected sets* or *docsets*. Using this notion, our previous findings translate to two strong properties of the instances of (MCIPP) we have to solve: First, every feasible instance has an optimal solution that is the sum of at most $f(k, \Delta)$ incidence vectors $\chi^S \in \{0,1\}^{V(G)}$, where S is a docset. Second, every docset S satisfies $\|W\chi^S\|_\infty \leq \Delta$.

Referring to the vertices whose variables appear with a nonzero coefficient in at least one of the extra constraints as *roots*, the second property above implies that roots cannot be arbitrarily distributed in the input graph. Roughly speaking, by carefully exploiting the structure of the instance, we will be able to guess $y(v)$ for each root v. Note that once all of these variables are fixed, the resulting IP becomes easy since its constraint matrix is totally unimodular. In fact, the guessing cannot be done for the whole graph at once and we will have to do it locally, and then combine the local optimal solutions via dynamic programming.

Our structural insights are based on the observation that our input graphs do not contain a *rooted $K_{2,t}$-minor*, where $t = 4k\Delta + 1$, provided that they are 2-connected. Here, the minors of a rooted graph (graph with a distinguished set of vertices called roots) are defined similarly as for usual graphs, with two differences: whenever some edge e is contracted we declare the resulting vertex as a root if and only if at least one of its ends is a root, and we have the possibility to remove a vertex from the set of roots. A rooted $K_{2,t}$ is said to be *properly rooted* if each one of the t vertices in the "large" side is a root. For the sake of simplicity, we say that a rooted graph contains a *rooted $K_{2,t}$-minor* if it has a rooted minor that is a properly rooted $K_{2,t}$, see Figure 1.

Our main structural result is a decomposition theorem for rooted graphs without rooted $K_{2,t}$-minor, see Theorem 2 below. It relies partly on several works about the structure of graphs excluding a minor, extending the original results of Robertson & Seymour within the graph minors project, more specifically on works by Diestel, Kawarabayashi, Müller & Wollan [DiKMW12], Kawarabayashi, Thomas

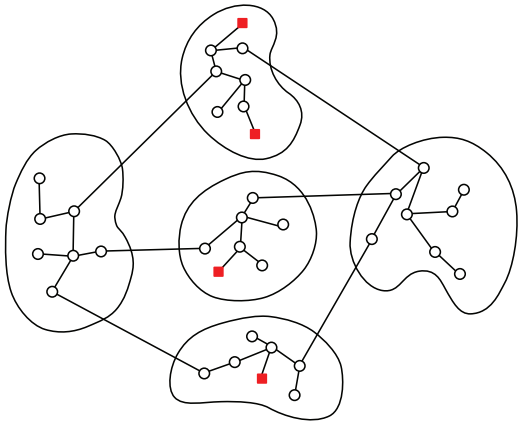


Figure 1: Subgraph containing a rooted $K_{2,3}$-minor. Roots are indicated with the red squares. Contracting all the edges in each of the five *branch sets* produces a properly rooted $K_{2,3}$.

& Wollan [KTW20], and Thilikos & Wiederrecht [TW22]. Furthermore, we use results of Böhme & Mohar [BM02] and Böhme, Kawarabayashi, Maharry & Mohar [BKMM08], to control the distribution of the roots in surface-embedded rooted graphs without rooted $K_{2,t}$-minors.

Our decomposition theorem is formulated in terms of a tree-decomposition of graph G. Recall that a *tree-decomposition* is a pair $(T, \mathcal{B})$ where T is a rooted tree (tree with a unique *root node*) and $\mathcal{B} = \{B_u : u \in V(T)\}$ is a collection of vertex subsets of G, called *bags*, such that for every vertex v of G the set of bags containing v induces a non-empty subtree of T, and for every edge e of G there is a bag that contains both ends of e. We define the *weak torso* of a bag B_u as the graph obtained from the induced subgraph $G[B_u]$ by adding a clique on $B_u \cap B_{u'}$ for each node $u' \in V(T)$ that is a child of u. Having stated these definitions, we are ready to state the (simplified version of our) decomposition theorem. See Figure 2 for an illustration.

Theorem 2 (simplified version)**.** *For every $t \in \mathbb{Z}_{\geq 1}$ there exists a constant $\ell = \ell(t)$ such that every 3-connected rooted graph G without a rooted $K_{2,t}$-minor admits a tree-decomposition $(T, \mathcal{B})$, where $\mathcal{B} = \{B_u : u \in V(T)\}$, with the following properties:*

(i) the bags B_u and $B_{u'}$ of two adjacent nodes $u, u' \in V(T)$ have at most ℓ vertices in common, and

(ii) for every node $u \in V(T)$, all but at most ℓ children $u' \in V(T)$ of u are leaves and the roots contained in the corresponding bags $B_{u'}$ are all contained in bag B_u, and

(iii) every node $u \in V(T)$ satisfies one of the following:

- *(a) bag B_u has at most ℓ vertices, or*
- *(b) u is a leaf and B_u has at most ℓ roots, all contained in the bag of the parent of u, or*
- *(c) after removing at most ℓ vertices, the weak torso of B_u becomes a 3-connected rooted graph that does not contain a rooted $K_{2,t}$-minor and has an embedding in a surface of Euler genus at most ℓ such that every face is bounded by a cycle, and all its roots can be covered by at most ℓ facial cycles.*

As we show, there is a polynomial-time algorithm that finds the tree-decomposition of Theorem 2 together with a polynomial-size collection $\mathcal{X}_u$ for each node $u \in V(T)$, containing all the possible intersections of a docset of G with the roots contained in bag B_u. This yields an efficient dynamic programming algorithm to solve the instances of (MCIPP) we are interested in, which proves Theorem 1.

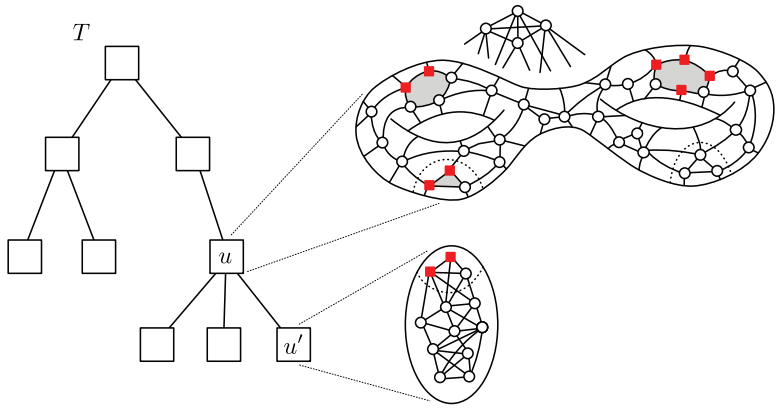


Figure 2: Illustrating the decomposition of Theorem 2. The decomposition tree T is shown on the left. The weak torsos of the two bags B_u and $B_{u'}$ are shown on the right. The top one satisfies (iii).(c), and the bottom one (iii).(b).

References

[AWZ17] Stephan Artmann, Robert Weismantel, and Rico Zenklusen. A strongly polynomial algorithm for bimodular integer linear programming. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 1206–1219. Association for Computing Machinery, 2017.

[BCC+10] Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: An $O(n^{1/4})$ approximation for densest k-subgraph. In *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10, pages 201–210, New York, NY, USA, 2010. Association for Computing Machinery.

[BDSE+14] Nicolas Bonifas, Marco Di Summa, Friedrich Eisenbrand, Nicolai Hähnle, and Martin Niemeier. On sub-determinants and the diameter of polyhedra. *Discrete & Computational Geometry*, 52(1):102–115, 2014.

[BFMRV14] Adrian Bock, Yuri Faenza, Carsten Moldenhauer, and Andres Jacinto Ruiz-Vargas. Solving the stable set problem in terms of the odd cycle packing number. In *34th International Conference on Foundation of Software Technology and Theoretical Computer Science*, volume 29 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 187–198. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014.

[BKK+24] Marcin Briański, Martin Koutecký, Daniel Král', Kristýna Pekárková, and Felix Schröder. Characterization of matrices with bounded graver bases and depth parameters and applications to integer programming. *Mathematical Programming*, 2024.

[BKMM08] Thomas Böhme, Ken-ichi Kawarabayashi, John Maharry, and Bojan Mohar. $K_{3,k}$-minors in large 7-connected graphs, 2008.

[BM02] Thomas Böhme and Bojan Mohar. Labeled $K_{2,t}$ minors in plane graphs. *Journal of Combinatorial Theory, Series B*, 84(2):291–300, 2002.

[CEH+21] Jana Cslovjecsek, Friedrich Eisenbrand, Christoph Hunkenschröder, Lars Rohwedder, and Robert Weismantel. Block-structured integer and linear programming in strongly polynomial and near linear time. In *Proceedings of the Thirty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '21, pages 1666–1681, USA, 2021. Society for Industrial and Applied Mathematics.

[CEP+21] Jana Cslovjecsek, Friedrich Eisenbrand, Michał Pilipczuk, Moritz Venzin, and Robert Weismantel. Efficient Sequential and Parallel Algorithms for Multistage Stochastic Integer Programming Using Proximity. In Petra Mutzel, Rasmus Pagh, and Grzegorz Herman, editors, *29th Annual European Symposium on Algorithms (ESA 2021)*, volume 204 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 33:1–33:14, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[CFH+20] Michele Conforti, Samuel Fiorini, Tony Huynh, Gwenaël Joret, and Stefan Weltge. The stable set problem in graphs with bounded genus and bounded odd cycle packing number. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2896–2915. SIAM, 2020.

[CFHW20] Michele Conforti, Samuel Fiorini, Tony Huynh, and Stefan Weltge. Extended formulations for stable set polytopes of graphs without two disjoint odd cycles. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 104–116. Springer, 2020.

[CGST86] William Cook, Albertus MH Gerards, Alexander Schrijver, and Éva Tardos. Sensitivity theorems in integer linear programming. *Mathematical Programming*, 34(3):251–264, 1986.

[CKL+24] Jana Cslovjecsek, Martin Koutecký, Alexandra Lassota, Michał Pilipczuk, and Adam Polak. *Parameterized algorithms for block-structured integer programs with large entries*, pages 740–751. SIAM, 2024.

[CKPW22] Marcel Celaya, Stefan Kuhlmann, Joseph Paat, and Robert Weismantel. Improving the Cook et al. proximity bound given integral valued constraints. In Karen Aardal and Laura Sanità, editors, *Integer Programming and Combinatorial Optimization*, pages 84–97, Cham, 2022. Springer International Publishing.

[Dad12] Daniel Dadush. *Integer programming, lattice algorithms, and deterministic volume estimation*. PhD thesis, Georgia Institute of Technology, 2012.

[DF94] Martin Dyer and Alan Frieze. Random walks, totally unimodular matrices, and a randomised dual simplex algorithm. *Math. Program.*, 64(1-3):1–16, 1994.

[DiKMW12] Reinhard Diestel, Ken ichi Kawarabayashi, Theodor Müller, and Paul Wollan. On the excluded minor structure theorem for graphs of large tree-width. *Journal of Combinatorial Theory, Series B*, 102(6):1189–1210, 2012.

[EHK+22] Friedrich Eisenbrand, Christoph Hunkenschröder, Kim-Manuel Klein, Martin Koutecký, Asaf Levin, and Shmuel Onn. An algorithmic theory of integer programming, 2022.

[EV17] Friedrich Eisenbrand and Santosh Vempala. Geometric random edge. *Math. Program.*, 164(1-2):325–339, 2017.

[EW19] Friedrich Eisenbrand and Robert Weismantel. Proximity results and faster algorithms for integer programming using the Steinitz lemma. *ACM Transactions on Algorithms (TALG)*, 16(1):1–14, 2019.

[FJWY22] Samuel Fiorini, Gwenaël Joret, Stefan Weltge, and Yelena Yuditsky. Integer programs with bounded subdeterminants and two nonzeros per row. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 13–24, 2022.

[Kan87] Ravi Kannan. Minkowski's convex body theorem and integer programming. *Mathematics of operations research*, 12(3):415–440, 1987.

[KS02] Stavros G. Kolliopoulos and George Steiner. Partially-ordered knapsack and applications to scheduling. In Rolf Möhring and Rajeev Raman, editors, *Algorithms — ESA 2002*, pages 612–624, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[KTW20] Ken-ichi Kawarabayashi, Robin Thomas, and Paul Wollan. Quickly excluding a non-planar graph. *arXiv:2010.12397*, 2020.

[LJ83] Hendrik W Lenstra Jr. Integer programming with a fixed number of variables. *Mathematics of operations research*, 8(4):538–548, 1983.

[Maa22] Nicolas El Maalouly. Exact matching: Algorithms and related problems. *arXiv:2203.13899*, 2022.

[Man17] Pasin Manurangsi. Almost-polynomial ratio eth-hardness of approximating densest k-subgraph. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 954–961, New York, NY, USA, 2017. Association for Computing Machinery.

[MVV87] Ketan Mulmuley, Umesh V Vazirani, and Vijay V Vazirani. Matching is as easy as matrix inversion. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 345–354, 1987.

[NNSZ23] Martin Nägele, Christian Nöbel, Richard Santiago, and Rico Zenklusen. Advances on strictly Δ-modular ips. In *Proceedings of the 24th Conference on Integer Programming and Combinatorial Optimization (IPCO '23)*, pages 393–407, 2023.

[NSZ22] Martin Nägele, Richard Santiago, and Rico Zenklusen. Congruency-constrained TU problems beyond the bimodular case. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2743–2790. SIAM, 2022.

[Pap81] Christos H Papadimitriou. On the complexity of integer programming. *Journal of the ACM (JACM)*, 28(4):765–768, 1981.

[PWW20] Joseph Paat, Robert Weismantel, and Stefan Weltge. Distances between optimal solutions of mixed-integer programs. *Mathematical Programming*, 179(1):455–468, 2020.

[RR23] Victor Reis and Thomas Rothvoss. The subspace flatness conjecture and faster integer programming. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 974–988. IEEE, 2023.

[Sey80] Paul D Seymour. Decomposition of regular matroids. *Journal of combinatorial theory, Series B*, 28(3):305–359, 1980.

[Tar86] Eva Tardos. A strongly polynomial algorithm to solve combinatorial linear programs. *Operations Research*, 34(2):250–256, 1986.

[TW22] Dimitrios M Thilikos and Sebastian Wiederrecht. Killing a vortex. *arXiv:2207.04923*, 2022.

Limit theorems for the Erdős–Rényi random graph conditioned on being a cluster graph*

Martijn Gösgens[†1], Lukas Lüchtrath[‡2], Elena Magnanini[§2], Marc Noy[¶3], and Élie de Panafieu [‖4]

[1]Eindhoven University of Technology
[2]Weierstrass Institute, Berlin
[3]Universitat Politènica de Catalunuya, Barcelona
[4]Nokia Bell Labs France, Massy, France

Abstract

We investigate the structure of the random graph $G(n,p)$ on n vertices with constant (not depending on n) connection probability p, conditioned on the rare event that every component is a clique. We show that a phase transition occurs at $p = 1/2$, contrary to the dense $G(n,p)$ model. Our proofs are based on probabilistic methods, generating functions and analytic combinatorics.

1 Introduction

A cluster graph is a graph that is the disjoint union of complete graphs. In this paper, we consider the Erdős–Rényi (ER) random graph $G(n,p)$ on n vertices with connection probability p, conditioned on the rare event of being a cluster graph; in our situation $p \in (0,1)$ does not depend on n. We refer to such a graph as a random cluster graph (RCG). The initial motivation for our study was the observation that a random cluster graph is a good candidate for a Bayesian prior distribution in the context of community detection [3], which is the task of partitioning the nodes of a network into communities.

Secondly, it is an interesting probabilistic object due to its rare event character. Forming a cluster graph is no standard behaviour of the ER random graph and it is fascinating how drastically its behaviour is effected by this conditioning; an evidence of this fact is that the random graph obtained after this conditioning overcomes a phase transition in p (that is not present in the dense ER model).

Finally, when ignoring the edges and only considering each cluster as a set, a cluster graph represents a partition of the whole vertex set. The case $p = 1/2$ then coincides with the uniform distribution over set partitions. Uniform set partitions are standard objects in enumerative and probabilistic combinatorics [4]. Varying the value of p is a natural way of weighting partitions and thus the RCG gives rise to more general, non-uniform underlying distributions.

After stating our main results, we briefly explain the proof techniques, based on probabilistic methods and analytic combinatorics [2]. We conclude with a sketch of further results and concluding remarks.

*This project has been initiated during the RandNET Summer School and Workshop on Random Graphs in Eindhoven in August 2022. It was supported by the RandNET project, MSCA-RISE - Marie Sklodowska-Curie Research and Innovation Staff Exchange Programme (RISE), Grant agreement 101007705

[†]Email: research@martijngosgens.nl
[‡]Email: luechtrath@wias-berlin.de
[§]Email: magnanini@wias-berlin.de
[¶]Email: marc.noy@upc.edu
[‖]Email: depanafieuelie@gmail.com

2 Main results

We let $\mathbf{CG}_{n,p}$ denote a random cluster graph with parameters n and p. Our main quantities of interest are the number of connected components (clusters) in $\mathbf{CG}_{n,p}$, denoted by $\mathbf{C}_{n,p}$, the number of edges denoted by $\mathbf{M}_{n,p}$, and the degree $\mathbf{D}_{n,p}$ chosen independent and uniformly at random from the vertex set. Our main results concerning these parameters are the following.

Theorem 1 (Number of clusters in the RCG)**.** *Consider the random cluster graph $\mathbf{CG}_{n,p}$ on $n \in \mathbb{N}$ vertices and ER edge probability $p \in (0,1)$ and the number of its clusters $\mathbf{C}_{n,p}$.*

1. *If $p > 1/2$, then*
$$\lim_{n\to\infty} \mathbb{P}(\mathbf{C}_{n,p} = 1) = 1.$$
Put differently, $\mathbf{CG}_{n,p} = K_n$ with high probability.

2. *If $p = 1/2$, then $\mathbf{C}_{n,p}$ obeys a central limit theorem. That is,*
$$\frac{\mathbf{C}_{n,p} - \mathbb{E}\mathbf{C}_{n,p}}{\sqrt{\mathrm{Var}(\mathbf{C}_{n,p})}} \longrightarrow \mathcal{N}(0,1),$$
in distribution, as $n \to \infty$. Moreover,
$$\mathbb{E}\mathbf{C}_{n,p} \sim \frac{n}{\log n} \quad \text{and} \quad \mathrm{Var}(\mathbf{C}_{n,p}) \sim \frac{n}{\log(n)^2}.$$

3. *If $p < 1/2$, then $\mathbf{C}_{n,p}$ obeys a central limit theorem. That is,*
$$\frac{\mathbf{C}_{n,p} - \mathbb{E}\mathbf{C}_{n,p}}{\sqrt{\mathrm{Var}(\mathbf{C}_{n,p})}} \longrightarrow \mathcal{N}(0,1),$$
in distribution, as $n \to \infty$. Moreover,
$$\mathbb{E}\mathbf{C}_{n,p} \sim \sqrt{\frac{\log(1-p) - \log p}{2}} \frac{n}{\sqrt{\log n}} \quad \text{and} \quad \mathrm{Var}(\mathbf{C}_{n,p}) = \Theta\left(\frac{n}{\log(n)^{3/2}}\right).$$

Theorem 2 (Number of edges in the RCG)**.** *Consider the random cluster graph $\mathbf{CG}_{n,p}$ on $n \in \mathbb{N}$ vertices and ER edge probability $p \in (0,1)$ and its number of edges $\mathbf{M}_{n,p}$.*

1. *If $p > 1/2$, then*
$$\lim_{n\to\infty} \mathbb{P}\Big(\mathbf{M}_{n,p} = \binom{n}{2}\Big) = 1.$$

2. *If $p = 1/2$, then $\mathbf{M}_{n,p}$ obeys a central limit theorem. That is,*
$$\frac{\mathbf{M}_{n,1/2} - \mathbb{E}\mathbf{M}_{n,1/2}}{\sqrt{\mathrm{Var}(\mathbf{M}_{n,1/2})}} \longrightarrow \mathcal{N}(0,1)$$
in distribution as $n \to \infty$. Moreover,
$$\mathbb{E}\mathbf{M}_{n,1/2} \sim n\log n \quad \text{and} \quad \mathrm{Var}(\mathbf{M}_{n,1/2}) = \Theta(n\log(n)^2).$$

3. *If $p < 1/2$, then $\mathbf{M}_{n,p}$ obeys a central limit theorem. That is,*
$$\frac{\mathbf{M}_{n,p} - \mathbb{E}\mathbf{M}_{n,p}}{\sqrt{\mathrm{Var}(\mathbf{M}_{n,p})}} \longrightarrow \mathcal{N}(0,1)$$
in distribution as $n \to \infty$. Moreover,
$$\mathbb{E}\mathbf{M}_{n,p} \sim n\sqrt{\frac{\log n}{2(\log(1-p) - \log p)}} \quad \text{and} \quad \mathrm{Var}(\mathbf{M}_{n,p}) = \Theta\left(n\log(n)^{3/2}\right).$$

Theorem 3 (Degree distribution of the RCG)**.** *Consider the random cluster graph* $\mathbf{CG}_{n,p}$ *on* $n \in \mathbb{N}$ *vertices and ER edge probability* $p \in (0,1)$ *and the degree* $\mathbf{D}_{n,p}$ *of a uniformly chosen vertex.*

1. *If* $p > 1/2$*, then*
$$\lim_{n\to\infty} \mathbb{P}(\mathbf{D}_{n,p} = n-1) = 1.$$
2. *If* $p = 1/2$*, then for a Poisson random variable* X_n *with parameter* $\log n - \log\log n + o(1)$*, we have*
   (a) *for all* $z \in \mathbb{C}$*,*
   $$\mathbb{E}z^{\mathbf{D}_{n,1/2}} \sim \mathbb{E}z^{X_n}.$$
   That is, the probability generating function of $\mathbf{D}_{n,1/2}$ *and the one of* X_n *are asymptotically the same.*
   (b) *Additionally,*
   $$\lim_{n\to\infty} \mathrm{d}_{TV}(\mathbf{D}_{n,1/2}, X_n) = 0.$$
3. *If* $p < 1/2$*, then* $\mathbb{E}\mathbf{D}_{n,p} = \Theta(\sqrt{\log n})$*. Moreover, for each* $\lambda \in [0,1)$ *there exists a subsequence* $(n_k)_{k\in\mathbb{N}}$ *such that*
   $$\mathbf{D}_{n_k,p} - \left\lfloor \sqrt{\frac{2\log n_k}{\log(1-p) - \log p}} - 1 - \frac{1}{\log(1-p) - \log p} \right\rfloor \longrightarrow X_\lambda$$
   in distribution as $k \to \infty$*, where* X_λ *is defined by*
   $$\mathbb{P}(X_\lambda = d) = \frac{\left(\frac{p}{1-p}\right)^{(d-\lambda)^2/2}}{\sum_{d'\in\mathbb{Z}} \left(\frac{p}{1-p}\right)^{(d'-\lambda)^2/2}}$$
   for all $d \in \mathbb{Z}$*.*

Notice that the fact that $\mathbf{D}_{n,p} = \Theta(\sqrt{\log n})$ when $p < 1/2$ follows directly from Theorem 2. However, to obtain the distribution full of $\mathbf{D}_{n,p}$ is technically quite involved.

3 Generating functions and analytic combinatorics

By conditioning $G(n,p)$ we loose the independence of the $G(n,p)$ model. To overcome this fact we use *counting* techniques. Let $\mathcal{F}$ be a class (invariant under isomorphims) of labelled graphs, and let $\mathcal{F}_{n,m}$ be the graphs in $\mathcal{F}$ with n vertices and m edges. We denote by $n(G)$ number of vertices of G, and by $m(G)$ the number of edges. The exponential generating function (EGF) associated to $\mathcal{F}$ is

$$F(w,z) = \sum_{G\in\mathcal{F}} w^{m(G)} \frac{z^{n(G)}}{n(G)!},$$

so that $|\mathcal{F}_{n,m}| = n![w^m z^n]F(w,z)$. In particular, the EGF of the class of non-empty cliques is

$$C(w,z) = \sum_{n\geq 1} w^{\binom{n}{2}} \frac{z^n}{n!}$$

From now on we use freely the symbolic method, as described in [2]. In particular, since a cluster graph is a *set* of cliques, its EGF is $\exp(uC(w,z))$, where the variable u marks components.

It is easy to see that the distribution of random cluster graphs is equal to

$$\mathbb{P}(\mathbf{CG}_{n,p} = G) = \frac{\left(\frac{p}{1-p}\right)^{m(G)}}{B_n(p/1-p)},$$

where the *partition function* $B_n(w)$ is given by $B_n(w) = n![z^n]e^{C(w,z)}$. We notice that $B_n(1)$ is the n-th Bell number, counting partitions of a set of size n. From here one easily obtains the probability generating functions (PGF) of the main parameters. Recall that the PGF of an integer-valued non-negative random variable X is defined as

$$\mathrm{PGF}_X(u) = \mathbb{E}(e^X) = \sum_{k\geq 0} \mathbb{P}(X = k)u^k.$$

Proposition 4. *Let $\mathbf{M}_{n,p}$, $\mathbf{C}_{n,p}$ and $\mathbf{D}_{n,p}$ as in Section 2. Set $B_n(w) = n![z^n]e^{C(w,z)}$ as before, and and $w = \frac{p}{1-p}$. The probability generating functions of these random variables are equal to*

$$\mathrm{PGF}_{\mathbf{M}_{n,p}}(u) = \frac{B_n(u\,w)}{B_n(w)},$$
$$\mathrm{PGF}_{\mathbf{C}_{n,p}}(u) = \frac{[z^n]e^{uC(w,z)}}{[z^n]e^{C(w,z)}},$$
$$\mathrm{PGF}_{\mathbf{D}_{n,p}}(u) = \frac{[z^n]C_1(w, u\,z)e^{C(w,z)}}{u[z^n]C_1(w, z)e^{C(w,z)}}.$$

In order to obtain limit theorems we use the moment generating function (alternatively, the characteristic function $\mathbb{E}(e^{itX})$)

$$\mathbb{E}(e^{tX}) = \mathrm{PGF}_X(e^t).$$

Our main tool is Levy's continuity theorem:

Theorem 5. *Let X_n and Y be real valued random variables. If $\mathbb{E}(e^{tX_n})$ converges pointwise for t in a neighborhood of 0 to $\mathbb{E}(e^{tY})$, then X_n converges in law to Y.*

In particular, if there exists μ_n and σ_n such that, pointwise for s in a neighborhood of 0

$$\mathrm{PGF}_{X_n}(e^{s/\sigma_n}) \sim e^{s\mu_n/\sigma_n}e^{s^2/2} \qquad \text{as } n \to \infty$$

then the renormalized random variables $X_n^\star = \frac{X_n-\mu_n}{\sigma_n}$ converges to the standard normal distribution.

In order to apply the previous result we need to estimate the corresponding PGFs as $n \to \infty$. This is not an easy task, due mainly to the quadratic exponent $\binom{n}{2}$ in the expression for $C(w, z)$. In fact, to compute moments, we need more generally to estimate the derivatives of $C(w, z)$ with respect to z. This is the most technical part of our work, involving Cauchy integrals, saddle-point methods, and the so-called Hayman admissible functions [2], among other tools.

We observe that the size of the largest block in the $p = 1/2$ regime is known to be $\Theta(\log n)$. When $p < 1/2$ it should be $\Theta(\sqrt{\log n})$ due to concentration, but we have not worked out the details.

4 Further results

In this final section, we collect further results on random cluster graphs.

The critical window when $p \downarrow \frac{1}{2}$. We know that when $p > 1/2$ the random cluster gaph $\mathbf{CG}_{n,p}$ is almost surely a single clique. If we let $p = p(n) > 1/2$, we are interested in the scale at which $\mathbf{CG}_{n,p}$ becomes a single clique.

Proposition 6. *Let $q \in (0, 1)$ and $p_n(q)$ defined by*

$$\mathbb{P}(\mathbf{C}_{n,p_n(q)} = 1) = q.$$

Then

$$p_n(q) = \frac{1}{2} + \frac{\log(n)}{2n} + O\left(\frac{\log\log n}{n}\right).$$

Notice that the precise value of q is not important, in fact it only appears in the error term.

In addition, we show that there exists no 'almost complete' regime. For instance, for any sequence $p_n \in [0,1]$ we have

$$\mathbb{P}(\mathbf{C}_{n,p_n(q)} = K_{n-1} \cup K_1) \to 0, \quad \text{as } n \to \infty,$$

and similarly for $\mathbf{C}_{n,p_n(q)} = K_{n-r} \cup \{ \text{ small cliques } \}$, for fixed $r > 0$.

The upercritical regime ($p > \frac{1}{2}$). In this regime we know that there is only one clique w.h.p. Our next result is an asymptotic expansion for $\mathbb{P}(\mathbf{C}_{n,p} = K_n)$. First notice that if $w = \frac{p}{1-p} > 1$ then $C(w,z) = \sum_{n\geq 1} w^{\binom{n}{2}} \frac{z^n}{n!}$ has zero radius of convergence. Using recent tools for estimating coefficients of divergent series [1] we show that

Proposition 7.

$$\mathbb{P}(\mathbf{CG}_{n,p} = K_n) = 1 + \sum_{m=1}^{R-1} w^{-mn} P_m(n) + O\left(w^{-Rn} n^R\right)$$

where $P_m(n)$ are certain polynomials and $R \geq 0$ is an integer

The first terms in the expansion are $\mathbb{P}(\mathbf{CG}_{n,p} = K_n) = 1 - nw \cdot w^{-n} + O(n^2 w^{-2n})$.

The sparse regime $p \to 0$. We focus on the case where p_n decreases like a monomial $p_n = n^{-\alpha+o(1)}$ for $\alpha > 0$. We prove that in this regime, the degree distribution concentrates around one or two values. We first show how α should be chosen to concentrate this distribution around a particular degree d:

Theorem 8. *Let $d \in \mathbb{N} \cup \{0\}$ and consider a limiting sequence $p_n = n^{-\frac{2}{(d+1)^2}+o(1)}$. Then*

$$\mathbb{P}(\mathbf{D}_n = d) \to 1.$$

Furthermore, for any other $d' \in \mathbb{N} \cup \{n\}$, the degree distribution satisfies

$$\mathbb{P}(\mathbf{D}_n = d') = n^{-\left(\frac{d'-d}{d+1}\right)^2 + o(1)}. \tag{1}$$

In the field of random graphs, the case $p_n = \lambda/n$ is one of the most interesting regimes, known as the *sparse regime*. The next lemma shows that in this regime, the degree distribution is concentrated around two values, rather than one:

Proposition 9. *Let $\lambda > 0$ and consider the sequence $p_n \sim \lambda/n$, then*

$$\mathbb{P}(\mathbf{D}_n = 0) \to \frac{\sqrt{4\lambda+1}-1}{2\lambda}, \quad \mathbb{P}(\mathbf{D}_n = 1) \to 1 - \frac{\sqrt{4\lambda+1}-1}{2\lambda},$$

In particular, the sequence $p_n \sim 1/n$ yields $\mathbb{P}(\mathbf{D}_n = 0) \to \rho^{-1}$, where $\rho = \frac{\sqrt{5}+1}{2}$ is the golden ratio.

Conditioning to other classes of graphs. For fixed $p \in (0,1)$, let $F(n,p)$ the random graph $G(n,p)$ conditioned to be a forest. $F(n,p)$ behaves like a random uniform forest, in the sense that the number of edges is linear and asymptotically Gaussian, and the number of components is asymptotically Poisson distributed; only the constants depend on p and thereis no phase transition. The same is true conditioning on being planar, or related classes of graphs.

In order to get a situation like for random cluster graphs, we believe that one should need to condition on classes of graphs admitting superlinear number of edges.

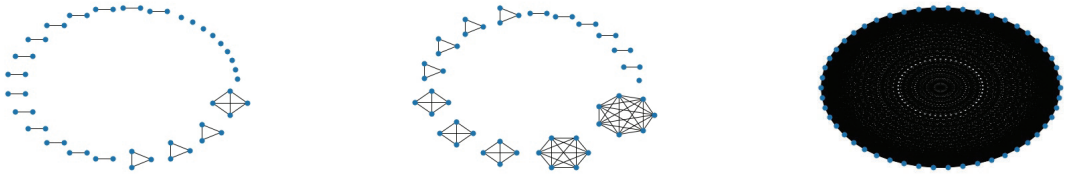
Sampling. How can we sample a random cluster graph $\mathbf{CG}_{n,p}$? Certainly not sampling with rejection, since the event $G(n,p)$ being a cluster graph is extremely rare. Instead we sample first the size of one clique and the rest by induction. Let $\mathbf{S}_{n,p}$ be the size of the clique containing vertex 1. Then we have

Proposition 10.

$$\mathbb{P}(\mathbf{S}_{n,p} = s) = \binom{n}{s-1} \left(\frac{p}{1-p}\right)^{\binom{s}{2}} \frac{B_{n-s}(p/(1-p))}{B_n(p/(1-p))},$$

where $B_n(w)$ is as in Section 3.

Once we sample the size s of the first clique according to the previous distribution, we can sample recursively on the remaining $n-s$ vertices. Below we show examples of this procedure for (from left to right) $p = 0.25, p = 0.51$ and $p = 0.53$.



Application to community detection. We come finally to the original motivation for our research. *Community detection* aims at partitioning the nodes of a network into *communities*: sets of vertices that are more strongly connected to each other than to the remainder of the network. A popular approach is to optimize a quantity known as *text*modularity over the set of partitions. A resolution parameter controls the granularity of the obtained clustering

Given a graph G and cluster graphCG representing a potential partition, and a resolution parameter γ, the modularity is defined as

$$M(G, CG, \gamma) = \frac{1}{m(G)} \left(m(G \cap CG) - \gamma \cdot m(CG)\right)$$

The main goal is to understand modularity better and how to choose γ. For that one can use $\mathbf{CG}_{n,p}$ as a model for a prior distribution. When the communities have sizes close to $\log n$, setting $p = 1/2$ will likely lead to detecting communities of the desired granularity. But when the communities are significantly smaller than $\log n$, one should choose $p > 1/2$. Preliminary inevstigations indicate that the choice of

$$p_n = \frac{1}{2} + \frac{\log n}{2n} + O\left(\frac{\log\log n}{n}\right)$$

leads to significantly better community-detection performance.

References

[1] S. Dovgal, K. Nurligareev. *Asymptotics for graphically divergent series: dense digraphs and 2-SAT formulae*, arXiv:2310.05282 (2023).

[2] P. Flajolet, R. Sedgewick. *Analytic Combinatorics*. Cambridge U. Press, 2009.

[3] S. Fortunato, M. Barthélemy. Resolution limit in community detection. *Proc. Nat. Acad. Sci.* 104.1 (2007), 36–41.

[4] V. N. Sachkov. *Probabilistic methods in combinatorial analysis*. Cambridge U. Press, 1997.

On the sum of several finite subsets in $\mathbb{R}^2$ *

Mario Huicochea[†1], René González-Martínez [‡2], Amanda Montejano[§3], and David Suárez[¶3]

[1]Universidad Autónoma de San Luis Potosí
[2]Universidad Autónoma de Zacatecas
[3]UMDI, Facultad de Ciencias, UNAM Juriquilla, Querétaro, México

Abstract

Let $h \geq 2$ be an integer, $\mathbf{u} \in \mathbb{R}^2 \setminus \{(0,0)\}$, and $A_1, A_2, \ldots, A_h$ be nonempty finite subsets of $\mathbb{R}^2$. For each $i \in \{1, 2, \ldots, h\}$, denote by m_i the number of lines parallel to the line generated by the vector $\mathbf{u}$ that intersect A_i. We show that

$$|A_1 + A_2 + \ldots + A_h| \geq \left(\left(\sum_{i=1}^{h} \frac{|A_i|}{m_i} \right) - (h-1) \right) \left(\left(\sum_{i=1}^{h} m_i \right) - (h-1) \right)$$

generalizing a statement of D. J. Grynkiewicz and O. Serra for $h = 2$. We also characterize the case of equality; that is, we describe the structure of finite 2-dimensional subsets of $\mathbb{R}^2$ which are extremal with respect to the inequality above. This also generalizes a result of G. A Freiman, D. Grynkiewicz, O. Serra and Y. V. Stanchescu.

1 Introduction

One of the most important problems in Additive Number Theory has been to determine nontrivial lower bounds for the cardinality of $A_1 + A_2 + \ldots + A_h = \{\mathbf{a}_1 + \mathbf{a}_2 + \ldots + \mathbf{a}_h \mid \mathbf{a}_i \in A_i \text{ for each } i \in \{1, 2, \ldots, h\}\}$, where $A_1, A_2, \ldots, A_h$ are nonempty finite subsets of an abelian group G, see for instance [2, 7, 9, 11, 12, 13]. Particularly, there is interesting recent progress concerning this problem in $\mathbb{R}^d$. In this work, we only focus our attention in $\mathbb{R}^2$. Given $\mathbf{u} \in \mathbb{R}^2 \setminus \{(0,0)\}$ and $A_1, A_2, \ldots, A_h$ nonempty finite subsets of $\mathbb{R}^2$, we give a lower bound of $|A_1 + A_2 + \ldots + A_h|$ in terms of the number of lines parallel to the line generated by $\mathbf{u}$ which intersect A_i, for each $i \in \{1, 2, \ldots, h\}$. This was already done for two sets ($h = 2$) by Grynkiewicz and Serra [10]. Moreover, Freiman, Grynkiewicz, Serra and Stanchescu characterized the extremal 2-dimensional sets attaining such lower bound [5]. We generalize both results for any integer $h \geq 2$.

Given $\mathbf{u} \in \mathbb{R}^2$, we denote by $\langle \mathbf{u} \rangle$ the subspace (line) generated by $\mathbf{u}$. Let $\phi_{\langle \mathbf{u} \rangle} : \mathbb{R}^2 \to \mathbb{R}^2/\langle \mathbf{u} \rangle$ the natural projection modulo $\langle \mathbf{u} \rangle$. For a finite subset A of $\mathbb{R}^2$, let $\phi_{\langle \mathbf{u} \rangle}(A) = \{\phi_{\langle \mathbf{u} \rangle}(\mathbf{a}) \,|\, \mathbf{a} \in A\}$. Thus, if $\mathbf{u} \neq (0,0)$, $|\phi_{\langle \mathbf{u} \rangle}(A)|$ is the number of lines parallel to $\langle \mathbf{u} \rangle$ that intersect A. As we already mentioned in the previous paragraph, Grynkiewicz and Serra were able to find a lower bound of $|A+B|$ for nonempty subsets A and B of $\mathbb{R}^2$ in terms of $|\phi_{\langle \mathbf{u} \rangle}(A)|$ and $|\phi_{\langle \mathbf{u} \rangle}(B)|$. Here we give the precise statement.

*Part of this research was supported by DGAPA PAPIIT IG100822.
†Email: dym@cimat.mx
‡Email: reneglzmtz@gmail.com
§Email: amandamontejano@ciencias.unam.mx.
¶Email: suardavid@hotmail.com

Theorem 1 (Grynkiewicz-Serra)**.** *Let A and B be nonempty finite subsets of $\mathbb{R}^2$. For every $\mathbf{u} \in \mathbb{R}^2 \setminus \{(0,0)\}$,*

$$|A+B| \geq \left(\frac{|A|}{m} + \frac{|B|}{n} - 1\right)(m+n-1), \tag{1}$$

where $m = |\phi_{\langle \mathbf{u} \rangle}(A)|$ and $n = |\phi_{\langle \mathbf{u} \rangle}(B)|$.

Proof. see [10, Thm. 1.3]. □

We generalize Theorem 1 for an arbitrary number of nonempty finite subsets of $\mathbb{R}^2$.

Theorem 2. *Let $h \geq 2$ be an integer, and let $A_1, \dots A_h$ be nonempty finite subsets of $\mathbb{R}^2$. For every $\mathbf{u} \in \mathbb{R}^2 \setminus \{(0,0)\}$,*

$$|A_1 + A_2 + \ldots + A_h| \geq \left(\left(\sum_{i=1}^{h} \frac{|A_i|}{m_i}\right) - (h-1)\right)\left(\left(\sum_{i=1}^{h} m_i\right) - (h-1)\right) \tag{2}$$

where $m_i = |\phi_{\langle \mathbf{u} \rangle}(A_i)|$ for each $i \in \{1, 2, \dots, h\}$.

In order to characterized the extremal sets attaining equality in (2) we need to define some sets called *trapezoids*. For the sake of clarity, we will use a definition that varies slightly from the one originally presented in [5].

Definition 3. *Let $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$ be an ordered base of $\mathbb{R}^2$, and let $m, h, c \in \mathbb{Z}$ with $(h-1) + (m-1)c \geq 0$. We say that a finite nonempty 2-dimensional set $A \subset \mathbb{R}^2$ is a* trapezoid of type $T_{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}(m, h, c)$ *if there is a vector $\boldsymbol{v} \in \mathbb{R}^2$ such that*

$$M_{\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle}(A) + \boldsymbol{v} = \{(x, y) \in \mathbb{R}^2 \,|\, 0 \leq y \leq m-1,\ 0 \leq x \leq (h-1) + cy\},$$

where $M_{(\boldsymbol{u}_1, \boldsymbol{u}_2)} : \mathbb{R}^2 \to \mathbb{R}^2$ is the linear mapping that leads the ordered base $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$ to the canonical ordered base $\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle$.

We shall note that the example showed in [5] as a standard trapezoid $T(6, 19, 2, -1)$ correspond to the translation of any trapezoid of type $T_{\langle (1,2),(0,1) \rangle}(19, 6, -3)$, see Figure 1. In general, a standard trapezoid $T(m, h, c, d)$ corresponds, after applying the linear transformation given by the matrix $M_{\langle (1,d),(0,1) \rangle} = \begin{pmatrix} -d & 1 \\ 1 & 0 \end{pmatrix}$, to a translation of any trapezoid of type $T_{\langle (1,d),(0,1) \rangle}(m, h, d-c)$.

Theorem 4. *Let $A_1 \dots, A_k$ be nonempty finite 2-dimensional subsets of $\mathbb{R}^2$, and let $\boldsymbol{u} \in \mathbb{R}^2$. If*

$$|A_1 + \cdots + A_k| = \left(\sum_{i=1}^{k} \left(\frac{|A_i|}{m_i}\right) - (k-1)\right)\left(\sum_{i=1}^{k} (m_i) - (k-1)\right), \tag{3}$$

where $m_i = |\phi_{\langle \boldsymbol{u} \rangle}(A_i)|$ for $1 \leq i \leq k$, then each A_i is a trapezoid of type $T_{\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle}(m_i, h_i, c)$ for some ordered base $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$, and some integers $h_1, \dots, h_k$, with common slope c.

The paper is organized as follows: Section 2 contains auxiliary results needed for proving Theorems 2 and 4. In particular, we present some properties of the technique known as (linear) compression. The proof of Theorem 2 is completed in Section 3. To prove Theorem 4 is a bit more technical. We present the strategy in Section 4.

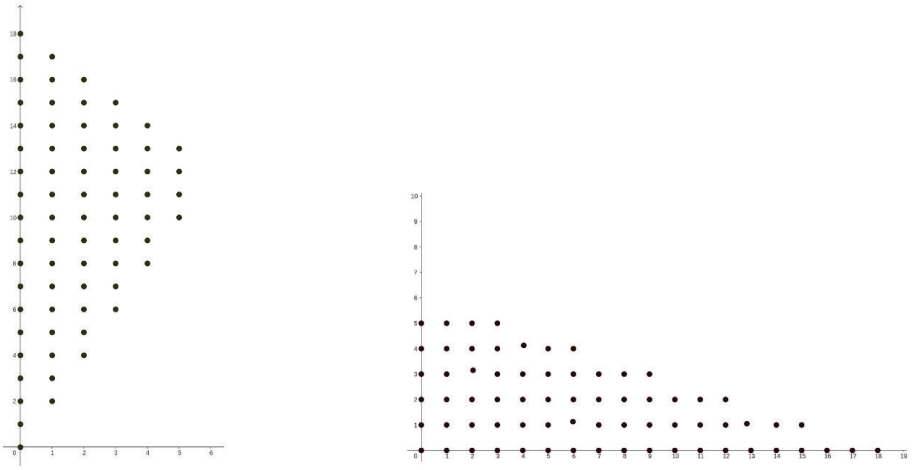


Figure 1: The standard trapezoid $T(6, 19, 2, -1)$ (depicted on the left) given as an example in [5] corresponds to the trapezoid of type $T_{\langle(1,2),(0,1)\rangle}(6, 19, -3)$ translated to the origin (depicted on the right). This can be seen by applying the linear transformation $M_{\langle(1,2),(0,1)\rangle}$ to $T(6, 19, 2, -1)$.

2 Preliminaries

Let $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ be an ordered basis of $\mathbb{R}^2$. For a finite subset $A \subset \mathbb{R}^2$ and $i \in \{1, 2\}$, the *linear compression* of A with respect to $\mathbf{u}_i$, denoted by $\mathbf{C}_i(A)$, is defined as follows. Take $j \in \{1, 2\} \setminus \{i\}$ and let $\mathbf{C}_i(A)$ be the set satisfying that for each $\mathbf{x} \in \langle \mathbf{u}_j \rangle$,

$$\phi_{\langle \mathbf{u}_j \rangle} \left(\mathbf{C}_i(A) \cap (\langle \mathbf{u}_i \rangle + \mathbf{x}) \right) = \{0, \mathbf{u}_i, 2\mathbf{u}_i, \cdots, (r-1)\mathbf{u}_i\} + \langle \mathbf{u}_j \rangle,$$

where $r = |A \cap (\langle \mathbf{u}_i \rangle + \mathbf{x})|$ and, if $r = 0$, we consider $\mathbf{C}_i(A) \cap (\langle \mathbf{u}_i \rangle + \mathbf{x}) = \emptyset$. The *compression of* A with respect to the ordered basis $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ is then defined by $\mathbf{C}(A) = \mathbf{C}_2(\mathbf{C}_1(A))$. Several properties of compression can be found in [1, 6, 8, 10]; we just need a few of them. Observe that we have by definition that

$$|A| = |\mathbf{C}(A)|, \tag{4}$$

and

$$|\phi_{\langle \mathbf{u}_1 \rangle}(A)| = |\phi_{\langle \mathbf{u}_1 \rangle}(\mathbf{C}(A))|. \tag{5}$$

One of the main properties of compression is the following.

Lemma 5. *For any nonempty finite subsets $A_1, A_2 \subset \mathbb{R}^2$, and an ordered basis $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$ of $\mathbb{R}^2$, it happens that $\mathbf{C}(A_1 + A_2) \supseteq \mathbf{C}(A_1) + \mathbf{C}(A_2)$. In particular, $|A_1 + A_2| \geq |\mathbf{C}(A_1) + \mathbf{C}(A_2)|$.*

Proof. See [6, Lemma 3.4]. □

We will also make use of the following well known fact.

Theorem 6 (Folklore)**.** *Let $A_1, \ldots, A_h$ be finite nonempty subsets of a torsion-free abelian group. Then*

$$|A_1 + A_2 + \ldots + A_h| \geq \left(\sum_{i=1}^{h} |A_i| \right) - (h - 1),$$

and the equality is achieved when $A_1, A_2, \ldots, A_h$ are arithmetic progressions with the same common difference.

Proof. See for instance [11, Theorem 1.4]. □

Let $A_1, A_2, \ldots, A_h$ be nonempty finite subsets of $\mathbb{R}^2$ and, for each $i \in \{1, 2, \ldots, h\}$, let $\mathbf{C}(A_i)$ be the compression of A_i with respect to the ordered basis $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle$. The projection $\phi_{\langle \mathbf{u}_1 \rangle}$ is a linear mapping, and the definition of $\mathbf{C}(A_i)$ implies that $\phi_{\langle \mathbf{u}_1 \rangle}(\mathbf{C}(A_i))$ is an arithmetic progression with difference $\mathbf{u}_2$ for each $i \in \{1, 2, \ldots, h\}$. From these facts and Theorem 6, it follows that

$$\left|\phi_{\langle \mathbf{u}_1 \rangle}(\mathbf{C}(A_1) + \mathbf{C}(A_2) + \ldots + \mathbf{C}(A_h))\right| = \left(\sum_{i=1}^{h} \left|\phi_{\langle \mathbf{u}_1 \rangle}(\mathbf{C}(A_i))\right|\right) - (h-1). \tag{6}$$

In order to prove Theorem 2 , we prove the next inequality.

Lemma 7. *Let $A_1, A_2, \ldots, A_h$ be nonempty finite subsets of $\mathbb{R}^2$, and let $\mathbf{C}(A_i)$ be the compression of A_i with respect to the ordered basis $\langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$. Then,*

$$|A_1 + A_2 + \ldots + A_h| \geq |\mathbf{C}(A_1) + \mathbf{C}(A_2) + \ldots + \mathbf{C}(A_h)|. \tag{7}$$

Proof. By induction on h taking $A = A_1 + A_2 + \ldots + A_{h-1}$ and A_h, with the use of Lemma 5. □

3 Proof of Theorem 2

Proof. We proceed by induction on h. If $h = 2$, the statement follows by Theorem 1. Consider now the sets $A = \mathbf{C}(A_1) + \mathbf{C}(A_2) + \ldots + \mathbf{C}(A_{h-1})$ and $B = \mathbf{C}(A_h)$ where, for each $i \in \{1, 2, \ldots, h\}$, $\mathbf{C}(A_i)$ is the compression of A_i with respect to the ordered basis $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = \langle \mathbf{u}, \mathbf{u}^{\perp} \rangle$. Set $m = |\phi_{\langle \mathbf{u} \rangle}(A)|$ and $n = |\phi_{\langle \mathbf{u} \rangle}(B)|$. Then

$$\begin{aligned} |A_1 + \ldots + A_h| &\geq |\mathbf{C}(A_1) + \ldots + \mathbf{C}(A_{h-1}) + \mathbf{C}(A_h)| && \Big(\text{by Lemma 7}\Big) \\ &= |A + B| \\ &\geq \left(\frac{|A|}{m} + \frac{|B|}{n} - 1\right)(m + n - 1). && \Big(\text{by Thm. 1}\Big) \end{aligned} \tag{8}$$

Hence, by definition and (5),

$$n = |\phi_{\langle \mathbf{u} \rangle}(B)| = |\phi_{\langle \mathbf{u} \rangle}(\mathbf{C}(A_h))| = |\phi_{\langle \mathbf{u} \rangle}(A_h)| = m_h, \tag{9}$$

and

$$\begin{aligned} m = |\phi_{\langle \mathbf{u} \rangle}(A)| &= |\phi_{\langle \mathbf{u} \rangle}(\mathbf{C}(A_1) + \mathbf{C}(A_2) + \ldots + \mathbf{C}(A_{h-1}))| \\ &= \left(\sum_{i=1}^{h-1} \left|\phi_{\langle \mathbf{u} \rangle}(\mathbf{C}(A_i))\right|\right) - (h-2) && \Big(\text{by (6)}\Big) \\ &= \left(\sum_{i=1}^{h-1} \left|\phi_{\langle \mathbf{u} \rangle}(A_i)\right|\right) - (h-2) && \Big(\text{by (5)}\Big) \\ &= \left(\sum_{i=1}^{h-1} m_i\right) - (h-2). \end{aligned} \tag{10}$$

Thus (9) and (10) yield that

$$m + n - 1 = \left(\sum_{i=1}^{h-1} m_i\right) - (h-2) + m_h - 1 = \left(\sum_{i=1}^{h} m_i\right) - (h-1). \tag{11}$$

By (4) we know $|B| = |\mathbf{C}(A_h)| = |A_h|$, and so, by (9), we get

$$\frac{|B|}{n} = \frac{|A_h|}{m_h}. \tag{12}$$

Now, since (4) and (5) imply $|\mathbf{C}(A_i)| = |A_i|$ and $|\phi_{\langle \mathbf{u}\rangle}(\mathbf{C}(A_i))| = |\phi_{\langle \mathbf{u}\rangle}(A_i)| = m_i$, it follows by definition and the induction hypothesis that

$$|A| = |\mathbf{C}(A_1) + \mathbf{C}(A_2) + \ldots + \mathbf{C}(A_{h-1})| \geq \left(\left(\sum_{i=1}^{h-1} \frac{|A_i|}{m_i}\right) - (h-2)\right)\left(\left(\sum_{i=1}^{h-1} m_i\right) - (h-2)\right). \quad (13)$$

Thus, (10) and (13) imply

$$\frac{|A|}{m} \geq \left(\sum_{i=1}^{h-1} \frac{|A_i|}{m_i}\right) - (h-2). \quad (14)$$

Finally, substituying (11), (12) and (14) in (8), we obtain

$$\begin{aligned} |A_1 + \ldots + A_h| &\geq \left(\frac{|A|}{m} + \frac{|B|}{n} - 1\right)(m+n-1) \\ &\geq \left(\left(\sum_{i=1}^{h-1} \frac{|A_i|}{m_i}\right) - (h-2) + \frac{|A_h|}{m_h} - 1\right)\left(\left(\sum_{i=1}^{h} m_i\right) - (h-1)\right), \\ &= \left(\left(\sum_{i=1}^{h} \frac{|A_i|}{m_i}\right) - (h-1)\right)\left(\left(\sum_{i=1}^{h} m_i\right) - (h-1)\right), \end{aligned}$$

and the prove is completed. □

4 Sketch of the proof of Theorem 4

Observe that, if A is trapezoid of type $T_{\langle \mathbf{u}_1, \mathbf{u}_2\rangle}(m, h, c)$ for some ordered base $\langle \mathbf{u}_1, \mathbf{u}_2\rangle$ of $\mathbb{R}^2$, and some integers m, h and c satisfying $(h-1) + (m-1)c \geq 0$ then, by definition, there is a vector $\mathbf{v}$ such that

$$M_{\langle \mathbf{u}_1, \mathbf{u}_2\rangle} A + \mathbf{v} = \bigcup_{i=0}^{m-1} \{(x, i) | 0 \leq x \leq (h-1) + ci\}.$$

Therefore,

$$|A| = m\left(h + \frac{c(m-1)}{2}\right). \quad (15)$$

With the use of (15) it is not hard to prove, by induction on k, the following.

Lemma 8. *Let $A_1, \ldots, A_k$ be trapezoids of type $T_{\langle \mathbf{u}_1, \mathbf{u}_2\rangle}(m_i, h_i, c)$ for some ordered base $\langle \mathbf{u}_1, \mathbf{u}_2\rangle$, integers m_i and h_i, for each $1 \leq i \leq h$, and a common slope c. Then $A_1 + \cdots + A_k$ is a trapezoid of type $T_{\langle \mathbf{u}_1, \mathbf{u}_2\rangle}(\sum_{i=1}^k m_i - (k-1), \sum_{i=1}^k h_i - (k-1), c)$.*

One of the key parts of the proof of Theorem 4 was to generalize a beautiful lemma which was used to prove Theorem 1 as well as the characterization of the extremal cases, see [10, 5]. For the sake of clarity, we present here only the statement for $k = 3$ and a sketch of its proof.

Lemma 9. *Let I, J, K be nonempty finite subsets of $\mathbb{R}$ with $min(|I|, |J|, |K|) \geq 2$. Let $a = \{a_i\}_I$, $b = \{b_j\}_J$ and $c = \{c_k\}_K$ sequences with $a_i, b_j, c_k > 0$ for $i \in I$, $j \in J$ and $k \in K$. For each $t \in I + J + K$, let $u_t(a, b, c) = max\{a_i + b_j + c_{k-i-j} : i \in I, j \in J, k - i - j \in K\}$. If*

$$\frac{1}{|I|}\sum_{i\in I} a_i + \frac{1}{|J|}\sum_{j\in J} b_j + \frac{1}{|K|}\sum_{k\in K} c_k \leq \frac{1}{|I| + |J| + |K| - 2} \sum_{t\in I+J+K} u_t(a, b, c). \quad (16)$$

If the equality holds then I, J and K are arithmetic progressions with common difference and the sequences a, b and c are also arithmetic progressions with common difference.

Proof. (sketch) For a sequence $\{x_i\}_{i\in L}$, denote by $\overline{x} = \frac{1}{|L|}\sum_{i\in L} x_i$ its average value. If $y = \{y_i\}_{i\in M}$ and $z = \{y_i\}_{i\in N}$ are also sequences, denote by $u^+(x,y,z)$ the subsequence of the $|L|+|M|+|N|-2$ elements in the sequence $u(x,y,z) = \{u_t(x,y,z) : t \in L+M+N\}$ which is well-defined in view of Theorem 6. Let $d = \{u_t(b,c) : t \in J+K\}$. First we shall prove that $u(a,b,c) = u(a,d)$ and then we need to prove that $\overline{u^+(a,d^+)} \leq \overline{u^+(a,d)}$, which will lead us to show that

$$\overline{u^+(a,b,c)} \leq \left(\frac{1}{|I|+|J|+|K|-2}\right) \sum_{t\in I+J+K} u_t(a,b,c). \tag{17}$$

From this position it is not hard to prove that

$$\frac{1}{|I|}\sum_{i\in I} a_i + \frac{1}{|J|}\sum_{j\in J} b_j + \frac{1}{|K|}\sum_{k\in K} c_k \leq \frac{1}{|I|+|J|+|K|-2} \sum_{t\in I+J+K} u_t(a,b,c) \tag{18}$$

Now, suppose that the equality holds if (16), we can see that I, J and K are arithmetic progressions with common difference. From here, one has to work to show that actually, the sequences a, b and c are arithmetic progression with a common difference. □

To prove Theorem 4 we define one set for each $1 \leq i \leq k$ as $I_i = \phi_{\langle \mathbf{u}\rangle}(A_i)$, and work to obtain the base $\langle \mathbf{u}_1, \mathbf{u}_2\rangle$ and the parameters of the trapezoids in terms of the differences of the arithmetic progression given by Lemma 9.

References

[1] B. Bollobás and I. Leader, *Sums in the grid*, Discrete Math. 162 (1996), 31-48.

[2] G. A. Freiman, *Foundations of a Structural Theory of Set Addition*, Transl. Math. Monogr. 37 Amer. Math. Soc. (1973).

[3] A. Mudgal, *Difference sets in higher dimensions*, Math. Proc. Cambridge Philos. Soc. 171 (2021), 467-480.

[4] D. Conlon and J. Lim,, *Difference sets in* $\mathbb{R}^d$, prepint available at arXiv:2110.09053 [math.CO].

[5] G. A. Freiman, D. Grynkiewicz, O. Serra, and Y.V. Stanchescu, *Inverse additive problems for Minkowski sumsets I*, Collectanea mathematica, 63(3) (2012), 261-286.

[6] R. J. Gardner and P. Gronchi, *A Brunn-Minkowski inequality for the integer lattice*, Trans. Amer. Math. Soc. 353 (2001), 3995-4024.

[7] A. Geroldinger and I. Ruzsa, *Combinatorial Number Theory and Additive Group Theory*, Birkhäuser (2009).

[8] B. Green and T. Tao, *Compressions, convex geometry and the Freiman-Bilu theorem*, Q. J. Math. 57 (2006), 495-504.

[9] D. J. Grynkiewicz, *Structural Additive Theory*, Developments in Mathematics 30 Springer (2013).

[10] D. J. Grynkiewicz and O. Serra, *Properties of two dimensional sets with small sumset*, J. Combin. Theory Ser. A 117 (2010), 164-188.

[11] M. Nathanson, *Additive Number Theory*, Graduate Texts in Mathematics 165 Springer (1996).

[12] I. Z. Ruzsa, *Sums of sets in several dimensions*, Combinatorica 14 (1994), 485-490.

[13] T. Tao and H. V. Vu, *Additive Combinatorics*, Cambridge Stud. Adv. Math. 105 Cambridge University Press (2006).

On a conjecture concerning the roots of Ehrhart polynomials of symmetric edge polytopes from complete multipartite graphs*

Max Kölbl†[1]

[1]Dept. of Pure and Applied Mathematics, Osaka University, 565-0871 Suita

Abstract

In [7], Higashitani, Kummer, and Michałek pose a conjecture about the symmetric edge polytopes of complete multipartite graphs and confirm it for a number of families in the bipartite case. We confirm that conjecture for a number of new classes following the authors' methods and we present a more general result which suggests that the methods in their current form might not be enough to prove the conjecture in full generality.

1 Introduction

This paper is an extended abstract of our recent work [8] for the Discrete Mathematics Day 2024. It contains the main results from Sections 3 and 4.

A *lattice polytope* is a convex polytope $P \subset \mathbb{R}^n$ which can be written as the convex hull of finitely many elements of $\mathbb{Z}^n$. Lattice polytopes arise naturally from attempts to endow combinatorial objects with a geometric structure. A family of lattice polytopes that has garnered some attention in recent years is that of *symmetric edge polytopes*, which are a type of graph polytopes. For graphs, we will henceforth use the notation $G = (V, E)$ where V denotes the set of *vertices* and E denotes the set of *edges* of G. Given a graph $G = (V, E)$, we thus define its symmetric edge polytope as follows

$$P_G = \operatorname{conv}\{\pm(e_v - e_w) \colon \{v, w\} \in E\} \subset \mathbb{R}^{|V|}.$$

Here, the vectors e_v are elements that form a lattice basis of $\mathbb{Z}^{|V|}$. For more context on symmetric edge polytopes, see e.g. [6, 9].

Next, we define the *lattice-point enumerator* of a set $S \subset \mathbb{R}^n$ as the function $E_S \colon \mathbb{N} \to \mathbb{N}$ via $E_S(k) = |kS \cap \mathbb{Z}^n|$. If S is a lattice polytope, E_S is a polynomial which we call the *Ehrhart polynomial* of S. The generating function of an Ehrhart polynomial is called *Ehrhart series* and can be written as

$$\operatorname{ehr}_P(t) = \sum_{k \geq 0} E_P(k) t^k = \frac{h^*(t)}{(1-t)^{d+1}},$$

where $h^*(t)$ is a polynomial with non-negative integer coefficients of degree d or less. We call this polynomial the *h^*-polynomial* of P. The Ehrhart polynomial and the h^*-polynomial hold valuable information about the underlying polytope, such as its (normalised) volume and the volume of its boundary. A specifically remarkable piece of information encoded by the h^*-polynomial is that of reflexivity: A lattice polytope is called *reflexive* if its polar dual is also a lattice polytope. By a result by Hibi [5], P is reflexive if and only if its h^*-polynomial is *palindromic*, i.e., $h^*_P(t) = \sum_{i=0}^{d} h^*_i(t)$ satisfies $h^*_i = h^*_{d-i}$ for all $0 \leq i \leq d$, and its degree is equal to $\dim P$.

*The full version of this work can be found in [8] and will be published elsewhere.

†Email: max.koelbl@ist.osaka-u.ac.jp

With some basic knowledge of generating functions (see e.g. [12]), one can check that knowing the Ehrhart polynomial of a lattice polytope amounts to knowing its h^*-polynomial. However, the converse is also true. Given the h^*-polynomial $h^*_P(t) = \sum_{i=0}^{d} h^*_i t^i$ of some lattice polytope P, the Ehrhart polynomial can be written as

$$E_P(x) = \sum_{i=0}^{d} h^*_i \binom{d+x-i}{d}.$$

For more context on Ehrhart theory, see e.g. [1]. One aspect of research in Ehrhart theory is the study of the *roots* of Ehrhart polynomials when their domain and range are extended from $\mathbb{N}$ to $\mathbb{C}$. For example in the case of reflexive polytopes, their Ehrhart polynomial roots exhibit symmetry not only across the real axis (i.e. if z is a root then so is its complex conjugate) but also, due to Ehrhart-Macdonald reciprocity and palindromicity of the h^*-polynomial, across the *canonical line*, i.e. set

$$\mathrm{CL} = \left\{ z \in \mathbb{C} \colon \Re(z) = -\frac{1}{2} \right\}$$

where $\Re(z)$ denotes the real part of $z \in \mathbb{C}$. This is to say that if z is a root then so is $-1-z$. Thus, it is natural to ask, what kind of polytopes have all of their Ehrhart polynomial roots *on* CL. First steps in this direction were made in [2] and [10], albeit in different contexts. In [9], the study of CL-*polytopes*, i.e., polytopes with all their Ehrhart polynomial roots on CL, has been initiated as a field of study in its own right. For low dimensions, a full classification was found in [4]. Some classes of examples include cross-polytopes, standard reflexive simplices, and root polytopes of type A.

For the rest of the paper, let $K_{a_1,\dots,a_k}$ denote the complete multipartite graph with k multipartite classes of of sizes a_1 through a_k. The Ehrhart polynomial of $P_{K_{a_1,\dots,a_k}}$ shall be denoted by $E_{a_1,\dots,a_k}$. In [7], the authors studied the roots of $E_{2,n}$ and $E_{3,n}$ and were able to prove that $P_{K_{2,n}}$ and $P_{K_{3,n}}$ are CL-polytopes. This extends the case of cross-polytopes, which are unimodularly equivalent to the symmetric edge polytopes of $K_{1,n}$. They accomplished that by using the technique of *interlacing polynomials*. Let f, g be polynomials of degree $d+1$ and d with roots $\left\{-\frac{1}{2} + i\, a_1, -\frac{1}{2} + i\, a_2, \dots, -\frac{1}{2} + i\, a_{d+1}\right\}$ and $\left\{-\frac{1}{2} + i\, b_1, -\frac{1}{2} + i\, b_2, \dots, -\frac{1}{2} + i\, b_d\right\}$ respectively for $a_j, b_j \in \mathbb{R}$. Then we say that g CL-*interlaces* f if

$$a_1 \le b_1 \le a_2 \le b_2 \le \cdots \le b_d \le a_{d+1}.$$

For more on the theory of interlacing polynomials, see [3]. The authors gave the following conjecture.

Conjecture 1 (Conjecture 4.10 in [7]). *(i) For any complete multipartite graph $K_{a_1,\dots,a_k}$ the Ehrhart polynomial $E_{a_1,\dots,a_k}$ has its roots on* CL.

(ii) Suppose $a_1 \le \cdots \le a_k$. The two Ehrhart polynomials $E_{a_1,\dots,a_k}$ and $E_{a_1,a_2,\dots,a_k-1}$ interlace on CL.

In Section 2, we will prove CL-ness of $E_{1,1,n}, E_{1,2,n}$, and $E_{1,1,1,n}$, as well as some conditional results (Theorem 9), using the techniques from [7]. In Section 3, we will investigate the connection between the γ-*vector* of the h^*-polynomial of an Ehrhart polynomial and the existence of recursive relations that generalise those in [7]. However, we also provide evidence for why their methods might not be enough to prove Conjecture 1 any further.

2 New recursive relations

In this section, we gather new evidence for Conjecture 1. First, we state the relevant h^*-polynomials.

Proposition 2 (Theorem 4.1 in [6]). *For all $a, b \ge 0$ let $h^*_{a,b}(t)$ denote the h^*-polynomial of the symmetric edge polytope of $K_{a+1,b+1}$. Then*

$$h^*_{a,b}(t) = \sum_{i=0}^{\min\{a,b\}} \binom{2i}{i}\binom{a}{i}\binom{b}{i} t^i (1+t)^{a+b+1-2i}.$$

Proposition 3. *The h^*-polynomials of the symmetric edge polytopes of the graphs $K_{1,m,n}$, $K_{1,1,1,n}$, and $K_{2,2,n}$, are given as follows.*

(a) $h^*_{1,m,n}(t) = \sum_{i=0}^{\min(m,n)} \binom{2i}{i}\binom{m}{i}\binom{n}{i} t^i (1+t)^{m+n-2i}$

(b) $h^*_{1,1,1,n}(t) = 3(n-1)n(1+t)^{n-2}t^2 + 2(2n+1)(1+t)^n t + (1+t)^{n+2}$

(c) $h^*_{2,2,n}(t) = 20\binom{n}{3}(1+t)^{n-3}t^3 + 2\binom{3n}{2}(1+t)^{n-1}t^2 + 2\binom{3n+1}{1}(1+t)^{n+1}t + (1+t)^{n+3}$

Since the proof is very technical, we will proceed directly to introducing a proposition which supplies a useful tool for checking CL-interlacing.

Proposition 4 (Lemmas 2.3, 2.4, 2.5 in [7])**.** *Let $f, g, h_1, \ldots, h_n$ be Ehrhart polynomials of reflexive polytopes such that $\deg f = \deg g + 1 = \deg h_i + 2$ for all $1 \leq i \leq n$. Assume the identity*

$$f(x) = (2x+1)\,\alpha\, g(x) + \sum_{i=1}^{n} \alpha_i\, h_i(x)$$

where $\alpha, \alpha_i > 0$ for all i. Then $\sum_{i=1}^n \alpha_i\, h_i$ CL-interlaces g if for every i, h_i CL-interlaces g. Also, the following are equivalent.

(a) $\sum_{i=1}^n \alpha_i\, h_i$ CL-*interlaces* g,

(b) g CL-*interlaces* f.

If this is the case, $(2x+1)\sum_{i=1}^n \alpha_i\, h_i$ CL-*interlaces* f.

An important class of reflexive polytopes is the class of *cross-polytopes* which are defined as the convex hull of the vectors $\pm e_1, \pm e_2, \ldots, \pm e_n \in \mathbb{R}^n$. As mentioned in the introduction, they are unimodularly equivalent to $P_{K_{1,n}}$. The Ehrhart polynomial of the n-dimensional cross-polytope (the n-th *cross-polynomial*) is given by

$$\mathcal{C}_n(x) = \sum_{k=0}^{n} \binom{n}{k}\binom{n+x-k}{n}.$$

Cross-polynomials are the first class of examples to showcase the usefulness of Proposition 4.

Proposition 5 (Example 3.3 in [7])**.** *For any $n \geq 2$, cross-polynomials satisfy the recursive relation*

$$\mathcal{C}_n(x) = \frac{1}{n}\,(2x+1)\,\mathcal{C}_{n-1}(x) + \frac{n-1}{n}\,\mathcal{C}_{n-2}(x).$$

Other classes of examples were found by Higashitani, Kummer, and Michałek in [7]. The authors found three recursive relations among Ehrhart polynomials $E_{1,n}, E_{2,n}, E_{3,n}$.

Proposition 6 (Proposition 4.5 in [7])**.** *The following relations hold:*

$$\begin{aligned}
E_{2,n}(x) &= \frac{1}{2}\,(2x+1)\,E_{1,n}(x) + \frac{1}{2}\,E_{1,n-1}(x),\\
E_{2,n}(x) &= \frac{1}{n}\,(2x+1)\,E_{2,n-1}(x) + \frac{1}{2n}\,\left(n\,E_{1,n-1}(x) + (n-2)\,(2x+1)\,E_{1,n-2}(x)\right),\\
E_{3,n+1}(x) &= \frac{(2x+1)(3n^2+13n+16)}{8(n^2+5n+6)}\,E_{2,n+1}(x)\\
&\quad + \frac{n^3+13n^2+18n}{8(n-1)(n^2+5n+6)}\,E_{2,n}(x) + \frac{4n^3+9n^2-13n-32}{8(n-1)(n^2+5n+6)}\,E_{1,n+1}(x).
\end{aligned}$$

Using this, the authors were able to prove the following result.

Proposition 7 (Lemmas 4.6-4.8, Theorem 4.9 in [7])**.** *The following statements hold for every positive integer n.*

(a) $E_{1,n}$ CL-*interlaces* $E_{1,n+1}$.

(b) $E_{1,n}$ *and* $(2x+1)E_{1,n-1}$ CL-*interlace* $E_{2,n}$.

(c) $E_{2,n}$ CL-*interlaces* $E_{2,n+1}$.

(d) $E_{2,n}$ CL-*interlaces* $E_{3,n}$.

In particular, for every positive integer n, the Ehrhart polynomial of $K_{m,n}$ is a CL-*polynomial if $m \leq 2$.*

To extend this result, we start by finding new recursive relations.

Proposition 8. *For every $n \geq 2$ there exist non-negative rational numbers $\alpha_1, \ldots, \alpha_{35}$ such that the following statements hold.*

$$
\begin{aligned}
E_{1,1,n}(x) &= \alpha_1 (2x+1) E_{1,n}(x) + \alpha_2 E_{1,n-1}(x),\\
E_{1,1,n+1}(x) &= \alpha_3 (2x+1) E_{1,1,n}(x) + \alpha_4 E_{1,1,n-1}(x) + \alpha_5 E_{1,n}(x),\\
E_{1,2,n}(x) &= \alpha_6 (2x+1) E_{1,1,n}(x) + \alpha_7 E_{1,1,n-1}(x) + \alpha_8 E_{1,n}(x),\\
E_{1,2,n+1}(x) &= \alpha_9 (2x+1) E_{1,2,n}(x) + \alpha_{10} E_{1,2,n-1}(x) + \alpha_{11} E_{1,1,n}(x) + \alpha_{12} E_{1,n+1}(x)\\
E_{1,1,1,n}(x) &= \alpha_{13} (2x+1) E_{1,1,n}(x) + \alpha_{14} E_{1,1,n-1}(x) + \alpha_{15} E_{1,n}(x)\\
E_{4,n}(x) &= \alpha_{16} (2x+1) E_{3,n}(x) + \alpha_{17} E_{3,n-1}(x) + \alpha_{18} E_{2,n}(x) + \alpha_{19} E_{1,n+1}(x),\\
E_{3,n+1}(x) &= \alpha_{20} (2x+1) E_{3,n}(x) + \alpha_{21} E_{3,n-1}(x) + \alpha_{22} E_{2,n}(x) + \alpha_{23} E_{1,n+1}(x),\\
E_{2,2,n}(x) &= \alpha_{24} (2x+1) E_{1,2,n}(x) + \alpha_{25} E_{1,2,n-1}(x) + \alpha_{26} E_{1,1,n}(x) + \alpha_{27} E_{1,n+1}(x),\\
E_{1,3,n}(x) &= \alpha_{28} (2x+1) E_{1,2,n}(x) + \alpha_{29} E_{1,2,n-1}(x) + \alpha_{30} E_{1,1,n}(x) + \alpha_{31} E_{1,n+1}(x),\\
E_{1,1,1,n+1}(x) &= \alpha_{32} (2x+1) E_{1,1,1,n}(x) + \alpha_{33} E_{1,1,1,n-1}(x) + \alpha_{34} E_{1,1,n}(x) + \alpha_{35} E_{1,n+1}(x).
\end{aligned}
$$

These relations can be obtained algorithmically. We explain the method using the first identity as an example. The identity holds if and only if it holds after replacing $E_{1,1,n}(x)$, $(2x+1)E_{1,n}(x)$, and $E_{1,n-1}(x)$ by their respective generating functions. All three of these can be obtained from h^*-polynomials given in Propositions 2 and 3. After dividing by the left-hand side, the right hand side becomes a rational function whose numerator polynomial has coefficients which are either constant or linear in α_1 and α_2. The left-hand side becomes 1. Thus, on the right-hand side, we can compare the coefficients of the numerator polynomial with those of the denominator polynomial and solve for α_1 and α_2. Note however, that in general there need not be a solution. A SAGEMATH [11] implementation of this algorithm is available on

`https://github.com/maxkoelbl/seps_multipartite_graphs/`.

We can state the main result of this section.

Theorem 9. *The following statements hold for every positive integer n.*

(a) $E_{1,n}$ CL-*interlaces* $E_{1,1,n}$.

(b) $E_{1,1,n}$ CL-*interlaces* $E_{1,1,n+1}$.

(c) $E_{1,1,n}$ CL-*interlaces* $E_{1,2,n}$.

(d) $E_{1,1,n}$ CL-*interlaces* $E_{1,1,1,n}$.

(e) $E_{3,n}$ CL-*interlaces* $E_{4,n}$ *if* $E_{1,n+1}$ CL-*interlaces* $E_{3,n}$.

(f) $E_{1,2,n}$ CL-*interlaces* $E_{1,3,n}$ *if* $E_{1,n+1}$ CL-*interlaces* $E_{1,2,n}$.

(g) $E_{1,2,n}$ CL-*interlaces* $E_{2,2,n}$ *if* $E_{1,n+1}$ CL-*interlaces* $E_{1,2,n}$.

In particular, for every positive integer n, $E_{x,y,z,n}$ is a CL-*polynomial for $x+y+z \leq 3$ and $x, y, z \geq 0$.*

3 Recursive relations and the γ-vector

Looking at the recursive relations in Propositions 6 and 8, we may notice that as the parameters $a_1, \ldots, a_{k-1}$ of the multipartite graphs increase, then so does the complexity of the identities involving them. The results of this section show that this is not a coincidence. The key object here is the *γ-vector* of the h^*-polynomial of an Ehrhart polynomial.

Definition 10. *Let h be a palindromic polynomial of degree d. We define the* γ-vector *as the polynomial $\sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} \gamma_i t^i$ such that $h(t) = \sum_{i=0}^{\lfloor \frac{d}{2} \rfloor} \gamma_i (1+t)^{d-2i} t^i$. We call the degree of the γ-vector the* γ-degree *of h.*

Proposition 11. *Let p be a polynomial of degree d and let h be a polynomial defined by*

$$h(t) = (1-t)^{d+1} \sum_{k \geq 0} p(k)\, t^k.$$

If h is a palindromic polynomial with γ-vector γ, we get

$$p(x) = \sum_{i=0}^{\deg \gamma} (-1)^i c_i \, \mathcal{C}_{d-2i}(x).$$

where $c_i = \sum_{j=i}^{\deg \gamma} \frac{1}{4^j} \binom{j}{i} \gamma_j$.

In the setting of Proposition 11, we call the γ-degree of h the *cross-degree* of p. It is the key ingredient of this section's main theorem.

Theorem 12. *Let f be a degree $d+1$ polynomial with cross-degree $m+1$, let g be a degree d polynomial with cross-degree m, and let h_i be degree $d-1$ polynomials with cross-degree i for $1 \leq i \leq m$. Then there exist unique real numbers $\alpha, \alpha_1, \alpha_2, \ldots, \alpha_m$ which satisfy*

$$f(x) = (2x+1)\, \alpha \, g(x) + \sum_{i=1}^{m} \alpha_i \, h_i(x).$$

For complete bipartite graphs, Proposition 2 shows that the γ-degree of the h^*-polynomial of $K_{m,n}$ is $\min\{m,n\} - 1$. Thus, we get the following an immediate corollary.

Corollary 13. *Let n be a positive integer. For $1 \leq m \leq n$ there exist unique $\alpha, \alpha_0, \alpha_1, \ldots, \alpha_{m-1}$ and $\beta, \beta_0, \beta_1, \ldots, \beta_{m-1}$ in $\mathbb{R}$ such that the following equations are satisfied.*

$$E_{m+1,n+1}(x) = (2x+1)\, \alpha \, E_{m,n+1}(x) + \sum_{i=0}^{m-1} \alpha_i \, E_{m-i,n+i}(x)$$

$$E_{m,n+1}(x) = (2x+1)\, \beta \, E_{m,n}(x) + \sum_{i=0}^{m-1} \beta_i \, E_{m-i,n+i-1}(x)$$

This corollary alone is not enough to prove Conjecture 1 for all $K_{m,n}$ for two crucial reasons. Firstly, as m increases, the number of interlacings having to be satisfied increases too, and they are between polynomials whose cross-degrees puts them outside the scope of Theorem 12. This is noticeable in the last four statements of Theorem 9 where the interlacing of cross-degree 3 polynomials by cross-degree 2-polynomials depends on the interlacing of a cross-degree 2-polynomial by a cross-degree 0 polynomial. Secondly, there is no guarantee that the coefficients $\alpha, \alpha_1, \ldots, \alpha_m$ are non-negative. In fact, for $m \geq 4$, explicit computations reveal that $\alpha_2, \ldots, \alpha_{m-2}$ are always negative. In the case $m = 4$, we get $\alpha_2 = \frac{n-n^3}{8(5n^3+39n^2+100n+96)}$. To see the parameters for every $1 \leq m \leq 10$, we refer once again to the corresponding SAGEMATH code in the previously mentioned github repository.

We end by presenting a conjecture.

Conjecture 14. *Let $a_1 \leq a_2 \leq \cdots \leq a_k \leq n$ be positive integers and let m denote the cross-degree of the Ehrhart polynomial of the symmetric edge polytope of $K_{a_1,a_2,\ldots,a_k}$. Then the inequalities*

$$\left\lfloor \frac{\sum_{i=1}^{k} a_i}{2} \right\rfloor \leq m+1 \leq \sum_{i=1}^{k} a_i.$$

hold. Furthermore, the Ehrhart polynomial of the symmetric edge polytope of the graph $K_{1^k,n}$ interlaces that of $K_{1^{k+1},n}$, where 1^k represents a list k of ones. For $k+n \leq 10$, this has been computationally confirmed.

Acknowledgements

I would like to express my gratitude to my advisor Akihiro Higashitani and to Rodica Dinu for many useful discussions and invaluable advice.

References

[1] Matthias Beck and Sinai Robins. *Computing the continuous discretely*. Undergraduate Texts in Mathematics. Integer-point enumeration in polyhedra. Springer, New York, 2007, pp. xviii+226.

[2] Daniel Bump, Kwok-Kwong Choi, Pär Kurlberg, and Jeffrey Vaaler. "A local Riemann hypothesis. I". In: *Math. Z.* 233.1 (2000), pp. 1–19.

[3] S. Fisk. "Polynomials, roots, and interlacing". In: *arXiv Mathematics e-prints* (2006), math–0612833.

[4] Gábor Hegedüs, Akihiro Higashitani, and Alexander Kasprzyk. "Ehrhart polynomial roots of reflexive polytopes". In: *Electron. J. Combin.* 26.1 (2019), Paper No. 1.38, 27.

[5] Takayuki Hibi. "Dual polytopes of rational convex polytopes". In: *Combinatorica* 12.2 (1992), pp. 237–240.

[6] Akihiro Higashitani, Katharina Jochemko, and Mateusz Michałek. "Arithmetic aspects of symmetric edge polytopes". In: *Mathematika* 65.3 (2019), pp. 763–784.

[7] Akihiro Higashitani, Mario Kummer, and Mateusz Michałek. "Interlacing Ehrhart polynomials of reflexive polytopes". In: *Selecta Math. (N.S.)* 23.4 (2017), pp. 2977–2998.

[8] Max Kölbl. "On a Conjecture Concerning the Roots of Ehrhart Polynomials of Symmetric Edge Polytopes from Complete Multipartite Graphs". In: *arXiv preprint arXiv:2404.02136* (2024).

[9] Tetsushi Matsui, Akihiro Higashitani, Yuuki Nagazawa, Hidefumi Ohsugi, and Takayuki Hibi. "Roots of Ehrhart polynomials arising from graphs". In: *J. Algebraic Combin.* 34.4 (2011), pp. 721–749.

[10] Fernando Rodriguez-Villegas. "On the zeros of certain polynomials". In: *Proc. Amer. Math. Soc.* 130.8 (2002), pp. 2251–2254.

[11] The Sage Developers et al. *SageMath, version 10.2*. 2023. URL: `http://www.sagemath.org`.

[12] Herbert S. Wilf. *generatingfunctionology*. Third. A K Peters, Ltd., Wellesley, MA, 2006, pp. x+245.

An Approximate Counting Version of the Multidimensional Szemerédi Theorem*

Natalie Behague[†1], Joseph Hyde[‡2], Natasha Morrison[§2], Jonathan A. Noel[¶2], and Ashna Wright[‖2]

[1]Mathematics Institute, University of Warwick, UK
[2]Department of Mathematics and Statistics, University of Victoria, Canada

Abstract

Given a finite set $X \subseteq \mathbb{N}^d$, a *non-trivial copy* of X is a set obtained by scaling X by a positive factor and translating it. The Multidimensional Szemerédi Theorem of Furstenberg and Katznelson [6] asserts that the largest cardinality of a subset of $[n]^d$ without a non-trivial copy of X, denoted $r_X(n)$, is $o(n^d)$. We prove that, for any X with $|X| \geq 3$, there exists $C_X > 1$ such that the number of subsets of $[n]^d$ without a non-trivial copy of X is at most $2^{C_X \cdot r_X(n)}$ for infinitely many n.

1 Introduction

Roth's Theorem [10] famously states that every subset of $[n] := \{1, \dots, n\}$ without three elements in arithmetic progression has cardinality at most $o(n)$. This was extended to arithmetic progressions of arbitrary length in the groundbreaking work of Szemerédi [12]. Szemerédi's Theorem can be seen as a very strong "density version" of the elementary van der Waerden Theorem [13] which says that every colouring of the natural numbers with finitely many colours contains monochromatic arithmetic progressions of arbitrary length.

A few years later, Furstenberg [5] reproved Szemerédi's Theorem using tools from ergodic theory. This new perspective turned out to be widely applicable, yielding several remarkably general results. One such example is the so-called "Multidimensional Szemerédi Theorem" of Furstenberg and Katznelson [6], which we explain next. Given a set $X \subseteq \mathbb{N}^d$, a *copy* of X is a set of the form

$$\vec{b} + r \cdot X = \{\vec{b} + r\vec{x} : \vec{x} \in X\}$$

where $\vec{b} \in \mathbb{R}^d$ and $r \geq 0$. It is said to be a *non-trivial* copy if $r > 0$. As an example, a 3-term arithmetic progression is nothing more than a copy of $X = \{1, 2, 3\}$. A set $A \subseteq \mathbb{N}^d$ is *X-free* if it does not contain a copy of X and $r_X(n)$ denotes the cardinality of the largest X-free subset of $[n]^d$.

Theorem 1 (Multidimensional Szemerédi Theorem [6])**.** *For any finite set $X \subseteq \mathbb{N}^d$, $r_X(n) = o(n^d)$.*

We focus on the closely related problem of counting the number of X-free subsets of $[n]^d$. An obvious lower bound is $2^{r_X(n)}$, which one can get by taking any subset of the largest X-free set. Our main result says that the exponent $r_X(n)$ is within a constant factor of being tight for infinitely many $n \in \mathbb{N}$.

*The full version of this work can be found in [3] and will be published elsewhere.

[†]Email: natalie.behague@warwick.ac.uk. Research of N. Behague supported by PIMS
[‡]Email: josephhyde@uvic.ca.
[§]Email: nmorrison@uvic.ca. Research of N. Morrison supported by NSERC and a university start-up grant.
[¶]Email: noelj@uvic.ca. Research of J. A. Noel supported by NSERC and a university start-up grant.
[‖]Email: ashnawright@uvic.ca. Research of A. Wright supported by NSERC.

Theorem 2. *For any finite set $X \subseteq \mathbb{N}^d$ with $|X| \geq 3$ there exists $C_X > 1$ such that the number of X-free subsets of $[n]^d$ is at most $2^{C_X \cdot r_X(n)}$ for infinitely many n.*

We note that Theorem 2 extends the work of Balogh, Liu and Sharifzadeh [1] who proved it for arithmetic progressions and Kim [9] who focused on the case that $X = \{\vec{0}, \vec{e}_1, \ldots, \vec{e}_d\}$ where $\vec{e}_i$ is the ith standard basis vector of $\mathbb{R}^d$.

2 Key Ideas and Challenges

To anyone who has followed recent developments on obtaining "counting versions" of important theorems from extremal combinatorics, it should be no surprise that our proof is an application of the container method developed by Saxton and Thomason [11] and Balogh, Morris and Samotij [2]. The container method provides a widely applicable "recipe" for bounding the number of independent sets in a "well-behaved" hypergraph from above.

In our setting, the choice of the hypergraph is straightforward; we let $\mathcal{H}$ be the hypergraph with vertex set $[n]^d$ in which every non-trivial copy of X forms a hyperedge. An independent set in $\mathcal{H}$ clearly corresponds to an X-free set, and so our goal is to bound the number of independent sets in $\mathcal{H}$. Note that every vertex of $\mathcal{H}$ is contained within $\Theta(n)$ hyperedges and that any pair of distinct vertices of $\mathcal{H}$ are only contained within $O(1)$ hyperedges together. In other words, the hypergraph $\mathcal{H}$ has very small "co-degrees." As it turns out, this implies that it is sufficiently "well behaved" to apply the method.

The key ingredient in most applications of the container method is a so-called "supersaturation bound." In the context of our problem, a supersaturation theorem is a lower bound on the number of copies of X within a subset of $[n]^d$ of cardinality greater than $r_X(n)$. Roughly speaking, our key supersaturation bound is as follows.

Lemma 3 (Supersaturation Bound, Roughly)**.** *Let $X \subseteq \mathbb{N}^d$ with $|X| \geq 3$. There exists $C_X > 1$ such that, for infinitely many $n \in \mathbb{N}$, every set $A \subseteq [n]^d$ with $|A| \geq C_X \cdot r_X(n)$ contains a "large" number of copies of X.*

Given a sufficiently strong supersaturation bound of the type described in Lemma 3, Theorem 2 can be deduced from many different forms of the Hypergraph Container Lemma; in particular, we apply a version from Saxton and Thomason [11]. Most of the work in the paper is devoted to proving Lemma 3.

One of the key challenges in establishing a supersaturation bound that applies to all sets of cardinality $C_X \cdot r_X(n)$ for some constant C_X is that the function $r_X(n)$ itself is not well-understood. Letting $r_k(n) := r_{\{1,\ldots,k\}}(n)$, the best known bounds on $r_3(n)$ (i.e. the case of 3-term arithmetic progressions) are currently

$$\frac{n}{2^{c\sqrt{\log_2(n)}}} \leq r_3(n) \leq \frac{n}{2^{\log(n)^\beta}} \tag{1}$$

where $0 < c < 2\sqrt{2}$ and $\beta > 0$ are constants. Both inequalities were proved recently by Hunter [7] (lower bound) and Kelly and Meka [8] (upper bound) and represent substantial breakthroughs in the field. In spite of these achievements, we still do not know $r_3(n)$ to within a constant factor, nor do we have precise asymptotics for $r_X(n)$ for other sets X of interest. Thus, we will need to find a way of proving a supersaturation result for sets of cardinality $C_X \cdot r_X(n)$ which does not require us to know any detailed information about the growth rate of $r_X(n)$.

To obtain the desired supersaturation result, a key idea from [1] is to apply a "crude" supersaturation bound which we "amplify" via a double-counting trick. For the crude bound, if we let $\Gamma_X(A)$ denote the number of copies of X in a set $A \subseteq [n]^d$, then

$$\Gamma_X(A) \geq |A| - r_X(n).$$

This is trivial to see by simply deleting an element of A within a copy of X (if one exists) and applying induction. This trivial idea becomes more powerful if one applies it within a random "subcube" of $[n]^d$.

To be a bit less vague, suppose that $M \ll n$ and take a random copy of $[M]^d$ within $[n]^d$ by scaling $[M]^d$ by a random prime p and translating it by a random vector $\vec{b}$ (where p and $\vec{b}$ are chosen with some constraints to make this copy a subset of $[n]^d$). Now, given any set $A \subseteq [n]^d$, we can use the crude bound to get that the number of copies of X contained within this copy of $[M]^d$ is at least the number of points of A within this subcube minus $r_X(M)$. We can also bound this quantity above in terms of $\Gamma_X(A)$; putting these two bounds together yields a lower bound on $\Gamma_X(A)$.

Unfortunately, as it turns out, the argument in the previous paragraph provides a lower bound on $\Gamma_X(A)$ in terms of $r_X(M)$, where $M = M(n)$ is a sublinear function of n. In fact, the bound obtained is a multiple of

$$\frac{|A|}{2n^d} - \frac{r_X(M)}{M^d}$$

which is clearly useless in cases where $\frac{|A|}{2n^d} \leq \frac{r_X(M)}{M^d}$. Recall that we need the supersaturation bound to work for any set A such that $|A| \geq C_X \cdot r_X(n)$ for some constant $C_X > 1$. Thus, in order for the bound to be useful, we need that $\frac{C_X \cdot r_X(n)}{2n^d}$ is significantly larger than $\frac{r_X(M)}{M^d}$. Again, we run into the same problem: we do not understand the growth rate of $r_X(n)$. In particular, if the ratio $\frac{r_X(n)}{n^d}$ fluctuates wildly as n tends to infinity, then we have no hope of obtaining a bound of this form that holds for all n. However, by combining the lower bound in (1) (in fact, a weaker bound of Behrend [4] from 1946 suffices) generalized to $r_X(n)$ with a simple limit argument based on the fact that $\sqrt{N+\sqrt{N}} - \sqrt{N}$ converges, we can get that such a bound holds for infinitely many n, allowing us to obtain Lemma 3.

References

[1] J. Balogh, H. Liu, and M. Sharifzadeh, *The number of subsets of integers with no k-term arithmetic progression*, Int. Math. Res. Not. IMRN (2017), no. 20, 6168–6186.

[2] J. Balogh, R. Morris, and W. Samotij, *Independent sets in hypergraphs*, J. Amer. Math. Soc. **28** (2015), no. 3, 669–709.

[3] N. Behague, J. Hyde, N. Morrison, J. A. Noel, and A. Wright, *An approximate counting version of the multidimensional Szemerédi theorem*, E-print arXiv:2311.13709v1, 2023.

[4] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*, Proc. Nat. Acad. Sci. U.S.A. **32** (1946), 331–332.

[5] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256.

[6] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979).

[7] Z. Hunter, *New lower bounds for $r_3(N)$*, E-print arXiv:2401.16106v2, 2024.

[8] Z. Kelley and R. Meka, *Strong bounds for 3-progressions*, E-print arXiv:2302.05537v4, 2023.

[9] Y. Kim, *The number of k-dimensional corner-free subsets of grids*, Electron. J. Combin. **29** (2022), no. 2, Paper No. 2.53, 19.

[10] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109.

[11] D. Saxton and A. Thomason, *Hypergraph containers*, Invent. Math. **201** (2015), no. 3, 925–992.

[12] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.

[13] B. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **19** (1927), 212–216.

A short proof of an inverse theorem in bounded torsion groups *

Pablo Candela[†1], Diego González-Sánchez[‡2], and Balázs Szegedy[§2]

[1]UAM and ICMAT, Ciudad Universitaria de Cantoblanco Madrid 28049 Spain
[2]HUN-REN Alfréd Rényi Institute of Mathematics, Reáltanoda u. 13-15. Budapest, Hungary, H-1053

Abstract

Let Z be a finite abelian group of bounded torsion m and $f : \mathrm{Z} \to \mathbb{C}$ a 1-bounded function. Jamneshan, Shalom, and Tao proved the following inverse theorem: If $\|f\|_{U^{k+1}} \geq \delta > 0$, then f correlates non-trivially with a polynomial phase function of degree bounded in terms of m and k. They also ask whether the same holds with polynomials of degree at most k. In this paper we use the nilspace approach to investigate this problem proving: *a)* an inverse theorem for bounded torsion abelian groups where we replace the polynomial phase functions of degree k by *projected* polynomial phase functions of degree k, a notion introduced by the third-named author. *b)* Relying on *a)* we give a short proof of the result of Jamneshan, Shalom, and Tao.

1 Introduction

This work is a shortened version of [6] and some parts have been taken from the latter directly.

Since their introduction in the seminal work of Gowers [9], the study of Gowers norms (denoted by $\|\cdot\|_{U^{k+1}}$) have been central in the area of higher-order Fourier analysis. An important question related to these norms is inverse theorems. Such results were initially proved for finite cyclic groups (or intervals of $\mathbb{Z}$) in [12] and state, loosely speaking, that if a function has a large U^{k+1}-norm then it must correlate with a *nil-function*. The precise notion of what a *nil-function* is depends on the type of abelian groups we are considering (e.g., cyclic groups, finite torsion vector spaces $\mathbb{F}_p^n$, etc.) We refer to [7, 10, 14, 16] for more background on these results. In this paper, we will focus on finite abelian groups with fixed finite torsion $m \geq 1$ (or m-torsion abelian groups). That is, abelian groups Z such that $mx = 0$ for all $x \in \mathrm{Z}$.

Our work is motivated by a recent paper of Jamneshan, Shalom, and Tao [15] where they prove an inverse theorem for m-torsion abelian groups where the *nil-function* mentioned above is a *polynomial phase function* of degree bounded in terms of m and k. Recall that given abelian groups Z, Z', a map $P : \mathrm{Z} \to \mathrm{Z}'$ and any $h \in \mathrm{Z}$, we may take the discrete derivative $\partial_h P : \mathrm{Z} \to \mathrm{Z}'$ defined by $\partial_h P(x) = P(x+h) - P(x)$. Then we say that P is *polynomial of degree at most k* if $\partial_{h_1} \cdots \partial_{h_{k+1}}(P)(x) = 0$ for all $x, h_1, \ldots, h_{k+1} \in \mathrm{Z}$. A *polynomial phase function* of degree at most k is then a *polynomial of degree* at most k where $\mathrm{Z}' = \mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\}$.

*The full version of this work can be found in [6] and will be published elsewhere. All authors used funding from project PID2020-113350GB-I00 of Spain's MICINN. The second-named author received funding from projects KPP 133921 and Momentum (Lendület) 30003 of the Hungarian Government. The research was also supported partially by the NKFIH "Élvonal" KKP 133921 grant and by the Hungarian Ministry of Innovation and Technology NRDI Office in the framework of the Artificial Intelligence National Laboratory Program.

[†]Email: pablo.candela@uam.es.
[‡]Email: diegogs@renyi.hu.
[§]Email: szegedyb@gmail.com.

Theorem 1. ([15, Theorem 1.12]) *Let k, m be positive integers and let $\delta > 0$. Then there exist constants $\varepsilon = \varepsilon(\delta, k, m) > 0$ and $C = C(k, m) > 0$ such that for every finite m-torsion abelian group Z and every 1-bounded function $f : \mathrm{Z} \to \mathbb{C}$ with $\|f\|_{U^{k+1}} > \delta$, there exists a polynomial phase $Q : \mathrm{Z} \to \mathbb{S}^1$ of degree at most C such that $|\mathbb{E}_{x\in\mathrm{Z}} f(x)\overline{Q(x)}| > \varepsilon$.*

This result is inspired by the special case $m = p$ a prime number, where much more is known. In fact in this case Z is just $\mathbb{F}_p^n$ for some integer n, and it is known (see [18, Theorems 1.9 and 1.10] and [19, Theorem 1.10]) that we can take $C(k, p) = k$.[1] Jamneshan, Shalom, and Tao ask whether this holds also in the case of m not being a prime [15, Question 1.9]. An important aspect related to the constant $C(k, m)$ is that if it equals k for any $m \in \mathbb{N}$, then this would be the optimal value. The proof of this fact follows from the following result (valid for any finite abelian group Z, see Proposition 9):

Lemma 2. *Let $\delta > 0$. For any 1-bounded function $f : \mathrm{Z} \to \mathbb{C}$ if[2] $|\langle f, Q\rangle| \geq \delta$ for Q a polynomial phase function of degree k then $\|f\|_{U^{k+1}} \geq \delta$.*

Any function Q (not necessarily a polynomial phase) satisfying the conclusion of Lemma 2 is called an *obstruction* to the U^{k+1} norm.[3] For $k' > k$ in general (and in particular, in the case of m-torsion groups), polynomial phases of degree at most k' are not *necessarily* obstructions to the U^{k+1} norm.

In this paper, we prove an inverse theorem for m-torsion abelian groups where the *nil-functions* appearing are a generalization of polynomial phase functions of degree k but nevertheless, they are obstructions to the U^{k+1} Gowers norm for m-torsion groups. This notion was introduced originally by the third-named author in the unpublished work [17]. We recall its definition (see [17, Definition 1.2]).

Definition 3. *Let Z be a finite abelian group and let $k \in \mathbb{N}$. A* projected phase polynomial of degree *k on Z is a 1-bounded function $\phi_{*\tau} : \mathrm{Z} \to \mathbb{C}$ of the following form. There is a finite abelian group Z', a surjective homomorphism $\tau : \mathrm{Z}' \to \mathrm{Z}$, and a polynomial phase function $\phi : \mathrm{Z}' \to \mathbb{C}$ of degree at most k, such that $\phi_{*\tau}(x) = \mathbb{E}_{y\in\tau^{-1}(x)}\phi(y)$ for every $x \in \mathrm{Z}$. If the torsions of Z and Z' are respectively m and m' we say that $\phi_{*\tau}$ has torsion (m, m'). We say it is rank-preserving if the rank of Z is equal to the rank of Z' (where the rank is the minimal number of generators).*

We can now state our first main result (see [6, Theorem 1.12] and the discussion below).

Theorem 4. *Let k, m be positive integers and let $\delta > 0$. Then there exists $\gamma = \gamma(k) \in \mathbb{N}$ and $\varepsilon = \varepsilon(\delta, k, m) > 0$ such that the following holds. For any m-torsion abelian group Z and any 1-bounded function $f : \mathrm{Z} \to \mathbb{C}$ with $\|f\|_{U^{k+1}} \geq \delta$, there exists a rank-preserving projected phase polynomial $\phi_{*\tau}$ of degree k and torsion (m, m^γ) on Z such that $|\langle f, \phi_{*\tau}\rangle| \geq \varepsilon$. Conversely, if for any $\delta' > 0$ we have $|\langle f, \phi_{*\tau}\rangle| \geq \delta'$ for some projected phase polynomial $\phi_{*\tau}$ of degree k then $\|f\|_{U^{k+1}} \geq \delta'$.*

As we can see, the first part of Theorem 4 is very similar to Theorem 1, but it replaces the polynomial phase functions of degree $C(k, m)$ by projected phase polynomials of degree k. The second main result of this paper shows that the former result is strictly stronger.

Theorem 5. *Theorem 1 can be deduced from Theorem 4.*

2 An inverse theorem with bounded-torsion nilspaces

In this paper, we rely on the theory of nilspaces (see e.g., [1, 2, 13] and references therein). In a few words, a nilspace is an algebraic object (you can also endow it with a natural topology) that generalizes abelian groups. An important feature of nilspaces is the concept of step: the class of 1-step nilspaces

[1] These results are qualitative. We refer to [10, 11] and references therein for quantitative results.

[2] For any finite abelian group Z and any pair of functions $f, g : \mathrm{Z} \to \mathbb{C}$ we denote $\langle f, g\rangle := \mathbb{E}_{x\in\mathrm{Z}} f(x)\overline{g(x)}$.

[3] To be fully precise, we need to fix a family of finite abelian groups $\mathcal{F}$ and $\delta > 0$. Then, given a family of functions $\mathcal{T} = \{Q : \mathrm{Z} \to \mathbb{C} : \mathrm{Z} \in \mathcal{F}\}$ we say that the functions of that family are obstructions to the U^{k+1} norm for $\mathcal{F}$ if there exists $\epsilon = \epsilon(\delta, \mathcal{F}) > 0$ such that for any $\mathrm{Z} \in \mathcal{F}$, any 1-bounded $f : \mathrm{Z} \to \mathbb{C}$, and any $Q \in \mathcal{T}$ if $|\langle f, Q\rangle| > \delta$ then $\|f\|_{U^{k+1}} > \epsilon$.

equals the class of (affine[4]) abelian groups. For higher step $k \geq 2$, nilspaces can be seen as a tower of k extensions (or bundles) by abelian groups, called structure groups (see [2, §1.2 and §3.2] for details). In this paper, we will be interested in the class of nilspaces such that all the abelian groups in these extensions are m-torsion abelian groups. We denote such nilspaces by *m-torsion* nilspaces (see [6]).

Similarly to abelian groups, where we have the concept of homomorphism $\varphi : \mathrm{Z} \to \mathrm{Z}'$ between abelian groups, for nilspaces, this generalizes to the concept of nilspace morphism $\varphi : \mathrm{X} \to \mathrm{Y}$. The importance of nilspaces and nilspace morphisms is that we can formulate the most general inverse theorem known at the time of writing this paper (valid for any compact abelian group or even any nilmanifold) as shown in [7, Theorem 1.6]. This result can be specialized in the case that we want to find inverse theorems among various classes of groups (see [7, Theorem 1.7] and [5, §6]). In this paper, we are interested in the class of finite abelian m-torsion groups. The following result is a version of [6, Theorem 2.3]:

Theorem 6. *For any $k, m \in \mathbb{N}$ and $\delta > 0$ there exists $C > 0$ such that the following holds. Let Z be a finite abelian m-torsion group, and let $f : \mathrm{Z} \to \mathbb{C}$ be a 1-bounded function with $\|f\|_{U^{k+1}} \geq \delta$. Then there is a finite m-torsion k-step nilspace X of cardinality $|\mathrm{X}| \leq C$, a morphism $\phi : \mathcal{D}_1(\mathrm{Z}) \to \mathrm{X}$, and a 1-bounded function $F : \mathrm{X} \to \mathbb{C}$, such that $\langle f, F \circ \phi \rangle \geq \frac{1}{2}\delta^{2^{k+1}}$.*

Hence, we can reduce the question of studying the inverse theorem for m-torsion groups to studying morphisms $\phi : \mathcal{D}_1(\mathrm{Z}) \to \mathrm{X}$. The nilspace $\mathcal{D}_1(\mathrm{Z})$ or more generally $\mathcal{D}_k(\mathrm{Z})$ for any $k \geq 1$ is a special type of nilspace constructed from any abelian group Z, see [2, §2.2.4] for the precise definition of $\mathcal{D}_k(\mathrm{Z})$. What matters for us is that these nilspaces are the *simplest* types of nilspaces. For example, by [4, Lemma A.2] we have that the morphisms $\varphi : \mathcal{D}_\ell(\mathrm{Z}) \to \mathcal{D}_k(\mathrm{Z}')$ are exactly the polynomials $\mathrm{Z} \to \mathrm{Z}'$ of degree at most $\lfloor k/\ell \rfloor$.

Let us now outline the rest of the proof idea:

(i) Given any m-torsion nilspace X there exists numbers $a_1, \dots, a_k$ and γ such that the following holds. There is a totally surjective bundle morphism[5] $\tilde{\varphi} : \prod_{i=1}^k \mathcal{D}_i(\mathbb{Z}^{a_i}) \to \mathrm{X}$ which is m^γ periodic. We let $\mathrm{Y} := \prod_{i=1}^k \mathcal{D}_i(\mathbb{Z}^{a_i}_{m^\gamma})$. This notion is analogous to the known fact for finite abelian groups that if A is abelian and m-torsion, then there exists a surjective m-periodic homomorphism $\mathbb{Z}^a \to A$ for some $a \in \mathbb{N}$.

(ii) Letting $\tilde{\tau} : \mathbb{Z}^r \to \mathrm{Z}$ be a surjective homomorphism there exists a morphism $g : \mathcal{D}_1(\mathbb{Z}^r) \to \mathrm{Y}$ such that $\varphi \circ g = \phi \circ \tilde{\tau}$. Moreover, the map g is $m^{\gamma'}$-periodic for some $\gamma' = \gamma'(k)$. Letting $B := \mathbb{Z}^r_{m^{\gamma'}}$ the situation can be better seen in the following commutative diagram:

$$\begin{array}{ccccc} & & g & & \\ \mathcal{D}_1(\mathbb{Z}^r) & \xrightarrow[p]{} & \mathcal{D}_1(B) & \xrightarrow[\psi]{} & \mathrm{Y} \\ & \searrow^{\tilde{\tau}} & \downarrow \tau & & \downarrow \varphi \\ & & \mathrm{Z} & \xrightarrow{\phi} & \mathrm{X}\,. \end{array} \tag{1}$$

Coming back to the case of m-torsion abelian groups, note that if we have $\mathrm{X} = A$ and $\mathrm{Y} = \mathbb{Z}^a_m$ then the map $\varphi \circ g$ is m-periodic (no need for a γ' power).

(iii) Recall that we wanted to study the map $F \circ \phi : \mathrm{Z} \to \mathbb{C}$. By (1) note that $F \circ \phi \circ \tau = F \circ \varphi \circ \psi$. But now the map ψ is relatively easy to understand as it consists of polynomials of degree at most k. Hence, by using regular Fourier analysis on the abelian group Y (yes, it is a nilspace, but you can see it also as an abelian group and do Fourier analysis on it) we can write $F \circ \varphi$ as a sum of a bounded number of harmonics which yields the result.

[4] There is no fixed 0 element, but after choosing any arbitrary element to be 0 these classes can be proved to be equal.
[5] See [2, Definition 3.3.1].

3 Proof idea of Theorem 4

The point (i) from the previous section relies on generalizations of two well-known results for finite abelian groups. The first is that letting A be a finite abelian group there exists $a \in \mathbb{N}$ and a surjective homomorphism $\mathbb{Z}^a \to A$. The second is that if we further assume that A is m-torsion, then the former surjective homomorphism can be proved to be m-periodic. We need generalizations of these results for m-torsion nilspaces. Generalizing the first result, namely, that for a finite k-step nilspace X there exists a fibration $\prod_{i=1}^{k} \mathcal{D}_i(\mathbb{Z}^{a_i}) \to \mathrm{X}$ is a non-trivial result shown by the authors, [4, Theorem 4.4] (see also [6, Corollary 5.4]).

For the second result, the periodicity of morphisms of the form $\prod_{i=1}^{k} \mathcal{D}_i(\mathbb{Z}^{a_i}) \to \mathrm{X}$ when X is m-torsion, let us prove here the core lemma (a version of [6, Lemma 5.1]) that leads to the full result (see [6, Corollary 5.4] for details).

Lemma 7. *For any positive integer k there exists $\alpha > 0$ such that the following holds. Let A be any m-torsion abelian group and let $\phi : \mathbb{Z} \to A$ be a polynomial of degree at most k. Then ϕ is m^α-periodic.*

Proof. By [4, Theorem A.6], any polynomial ϕ of the latter type has an expression of the form $\phi(x) = \sum_{i=1}^{k} a_i \binom{x}{i}$ for some $a_i \in A$. Let us prove now that $\binom{x+m^{k+1}}{i} - \binom{x}{i}$ is a multiple of m for any $x \in \mathbb{Z}$ and $i \in [k]$. For any prime $p|m$ suppose that $m = p^c m'$ where p and m' are coprime. If we prove that $\binom{x+m^{k+1}}{i} - \binom{x}{i}$ is a multiple of p^c then we are done (as we can then argue similarly for every prime dividing m). Using the identity $\binom{x}{i} = \frac{x(x-1)\cdots(x-i+1)}{i}$ we have $\binom{x+m^{k+1}}{i} - \binom{x}{i} = \frac{m^{k+1}}{i!} Q(x, m, i)$ for some integer-valued polynomial Q. If we prove that $\frac{m^{k+1}}{i!}$ is always a multiple of p^c then we will be done. Note that the largest power of p dividing m^{k+1} is precisely $c(k+1)$. On the other hand, in $i!$ we have at most $\sum_{j=1}^{\infty} \lfloor i/p^j \rfloor \leq \sum_{j=1}^{\infty} i/p^j = \frac{i}{p-1} \leq \frac{k}{p-1}$ factors of p. But as for any $c \in \mathbb{N}$ and p prime we have that $\frac{k}{p-1} + c \leq c(k+1)$ the result follows. □

To prove (ii), we again rely on generalizing a known result for abelian groups. Namely, let A, C be abelian groups, $\varphi : C \to A$ be a surjective homomorphism, and $q : \mathbb{Z}^n \to A$ be any homomorphism. Then there exists $\tilde{q} : \mathbb{Z}^n \to C$ such that $\varphi \circ \tilde{q} = q$. This result generalizes to nilspaces as shown in [5, Corollary A.6]. Thus, the diagram (1) follows.

Proof sketch of first part of Theorem 4. We apply Theorem 6 and let X be the resulting nilspace of torsion m, ϕ the resulting morphism $\mathcal{D}_1(\mathrm{Z}) \to \mathrm{X}$, and $F : \mathrm{X} \to \mathbb{C}$ the resulting 1-bounded function such that $\mathbb{E}_{x \in Z} f(x) F(\phi(x)) \geq \delta^{2^{k+1}}/2$. We construct now Diagram (1) as explained before. Let $h := F \circ \varphi : \mathrm{Y} \to \mathbb{C}$. Then $\mathbb{E}_{x \in Z} f(x) F(\phi(x)) = \mathbb{E}_{y \in B} f \circ \tau(y) h \circ \psi(y)$. By the Fourier decomposition of h on the finite abelian group Y, and the pigeonhole principle, there is a character $\chi \in \widehat{B}$ such that $\varepsilon \leq \mathbb{E}_{y \in B} f(\tau(y)) \chi(\psi(y)) = \mathbb{E}_{x \in Z} f(x) \mathbb{E}_{y \in \tau^{-1}(x)} \chi(\psi(y))$, which proves the result with $\phi := \chi \circ \psi$. □

The second part of Theorem 4 follows from the next results.

Lemma 8. *Let $\phi_{*\tau}$ be a projected phase polynomial of degree k on a finite abelian group. Then $\|\phi_{*\tau}\|_{U^{k+1}}^* \leq 1$ where $\|\cdot\|_{U^{k+1}}^*$ is the U^{k+1}-dual-norm.*

Proof. Recall the definition $\|\phi_{*\tau}\|_{U^{k+1}}^* = \sup_{g : \mathrm{Z} \to \mathbb{C} : \|g\|_{U^{k+1}} \leq 1} |\langle \phi_{*\tau}, g \rangle|$. Denoting by Z' the (abelian group) domain of τ, the map $\tau^{[\![k+1]\!]} : \mathrm{C}^{k+1}(\mathrm{Z}') \to \mathrm{C}^{k+1}(\mathrm{Z})$ defined by $\tau^{[\![k+1]\!]}(\mathrm{c}) : v \mapsto \tau(\mathrm{c}(v))$ is a surjective homomorphism. It follows that for every map $g : \mathrm{Z} \to \mathbb{C}$ we have $\|g \circ \tau\|_{U^{k+1}(\mathrm{Z}')} = \|g\|_{U^{k+1}(\mathrm{Z})}$. Then we have $|\langle \phi_{*\tau}, g \rangle_{\mathrm{Z}}| = |\mathbb{E}_{x \in \mathrm{Z}} \mathbb{E}_{y \in \tau^{-1}(x)} g \circ \tau(y) \phi(y)| = |\langle g \circ \tau, \phi \rangle_{\mathrm{Z}'}| \leq \|g \circ \tau\|_{U^{k+1}(\mathrm{Z}')} \|\phi\|_{U^{k+1}(\mathrm{Z}')}^* = \|g\|_{U^{k+1}(\mathrm{Z})} \|\phi\|_{U^{k+1}(\mathrm{Z}')}^*$. Therefore $\|\phi_{*\tau}\|_{U^{k+1}(\mathrm{Z})}^* \leq \|\phi\|_{U^{k+1}(\mathrm{Z})}^*$. Since ϕ is a phase polynomial of degree k, we have that $|\langle \phi, g \rangle| = \|\phi \overline{g}\|_{U^1} \leq \|\phi \overline{g}\|_{U^{k+1}} = \|g\|_{U^{k+1}}$, see [11, (2.1)]. Thus $\|\phi\|_{U^{k+1}(\mathrm{Z}')}^* \leq 1$. □

Proposition 9. *Let $\phi_{*\tau}$ be a projected phase polynomial of degree k on a finite abelian group Z, and suppose that $f : \mathrm{Z} \to \mathbb{C}$ satisfies $|\langle f, \phi_{*\tau} \rangle| \geq \delta$. Then $\|f\|_{U^{k+1}} \geq \delta$.*

Proof. By Lemma 8, $\delta \leq |\langle f, \phi_{*\tau} \rangle| \leq \|f\|_{U^{k+1}} \|\phi_{*\tau}\|_{U^{k+1}}^* \leq \|f\|_{U^{k+1}}$. □

4 Proof of Theorem 5

The idea is to prove that the projected phase polynomials appearing in Theorem 4 can be written as averages of polynomials of possibly larger degree. We shall prove this below and then apply it to give an alternative proof of [15, Theorem 1.12]. Given a surjective homomorphism $\tau : B \to \mathrm{Z}$, by a *polynomial cross-section* for τ we mean a map $\iota : \mathrm{Z} \to B$ which is polynomial and such that $\tau \circ \iota$ is the identity map on Z. The main result that we will use is the following.

Theorem 10. *Let $m, m' \in \mathbb{N}$. Then there exists a constant $C(m, m') \in \mathbb{N}$ such that the following holds. Let Z, B be finite abelian groups of torsion m and m' respectively and let $\tau : B \to \mathrm{Z}$ be a surjective homomorphism. Then there exists a polynomial cross-section $\iota : \mathrm{Z} \to B$ of degree at most $C(m, m')$.*

The full proof of Theorem 10 can split into several lemmas, see [6, §5.1] for details. Here we are only going to show the first one. See also [15, Lemma 8.2] for an alternative approach.

Lemma 11. *Let $d \geq s$ be positive integers and let p be a prime. Let $\varphi : \mathbb{Z}_{p^d} \to \mathbb{Z}_{p^s}$ be the map $x \mod p^d \mapsto x \mod p^s$. Let $\iota : \mathbb{Z}_{p^s} \to \mathbb{Z}_{p^d}$ be defined by $n \mod p^s \mapsto n \mod p^d$ for each $n \in [0, p^s - 1]$. Then ι is a polynomial cross-section for φ of degree at most $(d - s)p^s + 1$.*

The argument has similarities with the proof of [5, Proposition B.2]. We want to prove that if we take sufficiently many derivatives of ι then we obtain the 0 map. Without loss of generality, it suffices to take derivatives $\partial_a \iota(x) := \iota(x + a) - \iota(x)$ with respect to the generator $a = 1 \in \mathbb{Z}_{p^s}$. Note that $\partial_1 \iota(x) = 1$ if $x \neq p^s - 1$ and $\partial_1 \iota(p^s - 1) = 1 - p^s$. Taking one more derivative, $\partial_1^2 \iota(x) = 0$ if $x \neq p^s - 1$, $\partial_1^2 \iota(p^s - 2) = -p^s$ and $\partial_1^2 \iota(p^s - 1) = p^s$. To take derivatives of higher degree, as is standard, we can view the map $\partial_1^2 \iota$ as a vector in $\mathbb{Z}_{p^d}^{p^s}$ and take the derivatives by left-multiplying this vector by the forward difference matrix, i.e. the circulant matrix $C_{p^s} := \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix} \in M_{p^s \times p^s}(\mathbb{Z})$. Known results on circulant matrices imply the following fact.

Lemma 12. *For any prime p and any integer $s \geq 1$ all the entries of $C_{p^s}^{p^s}$ are multiples of p.*

Proof. By equation (8) in [8], for every $q \in \mathbb{N}$ we have $C_{p^s}^q = \sum_{j=0}^q \binom{q}{j} (-1)^j A_{p^s}^{q-j}$, where A_{p^s} is the cyclic permutation matrix (see [8]). Taking $q = p^s$, we claim that it suffices to prove that $\binom{p^s}{j} = \frac{p^s!}{j!(p^s-j)!}$ is a multiple of p if $0 < j < p^s$. In fact, the contributions for $j = 0$ and $j = p^s$ cancel each other if p is odd as $A_{p^s}^0 = A_{p^s}^{p^s} = \mathrm{id}_{p^s \times p^s}$ and thus $\binom{p^s}{0}(-1)^0 \mathrm{id}_{p^s \times p^s} + \binom{p^s}{p^s}(-1)^{p^s} \mathrm{id}_{p^s \times p^s} = 0$. If $p = 2$ we have $\binom{2^s}{0}(-1)^0 \mathrm{id}_{2^s \times 2^s} + \binom{2^s}{2^s}(-1)^{2^s} \mathrm{id}_{2^s \times 2^s} = 2\mathrm{id}_{2^s \times 2^s}$ which is a multiple of $p = 2$ as claimed. To see the general case $0 < j < p^s$, note first that the number of p factors in $j!$ is precisely $\sum_{i=1}^{s-1} \lfloor j/p^i \rfloor$. Thus, it suffices to prove that $\sum_{i=1}^{s-1} \lfloor j/p^i \rfloor + \sum_{i=1}^{s-1} \lfloor (p^s - j)/p^i \rfloor < 1 + p + \cdots + p^{s-1} = \frac{p^s - 1}{p-1}$, where the right hand side is the number of p factors of $p^s!$. The left hand side can be estimated using the bound $\sum_{i=1}^{s-1} \lfloor j/p^i \rfloor \leq \sum_{i=1}^{s-1} j/p^i = j \frac{p^{s-1}-1}{(p-1)p^{s-1}}$. Hence, the left side is bounded above by $j \frac{p^{s-1}-1}{(p-1)p^{s-1}} + (p^s - j) \frac{p^{s-1}-1}{(p-1)p^{s-1}} = \frac{p^s - p}{p-1}$, which is smaller than the number of p factors in $p^s!$. □

Proof of Lemma 11. Note that after two derivatives, the map $\partial_1^2 \iota$ has already a factor p^s. Each time that we differentiate p^s additional times we add (at least) a factor p by Lemma 12. It follows that $\partial_1^{kp^s+2} \iota(x)$ is a multiple of p^{k+s} for any $x \in \mathbb{Z}_{p^s}$. Thus, if $k + s = d$ then we have $\partial_1^{kp^s+2} \iota = 0 \mod p^d$. Hence ι is a polynomial of degree at most $(d - s)p^s + 1$. □

The rest of the proof of Theorem 10 follows by first taking the Sylow decomposition on Z and B, thus reducing the problem to the case of p-groups. And then, by proving that any surjective homomorphism between p-groups can be reduced (via isomorphisms and projections) to the case of Lemma 11.

Proof of Theorem 5. By Theorem 4, the function f correlates with a projected phase polynomial $(\chi\circ\psi)_{*\tau}$ of degree k and torsion $(m, m^{O_k(1)})$, for some homomorphism $\tau : B \to \mathrm{Z}$. By Theorem 10 there exists a polynomial cross-section $\iota : \mathrm{Z} \to B$ of degree $O_{m,k}(1)$. Moreover, for any $u \in \ker(\tau)$ we have that $\iota_u(x) := \iota(x) + u$ is clearly also a polynomial cross-section. Recall that $(\chi\circ\psi)_{*\tau}$ is the map $\mathbb{E}_{y\in\tau^{-1}(x)}\chi(\psi(y))$ defined for $x \in \mathrm{Z}$. However, for any $x \in \mathrm{Z}$ we have $\mathbb{E}_{y\in\tau^{-1}(x)}\chi(\psi(y)) = \mathbb{E}_{u\in\ker(\tau)}\chi(\psi(\iota_u(x)))$. Thus $\varepsilon < |\mathbb{E}_{x\in\mathrm{Z}}f(x)\mathbb{E}_{u\in\ker(\tau)}\chi(\psi(\iota_u(x)))| = |\mathbb{E}_{u\in\ker(\tau)}\mathbb{E}_{x\in\mathrm{Z}}f(x)\chi(\psi(\iota_u(x)))|$. Hence, $\varepsilon < |\mathbb{E}_{x\in\mathrm{Z}}f(x)\chi(\psi(\iota_u(x)))|$ for some $u \in \ker(\tau)$. Finally note that by [4, Lemma A.2], $\psi\circ\iota_u$ is in fact a polynomial map with degree bounded by $\deg(\iota)k = O_{m,k}(1)$. The result follows. □

References

[1] O. A. Camarena and B. Szegedy, *Nilspaces, nilmanifolds and their morphisms*, preprint (2010), `arXiv:1009.3825`.

[2] P. Candela, *Notes on nilspaces: algebraic aspects*, Discrete Anal. (2017), Paper No. 15, 59 pp.

[3] P. Candela, D. González-Sánchez and B. Szegedy *On nilspace systems and their morphisms*, Ergodic Theory Dynam. Systems **40** (2020), no. 11, 3015–3029.

[4] P. Candela, D. González-Sánchez, B. Szegedy, *Free nilspaces, double-coset nilspaces, and Gowers norms*, preprint (2023), `arxiv:2305.11233`.

[5] P. Candela, D. González-Sánchez, B. Szegedy, *On higher-order Fourier analysis in characteristic p*, Ergodic Theory Dynam. Systems **43** (2023), no. 12, 3971–4040.

[6] P. Candela, D. González-Sánchez, B. Szegedy, *On the inverse theorem for Gowers norms in abelian groups of bounded torsion*, preprint, arXiv:2311.13899.

[7] P. Candela and B. Szegedy, *Regularity and inverse theorems for uniformity norms on compact abelian groups and nilmanifolds*, J. Reine Angew. Math. **789** (2022), 1–42.

[8] J. Feng, *A note on computing of positive integer powers for circulant matrices*, Appl. Math. Comput. Appl. Math. Comput. **223** (2013), 472–475.

[9] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), no. 3, 465–588.

[10] W. T. Gowers, L. Milićević, *An inverse theorem for Freiman multi-homomorphisms*, preprint (2020) `arXiv:2002.11667` .

[11] B. Green, T. Tao, *An inverse theorem for the Gowers $U^3(G)$-norm*, Proc. Edinb. Math. Soc. (2) **51** (2008), no. 1, 73–153.

[12] B. Green, T. Tao and T. Ziegler, *An inverse theorem for the Gowers $U^{s+1}[N]$-norm*, Ann. of Math. (2) **176** (2012), no. 2, 1231–1372.

[13] Y. Gutman, F. Manners, P. P. Varjú, *The structure theory of nilspaces I*, J. Anal. Math. **140** (2020), no. 1, 299–369.

[14] A. Jamneshan, T. Tao, *The inverse theorem for the U^3 Gowers uniformity norm on arbitrary finite abelian groups: Fourier-analytic and ergodic approaches*, Discrete Anal., 2023:11, 48 pp.

[15] A. Jamneshan, O. Shalom, T. Tao, *The structure of totally disconnected Host–Kra–Ziegler factors, and the inverse theorem for the U^k Gowers uniformity norms on finite abelian groups of bounded torsion*, preprint (2023), `arXiv:2303.04860`.

[16] F. Manners, *Quantitative bounds in the inverse theorem for the Gowers U^{s+1}-norms over cyclic groups*, preprint (2018), `arXiv:1811.00718`.

[17] B. Szegedy, *Structure of finite nilspaces and inverse theorems for the Gowers norms in bounded torsion groups*, preprint (2010), `arXiv:1011.1057`.

[18] T. Tao and T. Ziegler, *The inverse conjecture for the Gowers norm over finite fields via the correspondence principle*, Anal. PDE **3** (2010), no. 1, 1–20.

[19] T. Tao and T. Ziegler, *The inverse conjecture for the Gowers norm over finite fields in low characteristic*, Ann. Comb. **16** (2012), no. 1, 121–188.

Complexity measures of trilean functions *

Sara Asensio[†1], Ignacio García-Marco[‡2], and Kolja Knauer[§3]

[1]Instituto de Investigación en Matemáticas (IMUVa), Universidad de Valladolid
[2]Instituto de Matemáticas y Aplicaciones (IMAULL), Universidad de La Laguna
[3]Universitat de Barcelona

Abstract

The study of the relations between different complexity measures of boolean functions led Nisan and Szegedy to state the sensitivity conjecture in 1994. This problem remained unsolved until 2019, when Huang proved the conjecture by means of an equivalent reformulation of the problem in graph theory. We wonder if the same type of results hold for functions defined on finite alphabets of cardinality greater than two. This work concerns functions over an alphabet with three symbols, which we call *trilean functions*. In this context we follow the steps of Nisan and Szegedy and extend most of the results for boolean functions. Also, we find an equivalent reformulation in graph theoretical terms of the trilean version of the sensitivity conjecture.

1 Introduction

One can measure how complex a given boolean function $f : \{0,1\}^n \to \{0,1\}$ is in many ways, and these different conceptions give rise to different complexity measures such as: the degree, the sensitivity, the block sensitivity, the decision tree complexity... In 1994, it was already known that all these complexity measures were polynomially related except the sensitivity. This led Nisan and Szegedy to the statement of the sensitivity conjecture [5], which claimed that the sensitivity was also polynomially related to the other measures. This conjecture was proved almost 30 years later by Huang [4], and his proof relies on an equivalence theorem due to Gotsman and Linial [3] which translated the sensitivity conjecture to an equivalent problem in graph theory. An extended summary of this story can be found in [1].

After this study, it is natural to ask whether there is a similar situation with functions $f : T^n \to T$, where T is a finite set of cardinality $m > 2$. Our work focuses on the case $m = 3$ and its corresponding functions, which we call *trilean functions*. For technical reasons, our set T will be the set $\{1, \varepsilon, \varepsilon^2\}$ where ε is a primitive cubic root of unity.

This abstract is divided in four sections. In the first one, we define two complexity measures of a trilean function f, namely the sensitivity $s(f)$ and the degree $\deg(f)$. In Proposition 4 we provide a quadratic upper bound for the sensitivity in terms of the degree. This was also known for boolean functions before 1994. Again, the most difficult question seems to be the construction of a polynomial bound in the other direction. In the second section we prove in Theorem 6 that providing an upper bound for the degree in terms of the sensitivity is equivalent to a graph theoretical problem that can be entirely stated in terms of a generalization of the n-dimensional hypercube. In the third section, by means of Theorem 6 we show that there are trilean functions such that $s(f) < \sqrt{2\deg(f)} + 1$. We conclude with a future work section.

*This research is supported by the Spanish MICINN through grant PID2022-137283NB-C22, and also by the "European Union NextGenerationEU/PRTR" through grant TED2021-130358B-I00 in the case of the first author

[†]Email: sara.asensio@uva.es
[‡]Email: iggarcia@ull.edu.es.
[§]Email: kolja.knauer@googlemail.com

2 Two complexity measures and the first relation between them

First of all, let's define two complexity measures of trilean functions which generalize those of boolean functions. The first one is the degree, whose definition relies on the following result:

Proposition 1. *Let $f : \{1, \varepsilon, \varepsilon^2\}^n \to \{1, \varepsilon, \varepsilon^2\}$ be a trilean function. Then, there is a unique polynomial $F \in \mathbb{C}[x_1, \ldots, x_n]$ of degree at most 2 in each variable that represents f (i.e., $F(\mathbf{x}) = f(\mathbf{x})$ at every $\mathbf{x} \in \{1, \varepsilon, \varepsilon^2\}^n$).*

Definition 2. *The* degree *of a trilean function $f : \{1, \varepsilon, \varepsilon^2\}^n \to \{1, \varepsilon, \varepsilon^2\}$ is the degree of the unique polynomial of degree at most 2 in each variable that represents f.*

The second complexity measure of trilean functions we are considering is the sensitivity.

Definition 3. *The* local sensitivity *of a trilean function f at $\mathbf{x} \in \{1, \varepsilon, \varepsilon^2\}^n$, $s_{\mathbf{x}}(f)$, is the number of elements $\mathbf{y} \in \{1, \varepsilon, \varepsilon^2\}^n$ which differ from $\mathbf{x}$ in exactly one entry and $f(\mathbf{x}) \neq f(\mathbf{y})$.*

The sensitivity *of f is $s(f) = \max_{\mathbf{x} \in \{1, \varepsilon, \varepsilon^2\}^n} \{s_{\mathbf{x}}(f)\}$.*

We have the following polynomial upper bound for the sensitivity of a trilean function in terms of its degree:

Proposition 4. *For every trilean function $f : \{1, \varepsilon, \varepsilon^2\}^n \to \{1, \varepsilon, \varepsilon^2\}$, $s(f) \leq 64 \deg(f)^2$.*

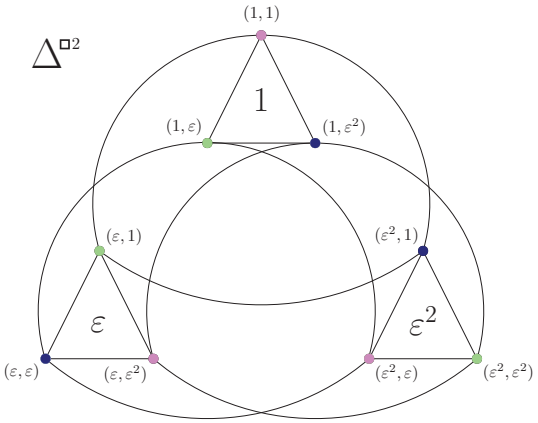


Figure 1: A proper 3-coloring of $\Delta^{\square 2}$.

3 The equivalence theorem

The proof of the sensitivity conjecture for boolean functions would not have been possible without the equivalence theorem by Gotsman and Linial [3], which stated that solving the sensitivity conjecture was equivalent to solving a combinatorial problem in graph theory. That problem consisted on finding a polynomial lower bound on n for the maximum between the maximum degree of every induced subgraph of the n-dimensional hypercube with not exactly half of its vertices and the maximum degree of its complementary.

We have obtained an equivalence theorem in the trilean case, which is somehow more involved than the boolean one. Before presenting its statement, we need to introduce some definitions and notation.

In the case of boolean functions, the set $\{0, 1\}^n$ where they are defined can be seen as the vertex set of the n-dimensional hypercube Q_n. In the context of trilean functions, a natural generalization is the graph G whose vertex set is $\{1, \varepsilon, \varepsilon^2\}^n$ and where two vertices are adjacent if and only if they differ in exactly one of their entries. This graph can be described in terms of the cartesian product of graphs:

Definition 5. *Given two graphs G and H, its* cartesian product $G\Box H$ *is the graph with vertex set $V(G) \times V(H)$ and whose edges are the pairs $\{(u,v),(u',v')\}$ such that*

- *$u = u'$ and $\{v,v'\}$ is an edge of H, or*
- *$v = v'$ and $\{u,u'\}$ is an edge of G.*

If we denote the complete graph on three vertices (the triangle graph) by Δ, then the generalization of the n-dimensional hypercube in the above sense is the graph $\Delta\Box \overset{n)}{\cdots} \Box\Delta = \Delta^{\Box n}$. We observe that the product of the entries of every vertex provides a proper 3-coloring of $\Delta^{\Box n}$ (see Figure 1 for a 3-coloring of $\Delta^{\Box 2}$).

We are going to denote each set of the resulting tripartition by C_i with $i \in \{1, \varepsilon, \varepsilon^2\}$; i.e.,

$$C_i = \left\{ \mathbf{x} = (x_1, \ldots, x_n) \in \{1,\varepsilon,\varepsilon^2\}^n \;\middle|\; \prod_{j=1}^{n} x_j = i \right\} .$$

Furthermore, given three induced subgraphs $H_1, H_\varepsilon, H_{\varepsilon^2}$ of $\Delta^{\Box n}$ whose vertex sets constitute a tripartition of $\{1,\varepsilon,\varepsilon^2\}^n$, we define three new induced subgraphs (see Figure 2):

- H'_1, with $V(H'_1) = (V(H_1) \cap C_1) \cup (V(H_{\varepsilon^2}) \cap C_\varepsilon) \cup (V(H_\varepsilon) \cap C_{\varepsilon^2})$.
- H'_ε, with $V(H'_\varepsilon) = (V(H_\varepsilon) \cap C_1) \cup (V(H_1) \cap C_\varepsilon) \cup (V(H_{\varepsilon^2}) \cap C_{\varepsilon^2})$.
- H'_{ε^2}, with $V(H'_{\varepsilon^2}) = (V(H_{\varepsilon^2}) \cap C_1) \cup (V(H_\varepsilon) \cap C_\varepsilon) \cup (V(H_1) \cap C_{\varepsilon^2})$.

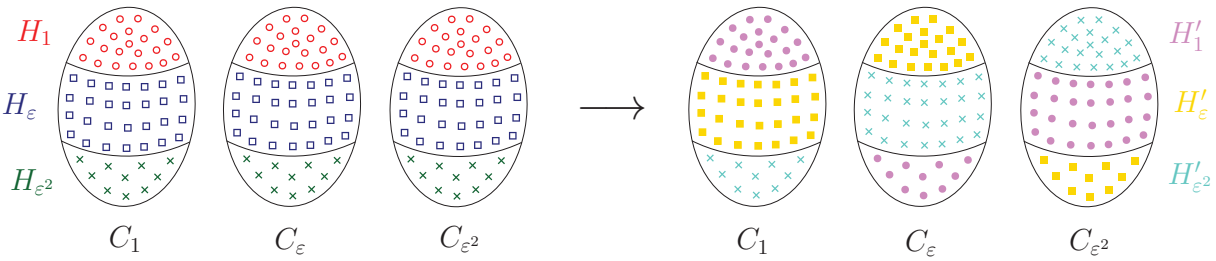


Figure 2: The vertices of $H'_1, H'_\varepsilon, H'_{\varepsilon^2}$ can be visualized as a certain cyclic rotation of the vertices of $H_1, H_\varepsilon, H_{\varepsilon^2}$.

Theorem 6 (Equivalence theorem)**.** *The following are equivalent for any function $h : \mathbb{N} \to \mathbb{R}$:*

1. *For any induced subgraphs H_1, H_ε and H_{ε^2} of $\Delta^{\Box n}$ such that their vertex sets constitute a tripartition of $V(\Delta^{\Box n})$ and the corresponding H'_1, H'_ε and H'_{ε^2} do not verify that $|V(H'_1)| = |V(H'_\varepsilon)| = |V(H'_{\varepsilon^2})| = 3^{n-1}$,*

$$\Gamma(H_1, H_\varepsilon, H_{\varepsilon^2}) = 2n - \min\{\delta(H_1), \delta(H_\varepsilon), \delta(H_{\varepsilon^2})\} \geq h(n) \ ,$$

   where $\delta(H_i)$ is the minimum degree of H_i for all $i \in \{1,\varepsilon,\varepsilon^2\}$.

2. *For any trilean function $f : \{1,\varepsilon,\varepsilon^2\}^n \to \{1,\varepsilon,\varepsilon^2\}$, $s(f) \geq h\left(\frac{1}{2}\deg(f)\right)$.*

Hence, if we could prove the first statement for a polynomial function h, we would have proved the sensitivity conjecture in the trilean case.

4 Searching for trilean functions with low sensitivity and high degree

In the case of boolean functions, Chung, Füredi, Graham and Seymour [2] proved that there exists an induced subgraph of the n-dimensional hypercube with $2^{n-1}+1$ vertices whose maximum degree is strictly smaller than $\sqrt{n}+1$. By Gotsman-Linial's equivalence theorem, this resulted in the existence of boolean functions with $s(f) < \sqrt{\deg(f)}+1$. Interestingly, Huang later proved that $s(f) \geq \sqrt{\deg(f)}$ for every boolean function f.

In the trilean case, we have proved that there are arbitrary large values of n such that there exist three induced subgraphs H_1, H_ε and H_{ε^2} of $\Delta^{\square n}$ with the following properties:

- their vertices constitute a tripartition of $\{1, \varepsilon, \varepsilon^2\}^n$,
- the corresponding H_1', H_ε' and H_{ε^2}' do not verify that $|V(H_1')| = |V(H_\varepsilon')| = |V(H_{\varepsilon^2}')| = 3^{n-1}$, and
- $\Gamma(H_1, H_\varepsilon, H_{\varepsilon^2}) = 2n - \min\{\delta(H_1), \delta(H_\varepsilon), \delta(H_{\varepsilon^2})\} < 2\sqrt{n}+1$.

This construction together with Theorem 6 yields the following:

Proposition 7. *For n sufficiently large there exist trilean functions $f : \{1, \varepsilon, \varepsilon^2\}^n \to \{1, \varepsilon, \varepsilon^2\}$ such that*

$$s(f) < \sqrt{2\deg(f)} + 1.$$

5 Future work

The main open problem now is to determine whether the sensitivity conjecture is true or not in the trilean case, which is essentially a combinatorial problem in graph theory thanks to the equivalence theorem. Furthermore, it would be interesting to generalize all of our results to the case of functions on finite sets of cardinality greater than three.

References

[1] S. Asensio, La conjetura de la sensibilidad y su resolución vía teoría de grafos, *La Gaceta de la RSME* **26** (2023), 131-148.

[2] F. R. K. Chung, Z. Füredi, R. L. Graham and P. Seymour, On induced subgraphs of the cube, *J. Combin. Theory Ser. A* **49** (1988), 180-187.

[3] C. Gotsman and N. Linial, The equivalence of two problems on the cube, *J. Combin. Theory Ser. A* **61** (1992), 142-146.

[4] H. Huang, Induced subgraphs of hypercubes and a proof of the sensitivity conjecture, *Ann. of Math.* **190** (2019), 949-955.

[5] N. Nisan and M. Szegedy, On the degree of Boolean functions as real polynomials, *Comput. Complexity* **4** (1994), 301-313.

Geometric quasi-cyclic low density parity check codes Discrete Mathematics Days 2024*

Simeon Ball[†1] and Tomas Ortega[‡2]

[1]Dept. of Mathematics, Universitat Politecnica Catalunya, 08034 Barcelona
[2]Dept. of Electrical Engineering and Computer Science, University of California, Irvine Irvine, CA 92697, USA.

Abstract

In this talk we present families of quasi-cyclic LDPC codes derived from the quasi-cyclic representation of the point-line incidence matrix of the classical finite generalised quadrangles. We detail how to explicitly calculate quasi-cyclic generator and parity check matrices for classical finite generalised quadrangles codes of length up to 400000. These codes cover a wide range of transmission rates, are easy and fast to implement and perform close to Shannon's limit.

1 Introduction

In many modern communication systems Low Density Parity Check (LDPC) codes are used. LDPC codes are those codes for which the number of ones in the check matrix is very small compared to the size of the matrix. A quasi-cyclic LDPC check matrix H can be described by a block size b, sometimes called the lifting degree or lifting factor, and a $(m/b) \times (n/b)$ matrix $\mathrm{H}^{\mathrm{rep}}$, whose entries are subsets H_{ij} of $\{1, \ldots, b\}$, where $i \in \{1, \ldots, (m/b)\}$ and $j \in \{1, \ldots, (n/b)\}$. Typically the subset H_{ij} is empty which corresponds to the $b \times b$ zero matrix in H in the (i, j) cell. A singleton subset $\mathrm{H}_{ij} = \{r\}$ indicates that in the (i, j) cell we have a copy of the $b \times b$ identity matrix shifted r bits (cyclically) to the right. A larger subset will involve a superposition of such shifts of the identity matrix. This representation of the quasi-cyclic LDPC check matrix H allows one to implement decoding algorithms, such as the sum-product algorithm, in an efficient manner.

The Tanner graph Γ is the bipartite graph with stable sets of size m and n, where there is a correspondence between the edges in Γ and a one entry in the matrix H. The decoding algorithms mentioned in the previous paragraph work well if the girth of Γ, the length of the shortest cycle, is large, and decode quickly if Γ has low diameter [6]. The diameter is the maximum distance between any two vertices. These conflicting objectives are optimised when the girth is twice the diameter. The graphs Γ which achieve this bound are the incidence matrices H of a generalised polygon. The rows of H are indexed by the points of the polygon and the columns are indexed by the lines, or vice-versa, where there is a one entry in the matrix H if and only if the point indexing the column and the line indexing the row are incident in the geometry. Finite generalised polygons have diameter 3, 4, 6 or 8, see [3], and are respectively called, projective planes, generalised quadrangles, generalised hexagons and generalised octagons. The LDPC code used in IEEE 802.3 standard (2048,1723) LDPC code for the 10-G Base-T Ethernet, is a quasi-cyclic LDPC code from an affine plane over $\mathbb{F}_{32}$ (a projective plane over the field of 32 elements with a line deleted) which has block size $b = 64$, length $n = 2048$

*The full version of this work can be found in [1] and [2]. This research is supported by the Spanish Ministry of Science, Innovation and Universities grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

†Email: simeon.michael.ball@upc.edu.
‡Email: tomaso@uci.edu.

and dimension $k = 1723$, see [7, Example 10.5]. The LDPC code used in the NASA Landsat Data Continuation is a quasi-cyclic LDPC code from a 3-dimensional affine space (a projective space with a plane deleted) which has block size $b = 511$, length $n = 8176 = 16b$ and dimension $k = 7154 = 14b$, see [7, Example 10.10]. In this talk, we will describe how to efficiently employ quasi-cyclic LDPC codes derived from classical generalised quadrangles. These codes are fast and efficient, can be extremely long, and perform favourably compared to commercially used codes with similar parameters.

2 Quasi cyclic generator and check matrices

It was proven in [5] that the classical generalised quadrangles, of which there are six types, have a quasi-cyclic representation. However, up until now, no description of these quasi-cyclic representations was known. Here, we detail a simple, explicit description of a quasi-cyclic representation for all the classical generalised quadrangles. Using this representation, we describe how to employ the corresponding quasi-cyclic LDPC code in an efficient manner. In four of the six types, we do not take the entire quadrangle but a carefully chosen large sub-structure, which allows us to increase the size of the blocks b. It is advantageous to have a large block size since this allows the implementation of significantly longer codes. As evidenced in the proof of Shannon's theorem, the implementation of long codes brings the performance of the code close to Shannon's limit.

Once we have described how to construct the quasi-cyclic representation of the check matrix H in a purely algebraic manner, we can compute H for classical generalised quadrangles LDPC codes of length up to 400000. These computations were performed on a standard laptop and one can compute quasi-cyclic check matrices and generator matrices for longer codes with more computational power. The complexity of these computations for each quadrangle is detailed in Table 1.

Let C denote the binary linear code whose check matrix is H. We shall refer to C as the *full code*. Efficient encoding can be implemented if one can find a generator matrix for the code in quasi-cyclic form, see [4]. However, such a generator matrix for the full code C does not generally exist, so we take a large subcode C' of C for which there is a generator matrix G in quasi-cyclic form. We will call C' the *implementable code*. A $k \times n$ generator matrix is in standard form if it has $k \times k$ submatrix which is an identity matrix. For each quadrangle and each q, we compute a generator matrix G in standard quasi-cyclic form for the implementable code C'. Note that H is also a check matrix for C'. The matrix G can be described by a $(k/b) \times ((n-k)/b)$ matrix $\mathrm{P}^{\mathrm{rep}}$, where the (i,j) entry of $\mathrm{P}^{\mathrm{rep}}$ is P_{ij}, a vector of $\{0,1\}^b$. Replacing each first row, with the full circulant $b \times b$ matrix one obtains a matrix P, where

$$\mathrm{G} = (\mathrm{P} \mid \mathrm{id}). \tag{1}$$

Here, id denotes the $k \times k$ identity matrix.

Given the matrix $\mathrm{P}^{\mathrm{rep}}$, a shift-register-adder-accumulator (SRAA) circuit with a b-bit feedback shift register can be implemented to calculate each block of b parity check bits of the encoded codeword, see [4, Figure 1]. In series, this gives an encoding circuit of $(n-k)/b$ SRAA circuits with a total of $2(n-k)$ flip-flops, $n-k$ AND gates, and $n-k$ two-input XOR gates. The encoding is completed in a time proportional to $n-k$, see [4, Figure 2]. An encoder which completes in $n-k$ clock cycles and k/b feedback shift registers, each with b flip-flops can be implemented when the circuits are put in parallel, see [4, Figure 3].

Thus, we have an efficient encoding and decoding of very long quasi-cyclic LDPC codes whose performance is close to Shannon's limit.

In the talk we will give a simple algebraic description of the quasi-cyclic representations which allow the construction of $\mathrm{H}^{\mathrm{rep}}$ for codes of extraordinary length. In some cases it is feasible to calculate $\mathrm{H}^{\mathrm{rep}}$ for codes of length 10^7. One can calculate $\mathrm{P}^{\mathrm{rep}}$, and thus an explicit generator matrix in standard quasi-cyclic form for codes of length up to 400000 and transmission rates covering a wide spectrum of possible rates.

Code	increased block size b	block size	approx length n	complex -ity $\mathrm{H}^{\mathrm{rep}}$	complex -ity $\mathrm{P}^{\mathrm{rep}}$	min. dist.	approx. rate
$W(3,q)$	q^2+1	q^2+1 (q even) $\frac{1}{2}(q^2+1)$ (q odd)	q^3	$O(q^4)$	$O(q^9)$	$\geqslant 2q$	$1-q^{-0.286}$ (q even) 0.5 (q odd)
$W(3,q)$ dual	q^2+1	q^2+1 (q even) $\frac{1}{2}(q^2+1)$ (q odd)	q^3	$O(q^4)$	$O(q^9)$	$\geqslant 2q$	$1-q^{-0.286}$ (q even) 0.5 (q odd)
$Q(5,q)$	q^3+1	q^2-q+1 ($q=0,1 \bmod 3$) $\frac{1}{3}(q^2-q+1)$ ($q=2 \bmod 3$)	q^5	$O(q^6)$	$O(q^{13})$	$\geqslant 2q$	$1-q^{-1}$
$Q(5,q)$ dual	q^3+1	q^2-q+1 ($q=0,1 \bmod 3$) $\frac{1}{3}(q^2-q+1)$ ($q=2 \bmod 3$)	q^4	$O(q^6)$	$O(q^{14})$	$\geqslant q^3$	q^{-1}
$H(4,q^2)$	$\frac{q^5+1}{q+1}$	$\frac{q^5+1}{q+1}$	q^8	$O(q^{11})$	$O(q^{22})$	$\geqslant 2q^2$	$1-q^{-1}$
$H(4,q^2)$ dual	$\frac{q^5+1}{q+1}$	$\frac{q^5+1}{q+1}$	q^7	$O(q^{11})$	$O(q^{23})$	$\geqslant q^5$	q^{-1}

Table 1: The block size, length, and complexity of constructing $\mathrm{H}^{\mathrm{rep}}$ and $\mathrm{P}^{\mathrm{rep}}$.

3 Quasi-cyclic LDPC codes from classical generalised quadrangles.

Although there are six types of classical generalised quadrangles, these come in pairs, where one is the dual of the other. To obtain the dual quadrangle one switches the role of the points and the lines. In practical terms this is achieved by replacing the check matrix H by its transpose. Thus, we only need to describe how to construct H for three of the six types. These are labelled in Table 1 as $W(3,q)$ (the symplectic quadrangle), $Q(5,q)$ (the elliptic quadrangle) and $H(4,q^2)$ (the Hermitian quadrangle).

We will describe how to find these quasi-cyclic representations by consider the geometries as subsets of certain field extensions. This will lead us to increased block sizes which are detailed in Table 1. I will also present data on how these codes perform with respect to Shannon's bound. It is not possible to simulate performance of codes of very long length without using a field programmable gate array. The very longest codes, for example the quasi-cyclic LDPC codes arising from Q(5,13) with $n = 371462$ and rate $R = 0.9172$ were implemented using a field programmable gate array. It was seen empirically that these codes work exceedingly well with low complexity decoding algorithms which require just a few iterations. This indicates that they may have a use in storing large amounts of data, where fast and reliable decoding can be employed on retrieval.

References

[1] S. Ball and T. Ortega, Quasi-cyclic LDPC codes based on generalised quadrangles, Patent Cooperation Treaty, International Application No. PCT/EP2023/062797, World Intellectual Property Organization WO/2023/218050 (2023).

[2] S. Ball and T. Ortega, Practical implementation of geometric quasi-cyclic low density parity check codes, preprint, 2024.

[3] W. Feit and G. Higman, The nonexistence of certain generalised polygons, *J. Algebra*, **1** (1964) 114–131.

[4] Z. Li, L. Chen, L. Zeng and S. Lin, Efficient encoding of quasi-cyclic low-density parity-check codes, *IEEE Trans. Inform. Theory*, **54** (2006) 71–81.

[5] Z. Liu and D. Pados, LDPC codes from generalised polygons, *IEEE Trans. Inform. Theory*, **51** (2005) 3890–3898.

[6] G. A. Margulis, Explicit constructions of graphs without short cycles and low-density codes, *Combinatorica*, **2** (1982) 71–78.

[7] W. E. Ryan and S. Lin, *Channel codes: Classical and Modern*, Cambridge University Press, 2009.

On additive codes over finite fields *

Simeon Ball, Michel Lavrauw, and Tabriz Popatia

Abstract

In this article we prove a Griesmer type bound for additive codes over finite fields. This new bound gives an upper bound on the length of fractional maximum distance separable codes, codes which attain the Singleton bound. We prove that this bound can be obtained in some cases, surpassing the length of the longest known codes in the non-fractional case. We also provide some exhaustive computational results over small fields and dimensions.

1 Introduction

The full version of this work can be found in [1].

Let $\mathbb{F}_q$ denote the finite field with q elements. An *additive code* of length n over $\mathbb{F}_{q^h}$ is a subset C of $\mathbb{F}_{q^h}^n$ with the property that for all $u, v \in C$ the sum $u + v \in C$. It is easy to prove that an additive code is linear over some subfield, which we will assume to be $\mathbb{F}_q$. An additive code is *linear over* $\mathbb{F}_{q^h}$ if $u \in C$ implies $\lambda u \in C$ for all $u \in C$ and all $\lambda \in \mathbb{F}_{q^h}$.

We use the notation $[n, r/h, d]_q^h$ code to denote an additive code of length n over $\mathbb{F}_{q^h}$, of size q^r and minimum distance d. For an additive code C, the minimum distance d is equal to the minimum weight, so this is equivalent to saying that each non-zero vector in C has at least d non-zero coordinates. We will be particularly interested in the case where r/h is not an integer, we will call these codes *fractional codes*, and codes such that $r/h \in \mathbb{N}$ *integral codes*.

2 Griesmer bound for additive codes

The Griesmer bound [4] for linear codes states that, if there is a $[n, k, d]_q$ linear code then

$$n \geqslant \sum_{j=0}^{k-1} \lceil \frac{d}{q^j} \rceil.$$

This bound can be reformulated as n $\geqslant k + d - m + \sum_{j=1}^{m-1} \lceil \frac{d}{q^j} \rceil$, where we chose $m \leqslant k-1$ such that $q^m < d \leqslant q^{m+1}$, or $m = k$ if $q^k < d$. We prove that a similar but weaker bound holds for additive codes over finite fields,

Theorem 1. *If there is a $[n, r/h, d]_q^h$ additive code then*

$$n \geqslant \lceil r/h \rceil + d - m - 2 + \lceil \frac{d}{f(q,m)} \rceil,$$

where $r = (k-1)h + r_0$, $1 \leqslant r_0 \leqslant h$,

$$f(q,m) = \frac{q^{mh+r_0}(q^h - 1)}{q^{mh+r_0} - 1}$$

*5 April 2024. The first and third author are supported by the Spanish Ministry of Science, Innovation and Universities grant PID2020-113082GB-I00 funded by MICIU/AEI/10.13039/501100011033.

and $0 \leqslant m \leqslant k-2$ *is such that*

$$q^{mh+r_0} < d \leqslant q^{(m+1)h+r_0}$$

or $m = k-2$ *if* $d > q^r$.

We will also show that simply replacing k by $\lceil r/h \rceil$ or $\lfloor r/h \rfloor$ in (2) does not lead to a valid bound for additive codes by constructing additive codes that invalidate these natural generalisations. Furthermore we will show that we can reach the bound given by Theorem 1.

An additive $[n, r/h, d]_q^h$ code is equivalent to the following geometric structure. We define $\mathcal{X}$ to be the set of n subspaces of dimension at most $h-1$ in $\mathrm{PG}(r-1, q)$ with the property that at most $n-d$ of the subspaces are contained in a hyperplane. This representation of additive codes allows us to look at the codes geometrically, which is instrumental in a lot of the proofs and intuitions in this article. This is especially true for constructions of additive codes and classifying their dual code.

3 Singleton bound for additive codes

We will be particularly interested in codes C which meet the Singleton bound

$$|C| \leq q^{n-d+1},$$

for linear codes the Singleton bound can be reformulated as

$$k \leqslant n - d + 1.$$

Codes which attain this bound are called maximum distance separable codes, or simply MDS codes. MDS codes are an important class of codes, which are implemented in many applications where we can allow q to be large.

The Griesmer bound gives two important bounds for linear MDS codes. These results are well-known, but we list these as theorems since we will obtain similar bounds for additive MDS codes.

Theorem 2. *If* $k \geqslant 2$ *and there is a* $[n, k, d]_q$ *linear MDS code then* $d \leqslant q$ *and* $n \leqslant q + k - 1$.

Theorem 3. *If* $n \geqslant k + 2$ *and there is a* $[n, k, d]_q$ *linear MDS code then* $k \leqslant q - 1$.

As observed in previous articles, [5, Theorem 10], the Singleton bound can be reformulated for additive codes as

$$k = \lceil r/h \rceil \leqslant n - d + 1,$$

since n and d are always integers. We call codes which attain this bound additive MDS codes. Interestingly, the restrictions from the above theorems do not carry over to additive codes. We will provide a version for these bounds for additive codes and examples which better these bounds.

From Theorem 1 we get the following theorem for additive MDS codes that is the equivalent to Theorem 2.

Theorem 4. *If there is an* $[n, r/h, d]_q^h$ *additive MDS code then*

$$d \leqslant q^h - 1 + \frac{q^h - 1}{q^{r_0} - 1}$$

and

$$n \leqslant k - 2 + q^h + \frac{q^h - 1}{q^{r_0} - 1},$$

where $r = (k-1)h + r_0 > h$ *and* $1 \leqslant r_0 \leqslant h$.

In Theorem 2 we have that $d \leqslant q^h$, so the bound in Theorem 4 is the same when $r_0 = h$ and slightly weaker when $r_0 \neq h$. Furthermore notice a fractional additive MDS code n can be longer than the linear code by a tail of $\frac{q^h-1}{q^{r_0}-1} - 1$. We also have a an equivalent result to Theorem 3 for additive codes given by

Theorem 5. *If there is an $[n, r/h, d]_q^h$ additive MDS code C and $n \geqslant \lceil r/h \rceil + 2$ then*

$$\lceil r/h \rceil \leqslant q^h - 1 + (r_0 - h)\frac{q^h - q^{h-1}}{q^h - 1}.$$

In the next theorem we construct some additive MDS codes which prove that the bound in Theorem 4 is attainable if r_0 divides h and $k = 2$. Observe that Theorem 2 implies that for linear codes with $k = 2$, $n \leqslant q^h + 1$, so the following construction exceeds the bound of linear codes.

Theorem 6. *If r_0 divides h then there is a $[n, 1 + (r_0/h), n - 1]_q^h$ additive MDS code where*

$$n = q^h + \frac{q^h - 1}{q^{r_0} - 1}.$$

A similar construction can be used to prove that the bound in Theorem 4 is also attainable when $k = 3$ and $q = 2$, given by

Theorem 7. *There is a $[2^{h+1}, 2 + 1/h, 2^{h+1} - 2]_2^h$ additive MDS code.*

Although the codes constructed in Theorem 6 and Theorem 7 have very small rate, it is of interest that their length is superior to that of their linear counterparts. In the article [1] we also classify additive MDS codes over small fields and observe, once more, that there are additive MDS codes whose length exceeds linear MDS codes.

In the previous two constructions, we have that k is small. In the following constructions, we look at the other extreme, when d is small.

Theorem 8. *Let π_0 and π_∞ be h dimensional subspace of $\mathbb{F}_q^{2h} = \pi_0 \oplus \pi_\infty$. If there are k r_i-dimensional subspace π_i (with $r_i < h$) such that*

$$\pi_i \cap \pi_j = \{0\}$$

for all distinct $i, j \in \{0, 1 \ldots, k, \infty\}$, and with

$$\sum_{i-1}^{k} r_i = r,$$

then there exists an $[\lceil r/h \rceil] + 2, r/h, 3]_q^h$ additive MDS code.

From Theorem 8 we get the following theorem, which implies that the bound in Theorem 5 is atainable when $r_0 = h - 1$, given by

Theorem 9. *There exists an $[\lceil r/h \rceil] + 2, r/h, 3]_q^h$ additive MDS code for all $\lceil r/h \rceil \leq q^h - 1$*

4 The dual of an additive code

The dual of an additive $[n, r/h, d]_q^h$ code C is defined by

$$C^\perp = \{v \in \mathbb{F}_{q^h}^n \mid \mathrm{tr}_{q^h \mapsto q}(u \cdot v) = 0, \text{ for all } u \in C\},$$

where $\mathrm{tr}_{q^h \mapsto q}$ denotes the trace function from $\mathbb{F}_{q^h}$ to $\mathbb{F}_q$, and $u \cdot v$ denotes the euclidean inner product.

Since an $\mathbb{F}_q$-basis for C defines r equations for the $\mathbb{F}_q$-subspace $C^\perp$, the dual code is an additive $[n, n-r/h, d^\perp]_q^h$ code. Note that as an $\mathbb{F}_q$-vector space the vector space $\mathbb{F}_{q^h}^n$ has dimension hn implying that $|C^\perp| = q^{nh-r}$.

Recall that in Section 2, we defined a set of subspaces $\mathcal{X}$ of $PG(r-1,q)$ from an additive code of size q^r. We say an additive code is *full* if all the elements of $\mathcal{X}$ are of rank h, i.e. projective $(h-1)$-dimensional subspaces of $PG(r-1,q)$ (and *non-full* otherwise). We can always extend the subspaces of $\mathcal{X}$ so that they have rank r. This can be done arbitrarily without the minimum distance decreasing. Thus, a non-full additive $[n, r/h, d]_q^h$ code C can always be converted in a full additive $[n, r/h, \geqslant d]_q^h$ code. Using the concept of full and non-full additive codes we get the following results.

Theorem 10. *C is a non-full additive $[n, r/h, d]_q^h$ code if and only if $C^\perp$ is an additive $[n, n-r/h, 1]_q^h$ code.*

Theorem 11. *C is a full additive $[n, r/h, d]_q^h$ code with $d \geqslant 2$ if and only if $C^\perp$ is a full additive $[n, r/h, d^\perp]_q^h$ code with $d^\perp \geqslant 2$.*

Theorem 12. *A fractional sub-code of a linear code C is non-full.*

For linear codes we have that that the dual of an MDS code is also MDS. For additive codes the following result from [3, Theorem 4.3] (see also [6, Theorem 3.3]) classifies a group of additive MDS codes with a dual that is also MDS.

Theorem 13. *The dual of an integral additive MDS code is an additive MDS code.*

In the fractional case, as pointed out in [6, Example 3.1] the dual of a fractional MDS code is not necessarily MDS. Here, we give a precise condition on when the dual of a fracitonal MDS code is also MDS,

Let J be a subset of $\{1, \ldots, n\}$. The *projection* of an additive code C at J is

$$C/J = \{u \in C \mid u_j = 0, \text{ for all } j \in J\}.$$

Geometrically, the code C/J can be obtained from the set $\mathcal{X}$ in the following way. Let $\mathcal{X}_J$ be the set of subspaces corresponding to the coordinates of J. By projecting from the subspace Σ spanned by X_J, we obtain a set of $n - |J|$ subspaces $\mathcal{X}/J$ which are the subspaces of $\mathcal{X} \setminus \mathcal{X}_J$ projected from Σ. Note that the operation of projection of a code is also known as shortening a code.

We say a projection is *non-obliterating* if the dimension of C/J is at least h, i.e. $|C/J| \geqslant q^h$. Note that an obliterating projection always yields a non-full additive code.

Theorem 14. *Let C be an additive MDS code. The dual code $C^\perp$ is an additive full MDS code if and only if every non-obliterating projection of C is full.*

We can now improve on our bound for n in the case that the MDS code is non-full.

Theorem 15. *If there is a non-full $[n, r/h, d]_q^h$ additive MDS code then*

$$n \leqslant q^h + k - 1 + \frac{q^h - q^{r_0+1}}{q^{r_0} - 1},$$

where $r = (k-1)h + r_0 > h$ and $1 \leqslant r_0 \leqslant h$.

References

[1] S. Ball, M. Lavrauw, and T. Popatia, On additive codes over finite fields, preprint (2024).

[2] S. Ball and M. Lavrauw, Arcs in finite projective spaces, EMS Surveys in Mathematical Science, 6 (2019) 133–172.

[3] S. Ball, G. Gamboa and M. Lavrauw, On additive MDS codes over small fields, *Adv. Math. Commun.*, **17** (2023) 828–844.

[4] J. H. Griesmer, A bound for error-correcting codes, *IBM Journal of Res. and Dev.*, **4** (1960) 532–542.

[5] W. Cary Huffman, On the theory of $\mathbb{F}_q$ linear $\mathbb{F}_{q^t}$-codes, *Adv. Math. Commun.*, **7** (2013) 349–878.

[6] M.Yadav and A. Sharma, Some new classes of additive MDS and almost MDS codes over finite fields, *Finite Fields Appl.* **95** (2024) 102394.

Increasing paths in the temporal stochastic block model Discrete Mathematics Days 2024

Sofiya Burova*[1,2], Gábor Lugosi†[2,3,5], and Guillem Perarnau‡[1,4]

[1]IMTECH and Departament de Matematiques, Universitat Politecnica de Catalunya, Spain
[2]ICREA, Pg. Lluis Companys 23, 08010 Barcelona, Spain
[3]Department of Economics and Business, Pompeu Fabra University, Barcelona, Spain
[4]Centre de Recerca Matematica, Barcelona, Spain
[5]Barcelona School of Economics, Barcelona, Spain

Abstract

We study random temporal graphs, in the context of the Stochastic Block Model. Temporal graphs naturally model time-dependent propagation processes, for instance social interactions or infection processes. In these graphs, every edge has a unique timestamp, chosen uniformly at random, and the notion of connectivity is limited to sequences of edges with timestamps that increase over time. Our goal is to understand the asymptotic behavior of the temporal diameter of these graphs, especially in the subcritical regime where the average number of connections per node is of order $c \log n$, with $c < 1$. We analyze the first-order asymptotics for various aspects, including the length of the longest increasing paths starting at a typical vertex, as well as the size of the set of vertices reachable from a typical vertex via increasing paths.

1 Introduction

The Stochastic Block Model (SBM) is a foundational random graph model used to understand and analyze the structural properties of networks. It models networks by dividing vertices into groups (blocks or communities) and specifying different connection probabilities between these groups. They are versatile tools used for clustering, community detection, anomaly detection, and link prediction. A Temporal Stochastic Block Model extends the traditional SBM by incorporating time into the edge formation process. Each edge in the network not only connects two nodes but also has an associated timestamp, indicating when the interaction occurred. The model can be extended to study dynamic processes over time, such as infection or diffusion processes. TSBMs can model the spread of diseases, where nodes represent individuals and edges represent interactions that could lead to transmission. Temporal information allows for the analysis of how infections spread over time. This perspective emphasizes the importance of *increasing paths* as a key object of interest.

In this work, we consider the stochastic block model of parameters $n \in \mathbb{N}$, $a, p_1, p_2, q \in [0, 1]$, denoted by $\mathbb{SBM}(n, a, p_1, p_2, q)$, comprised of two independent Erdős-Rényi *communities*, $G_1 \sim \mathbb{G}(\lfloor an \rfloor, p_1)$ and

*Email: sofiburova@gmail.com. Research of S.B. supported by the Spanish Agencia Estatal de Investigación under project PID2022-138268NB-100

†Email: gabor.lugosi@gmail.com. Research of G.L. supported by Ayudas Fundación BBVA a Proyectos de Investigación Científica 2021 and the Spanish Ministry of Economy and Competitiveness, Grant PID2022-138268NB-I00, financed by MCIN/AEI/10.13039/501100011033, FSE+MTM2015- 67304-P, and FEDER, EU

‡Email: guillem.perarnau@upc.edu. Research of G. P. supported by the Spanish Agencia Estatal de Investigación under projects PID2020-113082GB-I00 and the Severo Ochoa and Maria de Maeztu Program for Centers and Units of Excellence in R&D (CEX2020-001084-M)

$G_2 \sim \mathbb{G}(\lceil (1-a)n \rceil, p_2)$, where for all $u \in V(G_1), v \in V(G_2)$, the edge uv appears with probability q independently from all other edges. Additionally, to each edge e we assign a label $L(e)$, a uniform random variable on $[0,1]$ independent from all the others. Note that this is equivalent to taking a uniform permutation of $\{1,\dots,m\}$ where m is the total number of edges since we are only interested in the order of the edges with respect to one another. Let $G = (V, E, L) \sim \mathbb{SBM}(n, a, p_1, p_2, q)$ be a temporal stochastic block model where V is the vertex set, E is the set of edges and $(L(e))_{e\in E}$ is a family of iid[1] uniform random variables on $[0,1]$. We say a path $(w_1,\dots,w_k)$ is *increasing* if the sequence of labels $(L(v_iv_{i+1}))_{i\in[k-1]}$ is increasing and determine its length to be the number of edges in it. For $u, v \in V(G)$, we say u is reachable from v if there exists an increasing path from v to u. Denote by $B_\ell(v)$ where $\ell \in \mathbb{N}$, the set of reachable vertices from v via increasing paths of length ℓ. Let $B_n(v) = \cup_\ell B_\ell(v)$ be the reachable set of v in G.

The topic of random simple temporal graphs has been widely discussed in previous works ([2, 3, 4, 5]). It is known that the temporal Erdős-Rényi random graph $\mathbb{G}(n,p)$ undergoes a phase transition around $p = c\log n/n$. In particular, Casteigts et al. showed in [5] that the thresholds for the properties that a typical pair of vertices is connected, a typical vertex can reach all other vertices, and any pair of vertices is connected are, respectively, $\log n/n$, $2\log n/n$ and $3\log n/n$. Broutin, Kamčev and Lugosi, [4], recently extended these results providing tight bounds for the longest and shortest increasing paths in each regime, thus completing the study of the temporal diameter of $\mathbb{G}(n,p)$.

In this work, we generalize their approach to the stochastic block model as we suspect it exhibits similar behaviour. Set $p_1 = \lambda_1 \log n/n$, $p_2 = \lambda_2 \log n/n$ and $q = c\log n/n$. We focus on this regime as it is where the phase transition occurs in each community. We naturally assume that edges are more common within communities than between communities. In our setting, it suffices to assume that $c \le \sqrt{\lambda_1\lambda_2}$. The parameter of interest is

$$\theta = \theta(a,\lambda_1,\lambda_2,c) \coloneqq \frac{1}{2}(a\lambda_1 + (1-a)\lambda_2) + \frac{1}{2}\sqrt{(a\lambda_1-(1-a)\lambda_2)^2 + 4a(1-a)c^2}. \tag{1}$$

Notice that, the case $a = 1-a = 1/2$ and $\lambda_1 = \lambda_2 = c$ yields the $\mathbb{G}(n, c\log n/n)$ setting from [4]. We say that a sequence of events $(E_n)_{n\in\mathbb{N}}$ occurs *with high probability* (whp for brevity) if $\lim_{n\to+\infty} \mathbb{P}_n[E_n] = 1$. We are now ready to state our main results.

Theorem 1 (Reachability from a vertex)**.** *Let $w \in V(G)$. If $0 < c \le \sqrt{\lambda_1\lambda_2}$ and $\theta < 1$, then whp*

$$\frac{\log|B_n(w)|}{\log n} \to \theta.$$

We present a constructive algorithm for building the reachable set from a typical vertex via increasing paths. In particular, we show that one can embed into the temporal $\mathbb{SBM}$ a *uniform random recursive tree* (URRT) of size roughly n^θ, which is a well-understood random object, obtained by repeatedly attaching a leaf to a random vertex of the existing tree, thus providing deeper understanding of the structure of the graph. The main challenge is that the labels of the edges of the tree in the temporal stochastic block model depend on the number of vertices it contains from each community. We couple the evolution of these quantities with a Pólya urn process to keep track and use a result by Janson, [8], to understand their asymptotic behaviour. Furthermore, since the height of a random recursive tree of size m is known to be asymptotically $e\log m$ ([6, 9]), we can derive from this result a lower bound on the size of the longest path with a fixed starting point.

Theorem 2 (Longest increasing paths)**.** *If $0 < c \le \sqrt{\lambda_1\lambda_2}$ and $\theta < 1$, then whp*

(i) there are no increasing paths between two fixed vertices;

[1] independent and identically distributed

(ii) the size of the longest increasing path starting at a fixed vertex $w \in V(G)$, denoted by $\gamma_{\max}(w)$,

$$\frac{|\gamma_{\max}(w)|}{\log n} \to e\theta.$$

The upper bounds in both Theorem 1 and Theorem 2 are obtained by combinatoric computations and standard probabilistic methods, involving some cumbersome calculations, see Section 2. In Section 3, we prove the lower bounds through well-placed couplings, the most important of which is described in our key Lemma 10. To fit our setting, we extend a concentration inequality by Janson ([7]) on sums of independent exponential random variables of deterministic parameters to randomized ones, under some additional assumptions.

Remark 3. *Observe that $\theta = \theta(c)$ is increasing as a function of c, hence*

$$\theta(0) = \max\{a\lambda_1, (1-a)\lambda_2\} \leq \theta \leq a\lambda_1 + (1-a)\lambda_2 = \theta(\sqrt{\lambda_1\lambda_2})$$

since $0 \leq c \leq \sqrt{\lambda_1\lambda_2}$. It follows from the first inequality that the regime $\theta < 1$ renders both G_1 and G_2 subcritical.

Remark 4. *The threshold for the emergence of a giant component of linear size in the classic $\mathbb{SBM}$ is in the regime where $p_1 = \lambda_1/n, p_2 = \lambda_2/n, q = c/n$ and is known to be $\theta = 1$ ([1]). We expect the critical regime in the temporal setting to be delayed by a factor of $\log n$ and preserve the threshold $\theta = 1$ as is observed for $\mathbb{G}(n,p)$.*

2 Upper bounds: First moment method

2.1 Combinatorics of paths

Fix $u, v \in V(G)$ distinct. Denote by $\Lambda^k = \Lambda^k(u,v)$ the set of all self-avoiding paths on G of length $k \in \mathbb{N}$ from u to v. We partition

$$\Lambda_k = \bigcup_{\ell\in[k], h\in[k-\ell]} \Lambda^k_{\ell,h}$$

where ℓ is the number of inter-communal edges and h is the number of inner-communal edges of the community of origin used to construct the paths. By convention, for $i \in \{1,2\}$, we write $\bar{i} = 3 - i$.

Proposition 5. *Suppose $u \in V(G_i)$ and $v \in V(G_j)$, $i,j \in \{1,2\}$. Let $\ell \in [k], h \in [k-\ell]$ and set $s_1 = h + \lfloor \ell/2 \rfloor - \mathbb{1}_{j=i}$ and $s_2 = k - h - \lfloor \ell/2 \rfloor - \mathbb{1}_{j=\bar{i}}$. Then, the size of the set $\Lambda^k_{\ell,h}$ is given by*

$$|\Lambda^k_{\ell,h}| = |\Lambda^k_{\ell,h}(i,j)| = \begin{cases} \frac{n_i!}{(n_i-k-1)!} & \text{if } \ell = 0,\ h = k,\ i = j; \\ \frac{n_i!}{(n_i-s_1)!}\frac{n_{\bar{i}}!}{(n_{\bar{i}}-s_2)!}\binom{h+\lceil \frac{\ell+1}{2} \rceil - 1}{h}\binom{k-h-\lceil \frac{\ell+1}{2} \rceil}{k-\ell-h} & \text{if } \ell \in 2\mathbb{N}^* - |i-j|; \\ 0 & \text{otherwise.} \end{cases}$$

2.2 First moment method

Let $u \in V(G_i), v \in V(G_j)$ and ℓ, h be such that $\Lambda^k_{\ell,h} \neq \emptyset$. Then, for all $\gamma \in \Lambda^k_{\ell,h}$, define the event $A(\gamma) = \{\gamma \text{ is increasing and all its edges are present in } G\}$. It is straightforward that

$$\mathbb{P}[A(\gamma)] = \frac{p_i^h q^\ell p_{\bar{i}}^{k-\ell-h}}{k!} = \frac{(\log n)^k}{n^k k!}\lambda_i^h c^\ell \lambda_{\bar{i}}^{k-\ell-h}$$

where the $1/k!$ factor ensures that the path is increasing. Denote by X_k the number of increasing paths of length k in G. Denote by $Y_k(i)$ the number of increasing paths of length k that start at a fixed vertex in G_i and let $Y_k = Y_k(1) + Y_k(2)$. Similarly, denote by $Z_k(i)$ the number of increasing paths of length k whose two endpoints are fixed vertices from the same community G_i and by Z_k^* the number of increasing paths of length k whose endpoints are fixed vertices from different communities. Let $Z_k = Z_k(1) + Z_k(2) + Z_k^*$.

Lemma 6. *Let $k \in \mathbb{N}$. There exists a finite natural number $d \in \mathbb{N}$, not depending on k, such that*

$$\mathbb{E}[X_k] \leq O\left(\frac{nk^d}{k!}(\theta \log n)^k\right), \quad \mathbb{E}[Y_k] = O\left(\frac{k^d}{k!}(\theta \log n)^k\right), \quad \mathbb{E}[Z_k] = O\left(\frac{k^d}{nk!}(\theta \log n)^k\right).$$

Proof of the upper bound in Theorem 1 and Theorem 2. First, we prove (i). Fix two vertices $w, z \in V(G)$. Then, the probability that there exists an incresing path from w to z in G is given by

$$\mathbb{P}\left[\bigcup_{k\in\mathbb{N}} \bigcup_{\gamma\in\Lambda^k(w,z)} A(\gamma)\right] \leq \sum_{k\geq 1} \mathbb{P}[Z_k \geq 1] \leq C(\theta \log n)^d n^{\theta-1} = o(1)$$

since $\theta < 1$. For (ii), we proceed in a similar manner. Notice that, the existence of an increasing path starting at w of length k implies there are increasing paths starting at w of length ℓ for any $1 \leq \ell \leq k$. Thus, Markov's inequality implies

$$\mathbb{P}[|\gamma_{\max}(w)| \geq k] \leq \mathbb{P}[Y_k \geq 1] \leq \mathbb{E}[Y_k] \leq Ck^d \left(\frac{e\theta \log n}{k}\right)^k$$

where the last inequality follows by Lemma 6 and Stirling's approximation. Setting k to be $(1+\epsilon)e\theta \log n$ yields the desired result. For the reachable set of w, first notice that $|B_n(w)| \leq \sum_{k\geq 1} Y_k$. Hence, by Markov's inequality and Lemma 6,

$$\mathbb{P}[|B_n(w)| \geq n^{\alpha(1+\epsilon)}] \leq \frac{\sum_{k\geq 1} \mathbb{E}[Y_k]}{n^{\alpha(1+\epsilon)}} \leq C(\theta \log n)^d n^{-\alpha\epsilon} = o(1)$$

which concludes the proof. □

3 Lower bounds: Embedding a random recursive tree

We will couple $G = (V, E, L)$ with $G' = (V, E', L')$ defined as follows. To each pair of distinct vertices u, v, associate an exponential random variable W_{uv} independent of the others, of parameter p_{uv} where

$$p_{uv} \coloneqq \begin{cases} q & \text{if } u \in G_1, v \in G_2 \text{ or } u \in G_2, v \in G_1, \\ p_i & \text{if } u, v \in G_i \text{ for } i \in \{1, 2\}. \end{cases}$$

An edge $e = uv$ has label $L'(e) = W_{uv}$ and $e \in E'$ if and only if $W_{u,v} \leq 1$. Notice that

$$\mathbb{P}[e \in G'] = \mathbb{P}[W_{uv} \leq 1] = 1 - e^{-p_{uv}} \leq p_{uv} = \mathbb{P}[e \in G]$$

so we can couple G and G' in such a way that $E' \subseteq E$.

Lemma 7. *The total variation distance between L and L' restricted to the edges that appear in G'*

$$d_{TV}(L_{|E'}, L'_{|E'}) \to 0 \qquad \text{as } n \to \infty.$$

Corollary 8. *For n large enough, there is a coupling between G and G' such that for any path γ,*

$$\mathbb{P}_{G'}[A(\gamma)] \leq \mathbb{P}_G[A(\gamma)],$$

i.e. if γ is open and increasing in G', then it is open and increasing in G as well.

Janson proved the following concentration inequality for sums of independent exponential random variables of deterministic parameters (Theorem 5.1, [7]). Here, we extend the result by considering randomized parameters under mild additional assumptions. The last ingredient that we need to prove our key lemma is the following concentration inequality.

Lemma 9. *Let $(\xi_i)_{i\in\mathbb{N}}$ be a sequence of real random variables such that*

- *there exist $c_1, c_2 \in \mathbb{R}_+$ such that $c_1 i \leq \xi_i \leq c_2 i$ deterministically for all $i \in \mathbb{N}$;*
- *ξ_i/i converges almost surely to a constant θ as $i \to \infty$.*

Let $X_i \sim Exp(\xi_i \log n)$ be exponential random variables such that, conditional on $(\xi_i)_{i\in\mathbb{N}}$, $(X_i)_i$ are mutually independent. Then, for all $\epsilon \in (0,1)$, if $r \leq n^{(1-\epsilon)\theta}$, with high probability $\sum_{i=1}^{r} X_i < 1$.

Lemma 10 (Key Lemma)**.** *Fix a vertex $w \in V(G)$. Let $r = \lfloor n^{(1-\epsilon)\theta} \rfloor$ and let T_r be a random recursive tree on r vertices. Then, whp, there is a coupling between G' and the random recursive tree T_r such that there is a tree $\tilde{T}_r$ rooted at vertex w which consists of increasing paths from w and is isomorphic to T_r.*

Proof of the lower bound in Theorem 1 and Theorem 2. Fix $\epsilon > 0$. From Corollary 8 and Lemma 10, it follows that we can embed into G a tree T of size $r = n^{(1-\epsilon)\theta}$ that is isomorphic to a URRT T_r. The lower bound in Theorem 1 follows immediately. It is known that the height of T_r is whp $e \log r$ ([6, 9]). Hence, whp $|\gamma_{\max}| \geq e \log r = (1-\epsilon)e\theta \log n$. □

References

[1] E. Abbe, Community Detection and Stochastic Block Models, *Foundations and Trends® in Communications and Information Theory* **14(1-2)** (2018), 1-162.

[2] O. Angel, A. Ferber, B. Sudakov and V. Tassion, Long monotone trails in random edge-labellings of random graphs, *Combinatorics, Probability and Computing* **29(1)** (2020), 22–30.

[3] R. Becker, A. Casteigts, P. Crescenzi, B. Kodric, M. Renken, M. Raskin and V. Zamaraev, Giant components in random temporal graphs, arXiv preprint, 2022, `arXiv:2205.14888`.

[4] N. Broutin, N. Kamčev and G. Lugosi, Increasing paths in random temporal graphs, arXiv preprint, 2023, `arXiv:2306.11401`.

[5] A. Casteigts, M. Raskin, M. Renken and V. Zamaraev, Sharp thresholds in random simple temporal graphs, *In 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (2022), 319–326.

[6] L. Devroye, Branching processes in the analysis of the heights of trees, *Acta Informatica* **24** (1987), 277–298.

[7] S. Janson, Tail bounds for sums of geometric and exponential variables, *Statist. Probab. Lett.* **135** (2018), 1–6.

[8] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns, *Stochastic Processes and their Applications* **110(2)** (2004), 177–245.

[9] B. Pittel, Note on the height of random recursive trees and m-ary search trees, *Random Structures and Algorithms* **5** (1994), 337–347.

Ranges of polynomials control degree ranks of Green and Tao over finite prime fields*

Thomas Karam†[1]

[1]Mathematical Institute, University of Oxford, OX26GG United Kingdom

Abstract

Let p be a prime, let $1 \leq t < d < p$ be integers, and let S be a non-empty subset of $\mathbb{F}_p$. We establish that if a polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$ with degree d is such that the image $P(S^n)$ does not contain the full image $A(\mathbb{F}_p)$ of any non-constant polynomial $A : \mathbb{F}_p \to \mathbb{F}_p$ with degree at most t, then P coincides on S^n with a polynomial that in particular has bounded degree-$\lfloor d/(t+1) \rfloor$-rank in the sense of Green and Tao. Similarly, we prove that if the assumption holds even for $t = d$, then P coincides on S^n with a polynomial determined by a bounded number of coordinates.

Throughout this paper, the letter n will always denote a positive integer, and all our statements will be uniform in n. A full version of the present paper may be found at [7].

1 Degree ranks and ranges of polynomials

A landmark result of Green and Tao proved in 2007 [3] states that over a finite prime field $\mathbb{F}_p$ for some prime p, a multivariate polynomial with degree $1 \leq d < p$ that is not approximately equidistributed can be expressed as a function of a bounded number of polynomials each with degree at most $d-1$. More formally, we have the following statement.

Theorem 1 ([3], Theorem 1.7)**.** *Let p be a prime, and let $1 \leq d < p$ be an integer. Then there exists a function $K_{p,d} : (0,1] \to \mathbb{N}$ such that for every $\epsilon > 0$, if $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is a polynomial with degree d satisfying $|\mathbb{E}_{x \in \mathbb{F}_p^n} \omega_p^{sP(x)}| \geq \epsilon$ for some $s \in \mathbb{F}_p^*$, then there exist $k \leq K_{p,d}(\epsilon)$, polynomials $P_1, \ldots, P_k : \mathbb{F}_p^n \to \mathbb{F}_p$ with degree at most $d-1$ and a function $F : \mathbb{F}_p^k \to \mathbb{F}_p$ satisfying*

$$P = F(P_1, \ldots, P_k).$$

It has been known since at least the works of Janzer [5] and Milićević [9] that the conclusion can be made qualitatively more precise. Before stating this strengthening, let us define a notion of degree-d rank for polynomials.

Definition 2. *Let $\mathbb{F}$ be a field, and let $P : \mathbb{F}^n \to \mathbb{F}$ be a polynomial. Let $d \geq 1$ be an integer.*

We say that a polynomial P has degree-d rank at most 1 *if we can write P as a product of polynomials each with degree at most d.*

The degree-d rank *of P is defined to be the smallest nonnegative integer k such that there exist polynomials $P_1, \ldots, P_k$ each with degree-d rank at most 1, with degree at most the degree of P, and satisfying*

$$P = P_1 + \cdots + P_k.$$

We denote this quantity by $\mathrm{rk}_d P$.

*The full version of this work can be found in [7].

†Email: thomas.karam@maths.ox.ac.uk. Supported by ERC Advanced Grant 883810.

The zero polynomial in particular has degree-d rank equal to 0 for all d, and constant polynomials have degree-d rank at most 1 for all d.

We define this notion of degree-d rank in this way as doing so will be convenient for us, but it is worth pointing out that in the original paper [3] of Green and Tao, the notion referred to as the degree-d rank was slightly different: for instance the degree-$(d-1)$ rank was the largest possible k in Theorem 1. Nonetheless, it follows immediately from the definitions that for every polynomial $P : \mathbb{F}_p^n \to \mathbb{F}_p$ and every positive integer d, the degree-d rank of P in the sense of Green and Tao is at most d times the degree-d rank of P in our sense. Therefore, proving that a polynomial has bounded degree-d rank in our sense implies showing that it has bounded degree-d rank in the sense of Green and Tao.

The main qualitative refinement shown in the papers of Janzer [5] and Milićević [9] is that there exists some function $H_{p,d} : (0,1] \to \mathbb{N}$ such that under the assumptions of Theorem 1, we can find $k \leq H_{p,d}(\epsilon)$ and polynomials $Q_1, R_1, \dots, Q_k, R_k$ satisfying

$$\deg Q_i, \deg R_i \leq d-1 \text{ and } \deg Q_i + \deg R_i \leq d$$

for each $i \in [k]$ and such that

$$P = Q_1 R_1 + \cdots + Q_k R_k.$$

In other words, it was shown that

$$\mathrm{rk}_{d-1} P \leq H_{p,d}(\epsilon).$$

This is a bound on the degree-$(d-1)$ rank of P, and the numerous developments which arose out of Theorem 1 have to our knowledge entirely or almost entirely focused on the degree-$(d-1)$ rank of P: some extended the range of validity of the results (Kaufman and Lovett [8], Bhowmick and Lovett [2]), and others improved the quantitative bounds on the degree-$(d-1)$ rank, through the closely related question of comparing the partition rank to the analytic rank of tensors (Janzer [5], Milićević [9], Adiprasito, Kazhdan and Ziegler [1], Moshkovitz and Cohen [10], [11], Moshkovitz and Zhu [12]).

For the purposes of studying approximate equidistribution of polynomials this is unsurprising, since the notion of degree-$(d-1)$ rank is indeed by far the most relevant: for instance a random polynomial of the type

$$x_1 Q(x_2, \dots, x_n)$$

with $\deg Q = d-1$ has high degree-$(d-2)$ rank but is nonetheless not approximately equidistributed, since the probability that it takes the value 0 is approximately $2/p - 1/p^2 > 1/p$.

Rather than focus on the fact that for a degree-d polynomial, lack of equidistribution implies bounded degree-$(d-1)$ rank, we may ask for analogues of this statement involving much stronger properties in the assumption and in the conclusion. Correspondingly, the main motivations of this paper are twofold. In one direction, we ask what can be deduced about polynomials for which we know much more than lack of equidistribution. What can we say if we know that a polynomial does not take every value of $\mathbb{F}_p$, or has a smaller range still, in a sense to be made precise ? In the other direction, we can ask, for a fixed integer $1 \leq e \leq d-1$, whether there are any properties of the distribution of the values of a polynomial which would guarantee that its degree-e rank is bounded above. We will contribute to both directions simultaneously, by showing that if a polynomial P does not have full range, then it must have bounded degree-e rank, for some integer e that is determined by the degree of P and by the smallest degree of a non-constant *one-variable* polynomial that has a range contained in the range of P.

Theorem 3. *Let p be a prime, and let $1 \leq t \leq d < p$ be integers. There exists a positive integer $\gamma(p,d,t)$ such that the following holds. Let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial with degree at most d. Assume that the image $P(\mathbb{F}_p^n)$ does not contain the image of $\mathbb{F}_p$ by any non-constant polynomial $\mathbb{F}_p \to \mathbb{F}_p$ with degree at most t.*

1. *If $t \leq d-1$, then P has degree-$\lfloor d/(t+1) \rfloor$-rank at most $\gamma(p,d,t)$.*

2. If $t = d$ then P is a constant polynomial.

The value $\lfloor d/(t+1) \rfloor$ in the degree of the rank in Theorem 3 is optimal in general, as the following example shows.

Example 4. *Let p be a prime, let $1 \leq d < p$ be an integer, let $t, u \geq 1$ be integers such that $tu \leq d$. If Q is a random polynomial $Q : \mathbb{F}_p^n \to \mathbb{F}_p$ with degree u, then Q^t has degree at most d, the image $Q(\mathbb{F}_p^n)$ is contained in the set $\{y^t : y \in \mathbb{F}_p\}$ of t-th power residues mod-p, but the degree-$(u-1)$-rank of P is usually arbitrarily large as n tends to infinity, even if it is taken in the sense of Green and Tao.*

The last part follows from a counting argument: as n tends to infinity there are $p^{O(n^{u-1})}$ polynomials $\mathbb{F}_p^n \to \mathbb{F}_p$ with degree at most $u-1$, so for every $k \geq 1$, the number of polynomials of the type $F(P_1, \dots, P_k)$ with $P_1, \dots, P_k$ with degree at most $u-1$ and $F : \mathbb{F}_p^k \to \mathbb{F}_p$ a function is at most $p^{p^k} p^{O(kn^{u-1})} = p^{O(kn^{u-1})}$, whereas there are $p^{\Omega(kn^u)}$ polynomials $\mathbb{F}_p^n \to \mathbb{F}_p$ with degree u and hence at least $1/t$ times as many polynomials of the type Q^t above.

Powers of polynomials are not the only simple examples of polynomials that do not have full range in general. They can instead be viewed as a special case of a broader class of examples that arises from composition with a one-variable polynomial.

Example 5. *Let p be a prime, let $1 \leq d < p$ be an integer, let $t, u \geq 1$ be integers such that $tu \leq d$. If $Q : \mathbb{F}_p^n \to \mathbb{F}_p$ is a polynomial with degree u, and $A : \mathbb{F}_p \to \mathbb{F}_p$ is a polynomial with degree t, then the polynomial $A \circ Q$ has degree at most d, and the image $A \circ Q(\mathbb{F}_p^n)$ is contained in the image $A(\mathbb{F}_p)$.*

We stress that the main result from the approximate equidistribution regime will itself be an important black box that we will use in our proof of Theorem 3.

2 Variables with restricted range

We shall in fact prove results in a more general setting than that of Theorem 3, where we allow the assumption to be on the image $P(S^n)$ for some non-empty subset S of $\mathbb{F}_p$ rather than on the whole image $P(\mathbb{F}_p^n)$. On a first reading the set S may be taken to be $\{0, 1\}$. In the setting of restrictions to S^n, the approximate equidistribution statement was proved by Gowers and the author in [4]. Before stating it, let us recall from that paper two points to be aware of regarding restrictions of polynomials to S^n.

The first is that whereas an affine polynomial is either constant or perfectly equidistributed on $\mathbb{F}_p^n$, there is already something to say about the distribution of an affine polynomial P on S^n for general non-empty S: if $S \neq \mathbb{F}_p$ and P depends only on one coordinate, then $P(S^n)$ is not even the whole of $\mathbb{F}_p$. As a simple Fourier argument however shows ([4], Proposition 2.2), an affine polynomial depending on many coordinates is approximately equidistributed on S^n, provided that S contains at least two elements. The second is that we may no longer hope to conclude in general that a polynomial with degree d which is not approximately equidistributed on S^n must itself have bounded degree-$(d-1)$ rank: for instance, the polynomial

$$\sum_{i=1}^{n} x_i^2 - x_i$$

has degree-1 rank equal to n, but only takes the value 0 on $\{0, 1\}^n$ and is in particular not approximately equidistributed on $\{0, 1\}^n$. Nonetheless, the zero polynomial, with which this polynomial coincides on $\{0, 1\}^n$, itself has degree-1 rank equal to 0.

These two remarks motivate an extension of Definition 2.

Definition 6. *Let $\mathbb{F}$ be a field, and let $P : \mathbb{F}^n \to \mathbb{F}$ be a polynomial.*

The degree-0 rank *of P is defined to be the smallest nonnegative integer k such that we can write P as a linear combination of at most k monomials. We denote this quantity by* $\mathrm{rk}_0 P$.

If d is a nonnegative integer and S is a non-empty subset of $\mathbb{F}$ then we define the degree-d rank of P with respect to S *as the smallest value of* $\mathrm{rk}_d(P-P_0)$*, where the minimum is taken over all polynomials P_0 with degree at most the degree of P and satisfying $P_0(S^n)=\{0\}$. We denote this quantity by* $\mathrm{rk}_{d,S}\,P$.

We now recall a slight weakening of the main result of [4], Theorem 1.4 from that paper. (Although the full statement of that theorem is slightly more precise, the formulation below is slightly simpler to use and suffices for the purposes of the present paper.)

Theorem 7. *Let p be a prime, let $1 \le d < p$ be an integer, and let S be a non-empty subset of $\mathbb{F}_p$. There exists a function $H_{p,d,S} : (0,1] \to \mathbb{N}$ such that for every $\epsilon > 0$, if $P : \mathbb{F}_p^n \to \mathbb{F}_p$ is a polynomial with degree d satisfying $|\mathbb{E}_{x\in S^n}\omega_p^{sP(x)}| \ge \epsilon$ for some $s \in \mathbb{F}_p^*$, then $\mathrm{rk}_{d-1,S}\,P \le H_{p,d,S}(\epsilon)$.*

We note that if S has size 1, then Theorem 7 as well as many of the new results of the present paper hold for immediate reasons: the set S^n then also has size 1, so every polynomial coincides on S^n with a constant polynomial, so has degree-d rank at most 1 for every d.

When S is not the whole of $\mathbb{F}_p$, one important difference between the sets $\mathbb{F}_p^n$ and S^n is that the former is invariant under linear transformations, whereas the latter is not. We have already discussed one effect on this: the fact that x_1 does not take every value of $\mathbb{F}_p$ whereas $x_1+\cdots+x_n$ is approximately equidistributed for n large. The role of coordinates as opposed to general degree-1 polynomials will manifest itself further in the proofs and in the main results of this paper. For this purpose let us make one last definition.

Definition 8. *Let p be a prime, and let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial.*

For each $i \in [n]$, we say that P depends on x_i if the coordinate x_i arises in some monomial of P.

For k nonnegative integer, we will say that P is k-determined *if it depends on at most k coordinates.*

3 Statements of main results

Using Theorem 7 as a black box we will prove the following analogue of Theorem 3, where the assumption on the image is now on $P(S^n)$ rather than on $P(\mathbb{F}_p^n)$. The following theorem is the main result that we shall prove in the present paper.

Theorem 9. *Let p be a prime, let $1 \le t \le d < p$ be integers and let S be a non-empty subset of $\mathbb{F}_p$. Then there exists a positive integer $C(p,d,t)$ such that the following holds. Let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial with degree at most d. Assume that $P(S^n)$ does not contain the image of $\mathbb{F}_p$ by any non-constant polynomial $\mathbb{F}_p \to \mathbb{F}_p$ with degree at most t.*

1. *If $t \le d-1$, then P coincides on S^n with a polynomial that has degree-$\lfloor d/(t+1)\rfloor$-rank at most $C(p,d,t)$ and has degree at most d.*
2. *If $t = d$ then P coincides on S^n with a linear combination of at most $C(p,d,t)$ monomials with degrees at most d.*

Equivalently, in both cases we have

$$\mathrm{rk}_{\lfloor \frac{d}{t+1}\rfloor,S}\,P \le C(p,d,t).$$

The optimal bounds in Theorem 9 and in several of our other statements involving the set S may depend on the choice of S. However, to avoid heavy notation we will at many places avoid making this dependence explicit. (We may safely do so, since for each prime p there are only finitely many subsets of $\mathbb{F}_p$).

Let us look at the extreme cases of item 1 from Theorem 9, and at a situation where they are both simultaneously realised.

Corollary 10. *Let p be a prime, let $1 \leq t \leq d < p$ be integers and let S be a non-empty subset of $\mathbb{F}_p$. Let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial with degree at most d. Let $C_{(i)}(p,d) = C(p,d,1)$ and let $C_{(ii)}(p,d) = C(p,d,\lfloor d/2 \rfloor)$.*

(i) If $P(S^n) \neq \mathbb{F}_p$ then $\mathrm{rk}_{\lfloor d/2 \rfloor, S} P \leq C_{(i)}(p,d)$.

(ii) If $P(S^n)$ does not contain the image of any non-constant polynomial $\mathbb{F}_p \to \mathbb{F}_p$ with degree at most $\lfloor d/2 \rfloor$ then $\mathrm{rk}_{1,S} P \leq C_{(ii)}(p,d)$.

(iii) If $d = 3$ and $P(S^n) \neq \mathbb{F}_p$, then $\mathrm{rk}_{1,S} P \leq C_{(i)}(p,3) = C_{(ii)}(p,3)$.

Proof. Items (i) and (ii) follow from taking $t = 1$ and $t = \lfloor d/2 \rfloor$ in Theorem 9 respectively. Item (iii) follows from either of the items (i) and (ii). □

We now turn our attention to the case of degree-2 polynomials. Throughout the paper, we will write Q_p for the set $\{y^2 : y \in \mathbb{F}_p\}$ of mod-p quadratic residues. Provided that $p \geq 3$, this set has size $\frac{p+1}{2}$ and is in particular not the whole of $\mathbb{F}_p$. We say that a subset of $\mathbb{F}_p$ is an *affine translate* of Q_p if it can be written as $aQ_p + b$ for some $a \in \mathbb{F}_p^*$ and some $b \in \mathbb{F}_p$. In light of the preceding discussion we can formulate three basic constructions of a degree-2 polynomial P such that $P(S^n) \neq \mathbb{F}_p$.

(i) A polynomial of the type $A \circ L$ for some affine polynomial $L : \mathbb{F}_p^n \to \mathbb{F}_p$ and some degree-2 polynomial $A : \mathbb{F}_p \to \mathbb{F}_p$. (Equivalently, the sum of a multiple of L^2 and of a constant.)

(ii) A polynomial that depends only on a small number $r < \log p / \log |S|$ of coordinates, since $P(S^n)$ then necessarily has size at most $|S|^r$.

(iii) A polynomial that vanishes on S^n and has degree at most 2.

The first item of the following result can be interpreted as a converse which says that every example arises as a sum of these three examples, letting aside the value of the bound on the number of coordinates in the second example.

Proposition 11. *There exists an absolute constant $\kappa > 0$ such that the following holds. Let p be a prime, and let S be a non-empty subset of $\mathbb{F}_p$. Let $P : \mathbb{F}_p^n \to \mathbb{F}_p$ be a polynomial with degree 2.*

1. *If $P(S^n) \neq \mathbb{F}_p$, then there exists an affine polynomial $L : \mathbb{F}_p^n \to \mathbb{F}_p$, a degree-2 polynomial $A : \mathbb{F}_p \to \mathbb{F}_p$, and a κp^{15}-determined polynomial J with degree at most 2 such that P coincides on S^n with $A \circ L + J$. (Equivalently, with $AL^2 + J$ for some $A \in \mathbb{F}_p$, with J changed by a constant.)*

2. *If furthermore $P(S^n)$ does not contain any affine translate of Q_p, then P coincides on S^n with a κp^{15}-determined polynomial that has degree at most 2.*

Item 1 from Proposition 11 is significantly stronger than the conclusion that item 1 from Theorem 9 gives in the corresponding case $d = 2$ and $t = 1$: the latter is merely that P has bounded degree-1 rank with respect to S, which we already know by Theorem 7. The proof of Proposition 11 will instead use different techniques which do not appear to generalise well to higher-degree polynomials.

In the more general case where P has general degree $2 \leq d \leq p-1$, one may ask whether just as with item 1 from Proposition 11, it is the case that provided that $P(S^n) \neq \mathbb{F}_p$ we can always obtain a decomposition $P = A \circ Q + J$ with $Q : \mathbb{F}_p^n \to \mathbb{F}_p$, $A : \mathbb{F}_p \to \mathbb{F}_p$ polynomials satisfying $\deg Q \deg A \leq d$, $\deg A \geq 2$ and with J a polynomial determined by a bounded number of coordinates and with degree at most d. This is however not the case in general, as a wider class of examples comes in: for instance, if $d = p-1$, then the polynomial $A : x \to x^{p-1}$ satisfies $A(\mathbb{F}_p) = \{0,1\}$, so if $L_1, \ldots, L_{p-2}$ are arbitrary affine polynomials then the image of $\mathbb{F}_p^n$ by the polynomial

$$P = A \circ L_1 + \cdots + A \circ L_{p-2}$$

does not contain $p-1$. For $d=2$, such a situation cannot occur, because the Cauchy-Davenport theorem, which will play some role in the proof of Proposition 11, shows that the sumset of any two affine translates of Q_p is the whole of $\mathbb{F}_p$. (However, this is by no means the only or even the main specificity of the case $d=2$ that allows us to say more there than for general $d<p$.)

4 Techniques for the proof of Theorem 9

The basic strategy which we will use to prove Theorem 9 will be essentially as follows: because P has degree d and is not approximately equidistributed, Theorem 7 shows that P coincides on S^n with some polynomial with degree at most d and of the type

$$M \circ (P_1, \dots, P_k)$$

where k is bounded and M is some polynomial. One of the following is always true: either the polynomials $P_1, \dots, P_k$ are approximately jointly equidistributed, in which case the image $(P_1, \dots, P_k)(S^n)$ is the same as if the polynomials $P_1, \dots, P_k$ were jointly equidistributed, or they are not, in which case at least one non-trivial linear combination of the polynomials $P_1, \dots, P_k$ has bounded degree-$(d'-1)$ rank with respect to S, where

$$d' = \max(\deg P_1, \dots, \deg P_k),$$

and we may hence without loss of generality assume that P coincides on S^n with some polynomial with degree at most d and of the type

$$M' \circ (P_1, \dots, P_{k-1}, Q_1, \dots, Q_{k'})$$

where k' is bounded, $Q_1, \dots, Q_{k'}$ are polynomials with degree strictly smaller than the degree of P_k, and M' is some polynomial. This second step, in turn, can only be performed a bounded number of times, which will conclude the argument.

References

[1] K. Adiprasito, D. Kazhdan, and T. Ziegler, *On the Schmidt and analytic ranks for trilinear forms*, arXiv:2102.03659 (2021).

[2] A. Bhowmick and S. Lovett, *Bias vs structure of polynomials in large fields, and applications in effective algebraic geometry and coding theory*, IEEE Trans. Inf. Theory, arXiv:1506.02047 (2015).

[3] B. Green and T. Tao, *The distribution of polynomials over finite fields, with applications to the Gowers norms.* Contr. Discr. Math., **4** (2009), no. 2, 1-36.

[4] W. T. Gowers and T. Karam, *Equidistribution of high-rank polynomials with variables restricted to subsets of* $\mathbb{F}_p$, arXiv:2209.04932 (2022).

[5] O. Janzer, *Polynomial bound for the partition rank vs the analytic rank of tensors*, Discrete Anal. **7** (2020), 1-18.

[6] T. Karam, *High-rank subtensors of high-rank tensors*, arXiv:2207.08030 (2022).

[7] T. Karam, *Ranges of polynomials control degree ranks of Green and Tao over finite prime fields*, arXiv:2305.11088 (2023).

[8] T. Kaufman and S. Lovett, *Worst case to average case reductions for polynomials*, 49th Annual IEEE Symposium on Foundations of Computer Science (2008), 166-175.

[9] L. Milićević, *Polynomial bound for partition rank in terms of analytic rank*, Geom. Funct. Anal. **29** (2019), 1503-1530.

[10] A. Cohen and G. Moshkovitz, Structure vs. randomness for bilinear maps, Discrete Anal. **12** (2022).

[11] A. Cohen and G. Moshkovitz, Partition and analytic rank are equivalent over large fields, (2022).

[12] G. Moshkovitz and D. G. Zhu, *Quasi-linear relation between partition and analytic rank*, arXiv:2211.05780 (2022).

Product representation of perfect cubes*

Zsigmond György Fleiner[†1], Márk Hunor Juhász[‡1], Blanka Kövér[§1], Péter Pál Pach[¶2,3], and Csaba Sándor[‖2,3,4]

[1]ELTE Eötvös Loránd University Faculty of Science, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary
[2]Department of Computer Science and Information Theory, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[3]MTA-BME Lendület Arithmetic Combinatorics Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary
[4]Department of Stochastics, Institute of Mathemetics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary

Abstract

Let $F_{k,d}(n)$ be the maximal size of a set $A \subseteq \{1, 2, \ldots, n\}$ such that the equation

$$a_1 a_2 \ldots a_k = x^d,\ a_1 < a_2 < \ldots < a_k$$

has no solution with $a_1, a_2, \ldots, a_k \in A$ and integer x. Erdős, Sárközy and T. Sós studied $F_{k,2}$, and gave bounds when $k = 2, 3, 4, 6$ and also in the general case. We study the problem for $d = 3$, and provide bounds for $k = 2, 3, 4$ and 6, furthermore, in the general case as well. In particular, we refute an 18-year-old conjecture of Verstraëte.

We also introduce another function $f_{k,d}$ closely related to $F_{k,d}$: While the original problem requires $a_1, \ldots, a_k$ to all be distinct, we can relax this and only require that the multiset of the a_i's cannot be partitioned into d-tuples where each d-tuple consists of d copies of the same number.

1 Introduction

The problem of the solvability of equations of the form

$$a_1 a_2 \ldots a_k = x^2,\ a_1 < a_2 < \ldots < a_k$$

in a set $A \subseteq [n] = \{1, 2, \ldots, n\}$ first appeared in a 1995 paper of Erdős, Sárközy and T. Sós [3]. They investigated the maximal size of a set A such that the equation cannot be solved in A, that is, there are no distinct $a_1, \ldots, a_k \in A$ whose product is a perfect square. This motivates the following definitions:

Let $F_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that

$$a_1 a_2 \ldots a_k = x^d,\ a_1 < a_2 < \ldots < a_k \tag{1}$$

*The full version of this work can be found in [4] and will be published elsewhere.

[†]Email: zsgyfleiner@gmail.com
[‡]Email: markh.shepherd@gmail.com
[§]Email: koverblanka@gmail.com
[¶]Email: pach.peter@vik.bme.hu. Research of P. P. P. supported by the Lendület program of the Hungarian Academy of Sciences (MTA) and by the National Research, Development and Innovation Office NKFIH (Grant Nr. K146387)

[‖]Email: sandor.csaba@ttk.bme.hu. Research of C. S. supported by the Lendület program of the Hungarian Academy of Sciences (MTA) and by the National Research, Development and Innovation Office NKFIH (Grant Nr. K146387)

has no solution with $a_1, a_2, \ldots, a_k \in A$ and integer x. Similarly, let $f_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that

$$a_1 a_2 \ldots a_k = x^d \tag{2}$$

has no solution with $a_1, a_2, \ldots, a_k \in A$ and integer x, except *trivial* solutions that we specify below. If we allow some of the a_i's in equation (2) to coincide, some trivial solutions do arise: It is clear, for instance, that $a_1 = \ldots = a_d$ will yield a solution to the equation $a_1 \ldots a_d = x^d$. Let us call a solution trivial if the multiset of the a_i's can be partitioned into d-tuples where each d-tuple consists of d copies of the same number: see for example $(a_1a_1a_1)(a_2a_2a_2)(a_3a_3a_3) = x^3$ for $k = 9$, $d = 3$. Note that trivial solutions arise only if $d \mid k$. Let $f_{k,d}(n)$ be the maximal size of a set $A \subseteq [n]$ such that the equation $a_1 a_2 \ldots a_k = x^d$ does not have any nontrivial solution with $a_1, a_2, \ldots, a_k \in A$. Note that $f_{k,d} \leq F_{k,d}$.

With our notation, Erdős, Sárközy and T. Sós [3] proved the following results (and also gave bounds for $F_{k,2}$ for every k):

Theorem 1 (Erdős, Sárközy, T. Sós)**.** *For every $\ell \in \mathbb{Z}^+$, we have*

(i) $F_{2,2}(n) = \left(\frac{6}{\pi^2} + o(1)\right) n$;

(ii) $\frac{n^{3/4}}{(\log n)^{3/2}} \ll F_{4,2}(n) - \pi(n) \ll \frac{n^{3/4}}{(\log n)^{3/2}}$;

(iii) $\frac{n^{2/3}}{(\log n)^{4/3}} \ll F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right) \ll n^{7/9} \log n$.

Later Győri [5] and the fourth named author [7] improved the upper bound for $F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right)$. The current best upper bound is

$$F_{6,2}(n) - \left(\pi(n) + \pi\left(\frac{n}{2}\right)\right) \ll n^{2/3} (\log n)^{2^{1/3} - 1/3 + o(1)}.$$

For general cases, the current best lower bound estimates have been proved recently by the fourth named author and Vizer [8].

Note that the case $2 \mid k$ is closely related to (generalized) multiplicative Sidon sets, as a solution to the (multiplicative) Sidon equation $a_1 \ldots a_k = b_1 \ldots b_k$ provides a solution $a_1 \ldots a_k b_1 \ldots b_k = x^2$. However, the case $2 \nmid k$ seems to be much more difficult. Erdős, Sárközy and T. Sós proved the following results:

Theorem 2 (Erdős, Sárközy, T. Sós)**.** *For every $\ell \in \mathbb{Z}^+$ and $\varepsilon > 0$, we have*

(i) $\frac{n}{(\log n)^{1+\varepsilon}} \ll n - F_{3,2}(n) \leq n - f_{3,2}(n) \ll n (\log n)^{\frac{e \log 2}{2} - 1 + \varepsilon}$;

(ii) $\liminf_{n \to \infty} \dfrac{F_{2\ell+1,2}(n)}{n} \geq \log 2 = 0.69 \ldots$;

(iii) $\frac{n}{(\log n)^2} \ll n - F_{2\ell+1,2}(n)$.

Note that similar bounds can be proved for the functions $f_{k,2}(n)$.

It remained an interesting problem to find the right shape of the function $F_{2\ell+1,2}$ for $\ell \geq 2$. Very recently, Tao [10] proved that for every $k \geq 4$ there exists some constant $c_k > 0$ such that $F_{k,2}(n) \leq (1 - c_k + o(1))n$ as $n \to \infty$.

Based on the work of Erdős, Sárközy, and T. Sós, Verstraëte [11] studied a similar problem: He aimed to find the maximal size of a set $A \subseteq [n]$ such that no product of k distinct elements of A is in the value set of a given polynomial $f \in \mathbb{Z}[x]$. He showed that for a certain class of polynomials the answer is $\Theta(n)$, for another class it is $\Theta(\pi(n))$, and conjectured that these are the only two possibilities:

Conjecture 3. *Let $f \in \mathbb{Z}[x]$ and let k be a positive integer. Then, for some constant $\rho = \rho(k, f)$ depending only on k and f, the maximal size of a set $A \subseteq [n]$ such that no product of k distinct elements of A is in the value set of f is either $(\rho + o(1))n$ or $(\rho + o(1))\pi(n)$ as $n \to \infty$.*

For further related results, see [6, 9].

2 Our results

We investigated the original problem in the case $d = 3$, and provided bounds for both $F_{k,3}$ and $f_{k,3}$. As expected, several additional difficulties arise compared to the case $d = 2$. To overcome these, various new ideas are needed of combinatorial and number theoretic nature. We summarize our results below.

For $k = 2$, the following bounds hold:

Theorem 4. *There exist positive constants c_1 and c_2 such that*

$$c_1 n^{2/3} < n - F_{2,3}(n) \leq n - f_{2,3}(n) < c_2 n^{2/3}.$$

For the case $k = 3$ we prove that $f_{3,3}(n)/n$ converges to a constant $c_{3,3} \in (0, 1)$, which we can approximate (theoretically to arbitrary precision):

Theorem 5. *There exists a constant $0.6224 \leq c_{3,3} \leq 0.6420$ such that*

$$f_{3,3}(n) = (c_{3,3} + o(1))n.$$

(An analogous result holds for $F_{3,3}(n)$, as well.)

In the case $k = 4$ we show that for large n, the following bounds hold. Our proofs generalize and extend ideas from [3] used for the estimation of $F_{3,2}(n)$.

Theorem 6. *Let $\varepsilon > 0$. There exists some $n_0(\varepsilon)$ such that for every $n \geq n_0(\varepsilon)$ we have*

$$\frac{n}{(\log n)^{2+\varepsilon}} < n - F_{4,3}(n) \leq n - f_{4,3}(n) < \frac{n}{(\log n)^{1-\frac{e \log 3}{2\sqrt{3}}-\varepsilon}}.$$

For $k = 6$ we obtained the following results:

Theorem 7. *There exist positive constants c_1 and c_2 such that*

$$c_1 \frac{n^{3/4}}{(\log n)^{3/2}} < f_{6,3}(n) - \pi(n) < c_2 \frac{n^{3/4}}{(\log n)^{3/2}}.$$

Theorem 8. *For $F_{6,3}(n)$ the following holds:*

$$F_{6,3}(n) = (1 + o(1)) \frac{n \log\log n}{\log n}.$$

Note that Theorem 8 refutes Conjecture 3 of Verstraëte [11].

We also give bounds for larger values of k, all the results and proofs are contained in the preprint [4].

3 Proof ideas

The different behaviours of the function $f_{k,3}$ (and $F_{k,3}$) can be illustrated by the cases $k = 2, 3, 4, 6$. Here we give a brief outline of the proof ideas in these cases.

3.1

For proving Theorem 4 we shall notice that $a_1 a_2 = x^3$ holds if and only if the product of the cubefree parts of a_1 and a_2 is a perfect cube. That is, if the cubefree part of a_1 is uv^2 (where uv is squarefree), then in a solution the cubefree part of a_2 has to be $u^2 v$. With the help of this observation one can show the exact result

$$f_{2,3}(n) = n - \sum_{\substack{1 \leq u < v \\ \gcd(u,v)=1 \\ uv^2 \leq n \\ u,v \text{ squarefree}}} \left\lfloor \sqrt[3]{\frac{n}{uv^2}} \right\rfloor,$$

for getting the claimed bound we have to estimate this sum. (Also, note that $F_{2,3}(n) = f_{2,3}(n) + 1$.)

3.2

For getting the bound in Theorem 5 let r be a fixed positive integer and let p_i denote the ith prime. Each cubefree positive integer a can be written as

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} a',$$

where $\alpha_1, \dots, \alpha_r \in \{0,1,2\}$ and a' is cubefree satisfying $\gcd(a', p_1 p_2 \dots p_r) = 1$. Here $p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r}$ is the p_r-*smooth* and a' is the p_{r+1}-*rough* part of the number a. Observe that the product of three integers is a perfect cube if and only if so are the product of their p_r-smooth parts and the product of their p_{r+1}-rough parts. In particular, for a fixed a' there cannot be three elements in A with p_{r+1}-rough part a' such that the product of their p_r-smooth parts is a perfect cube. Note that the product of three p_r-smooth numbers is a cube if and only if the sum of their exponent vectors $(\alpha_1, \alpha_2, \dots, \alpha_r)$ add up to $(0,0,\dots,0)$ calculating coordinate-wise modulo 3. Alternatively, if we consider the exponent vectors as elements of $\mathbb{F}_3^r$, they form a nontrivial 3-term arithmetic progression (3AP). Let $L_r(i)$ be the set of p_r-smooth cubefree integers up to i:

$$L_r(i) := \{p_1^{\alpha_1} p_2^{\alpha_2} \dots p_r^{\alpha_r} : \ \alpha_1, \dots, \alpha_r \in \{0,1,2\}\} \cap [i],$$

and let $s_r(i)$ denote the largest possible size of a subset of $L_r(i)$ avoiding nontrivial solutions to $a_1 a_2 a_3 = x^3$. Note that $s_r(i)$ is the size of the largest 3AP-free subset of

$$\{(\alpha_1, \dots, \alpha_r) \in \{0,1,2\}^r : \ \alpha_1 \log p_1 + \dots + \alpha_r \log p_r \le \log i\},$$

if we consider this set as a subset of $\mathbb{F}_3^r$. Clearly, for every $i \ge p_1^2 \dots p_r^2$, we have $s_r(i) = s_r(p_1^2 \dots p_r^2)$ (whose common value is $r_3(\mathbb{F}_3^r)$, the largest possible size of a 3AP-free subset of $\mathbb{F}_3^r$). For getting good numerical bounds we shall calculate these $s_r(i)$ values, for which we used IP solvers. Note that the exact value of $r_3(\mathbb{F}_3^r)$ is known only for $r \le 6$, thus significantly improving our numerical bounds is a very difficult task.

3.3

First we sketch the proof of the lower bound in Theorem 6 (which provides upper bounds for $f_{4,3}$ and $F_{4,3}$).

Let $A \subseteq \{1,2,\dots,n\}$ be a subset such that $a_1 a_2 a_3 a_4 \ne x^3$ if $a_i \in A$, $a_1 < a_2 < a_3 < a_4$ and let $D = \{d_1, \dots, d_t\}$ be the set of all positive integers d such that $d \le n^{1/3}$ and $\Omega(d) \le \frac{1}{3} \log\log n$, where $\Omega(d)$ denotes the number of prime factors of d (counted by multiplicity). A calculation yields that

$$t = |D| > \frac{n^{1/3}}{(\log n)^{1+\frac{1}{3}\log\frac{1}{3}-\frac{1}{3}+\frac{\varepsilon}{3}}}.$$

Let H be the 3-uniform hypergraph on the vertex set $\{P_1, \dots, P_t\}$ such that $\{P_i, P_j, P_k\}$ is an edge in H if and only if $d_i d_j d_k \in A$. Let M be the set of those $m \in [n]$ such that $m \notin A$ and $m = d_i d_j d_k$ for some $1 \le i < j < k \le t$, then $|A| \le n - |M|$.

For a fixed $m \in M$ let $h(m)$ denote the number of triples (d_i, d_j, d_k) such that $m = d_i d_j d_k$, $1 \le i < j < k \le t$. If $m = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r} \in M$, then

$$\Omega(m) = \Omega(d_i) + \Omega(d_j) + \Omega(d_k) \le \log\log n,$$

hence

$$h(m) \le \tau_3(m) = \prod_{i=1}^{r} \binom{k_i+2}{2} \le \prod_{i=1}^{r} 3^{k_i} = 3^{\Omega(m)} \le 3^{\log\log n} = (\log n)^{\log 3},$$

where $\tau_3(m)$ denotes the number of triples (a,b,c) with $a,b,c \in \mathbb{Z}^+$ such that $m = abc$.

If H contains a K_4^3 (a subhypergraph G with vertex set $V = \{P_{i_1}, P_{i_2}, P_{i_3}, P_{i_4}\}$ such that $V \setminus \{P_{i_j}\}$ is an edge in G for every $j \in \{1,2,3,4\}$), then for some $d_{i_1} < d_{i_2} < d_{i_3} < d_{i_4}$ and

$$a_1 = d_{i_1}d_{i_2}d_{i_3}, \quad a_2 = d_{i_1}d_{i_2}d_{i_4}, \quad a_3 = d_{i_1}d_{i_3}d_{i_4}, \quad a_4 = d_{i_2}d_{i_3}d_{i_4}$$

we have $a_1 < a_2 < a_3 < a_4$, $a_1, a_2, a_3, a_4 \in A$ and $a_1a_2a_3a_4 = (d_{i_1}d_{i_2}d_{i_3}d_{i_4})^3$. Therefore, H does not contain any K_4^3. Hence, by a result of de Caen [1] there exists a constant $\delta > 0$ such that there at least δt^3 triples (i,j,k), $1 \le i < j < k \le t$ such that $\{P_i, P_j, P_k\}$ is not an edge in H.

Let $h = \max_{m \in M} h(m) \le (\log n)^{\log 3}$. If $\{P_i, P_j, P_k\} \notin H$, $1 \le i < j < k \le t$, then $m = d_i d_j d_k$ has at most h decompositions as a product of three positive integers, which gives the following bound on M:

$$|M| \ge \frac{\delta t^3}{h} \gg \frac{n}{(\log n)^{3+\log\frac{1}{3}-1+\varepsilon} \cdot (\log n)^{\log 3}} = \frac{n}{(\log n)^{2+\varepsilon}},$$

which completes the proof of the lower bound.

The construction providing the upper bound is the set of the integers a such that

(i) $\frac{n}{\log n} \le a \le n$,

(ii) $d^2 \mid a$ implies $d \le \log n$, and

(i)ii a cannot be written in the form $a = uvw$ with integers u, v, w such that $\frac{\sqrt[3]{n}}{(\log n)^{16}} \le u, v, w \le \sqrt[3]{n}(\log n)^{16}$.

Here, we omit the details.

3.4

The proof of Theorem 7 is a modification of the similar bounds for multiplicative 3-Sidon sets, that is, for sets avoiding solutions to the equation $a_1a_2a_3 = b_1b_2b_3$. (Note that the main term for multiplicative 3-Sidon sets is larger, $\pi(n) + \pi(n/2)$, so neither bound is a corollary, instead the methods should be adapted to this slightly different setting.)

The set achieving the asymptotically largest possible size for Theorem 8 is

$$A = \left\{ m : \ m = pq, \ \frac{n}{\log n} < m \le n, \ p, q \text{ primes}, \ p < \frac{q}{\log n} \right\}.$$

The upper bound is a consequence of [2, Theorem 3], since, according to this result, if n is large enough, there exist distinct $a_1, a_2, \ldots, a_6 \in A$ such that

$$a_1a_2 = a_3a_4 = a_5a_6,$$

however, then $a_1a_2a_3a_4a_5a_6$ is a perfect cube.

4 Concluding remarks and open problems

We gave bounds for the functions $F_{k,3}(n)$ and $f_{k,3}(n)$.

Finally, we pose some problems for further research.

Problem 1. *Let us suppose that $1 < k < d$. Is it true that*

$$n^{k/d} \ll n - F_{k,d}(n) \le n - f_{k,d}(n) \ll n^{k/d}?$$

Problem 2. *Is it true that there exists a constant c such that*

$$f_{2,3}(n) = n - (c + o(1))n^{2/3}?$$

Problem 3. *Let $d \geq 4$. Is it true that*

$$f_{d+1,d}(n) = (1 - o(1))n?$$

As a corollary of the above theorems we get the following result:

Corollary 9. *For $d = 2, 3$ and $k > d$, $d \mid k$, there exist constants $c_{k,d} > 0$ and $C_{k,d} \in \mathbb{Z}^+$ such that*

$$F_{k,d}(n) = (c_{k,d} + o(1))\pi_{C_{k,d}}(n),$$

where $\pi_r(n)$ denotes the number of positive integers up to n which have exactly r prime factors (counted with multiplicity)

Problem 4. *Is it true that for any $d \geq 4$ and $k > d$, $d \mid k$, there exist constants $c_{k,d} > 0$ and $C_{k,d} \in \mathbb{Z}^+$ such that*

$$F_{k,d}(n) = (c_{k,d} + o(1))\pi_{C_{k,d}}(n)?$$

References

[1] D. de Caen, Extensions of a theorem of Moon and Moser on complete subgraphs, *Ars Combin.* **16** (1983), 5–10.

[2] P. Erdős, On the multiplicative representation of integers, *Israel Journal of Mathematics*, **2** (4) (1964) 251–261.

[3] P. Erdős, A. Sárközy, V. T. Sós, On product representations of powers, I., *European Journal of Combinatorics* **16 (6)** (1995) 567–588.

[4] Zs. Gy. Fleiner, M. H. Juhász, B. Kövér, P. P. Pach, Cs. Sándor, Product representation of perfect cubes, arXiv: 2405.12088

[5] E. Győri, C_6-free bipartite graphs and product representation of squares, *Discrete Mathematics* **165** (1997) 371–375.

[6] P. P. Pach, Generalized multiplicative Sidon sets, *Journal of Number Theory* **157** (2015) 507–529.

[7] P. P. Pach, An improved upper bound for the size of the multiplicative 3-Sidon sets, *Int. J. Number Theory* **15 (8)** (2019) 1721–1729.

[8] P. P. Pach, M. Vizer, Improved lower bounds for multiplicative square-free sequences, *Electronic Journal of Combinatorics* **30 (4)** (2023) Article Number P4.31.

[9] G. N. Sárközy, Cycles in bipartite graphs and an application in number theory, *Journal of Graph Theory* **19 (3)** (1995) 323–331.

[10] T. Tao, On product representations of squares, arXiv: 2405.11610

[11] J. Verstraëte, Product representations of polynomials, *E*uropean Journal of Combinatorics, **27 (8)** (2006) 1350–1361.

Poster presentations

Computing 2-homogeneous equitable partitions of graphs with a unique tree representation*

Aida Abiad[†1] and Sjanne Zeijlemaker [‡1]

[1]Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands

Abstract

Equitable partitions have been applied to a wide variety of topics, ranging from algebraic graph theory to clustering. An equitable partition of a graph is called *k-homogeneous* if each cell has size k. Abiad et al. showed that the existence of a 2-homogeneous equitable partition can be decided in polynomial time for cographs. In this work, we provide an alternative proof of this result, which in turn gives rise to a more general algorithm for graph classes which admit a unique tree representation. We show how the result on cographs follows from our method, and explore further applications to other graph classes.

1 Introduction

Equitable partitions are a versatile tool that have been used in many different fields of mathematics, for example for deriving sharp eigenvalue bounds on the independence number [7], constructing self-orthogonal codes [6] and clustering in various types of networks [10, 12]. A natural question is therefore, how efficiently such partitions can be computed. However, little is known about the complexity of computing equitable partitions in general [11] and the few known results focus on particular kinds of partitions or graphs. For instance, the coarsest equitable partition of a graph can be computed in polynomial time, see for example Corneil and Gotlieb [4] and Bastert [2]. Abiad et al. [1] showed that determining the existence of a 2-homogeneous equitable partition is NP-hard in general, but can be done in quadratic time for cographs. Cographs are known to have a tree representation which is unique up to isomorphism [5], a property that extends to many other graph classes. In this work, we derive a more general algorithm to compute 2-homogeneous equitable partitions in graphs that have a unique tree representation, implying the known result from [1] on cographs. Next, we explore for which other graph classes our more method can efficiently compute 2-homogeneous equitable partitions. We propose several graph classes that fall under the more general framework, and briefly discuss their structural differences and similarities with cographs.

2 Preliminaries

We consider undirected, simple and loopless graphs. A graph is denoted by $G = (V, E)$ and its number of vertices by n. The set $\{1, 2, \ldots, n\}$ is abbreviated as $[n]$. Let $G = (V, E)$ be a graph and let $\mathcal{P} = \{V_1, V_2, \ldots, V_m\}$, with $m \in [n]$, be a partition of the vertex set V. We refer to the subsets V_i

*The full version of this work will be published elsewhere.

[†]Email: a.abiad.monge@tue.nl. Research of A. A. supported by NWO (Dutch Research Council) through the grant VI.Vidi.213.085.

[‡]Email: s.zeijlemaker@tue.nl

as *cells*. A partition is called *equitable* (or *regular*) if for all i, j, each vertex in V_i has the same number of neighbors in V_j. We call a partition *k-homogeneous* if every cell has size k.

An automorphism ϕ of a graph is called an *involution* if it has order two. It is *fixed-point-free* if no vertex is mapped to itself. The following Lemma by Abiad et al. establishes a one-to-one correspondence between 2-homogeneous equitable partitions and automorphisms with the aforementioned properties.

Lemma 1 ([1, Lemma 17])**.** *Let G be a graph on n vertices. Then G has an automorphism being an involution without fixed points if and only if G admits an equitable partition with $\frac{n}{2}$ cells each having size 2.*

3 An efficient algorithm for computing 2-homogeneous equitable partitions

In general, it is NP-hard to decide whether an arbitrary graph admits a 2-homogeneous equitable partition, see [1]. Nevertheless, there exist graph classes for which the existence of a 2-homogeneous equitable partition can be established in polynomial time. Lemma 1 was also used in [1] to show that this is the case for the class of cographs (see Section 4.1 for a definition). This was done using a characterization of cographs in terms of twin classes. However, cographs can be characterized in many other ways, most notably by a tree representation which is unique up to isomorphism [5]. In this section, we derive a more general algorithm to compute 2-homogeneous equitable partitions for graphs that allow a unique tree representation, and show how the result on cographs follows from our method.

For our algorithm, we consider the following class of graphs.

Definition 2. *Let $\mathcal{G}$ be the class of all graphs satisfying the following conditions.*

(C1) There exists a polynomial-time computable rooted tree representation T which uniquely characterizes the graphs of this class up to isomorphism. Each leaf of the tree corresponds to a graph on a subset of vertices from the original graph and each internal vertex to a graph operation on the subgraphs represented by the subtrees rooted at its (unordered) children.

(C2) A fixed-point-free involution on a graph G from this class corresponds one-to-one with a label-preserving automorphism ϕ on T plus a sequence of automorphisms $\psi_1, \ldots, \psi_m$ on the subgraphs corresponding to its leaves such that

- *ϕ is an involution on the leaves of T,*
- *only maps leaf i to itself if ψ_i is a fixed-point-free involution.*

(C3) The existence of a fixed-point-free involution should be decidable in polynomial time for the subgraphs represented by the leaves.

(C4) Isomorphism of the subgraphs represented by the leaves should be decidable in polynomial time.

In (C1), (C3) and (C4), 'polynomial' means polynomial in the number of vertices of the graph.

In general, the automorphism problem on rooted labeled trees is known to be polynomial-time solvable, see Colbourn and Booth [3]. However, in the context of Lemma 1 we need to determine the existence of a particular type of automorphism which is also an involution without fixed points. In terms of the unique tree representation, this means that there should be an automorphism which swaps the leaves pairwisely or leaves them in place. However, if a leaf is not swapped, its associated subgraph must admit a fixed-point-free involution internally. In Algorithm 1, we propose a recursive procedure which determines the existence of such a mapping for labeled rooted trees representing graphs in $\mathcal{G}$. Here T_v denotes the subtree of a tree T rooted at vertex v and we define G_v to be the induced subgraph of G corresponding to the subgraph associated with leaf v in T. Since it is assumed that isomorphism can be determined efficiently for the graphs associated with the leaves of T (see Condition (C4)), we

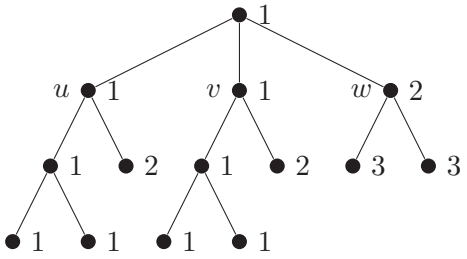


Figure 1: A j-numbered tree which admits a fixed-point-free involution on the leaves.

Algorithm 1: `hasNiceAutomorphism`

Input : A (labeled) rooted tree T representing a graph $G \in \mathcal{G}$ with root r and j-numbers assigned to each vertex, following the procedure from [3, Lemma 2.1]

Output: Does T admit an automorphism which is a fixed-point-free involution on the leaf vertices?

```
hasNiceAutomorphism(T, r)
    for each child v of r with distinct j-number do
        let k_v be the number of children with the same j-number as v
        if k_v is odd then
            if v is a leaf and G_v does not admit a fixed-point-free involution then
                return false
            else
                if hasNiceAutomorphism(T_v, v) = false then
                    return false
    return true
```

will assume that the leaves are labeled such that two leaves have the same label if and only if their associated subgraphs are isomorphic. The algorithm assumes that the input tree has been labeled using the j-numbering procedure by Colbourn and Booth [3]. These numbers are assigned in a top-down fashion to each vertex of the tree and, together with the depth of a vertex, partition the tree into its orbits under the automorphism group. A j-numbering can be computed in linear time, hence the running time of Algorithm 1 is polynomial.

Before we show correctness of Algorithm 1, we provide some intuition. Consider the rooted tree T given in Figure 1. It is clear that an automorphism of T could swap the subtrees rooted at u and v, as they are isomorphic and have the same parent. This is a fixed-point-free involution on the leaves that descend from u and v. The subtree of w cannot be mapped to another part of the graph in its entirety, but if we go one level down, we see that interchanging the children of w also results in a fixed-point-free involution on the remaining leaves.

Lemma 3. *Let $G \in \mathcal{G}$ and let T be the unique (labeled) rooted tree representing G. Algorithm 1 returns "true" if and only if T admits an automorphism which*

(i) is an involution on the leaves;

(ii) only maps a leaf to itself if the associated subgraph admits a fixed-point-free involution.

The running time of the algorithm is $O(m(m+p(n))$, where p is a polynomial and m and n denote the number of vertices of T and G respectively.

If m is polynomial in n, the running time of Algorithm 1 is also polynomial in n. Note that the number of vertices in a tree equals at most twice the number of leaves, so in order for this to hold, we only need the number of leaves of T to be polynomial in n. Combining this with Lemma 3, we obtain our main result.

Theorem 4. *Let $G \in \mathcal{G}$ be a graph with n vertices such that the number of leaves of its unique tree representation T is polynomial in n. Then the problem of deciding whether G admits an equitable partition with $\frac{n}{2}$ cells of size 2 can be solved in poly(n) time.*

Note that Algorithm 1 can easily be modified to keep track of the partial mappings and return an automorphism ψ of T satisfying conditions (i) and (ii). If we additionally compute a fixed-point-free involution for each G_v corresponding to a leaf v with odd k_v, we obtain a fixed-point-free involutionary automorphism of G, whose orbits form a 2-homogeneous equitable partition of G. Hence computing equitable partitions with $\frac{n}{2}$ cells of size 2 can also be done in polynomial time.

4 Applications

4.1 Cographs

In [1], it was shown that one can determine the existence of a 2-homogeneous equitable partition in quadratic time for the class of cographs. In this section, we provide an alternative proof of this result using Algorithm 1.

A *cograph* is a graph which does not have a P_4 as an induced subgraph. Alternatively, it can be characterized as follows.

Proposition 5 (Corneil et al. [5])**.** *A cograph is defined recursively using the following three rules.*

(i) A graph on a single vertex is a cograph.

(ii) If $G_1, \ldots, G_k$ are cographs, then so is $G_1 \cup \cdots \cup G_k$.

(iii) If G is a cograph, then so is its complement $\overline{G}$.

Note that we may equivalently replace (iii) by the condition that the join of two cographs is again a cograph. Using this characterization, the structure of a cograph can uniquely be represented by a rooted tree.

Let G be a cograph. A *cotree* of G is a rooted tree whose inner vertices each have a label 0 or 1. A leaf vertex corresponds to an induced subgraph on a single vertex and the subtree rooted at a vertex with label 0 or 1 corresponds to the union or join respectively of the subgraphs represented by its children. Note that two vertices form an edge in G if and only if their least common ancestor in the cotree has label 1. If we require the labels on a root-leaf path to be alternating, this tree is unique, see Corneil et al. [5]. The same authors observed that the graph isomorphism problem is therefore polynomial-time solvable for cographs.

Since isomorphism can be detected in polynomial time for cographs by studying the cotree, it makes sense that an automorphism of a cograph should correspond to a certain automorphism of its cotree. We make this intuition more precise in the following lemma.

Lemma 6. *Let G be a cograph with unique cotree T with alternating 0/1-labels. Then, $\phi\colon V \to V$ is an automorphism of G if and only if there exists an automorphism ψ on T such that $\psi|_V = \phi$ and which respects the 0/1-labeling.*

Lemma 6 implies that cographs satisfy Condition (C2). Combined with the uniqueness of the cotree, this implies that cographs form a subclass of $\mathcal{G}$. From Theorem 4, we then obtain an alternative proof for the following complexity result from [1].

Corollary 7 ([1, Theorem 25])**.** *The problem of deciding whether a cograph admits an equitable partition with $\frac{n}{2}$ cells of size 2 can be solved in $O(n^2)$ time.*

4.2 Other graph classes with tree representations

In the previous section, we showed how Algorithm 1 can be used to efficiently determine the existence of 2-homogeneous equitable partition in cographs. Several generalizations of cographs have been proposed in the literature, as well as other graph classes with a unique tree representation. Therefore, a natural direction for future research is to examine to which of these graphs our algorithm can be applied. Below, we highlight three promising graph classes, and indicate which further steps need to be taken to apply our approach.

Tree-cographs As we have seen in Proposition 5, cographs can be defined recursively by starting with a set of isolated vertices and repeatedly applying disjoint union and complementation operations. A *tree-cograph* is a graph which can be obtained by applying the same operations to a set of trees. Tree-cographs generalize both the class of trees and cographs. It was shown by Tinhofer [13] that these graphs can be uniquely characterized by a tree representation. Contrary to cographs, the inner vertices of this tree represent the disjoint union and complementation operations and the leaves each correspond to a tree. It is easy to see that this class of graphs satisfies Conditions (C1), (C3) and (C4). However, as the proof of Lemma 6 makes use of the properties of the join operator and a characterization of adjacency in cographs, an alternative proof is needed to show that tree-cographs are a subclass of $\mathcal{G}$.

P_4-sparse graphs Cographs are the class of graphs which contain no induced P_4. A natural generalization of this concept is the class of P_4-sparse graphs. A graph is *P_4-sparse* if every set of five vertices induces at most one P_4. Jamison and Olariu [8] showed that a graph is P_4-sparse if and only if it can be constructed from a set of isolated vertices through applying the disjoint union, join and a third operator denoted by ② (see [8] for more details). Moreover, this constructive characterization gives rise to a unique tree representation. Once again, an alternative to Lemma 6 is needed to show inclusion in $\mathcal{G}$, as Condition (C2) is not trivially satisfied.

Interval graphs An *interval graph* is the intersection graph of a set of intervals on the line. Interval graphs can be represented by a *PQ-tree* [9], whose leaves correspond to the maximal cliques of the associated graph. Colbourn and Booth [3] proved that an automorphism of an interval graph correspond one-to-one with an automorphism of its PQ-tree and a certain permutation of the vertices. This gives us a direct link between the fixed-point-free involutions of an interval graph and automorphisms of its PQ-tree, as needed for Condition (C2). However, to apply Algorithm 1, two additional properties of the PQ-tree need to be taken into consideration. Firstly, some vertices of the PQ-tree have ordered children which may not be exchanged arbitrarily. These restrictions can easily be incorporated into the algorithm. Secondly, the maximal cliques of an interval graph, and hence the subgraphs associated with the leaves of the PQ-tree, may not be disjoint. Algorithm 1 traverses a given tree in a top-down manner and attempts to pair up each subtree individually to obtain a fixed-point-free automorphism. For PQ-trees, consistency needs to be ensured for vertices appearing in multiple subtrees.

References

[1] A. Abiad, C. Hojny, and S. Zeijlemaker. Characterizing and computing weight-equitable partitions of graphs. *Linear Algebra and its Applications*, 645:30–51, 2022.

[2] O. Bastert. Computing equitable partitions of graphs. *Match*, 40:265-–272, 1999.

[3] C. J. Colbourn and K. S. Booth. Linear time automorphism algorithms for trees, interval graphs, and planar graphs. *SIAM Journal on Computing*, 10(1):203–225, 1981.

[4] D. G. Corneil and C. C. Gotlieb. An efficient algorithm for graph isomorphism *Journal of the ACM*, 17(1):51–64, 1970.

[5] D. G. Corneil, H. Lerchs, and L. S. Burlingham. Complement reducible graphs. *Discrete Applied Mathematics*, 3(3):163–174, 1981.

[6] D. Crnković, S. Rukavina, and A. Švob. Self-orthogonal codes from equitable partitions of association schemes. *Journal of Algebraic Combinatorics*, 55:157–171, 2022.

[7] A. J. Hoffman. On eigenvalues and colorings of graphs. In *Selected Papers of Alan J Hoffman: With Commentary*, 407-–419. World Scientific, 2003.

[8] B. Jamison and S. Olariu. A tree representation for P_4-sparse graphs. *Discrete Applied Mathematics*, 35(2):115–129, 1992.

[9] G. S. Lueker and K. S. Booth. A linear time algorithm for deciding interval graph isomorphism. *Journal of the ACM*, 26(2):183–195, 1979.

[10] M. Mattioni and S. Monaco. Cluster partitioning of heterogeneous multi-agent systems. *Automatica*, 138: 110136, 2022.

[11] B. D. McKay. Complexity of equitable partitions. URL: `https://mathoverflow.net/q/96858` (version: 2012-05-14)

[12] B. Prasse, K. Devriendt, and P. Van Mieghem. Clustering for epidemics on networks: a geometric approach. *Chaos: an Interdisciplinary Journal of Nonlinear Science*, 31(6):063115, 2021.

[13] G. Tinhofer. Strong tree-cographs are Birkhoff graphs. *Discrete Applied Mathematics*, 22(3):275–288, 1988.

Categorification of Flag Algebras*

Aldo Kiem[1,2], Sebastian Pokutta[1,2], and Christoph Spiegel[1,2]

[1]Technische Universität Berlin, Institute of Mathematics
[2]Zuse Institute Berlin, Department AIS2T, *lastname*@zib.de

Abstract

Razborov introduced the notion of flag algebras in 2007. Since then, they have become an important computational and theoretical tool in extremal combinatorics. Originally phrased in terms of universally quantified first-order theories and often presented purely combinatorially when applied, we propose a category theoretic foundation for flag algebras that unifies these previous approaches. This allows us to obtainin some new foundational results for flag algebras, such as a partial classification of linear and order-preserving maps between them and higher-order differential methods.

1 Introduction

Razborov introduced flag algebras as a unifying language to connect a distinct set of connected problems in extremal combinatorics in 2007 [13]. This was phrased in terms of universally quantified first-order theories and, roughly speaking, is applicable whenever any subset of a structure induces a substructure. Flag algebras also have a strong connection to combinatorial limit objects [4]. Razborov also introduced some fundamental techniques, such as ways to relate different algebras in the form of a downward- and upward uperator, the differential method, a Cauchy-Schwarz-type inequality and the prerequisites for the sum-of-squares method.

While purely theoretical applications of flag algebras have been important, such as the application of the differential method to resolve the minimal density of triangles in graphs [14], the computational approach stemming from the sum-of-squares method has had the most significant impact, covering results relating to 3-edge-colored triangles [5], 4-edge-colored triangles [7], 3-edge-colored rainbow triangles [3], pentagons in triangle-free graphs [6], as well as Turán problems [15], [1], [12] and[10]. These applications often present their own purely combinatorial derivation of flag algebras and their relevant properties. However, some of these ad-hoc derivations, while valid from a combinatorial point of view, are not covered by Razborov's original first-order framework; in particular problems relating to hypercubes [2], [1] and finite vector spaces [11] do not fulfill the property that every subset induces a substructure.

Partially to address this issue, we present a category theoretic foundation for flag algebras. This also answers Coregliano and Razborov's [4] inquiry for a more in-depth study of the category `Int` whose objects are the universally quantified first-order theories with all total interpretations as arrows. It also ties into previous efforts of obtaining a categorical generalisation of combinatorial phenomena as in part the study of graph limits [8] (see chapter 23.4 and the discussion concerning categories and flag algebras). The category theoretic view has the added benefit of resulting in an overall cleaner presentation. Finally, this framework also allows us to obtain some new results regarding the theory

*The full version of this work is currently under preparation. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689).

of flag algebras. In particular, we obtain a partial classification of linear and order-preserving maps between flag algebras, complementing previous efforts in Razborov's original paper, and we formulate higher-order vertex differential methods.

We start by giving the category theoretic formulation of flag algebras and its relation to `Int` in Section 2. We establish their basic properties within this framework in Section 3 and cover some novel statements in Section 4. We conclude with a short discussion in Section 5.

2 Construction

We let `FinInj` denote the category that has all finite sets as its objects and injections between these sets as arrows. Our first observation is that the category `Int` is a full subcategory of the category of presheaves on `FinInj`.

Observation 1. *Given a category $\mathfrak{A}$, denote by* $\texttt{FinPsh}(\mathfrak{A})$ *the category of finite presheaves on $\mathfrak{A}$. Let us show how to embedd* `Int` *as a full subcategory of* `FinSet`*. For an object T of* `Int`*, define the finite presheaf $F_T : \texttt{FinInj}^{\mathrm{op}} \to \texttt{FinSet}$ where $F_T(S) = \{M$ is a T-model with ground set $S\}$. For an arrow $\alpha : R \to S$ in* `FinInj`*, we let $F_T(\alpha)(M)$ be the unique model with ground set R so that α is an embedding $F_T(\alpha)(M) \to M$. The association $T \to F_T$ is then a full functor* $\texttt{Int} \to \texttt{FinPsh}(\texttt{FinSet})$.

In other words, every total interpretation corresponds to a natural transformation of presheaves and vice versa. In order to get a category theoretic definition of flag algebras we need the technical definition of an Archimedean partially-ordered vector space.

Definition 2. *An $\mathbb{R}$-vector space V with a preordering $\leq$ is* Archimedean *when for every $v, w \in V$ we have that $v \leq r\, w$ for every $r \in \mathbb{R}_{>0}$ implies $v \leq 0$. A linear map $f : V \to W$ between two Archimedean vector spaces is* order-preserving *if $v \leq v'$ in V implies that $f(v) \leq f(v')$ in W. Let* `Arch` *denote the category with all (possibly infinite) Archimedean preordered $\mathbb{R}$-vector spaces as its objects and all order-preserving linear maps as its arrows.*

Definition 3. *The* powering functor $\mathbb{R}^{\cdot} : \texttt{Set} \to \texttt{Arch}$ *is given by mapping any object S of* `Set` *to the Archimedean space $\mathbb{R}^S$ and any arrow $f : R \to S$ of* `Set` *to the map $\mathbb{R}^S \to \mathbb{R}^R$ that sends a function $g : S \to \mathbb{R}$ to $g \circ f$.*

Momentarily disregarding their multiplicative structure, the flag algebras as defined by Razborov in [13, 4] can be seen as a functor $\mathcal{A} : \texttt{Int} \to \texttt{Arch}$. The colimit $\mathcal{A}[F] = \operatorname{colim} \mathbb{R}^F$, which can be shown to exist for any $F \in \texttt{FinPsh}(\mathfrak{A})$ for any category $\mathfrak{A}$, lifts this to a functor $\texttt{FinPsh}(\mathfrak{A}) \to \texttt{Arch}$. In particular, for every universally quantified first-order theory T, the flag algebra $\mathcal{A}[T]$ as defined by Razborov is isomorphic as an Archimedean vector space to $\mathcal{A}[F_T]$ as just defined. The space $\lim \mathbb{R}[F_T]$ corresponds to $\mathbb{R}$-linear combinations of positive homomorphisms.

It remains to establish under what conditions we can define a multiplicative structure on $\mathcal{A}[F]$ in such a way that it has the same properties as the originally defined flag algebras. We note that we choose to put all structural requirements on the base category $\mathfrak{A}$ and none on the nature of the presheaf F, so that we can define an algebra for any $F \in \texttt{FinPsh}(\mathfrak{A})$ assuming the right conditions on $\mathfrak{A}$.

Definition 4. *For a given category $\mathfrak{A}$, we write $x \leq y$, whenever there exists a morphism from x to y. We say that $\mathfrak{A}$ is a* density category *if:*

1. *For any two objects x, y of $\mathfrak{A}$, there are finitely many arrows in $\mathfrak{A}(x, y)$;*
2. *For any two arrows $\alpha, \beta \in \mathfrak{A}(x, y)$, there exists an isomorphism $\gamma \in \mathfrak{A}(y, y)$ such that $\gamma \circ \alpha = \beta$;*
3. *There exists an increasing countable sequence of objects $x_1, x_2, \ldots$ so that for any object x there exists some index i s.t. $x \leq x_i$;*

4. *For any two objects $y, z \in \mathfrak{A}$ there exist $r \geq 0$ so that for every $\epsilon > 0$ there exists $x_0 \in \mathfrak{A}$ so that for all $x_0 \leq x$ the fraction $|\mathfrak{A}(y,x)|/|\mathfrak{A}(z,x)|$ is well-defined and does not differ from r by more than ϵ.*

5. *For any finite tuple of objects P there exists another finite tuple of objects $\mathrm{co}(P)$ together with arrows $\alpha_{x,y} : x \to y$ for any $x \in P$ and $y \in \mathrm{co}(P)$, which, for any $z \in \mathrm{Obj}(\mathfrak{A})$, induce an isomorphism of sets*
$$\prod_{x \in P} \mathfrak{A}(x,z) \simeq \coprod_{y \in \mathrm{co}\, P} \mathfrak{A}(y,z).$$

We use the term 'density categories', because these are precisely the properties that are needed in order to make sense of the usual notion of densities between structures as well as products of densities. When $\mathfrak{A}$ is a density category, $F \in \mathtt{FinPsh}(\mathfrak{A})$ as well as $f \in F(x)$ and $g \in F(y)$, we can define an equivalent notion to Razborov's homomorphism density through the quotient

$$p(f;g) = |\{\alpha \in \mathfrak{A}(x,y) | F(\alpha)(g) = f\}|/|\mathfrak{A}(x,y)|.$$

When $f \in \mathbb{R}^F(x)$ and $g \in \mathbb{R}F(y)$ are linear combinations, we can extend this density bilinearly to $p(f;g)$. We now define a multiplication in $\mathcal{A}[F]$.

Definition 5. *Let $F \in \mathtt{FinPsh}(\mathfrak{A})$. Given a tuple $f_1 \in F(x_1), \ldots, f_m \in F(x_m)$ and $P = (x_1, \ldots, x_m)$, define their product as*
$$f_1 \cdots f_m = \sum_{y \in \mathrm{co}(P)} g_y \lim_z \frac{|\mathfrak{A}(y,z)|}{\prod_{x \in P} |\mathfrak{A}(x,z)|} \in \mathcal{A}[F]$$
where $g_y \in \mathbb{R}^{F(y)}$ is the unweighted sum of all $g \in F(y)$ so that for $i = 1, \ldots m$, we have $f_i = g \circ F(a_{x_i,y})$.

We let $1 \in \mathcal{A}[F]$ denote the sum of all flags corresponding to an arbitrary but fixed source object y. Note that the definition of $1 \in \mathcal{A}[F]$ is independent of the choice of the source object y and multiplication in $\mathcal{A}[F]$ is associative, commutative, and has 1 as its unit. Furthermore, squares are positive in $\mathcal{A}[F]$.

Examples for density categories include the aforementioned `FinSet` as well as finite vector spaces $\mathtt{FinVec}_q$ over some finite field $\mathbb{F}_q$ with injective arrows and the category `HyperCube` of hypercubes with morphisms the injective face maps. In Observation 1 we showed how to convert a combinatorially meaningful object T in `Int` to an object of $\mathtt{FinPsh}(\mathtt{FinInj})$. The same theme can be used to convert previously studied theories over $\mathtt{FinVec}_q$ and `HyperCube` to finite presheaves. For example, c-vertex colorings of the objects of $\mathtt{FinVec}_q$ are represented by that $F \in \mathtt{FinPsh}(\mathtt{FinVec}_q)$ which maps a vector space V to the set $F(V)$ of all c-vertex colorings of V, not up to isomorphism. Similarly, c-edge colorings of the objects of `HyperCube` are represented by that $F \in \mathtt{FinPsh}(\mathtt{HyperCube})$ mapping a hypercube C to the set $F(C)$ of all c-edge colorings of the edges of C, not up to isomorphism.

3 Properties

All of the flag algebra calculus remains true when we are considering a finite presheaf $F \in \mathtt{FinPsh}(\mathfrak{A})$ over a density category. In particular, there are downward operators. Since squares are always positive, this means that the sum of squares method works for any such F.

Definition 6. *Let $P : \mathfrak{A} \to \mathfrak{B}$ be a functor between density categories. The* downward operator *between two flag algebras $[\![\cdot]\!]_P : \mathcal{A}[F \circ P] \to \mathcal{A}[F]$ is defined as the map induced by sending any $f \in (F \circ P)(x)$ to the same element in $F(P(x))$.*

The notion of types for flag algebras is essential and the basis for the sum-of-squares method. What they allow for is an amalgamated multiplication of flag aglebra elements. This produces elements in

the positive cone of $\mathcal{A}[F]$ that would be extremely difficult to detect otherwise. In the context of the categorification of flag algberas, types are given by coslice categories. Let $\mathfrak{A}$ be a density category and $x \in \mathrm{Obj}(\mathfrak{A})$ so that the under category $x/\mathfrak{A}$ is again a density category. Then, denote by U_x the forgetful functor $x/\mathfrak{A} \to \mathfrak{A}$. The algebra at x is $\mathcal{A}^x[F] = \mathcal{A}[F \circ U_x]$ and the downward operator between the algebras $\mathcal{A}^x[F] \to \mathcal{A}[F]$ is given by Definition 6. Similarly, when $\alpha : x \to y$ is a morphism, there is an upward operator $\pi^\eta : \mathcal{A}^x[F] \to \mathcal{A}^y[F]$. It is defined by multiplying an element $f \in \mathcal{A}^x[F]$ with a projection of the unit of $\mathcal{A}^y[F]$ and interpreting the result as an element of $\mathcal{A}^y[F]$ again. The upward and downward operators are compatible in the same ways as in [13].

Problems in extremal combinatorics usually speak of a certain limit of combinatorial structures that minimizes or maximizes an objective function. In terms of Razborov's flag algebras, these limit sequences are represented by order-preserving algebra homomorphisms $\mathcal{A}[F_T] \to \mathbb{R}$ and vice versa. Several of the properties we require of a density category have the sole purpose of ensuring that this correspondence remains true for $F \in \mathtt{FinPsh}(\mathfrak{A})$.

Definition 7. *For an increasing sequence $x_1, x_2, \ldots$ of objects in $\mathfrak{A}$ as in Item 3 and elements $u_n \in \mathbb{R}F(x_n)$, we say that u_n is convergent if for every flag f of $\mathcal{A}[F]$ the limit of $p(f; u_n)$ as $n \to \infty$ exists.*

Then, by our definition of density categories we get the following equivalent statement to Razborov's Theorem 3.3 from [13]. For two partially ordered algebras A and B, let $\mathrm{Hom}^+(A, B)$ denote the set of order-preserving maps from A to B.

Theorem 8. *Let $x_1, x_2, \ldots$ be a countable sequence of objects in $\mathfrak{A}$ as in Item 3. There is a surjective correspondence between $\mathrm{Hom}^+(\mathcal{A}[F], \mathbb{R})$ and convergent sequences $u_n \in F(x_n)$.*

4 Results

In this section, we investigate what order-preserving, not necessarily multiplicative, morphisms $\mathcal{A}[F] \to \mathcal{A}[G]$ there exist when $F, G \in \mathtt{FinPsh}(\mathfrak{A})$. Some of these maps are already known, like the downward operators or the the image under $\mathcal{A}$ of a natural transformation $G \to F$ of presheaves. We will focus on showing how to classify the particular case of natural transformations $\mathbb{R}^F \to \mathbb{R}^G$. Interestingly, when $\mathfrak{A} = \mathtt{FinInj}$, the presheaf G is G_T for some theory T like edge colored u-uniform hypergraphs without forbidden substructures, these natural transformations $\mathbb{R}^F \to \mathbb{R}^G$ coincide bijectively with the elements of $\mathrm{Hom}^+(\mathcal{A}_{[G,F]}, \mathbb{R})$ where $[G, F]$ is the internal hom of presheaves. In general however, $\mathrm{Hom}^+(\mathcal{A}_{[G,F]}, \mathbb{R})$ does not classify the natural transformations $\mathbb{R}^F \to \mathbb{R}^G$ for arbitrary presheaves F and G, even over $\mathfrak{A} = \mathtt{FinInj}$.

First we show that there are many ways to represent every natural transformation $\mathbb{R}^F \to \mathbb{R}^G$ injectively as a convergent sequence in the presheaf $\mathbb{R}[F \times G]$. Intuitively, we think of $\mathbb{R}[F \times G]$ as the space in which we represent the graph of a natural transformation $\mathbb{R}^F \to \mathbb{R}^G$, just as it would be when we consider the graph in $A \times B$ of a function $A \to B$ when A and B are sets.

Remark 9. *For a natural transformation $L : \mathbb{R}^F \to \mathbb{R}^G$, denote by $L^\vee$ the dual natural transformation $\mathbb{R}[G] \to \mathbb{R}[F]$. Let $\psi \in \mathrm{Hom}^+(\mathcal{A}[G], \mathbb{R})$ be a positive homomorphism that is non-zero on every element of G and let $u_i \in G(x_i)$ be a sequence that converges to ψ.*

Consider the convergent sequence in $\mathbb{R}[F \times G]$

$$w_i = L^\vee(u_i) \times u_i.$$

From such a sequence w_i, we can uniquely recover the values $L^\vee(g)$ for $g \in G(x)$ through the formula

$$L^\vee(g) = \tfrac{1}{\psi(g)} \sum_{f \in F(x)} \left(\lim_i p(f \times g; w_i) \right) f. \tag{1}$$

This is a consequence of the fact that L is a natural transformation. Therefore, given ψ, every natural transformation $L : \mathbb{R}^F \to \mathbb{R}^G$ corresponds to a unique convergent sequence in $\mathbb{R}[F \times G]$.

The problem is that given $\psi \in \mathrm{Hom}^+(\mathcal{A}[G], \mathbb{R})$, we do not know which convergent sequences in $\mathbb{R}[F \times G]$ correspond to natural transformations L. Therefore, we introduce the technical notion of a vertex extension property for ψ which guarantees a converse.

Theorem 10. *Assume that $\psi \in \mathrm{Hom}^+(\mathcal{A}[G], \mathbb{R})$ has the vertex extension property for all elements of G and let $u_i \in G(x_i)$ be a sequence that converges to ψ. Then for every sequence $v_i \in \mathbb{R}[F(x_i)]$ so that $v_i \times u_i$ converges in $\mathbb{R}[F \times G]$, Equation (1) defines a natural transformation.*

The precise definition of the vertex extension property is given in the following. Example homomorphisms that fulfill this condition are the u-uniform random hypergraphs $\psi \in \mathrm{Hom}^+(\mathcal{A}[F_{u\text{-Hyper}}], \mathbb{R})$. Therefore, we get a complete description of all natural transformations $\mathbb{R}^F \to \mathbb{R}^{F_{u\text{-Hyper}}}$ when $F \in \mathtt{FinPsh}(\mathtt{FinInj})$.

Definition 11. *Assume that $\mathfrak{A}$ and $x/\mathfrak{A}$ are density categories and let $f \in F(x)$. Let $\psi \in \mathrm{Hom}^+(\mathcal{A}[F], \mathbb{R})$ with $\psi(f) > 0$. Recall that we denote the forgetful functor $x/\mathfrak{A} \to \mathfrak{A}$ by U_x. For any y and $\eta : x \to y$ as well as $g \in F(y)$ denote by $\langle g \rangle_\eta \in (F \circ U_x)(\eta)$ the same object g but viewed as an element in a presheaf over a coslice category.*

We say that ψ has the vertex extension property for f if there exists a sequence $u_i \in F(x_i)$ converging to ψ so that for every $\alpha : x \to y$ and $h \in F \circ U_x(\alpha)$ with $F(\alpha)(\llbracket h \rrbracket_{U_x}) = f$, the sequence of random variables $E_i \circ \beta_i$

$$E_i \circ \beta_i = p(h; \langle u_i \rangle_{\beta_i}) - \tfrac{\psi(\llbracket h \rrbracket_{U_x})}{\psi(f)} p(f; F(\beta_i)(u_i))$$

with uniformly independent $\beta_i \in \mathfrak{A}(x, x_i)$, converges almost surely to 0 as i tends to infinity.

Finally, we study a special case of linear and order-preserving maps that give rise to the vertex differential method. When $\mathfrak{A} = \mathtt{FinInj}$ and the presheaf is F_T, the map arises from the notions we have just developed as follows. Denote by M_n the multiplication endofunctor $\mathtt{FinInj} \to \mathtt{FinInj}$ that maps $S \mapsto S \times [n]$. The vertex differential method then arises from a composition

$$\mathcal{A}[F_T] \to \mathcal{A}[F_T \circ \mathrm{M}_n] \to \mathcal{A}[F_T]$$

where the first arrow is induced by a natural transformation $\mathbb{R}^{F_T} \to \mathbb{R}^{F_T \circ \mathrm{M}_n}$ and the second is the downward operator. We will however simply write out the defining formula of this linear and order-preserving map directly.

Lemma 12. *We work over $\mathfrak{A} = \mathtt{FinInj}$ and denote by $\mathbf{1} \in \mathtt{FinInj}$ the set $\{1\}$. Let T be a universally quantified first-order theory. Choose any $h \in \mathcal{A}^{\mathbf{1}}[F_T]$ with $h \geq -1$ and $\llbracket h \rrbracket_{U_{\mathbf{1}}} = 0$. For $f \in F_T(y)$ define*

$$V(f) = \Big\llbracket \langle f \rangle_{\mathrm{id}_y} \prod_{\alpha \in \mathfrak{A}(\mathbf{1}, y)} \big(1 + \pi^\alpha(h)\big) \Big\rrbracket_{U_y}.$$

Then, V is linear order-preserving and its derivatives as $h \to 0$ are the (higher order) differential methods.

The fact that V is order-preserving follows from $h \geq -1$. The first-order method was already discovered by Razborov in [13]. All higher-order methods are novel. It is also possible to get higher-order edge-differential methods. To this end we must consider the functors $\mathrm{B}_u : \mathtt{FinInj} \to \mathtt{FinInj}$ which map $S \mapsto \binom{S}{u}$.

5 Discussion

We have shown that the theory of flag algebras very naturally carries over to the setting of density categories. However, several interesting avenues of exploration remain. It would for example be interesting to explore, if one can fully classify the homomorphism $\mathrm{Hom}^+(\mathcal{A}[F], \mathbb{R})$ as Razborov and Coregliano [4]

did for canonical universally quantified theiories. Furthermore, Razborov [13] was able to show that so called *open* interpretations give a class of maps between certain localizations of flag algebras. It would be interesting to see how our observations regarding the classification of natural transformations $\mathbb{R}^F \to \mathbb{R}^G$ would carry over to the case of localizations. In fact, one can use the endofunctor M_n to construct some basic maps into localizations, but we have not pursued this line of thought any further at this point.

The fact that we have chosen our presheaves to take values in finite sets is essential for the basic features of Flag Algebras. We are grateful to an anonymous referee for directing our attention to the structural limits of [9]. Stated briefly in our language, given a countable signature λ, define a measure space valued presheaf F on `FinInj`

$$F(S) = (\{(\mathbf{A}, v) \mid \mathbf{A} \text{ is a finite } \lambda\text{-structure on } S \text{ and } v : [n] \to S\}, D, \mu_{\text{count}})$$

where D is the σ-algebra generated by all subsets of $F(S)$ that are definable by $\mathrm{FO}[\lambda]$ formulas with free variables in $\{x_s \mid s \in S\}$. The measure μ_{count} is the counting measure that gives weight 1 to each equivalence class of the $(\mathbf{A}, v)$. Then, the basic objects of study of [9] are $\lim L^1(F)$, corresponding to the finite λ-structures, the flag algebra $\mathcal{A}[F]$ on $\operatorname{colim} L^\infty(F)$ corresponding to the Lindenbaum-Tarski algebra and its dual $\lim \mathrm{ba}(F)$.

References

[1] Rahil Baber. Turán densities of hypercubes. *arXiv preprint arXiv:1201.3587*, page 161171, 2012.

[2] József Balogh, Ping Hu, Bernard Lidický, and Hong Liu. Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube. *European Journal of Combinatorics*, 35:75–85, 2014. Selected Papers of EuroComb'11.

[3] József Balogh, Ping Hu, Bernard Lidický, Florian Pfender, Jan Volec, and Michael Young. Rainbow triangles in three-colored graphs. *Journal of Combinatorial Theory, Series B*, 126:83–113, 2017.

[4] Leonardo Nagami Coregliano and Alexander Razborov. Semantic limits of dense combinatorial objects. *Russian Mathematical Surveys*, 75(4):627, 2020.

[5] James Cummings, Daniel Král', Florian Pfender, Konrad Sperfeld, Andrew Treglown, and Michael Young. Monochromatic triangles in three-coloured graphs. *Journal of Combinatorial Theory, Series B*, 103(4):489–503, 2013.

[6] Hamed Hatami, Jan Hladkỳ, Daniel Král', Serguei Norine, and Alexander Razborov. On the number of pentagons in triangle-free graphs. *Journal of Combinatorial Theory, Series A*, 120(3):722–732, 2013.

[7] Aldo Kiem, Sebastian Pokutta, and Christoph Spiegel. The four-color ramsey multiplicity of triangles. *arXiv preprint arXiv:2312.08049*, 2023.

[8] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[9] Jaroslav Nešetřil and Patrice Ossona de Mendez. *A unified approach to structural limits and limits of graphs with bounded tree-depth*, volume 263. American Mathematical Society, 2020.

[10] Olaf Parczyk, Sebastian Pokutta, Christoph Spiegel, and Tibor Szabó. New Ramsey multiplicity bounds and search heuristics. *arXiv preprint arXiv:2206.04036*, 2022.

[11] Juanjo Rue Perna and Christoph Spiegel. The rado multiplicity problem in vector spaces over finite fields. In *Proceedings of the 12th European Conference on Combinatorics, Graph Theory and Applications, EUROCOMB'23*, pages 784 – 789, 2023.

[12] Oleg Pikhurko and Emil R Vaughan. Minimum number of k-cliques in graphs with bounded independence number. *Combinatorics, Probability and Computing*, 22(6):910–934, 2013.

[13] Alexander Razborov. Flag algebras. *The Journal of Symbolic Logic*, 72(4):1239–1282, 2007.

[14] Alexander Razborov. On the minimal density of triangles in graphs. *Combinatorics, Probability and Computing*, 17(4):603–618, 2008.

[15] Alexander Razborov. On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM Journal on Discrete Mathematics*, 24(3):946–963, 2010.

The rectilinear convex hull of disks

Carlos Alegría [1], Justin Dallant [2], Jean-Paul Doignon [2], Pablo Pérez-Lantero [3], and Carlos Seara [4]

[1]Dip. Ingegneria, Università Roma Tre, carlos.alegria@uniroma3.it
[2]Dept. Computer Science, Université libre de Bruxelles, {justin.dallant,paul.doignon}@ulb.be
[3]Depto. Matemática y Computación, Universidad de Santiago de Chile, pablo.perez.l@usach.cl.
[4]Dept. Matemàtiques, Universitat Politècnica de Catalunya, Spain, carlos.seara@upc.edu

Abstract

We explore an extension to orthogonal convexity of the classic problem of computing the convex hull of a collection of planar disks. Namely, we enumerate all the changes to the boundary of the rectilinear convex hull of a collection of n planar disks, while the coordinate axes are simultaneously rotated by an angle that goes from 0 to 2π. Our algorithm takes $\Theta(n \log n)$ time and $\Theta(n)$ space.

1 Introduction

Let D denote a collection of n closed planar disks. The convex hull of D, which we denote by $\mathcal{CH}(D)$, is the region obtained by removing from the plane all the open half-planes whose intersection with D is empty. The *rectilinear convex hull* of D, which we denote by $\mathcal{RCH}(D)$, is instead the region obtained by removing from the plane all the axis-aligned open wedges of aperture angle $\frac{\pi}{2}$, whose intersection with D is empty (a formal definition is given in Section 2). See Figure 1.

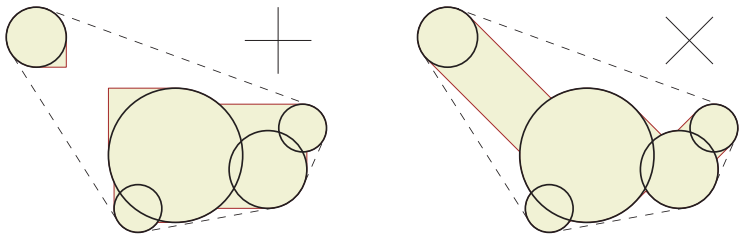


Figure 1: A set D of closed disks and $\mathcal{RCH}(D)$ for two orientations of the coordinate axes. The axes are shown in the top-right corner of each figure. The edges of $\mathcal{CH}(D)$ are shown with dashed lines. The interior and the edges of $\mathcal{RCH}(D)$ are shown respectively, in light and dark brown. On the left, $\mathcal{RCH}(D)$ has two connected components. On the right, $\mathcal{RCH}(D)$ has a single connected component.

The rectilinear convex hull is the analog of the (standard) convex hull on a non-traditional notion of convexity called *orthogonal convexity* [6]. In this notion of convexity, convex sets are restricted to those whose intersection with any horizontal or vertical line is either empty, a point, or a line segment. The rectilinear convex hull introduces two important differences with respect to the standard convex hull. On one hand, note that $\mathcal{RCH}(D)$ may be a simply connected set, yielding an intuitive and appealing structure. However, if the union of D is disconnected, then $\mathcal{RCH}(D)$ may have several simply connected components. On the other hand, observe that the orientation of the empty wedges changes along with the orientation of the coordinate axes, changing the shape of $\mathcal{RCH}(D)$ as well. The former property has shown to be useful to better separate finite sets of points that are not separable by a line or even by a standard convex hull [1]. The later property has been used to explore a family of problems in which the rotation angle of the coordinate axes is used as a search space for optimization criteria [2].

Let $\mathcal{RCH}_\theta(D)$ denote the rectilinear convex hull of D computed after simultaneously rotating the coordinate axes by an angle θ (a formal definition is given in Section 2). Let $\partial(S)$ denote the boundary of a planar set S. In this paper we describe an $\Theta(n \log n)$-time and $\Theta(n)$-space algorithm that enumerates the changes to $\partial(\mathcal{RCH}_\theta(D))$ while θ is increased from 0 to 2π. Notably, despite the introduction of rotations to the coordinate axes, our algorithm successfully achieves the complexities of the well known algorithm to compute the standard convex hull of a collection of planar disks [5].

2 Preliminaries

The *orientation* of a line is the smallest of the two possible angles it makes with the X^+ positive semiaxis. A *set of orientations* is a set of lines with different orientations passing through some fixed point. Hereafter we consider a set of orientations formed by two orthogonal lines. For the sake of simplicity, we assume that both lines are passing through the origin and are parallel to the coordinate axes. We denote such an orientation set with $\mathcal{O}$. We say that a planar region is $\mathcal{O}$*-convex*, if its intersection with a line parallel to a line of $\mathcal{O}$ is either empty, a point, or a line segment.

Let ρ_1 and ρ_2 be two rays with a common apex point $x \in \mathbb{R}^2$ such that, after rotating ρ_1 around x by an angle of $\omega \in [0, 2\pi)$, we obtain ρ_2. We refer to the two open regions in the set $\mathbb{R}^2 \setminus (\rho_1 \cup \rho_2)$ as *wedges*. We say that both wedges have vertex x and sizes ω and $2\pi - \omega$, respectively. A *quadrant* is a wedge of size $\frac{\pi}{2}$ whose rays are parallel to the lines of $\mathcal{O}$. We say that a planar region is *free of points of* D, or D*-free* for short, if it contains no point of a disk of D. Let $\mathcal{O}_\theta$ denote the set resulting after simultaneously rotating the lines of $\mathcal{O}$ in the counterclockwise direction by an angle of θ. Let $\mathcal{Q}_\theta$ be the set of all (open) D-free quadrants of the plane whose rays are parallel to the lines of $\mathcal{O}_\theta$.

Definition 1. *The rectilinear convex hull of* D with respect to $\mathcal{O}_\theta$*, is the closed and* $\mathcal{O}_\theta$*-convex set*

$$\mathcal{RCH}_\theta(D) = \mathbb{R}^2 \setminus \bigcup_{q \in \mathcal{Q}_\theta} q.$$

We assume that $\mathcal{O}_0 = \mathcal{O}$ *and* $\mathcal{RCH}_0(D) = \mathcal{RCH}(D)$.

As mentioned in Section 1, there are two main differences between the standard and the rectilinear convex hull. First, for any fixed value of θ we have that $\mathcal{RCH}_\theta(D)$ may be non-convex and may even be formed by several connected components; see again Figure 1. Each component is closed, simply connected, $\mathcal{O}_\theta$-convex, and is either a disk (if D is a singleton) or a region bounded by a *curvilinear polygon*, which is a simple polygon whose edges are either line segments or circular arcs. If an edge is a line segment, then it belongs to the boundary of a D-free quadrant. If it is instead a circular arc, then it belongs to the boundary of a disk of D. The second difference is a property we call *orientation dependency*: except for particular values of an angle α, such as multiples of $\frac{\pi}{2}$, we have that $\mathcal{RCH}_\theta(D) \neq \mathcal{RCH}_{\theta+\alpha}(D)$, $\alpha \in [0, 2\pi)$.

From an algorithmic point of view, $\mathcal{CH}(D)$ is described by a circular list of (possibly repeated) disks of D, sorted by appearance as we traverse $\partial(\mathcal{CH}(D))$ in counterclockwise direction. From this list we can trivially obtain $\partial(\mathcal{CH}(D))$ in linear time; see [5] for more details. We describe $\partial(\mathcal{RCH}_\theta(D))$ for fixed values of θ in a similar way. Instead of a single list, we use four disjoint lists containing each a set of circular arcs of the disks of D, sorted by appearance as we traverse $\partial(\mathcal{RCH}_\theta(D))$. To formally describe these lists, we use a simple (yet crucial) observation that derives from Definition 1. An ω-wedge is a wedge of size at least ω. We say a point $x \in \mathbb{R}^2$ is ω*-wedge* D*-free*, if there exists a D-free ω-wedge with apex at x.

Observation 2. *Consider a fixed value of* θ*. A point* x *of (a disk of)* D *lies on the boundary of* $\mathcal{RCH}_\theta(D)$ *if, and only if, it is* $\frac{\pi}{2}$*-wedge* D*-free.*

The *N-orientation* (for North-orientation) through a point $x \in \mathbb{R}^2$ is the ray pointing upwards starting at x. The *S-orientation*, *E-orientation*, and *W-orientation* through a point are defined analogously. A

point $x \in \mathbb{R}^2$ is *ω-wedge D-free* with respect to the N-orientation, if there is a D-free ω-wedge with apex at x that contains the N-orientation through x. The same definition can be analogously given for the remaining three orientations.

When computing $\partial(\mathcal{RCH}_\theta(D))$, there are two types of changes as θ is increased from 0 to 2π: *combinatorial changes*, where there is a change on the ordered list of circular arcs of D on $\partial(\mathcal{RCH}_\theta(D))$; and *geometric changes*, where there is no change on the list, but only on the coordinates of the endpoints of circular arcs. Geometric changes between two combinatorial changes can be accounted for by representing the circular arc endpoints as (known) functions of θ, instead of fixed values. We thus focus on computing combinatorial changes.

3 Rectilinear convex hull of a set of disks

Let $\mathcal{N}(D)$ denote the *upper envelope* of D, that is, the set of points of D seen from the north infinity. Analogously, let $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$ denote respectively, the envelopes of D seen from the South, East, and West infinities. To compute and maintain $\partial(\mathcal{RCH}_\theta(D))$ we proceed as follows. We first compute $\mathcal{N}(D)$, $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$. From the four envelopes we compute the points of D that are $\frac{\pi}{2}$-wedge D-free with respect to the N-, S-, W-, and E- orientations. We combine the information computed into a data structure to compute the four fronts for any value of θ. Then, we traverse this data structure increasing θ from 0 to 2π while computing all the combinatorial changes in $\partial(\mathcal{RCH}_\theta(D))$.

Computing the envelopes. The complexity of each envelope is $O(n)$ since the union of the disks of D has linear complexity [4]. Each envelope can be computed in $O(n \log n)$ time and $O(n)$ space [3].

The set of points that are $\frac{\pi}{2}$-wedge D-free. Consider in the following the envelope $\mathcal{N}(D)$; see Figure 2. The remaining three envelopes are similarly processed. The envelope is formed by a sequence of x-monotone circular arcs sorted by appearance while sweeping the plane from left to right. Two consecutive arcs may share and endpoint, and no vertical line intersects the interior of two arcs.

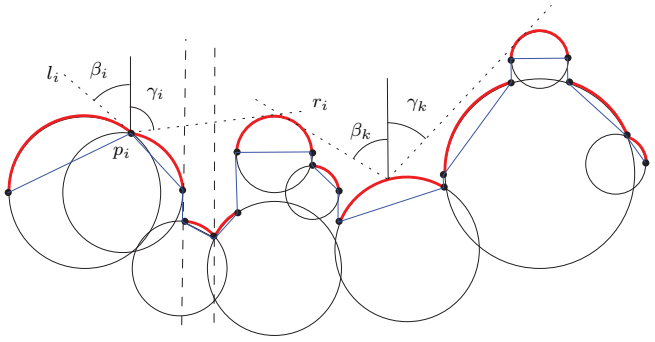


Figure 2: The upper envelope $\mathcal{N}(D)$ of a set of circles.

We first show how to decide which endpoints of the arcs of $\mathcal{N}(D)$ are $\frac{\pi}{2}$-wedge D-free. Let a_i denote the ith arc of $\mathcal{N}(D)$, and p_i, p_{i+1} denote the endpoints of a_i. For each endpoint p_i, we compute the wedge with apex at p_i that contains the N-orientation through p_i, and has the biggest possible size ω_i. We keep the endpoints for which $w_i \geq \frac{\pi}{2}$, as well as the corresponding w_i-wedge. Let $\leq$ denote the weak order on the endpoints induced by the values of their abscissas. We proceed as follows.

1. Sweep $\mathcal{N}(D)$ from left to right doing the following while visiting each endpoint p_i. Let D_i denote the subset of $\mathcal{N}(D)$ whose arcs have endpoints p_j such that (i) $p_j \leq p_i$ if p_i is a right endpoint of an arc, and (ii) $p_j < p_i$ otherwise. Update $\mathcal{CH}(D_i)$ in $O(\log n)$ time, and add p_i to D_i.

Notice that we have to update $\mathcal{CH}(D_i)$ from $\mathcal{CH}(D_{i-1})$ by either: (i) adding an arc $a_i = (p_{i-1}, p_i)$ to $\mathcal{CH}(D_{i-1})$ and computing the corresponding bridge, or (ii) adding a point p_i to $\mathcal{CH}(D_{i-1})$ by computing the supporting line to $\mathcal{CH}(D_{i-1})$ from p_i.

We do this computation in $O(n)$ overall amortized time since we use the order of the arcs in $\mathcal{N}(D)$ and walk along $\mathcal{N}(D)$ till we find the arcs to define the bridge contained in the supporting line. Once we find the arcs, the bridge can be computed in constant time using elementary geometry.

2. Compute the supporting line l_i from p_i to $\mathcal{CH}(D_{i-1})$ by traversing the boundary of $\mathcal{CH}(D_{i-1})$ until we find the tangent vertex of $\mathcal{CH}(D_{i-1})$. The tangent vertex can be either an endpoint in $\mathcal{N}(D)$ or a point in an arc in $\mathcal{N}(C)$, in the last case compute this point and the supporting line in constant time. At the end of the sweep, this process amortizes to $O(n)$ in time and space. We also compute the angle β_i formed by l_i and the line with N-orientation passing through p_i.

3. We can proceed doing the same computation but considering the right to left sorting, and maintaining D'_{i-1} which is defined symmetrically. Then, we compute the supporting line r_i from p_i to $\mathcal{CH}(D'_{i-1})$, and compute the angle γ_i formed by r_i and the line with N-orientation passing through p_i. Again, at the end of the sweep this process amortizes to $O(n)$ in time and space.

4. Compute $\omega_i = \beta_i + \gamma_i$, and check whether $\omega_i \geq \frac{\pi}{2}$. In the affirmative, let ω_i be the *angular interval* associated with p_i. For each p_i such that $\omega_i \geq \frac{\pi}{2}$, we form the angular interval for p_i as the intersection with the unit circle of the image of the wedge by mapping p_i to the origin.

Since there are $O(n)$ endpoints, the complexity of these steps is $O(n)$ time and space. We proceed analogously with the envelopes $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$. Thus, the total complexity for this process for the four envelopes is $O(n)$ time and space. We are considering the at most $O(n)$ endpoints p_i of the four envelopes and computing which of these endpoints p_i have $\omega_i \geq \frac{\pi}{2}$ for some of the four orientations above. For each such point and orientation, we record its angular interval (as it is defined a above in the case of the northern orientation). From the discussion above we have the following result.

Theorem 3. *The $O(n)$ endpoints p_i of the envelopes $\mathcal{N}(D)$, $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$ having an angle $\omega_i \geq \frac{\pi}{2}$, their angles ω_i, and their angular intervals can be computed in $O(n \log n)$ time and $O(n)$ space.*

We next show how to compute the interior points of the arcs of D that belong $\partial(\mathcal{RCH}_\theta(D))$. For each arc $a_i = (p_{i-1}, p_i)$ of $\mathcal{N}(D)$, we will show how to compute the angles β and γ for the interior points of a_i, and therefore, how to compute the parts of a_i (if any) whose interior points have angles β and γ such that $\beta + \gamma = \omega \geq \frac{\pi}{2}$.

Given an arc $a_i = (p_{i-1}, p_i)$ of $\mathcal{N}(D)$, we proceed as in the item 1. For the endpoint p_i of a_i we update in $O(\log n)$ time the $\mathcal{CH}(D_i)$, where D_i is the set of arcs in $\mathcal{N}(D)$ with the endpoints p_j such that (i) $p_j \leq p_i$ if p_i is a right endpoint of an arc, or (ii) $p_j < p_i$ otherwise adding p_i to D_i.

Proceeding with the endpoints of the arc $a_i = (p_{i-1}, p_i)$, assume that we have computed the left support line l_{i-1} and the right support line r_{i-1} from the endpoint p_{i-1}, and analogously, l_i and r_i from the endpoint p_i. Then, because we maintain the $\mathcal{CH}(D_i)$ and $\mathcal{CH}(D'_i)$ for a_i, we can split a_i into sub-arcs such that from the points inside each sub-arc we have a unique left (resp. right) supporting arc in $\mathcal{N}(D)$ for computing the respective supporting lines. In each sub-arc we know the left and right arcs that support the left (resp. right) supporting lines from the endpoints of a sub-arc in a_i. Next, we determine whether and where the points inside these sub-arcs having angle $w \geq \frac{\pi}{2}$. We parameterize the calculus of the angle ω in each sub-arc as a function of the angle θ that it is illustrated in Figure 3.

The blue curve in Figure 3 Right is a Limaçon curve, see Sánchez-Ramos et al. [7], also known as a Limaçon of Pascal or Pascal's Snail. The cardioid is a special case. We have drawn this curve for two circles in Figure 3 Right. The part of the Limaçon curve that we are interesting in is the part of the curve that is in between the two red circles, say c_j and c_k, and defined by the intersection point of the perpendicular lines which are tangents lines to c_j and c_k, as illustrated in Figure 3 Left.

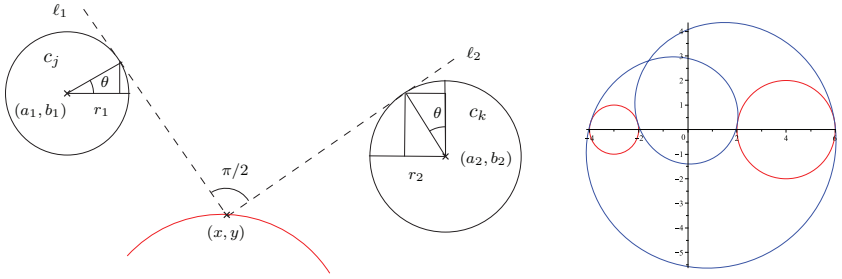


Figure 3: Left: the angle $\omega = \pi/2$ in a point (x, y) of a sub-arc a_i. Right: The Limaçon curve with angle $\pi/2$, in blue. The red circles contain the two arcs of the envelope $\mathcal{N}(D)$. The left circle has center $(-3, 0)$ and radius 1 and the right circle has center $(4, 0)$ and radius 2.

We compute the solutions of the equations and determine the intervals. In constant time, we can check whether the values of ω inside the computed intervals verify that $\omega \geq \frac{\pi}{2}$, and thus, we determine the constant number of intervals where $\omega \geq \frac{\pi}{2}$. We can proceed with the other envelopes $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$. Therefore, the total complexities for all together are $O(n \log n)$ time and $O(n)$ space.

Theorem 4. *The $O(n)$ intervals in the arcs of circles in the envelopes $\mathcal{N}(D)$, $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$ whose interior points have an angle $\omega \geq \frac{\pi}{2}$, their angles ω, and the corresponding angular intervals can be computed in $O(n \log n)$ time and $O(n)$ space.*

The data structure. We translate all the angular intervals for all the arcs of D in $\mathcal{N}(D)$, $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$, to angular intervals inside $[0, 2\pi]$ on the real line. In this way, we can do a line sweep with four vertical lines corresponding to angles θ, $\theta + \frac{\pi}{2}$, $\theta + \pi$, and $\theta + \frac{3\pi}{2}$ in a circular way (i.e., completing a $[0, 2\pi]$ round with each line), and then inserting and deleting the changes of the arcs (or part of them) that belong to $\partial(\mathcal{RCH}_\theta(D))$ as θ changes in $[0, 2\pi)$.

Now, considering that an endpoint or an interior point of an arc in any of the envelopes $\mathcal{N}(D)$, $\mathcal{S}(D)$, $\mathcal{E}(D)$, and $\mathcal{W}(D)$ can be the apex of a D-free $\frac{\pi}{2}$-wedge, from Theorems 3 and 4, we conclude that we can compute the points of disks of D that belong to $\partial(\mathcal{RCH}_\theta(D))$ as $\theta \in [0, 2\pi]$ in $O(n \log n)$ time and $O(n)$ space. From this discussion we obtain the main result of our paper.

Theorem 5. *Given a set D of n closed disks in the plane, computing and maintaining $\partial(\mathcal{RCH}_\theta(D))$ as θ is increased from 0 to 2π can be done in $O(n \log n)$ time and $O(n)$ space.*

References

[1] Alegría, C., Orden, D., Seara, C., Urrutia, J.: Separating bichromatic point sets in the plane by restricted orientation convex hulls. J. Global Optim. **85**, 1003–1036 (2023)

[2] Alegría-Galicia, C., Orden, D., Seara, C., Urrutia, J.: Efficient computation of minimum-area rectilinear convex hull under rotation and generalizations. J. Global Optim. **70**(3), 687–714 (2021)

[3] Devillers, O., Golin, M.J.: Incremental algorithms for finding the convex hulls of circles and the lower envelopes of parabolas. Information Processing Letters **56**(3), 157–164 (1995)

[4] Kedem, K., Livne, R., Pach, J., Sharir, M.: On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles. Discrete Comput. Geom. **1**(1), 59–71 (1986)

[5] Rappaport, D.: A convex hull algorithm for discs, and applications. Computational Geometry **1**(3), 171–187 (1992)

[6] Rawlins, G.J., Wood, D.: Ortho-convexity and its generalizations. In: Computational Morphology, Machine Intelligence and Pattern Recognition, vol. 6, pp. 137–152. North-Holland (1988)

[7] Sánchez Ramos, I., Meseguer-Garrido, F., Aliaga, J.J., Grau, J.: Generalization of the pedal concept in bidimensional spaces. Application to the limaçon of Pascal. DYNA **88**, 196–202 (2021)

Characterization of the equality in some discrete isoperimetric and Brunn-Minkowski type inequalities*

David Iglesias[†1] and Eduardo Lucas[‡2]

[1]Departamento de Matemáticas, Universidad de Murcia, Campus de Espinardo, 30100 Murcia, Spain
[2]Departamento de Ciencias, Centro Universitario de la Defensa, Universidad Politécnica de Cartagena, 30720 Santiago de la Ribera, Murcia, Spain

Abstract

Lattice cubes are optimal sets with respect to several recent inequalities in Discrete Geometry, including analogues - for the cardinality - of both the Brunn-Minkowski inequality and the isoperimetric inequality. While a general characterization of the equality case has not been obtained thus far, we show that when the cardinality is suitable lattice cubes do comprise the unique optimal sets with respect to the aforementioned discrete inequalities, and discuss the underlying techniques of the proof.

1 Introduction

The isoperimetric inequality is one of the most classical results in Geometry, dating back to antiquity in the case of the planar version. Informally, it asserts that Euclidean balls minimize the ratio between the surface area and the volume. More rigorously, given a bounded set with non-empty interior $M \subset \mathbb{R}^n$ and the unit Euclidean ball of dimension n, B_n, one has

$$\frac{\mathcal{S}(M)^n}{\mathrm{vol}(M)^{n-1}} \geq \frac{S(B_n)^n}{\mathrm{vol}(B_n)^{n-1}}, \tag{1}$$

where $\mathrm{vol}(\cdot)$ is the volume (Lebesgue measure) and $\mathcal{S}(\cdot)$ is the surface area measure. Equivalently, since $\mathcal{S}(B_n) = n\,\mathrm{vol}(B_n)$, one has

$$\mathcal{S}(M) \geq n\,\mathrm{vol}(M)^{1-\frac{1}{n}}\,\mathrm{vol}(B_n)^{\frac{1}{n}}.$$

If one restricts the inequality to convex sets only, then Euclidean balls characterize inequality (1).

On the other hand, the classic Brunn-Minkowski inequality provides a relationship between the volume and the addition of sets. Namely, given two non-empty compact sets $K, L \subset \mathbb{R}^n$, it states that

$$\mathrm{vol}(K+L)^{1/n} \geq \mathrm{vol}(K)^{1/n} + \mathrm{vol}(L)^{1/n}, \tag{2}$$

with equality if and only if K and L are either homothetic or lie in parallel hyperplanes. Here, the addition being employed is the standard pointwise addition - or *Minkowski addition* - of sets.

In other words, the functional $\mathrm{vol}(\cdot)^{1/n}$ is concave in the family of non-empty compact sets. In fact, the result holds true even when K, L and $K+L$ are just Lebesgue measurable.

*The full version of this work was published in *Electron. J. Combin.* in 2023 [15]. This research is supported by "Comunidad Autónoma de la Región de Murcia a través de la convocatoria de Ayudas a proyectos para el desarrollo de investigación científica y técnica por grupos competitivos, incluida en el Programa Regional de Fomento de la Investigación Científica y Técnica (Plan de Actuación 2022) de la Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia, REF. 21899/PI/22".

[†]Email: david.iglesias@um.es.
[‡]Email: eduardo.lucas@cud.upct.es.

The Brunn-Minkowski inequality gave way to the development of a rich theory of related results and generalizations to other contexts, such as functional analogues (e.g. [6, 7, 21, 22]); extensions for other measures, like the mixed volumes (see [1]); a reverse form [20]; or applications to the concentration of measure (see, e.g., [2]), among others. Notably, the isoperimetric inequality follows easily from the Brunn-Minkowski inequality (one may also derive the isodiametric inequality, and even its stronger counterpart for the mean width, Urysohn's inequality). For a thorough treatment of the Brunn-Minkowski inequality we refer to the comprehensive surveys [3, 9] and the excellent monographs [12, 26].

In recent decades, there has been considerable effort translating some of these results, and many others in the context of Convex Geometry, to the discrete setting. For the isoperimetric inequality, see e.g. [4, 5, 8, 16, 23, 28]; for the Brunn-Minkowski inequality, see e.g. [10, 11, 13, 14, 16, 17, 18, 19, 24, 25, 27]. Most commonly, one either considers finite subsets of lattices equipped with the cardinality measure, or regular bounded sets equipped with the lattice point enumerator measure. In this work we will focus on the former approach.

In the next section we will highlight a few of the inequalities discussed above, for which the announced characterizations have been obtained. In the final section we will provide a general outlook of the proof and its underlying ideas.

2 Preliminaries and overview

In [23], building on previous ideas from [5, 28], the authors obtained an isoperimetric inequality for the cardinality in the setting of $\mathbb{Z}^n$ and $\mathbb{N}^n$. In order to introduce it, we first need to recall their notion of boundary of a set in this context (see also Figure 1)

Definition 1. *The boundary of a discrete set $A \subset \mathbb{Z}^n$ is $\partial(A) = (A + \{-1, 0, 1\}^n) \setminus A$.*

Note that this setting can be interpreted as considering the lattice endowed with the L_∞ norm. The isoperimetric problem for the cardinality can then be reformulated as finding the sets with a given fixed cardinality which minimize $|\partial(\cdot)|$. By construction, this is equivalent to simply finding the minimizers for the functional $|A + \{-1, 0, 1\}^n|$.

In order to describe said minimizers, the authors defined a complete order in $\mathbb{Z}^n$. They then showed that the *initial segments* $\mathcal{I}_r \subset \mathbb{Z}^n$ (i.e., the first r points in the order, $r \in \mathbb{N}$; see Figure 2) constitute an infinite family of minimizers, one for each cardinality.

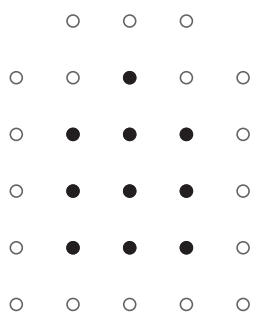


Figure 1: Boundary of a discrete set.

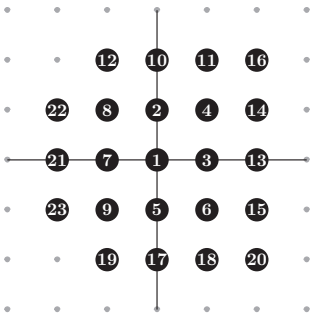


Figure 2: The set $\mathcal{I}_{23}$ with the ordering specified.

Observe that, due to them being initial segments, they are additionally all nested. We will not delve into the details of this order here, for which we instead refer to [23], as well as to [5, 10, 28], which use a similar approach with an order known as the *simplicial order*. As an example, the set from Figure 1 is $\mathcal{I}_{10}$, and the origin is $\mathcal{I}_1$. The authors obtained the following result:

Theorem 2. *[23, Theorem 1] Let $A \subset \mathbb{Z}^n$ with $r = |A| > 0$. Then*

$$\left|A + \{-1, 0, 1\}^n\right| \geq \left|\mathcal{I}_r + \{-1, 0, 1\}^n\right|. \tag{3}$$

They also considered the restriction of the order to $\mathbb{N}^n$, which gives rise to the corresponding initial segments $\mathcal{J}_r \subset \mathbb{N}^n$ of cardinality r, and derived the following analogous result:

Theorem 3. *[23, Corollary 1] Let $A \subset \mathbb{N}^n$ with $r = |A| > 0$. Then*

$$\left|\left(A + \{-1, 0, 1\}^n\right) \cap \mathbb{N}^n\right| \geq \left|\left(\mathcal{J}_r + \{-1, 0, 1\}^n\right) \cap \mathbb{N}^n\right|. \tag{4}$$

The authors noted, however, that these sets do not characterize the equality. Indeed, in the set from Figure 1, one may translate the outermost point one unit to the left or to the right, and the cardinality of the boundary will remain constant, despite these new sets not being initial segments.

It is relevant to mention now that both $\mathcal{I}_{\rho^n}$ and $\mathcal{J}_{\rho^n}$, for $\rho \in \mathbb{N}, \rho > 0$, are lattice cubes with side length ρ. Therefore, all lattice cubes are minimizers of this problem. Moreover, due to the nestedness, it follows that any initial segment is contained between two consecutive lattice cubes, that is, it is composed by a lattice cube and some additional points in its outermost layer (again, cf. Figure 1). As before, it is easy to find additional minimizers (which are not initial segments) by shifting points in this outermost layer. Even though other configurations do exist, this nevertheless suggests to study the special case of lattice cubes.

We show that, indeed, lattice cubes characterize the equality on both Theorem 2 and Theorem 3, whenever the cardinality involved is suitable:

Theorem 4. *[15, Theorem 1] Let $A \subset \mathbb{Z}^n$ with $|A| = (\rho + 1)^n$ for some $\rho \in \mathbb{N}$. Then equality holds in (3) if and only if A is a lattice cube.*

Theorem 5. *[15, Theorem 2] Let $A \subset \mathbb{N}^n$ with $|A| = (\rho + 1)^n$ for some $\rho \in \mathbb{N}$. Then equality holds in (4) if and only if $A = \{0, \ldots, \rho\}^n$.*

It can be shown that, in fact, both settings ($\mathbb{Z}^n$ and $\mathbb{N}^n$) are equivalent, and therefore, it suffices to focus on one of them. The above results both follow from the more general result below:

Theorem 6. *[15, Theorem 17] Let $A \subset \mathbb{N}^n$ with $|A| = (\rho+1)^n$ for some $\rho \in \mathbb{N}$ and let $s \in \mathbb{N}, s > 0$. If*

$$\left|A + \{0, \ldots, s\}^n\right| = \left|\mathcal{J}_{(\rho+1)^n} + \{0, \ldots, s\}^n\right|,$$

then A is a lattice cube.

Though we will not delve into it, the above result can also be applied in the context of the lattice point enumerator. In particular, it implies a characterization of the equality case in a discrete isoperimetric inequality for the lattice point enumerator recently obtained (see [15, Theorem 3]).

As for the Brunn-Minkowski inequality, we highlight a discrete analogue for the cardinality obtained in [18]. The authors proved the following inequality:

Theorem 7. *[18, Theorem 3.2] For every non-empty finite sets $A, B \subset \mathbb{Z}^n$,*

$$\left|A + B + \{0, 1\}^n\right|^{1/n} \geq |A|^{1/n} + |B|^{1/n}. \tag{5}$$

The inequality is sharp (e.g., for lattice cubes), and in fact, the authors showed that it is equivalent to the classic version for the volume (2).

One may wonder whether lattice cubes also characterize the above inequality when the cardinalities are suitable. While this is not true (see Figure 3), if we also know that one of the sets is a lattice cube, then the equality does imply that the other set must be a lattice cube as well. More specifically, as a consequence of Theorem 6, we obtain the following characterization of Theorem 7:

Corollary 8. *[15, Corollary 35] Let $A \subset \mathbb{Z}^n$ be a finite set with $|A| = (\rho + 1)^n$ for some $\rho \in \mathbb{N}$ and let B be a lattice cube. Then*

$$\left|A + B + \{0, 1\}^n\right|^{1/n} = |A|^{1/n} + |B|^{1/n}$$

if and only if A is a lattice cube.

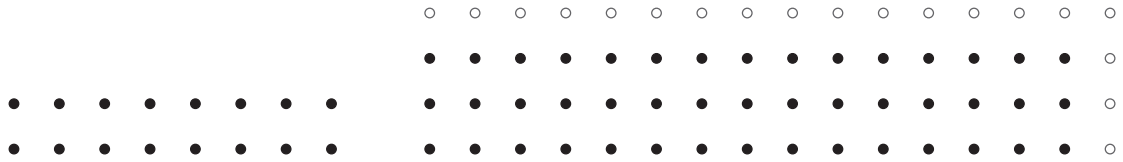


Figure 3: The set $A = \{0, \dots, 7\} \times \{0, 1\} \subset \mathbb{Z}^2$ (left) satisfies $|A| = 16$, and $A + A + \{0, 1\}^2$ (right) satisfies the equality in (5): $|A + A + \{0, 1\}^2|^{1/2} = 8 = 2|A|^{1/2}$.

3 Sketch of the proof

The overall idea is a refinement of the approach used in [23] to prove Theorems 2 and 3. As discussed, we can restrict our attention to $\mathbb{N}^n$, since the results follow for $\mathbb{Z}^n$. First, we define a transformation on $\mathbb{N}^n$ we shall refer to as *normalization*.

To normalize a given set $A \subset \mathbb{N}^n$ in the direction of the canonical vector e_i, $i = 1, \dots, n$, we consider the non-empty $(n-1)$-dimensional sections of A orthogonal to e_i and we perform the following operations (see Figure 4), which we proceed to describe informally:

1. Replace each section by the $(n-1)$-dimensional initial segment of the same cardinality.
2. Rearrange the sections in decreasing order of cardinality, such that the largest section is at the origin.
3. Translate points from the upper sections to the lower ones in a suitable way that preserves the first two properties, that is, the sections will remain initial segments ordered by cardinality.

The mathematical details missing in the above description, specially those pertaining to the third and final step, are rather technical, and therefore, we instead refer the reader to [15, Definition 28] for the precise and rigorous description.

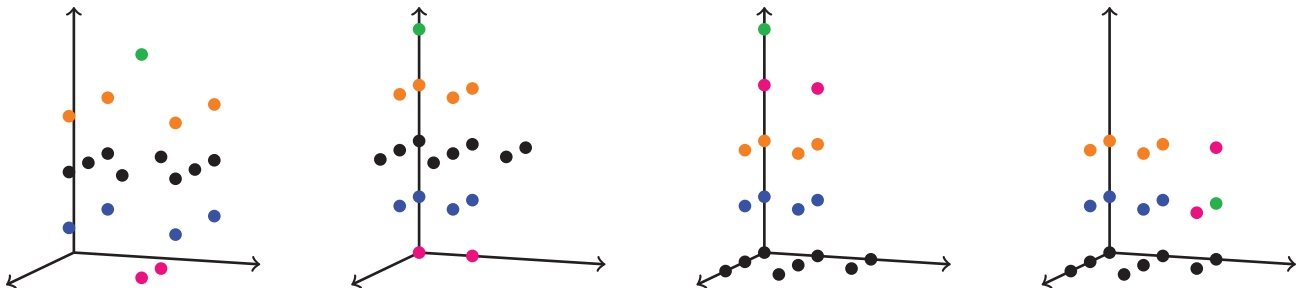


Figure 4: From left to right: a finite set, together with the same set after each step of the normalization is applied.

We will say that a set is normalized with respect to e_i if it coincides with its normalization in said direction. If one normalizes a set with respect to e_i and then with respect to e_j, for some $i, j = 1, \dots, n$, $i \neq j$, it is not guaranteed that the result will remain normalized with respect to e_i. However, it can be proved that the process is finite, in the sense that by repeatedly normalizing a set with respect to all (canonical) directions one eventually reaches a set which is normalized with respect to e_i for all $i = 1, \dots, n$ simultaneously (see [15, Lemma 30]).

Definition 9. *A set $A \subset \mathbb{N}^n$ is stable if it is normalized with respect to e_i for all $i = 1, \dots, n$ simultaneously.*

Moreover, we show that the normalization process does not increase the cardinality of the boundary (see [15, Lemma 31]), and therefore, in particular, the normalization of a minimizer is still a minimizer.

The approach then consists on reducing the study to stable minimizers, since any arbitrary minimizer can be transformed into a stable one via a finite number of normalizations.

We prove that any stable minimizer is a lattice cube (see [15, Lemma 34]). Finally, and crucially, in the proof of the main theorem it is shown that any minimizer which is normalized into a lattice cube must be a lattice cube itself. Putting everything together yields Theorem 6, and thus Theorems 4 and 5. We finish with two additional observations regarding the proofs of the results stated above.

First, the proof of the fact that any minimizer which normalizes into a lattice cube must be a lattice cube itself relies on induction on the dimension, which is possible since the $(n-1)$-dimensional sections of a minimizer are minimizers in $\mathbb{N}^{n-1}$ (see [15, Corollary 26]).

Second, the normalization process is so restrictive that a very explicit and precise description for stable sets can be obtained. The proof of the fact that stable minimizers are lattice cubes heavily depends on this description, which allows to perform direct computations on the stable set.

Namely, any stable set can be decomposed as the disjoint union of a lattice box and two additional, $(n-1)$-dimensional sets, as described in the following lemma (see also Figure 5).

Lemma 10. *Let $n \geq 3$, $\rho \geq 1$ and let $X \subset \mathbb{N}^n$ be a non-empty finite set with $|X| = (\rho+1)^n$. If X is stable, then there exist $A, B \subset \mathbb{N}^n$ such that*

$$A \subset \{0, \dots, \rho-1\}^{n-1} \times \{\rho+1\}, \qquad \emptyset \neq B \subset \{\rho\} \times \{0, \dots, \rho\}^{n-1}$$

and

$$X = A \cup B \cup \left(\{0, \dots, \rho-1\} \times \{0, \dots, \rho\}^{n-1}\right).$$

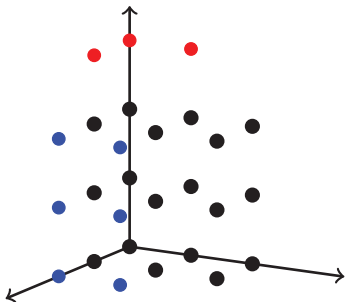


Figure 5: A stable set $X \subset \mathbb{N}^3$, highlighting the three subsets from the decomposition of Lemma 10.

Regarding the discrete Brunn-Minkowski analogue for the cardinality, we finish by obtaining Corollary 8 as a quick consequence of Theorem 6.

Proof of Corollary 8. Assume that $\left|A + B + \{0,1\}^n\right|^{1/n} = |A|^{1/n} + |B|^{1/n}$, and let $B = \{0, \dots, s\}^n$ for some $s \in \mathbb{N}$. Then, by the minimality of initial segments (see also [15, Corollary 11]), we have

$$(\rho+s+2)^n = |A + B + \{0,1\}^n| \geq |\mathcal{J}_{(\rho+1)^n} + B + \{0,1\}^n| = (\rho+s+2)^n.$$

Thus, $|A + \{0, \dots, s+1\}^n| = |\mathcal{J}_{(\rho+1)^n} + \{0, \dots, s+1\}^n|$, and Theorem 6 implies A must be a lattice cube, as desired. □

References

[1] A. D. Aleksandrov, *Selected Works. Part I: Selected Scientific Papers.* (Yu. G. Reshetnyak, S. S. Kutateladze, eds), trans. from the Russian by P. S. Naidu., Classics of Soviet Mathematics **4**, Gordon and Breach, Amsterdam, 1996.

[2] K. M. Ball, *An elementary introduction to modern convex geometry*, Cambridge University Press, Cambridge, 1997.

[3] F. Barthe, Autour de l'inégalité de Brunn-Minkowski, *Ann. Fac. Sci. Toulouse Math. (6)* **12** (2) (2003), 127–178.

[4] S. L. Bezrukov, Isoperimetric problems in discrete spaces. In: *Extremal problems for finite sets* (P. Frankl, Z. Füredi, G. Katona, D. Miklos eds.), J. Bolyai Soc. Math. Stud. 3, Budapest, 1994, 59–91.

[5] B. Bollobás, I. Leader, Compressions and isoperimetric inequalities, *J. Comb. Theory A* **56** (1991), 47–62.

[6] C. Borell, Convex set functions in d-space, *Period. Math. Hungar.* **6** (2) (1975), 111–136.

[7] H. J. Brascamp, E. H. Lieb, On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation, *J. Funct. Anal.* **22** (4) (1976), 366–389.

[8] A. Dall, F. von Heymann, and B. Vogtenhuber, Sets with small neighborhood in the integer lattice. In M. Noy and J. Pfeifle, editors, DocCourse combinatorics and geometry 2009: discrete and computational geometry, volume 5.3 of CRM Documents, 163–181. Centre de Recerca Matemàtica, 2010.

[9] R. J. Gardner, The Brunn-Minkowski inequality, *Bull. Amer. Math. Soc.* **39** (3) (2002), 355–405.

[10] R. J. Gardner and P. Gronchi, A Brunn-Minkowski inequality for the integer lattice, *Trans. Amer. Math. Soc.* **353** (10) (2001), 3995–4024.

[11] B. Green and T. Tao, Compressions, convex geometry and the Freiman-Bilu theorem, *Q. J. Math.* **57** (4) (2006), 495–504.

[12] P. M. Gruber, *Convex and Discrete Geometry*, Springer, Berlin Heidelberg, 2007.

[13] D. Halikias, B. Klartag and B. A. Slomka, Discrete variants of Brunn-Minkowski type inequalities, To appear in *Ann. Fac. Sci. Toulouse Math.*

[14] M. A. Hernández Cifre, D. Iglesias and J. Yepes Nicolás, On a discrete Brunn-Minkowski type inequality, *SIAM J. Discrete Math.* **32** (2018), 1840–1856.

[15] D. Iglesias and E. Lucas, On a characterization of lattice cubes via discrete isoperimetric inequalities, *Electron. J. Combin.* **30(1)** (2023), #P1.21.

[16] D. Iglesias, E. Lucas and J. Yepes Nicolás, On discrete Brunn-Minkowski and isoperimetric type inequalities, *Discrete Math.* **345** (1) (2022), 112640.

[17] D. Iglesias and J. Yepes Nicolás, On discrete Borell-Brascamp-Lieb inequalities, *Rev. Mat. Iberoam.* **36** (3) (2020), 711–722.

[18] D. Iglesias, J. Yepes Nicolás and A. Zvavitch, Brunn-Minkowski type inequalities for the lattice point enumerator, *Adv. Math.* **370** (2020), 107193.

[19] B. Klartag and J. Lehec, Poisson processes and a log-concave Bernstein theorem, *Stud. Math.* **247** (1) (2019), 85–107.

[20] V. D. Milman, Inégalité de Brunn-Minkowski inverse et applications à la théorie locale des espaces normés (An inverse form of the Brunn-Minkowski inequality, with applications to the local theory of normed spaces), *C. R. Acad. Sci. Paris Ser. I Math.* **302** (1) (1986), 25–28.

[21] A. Prékopa, Logarithmic concave measures with application to stochastic programming, *Acta Sci. Math.* **32** (1971), 301–316.

[22] A. Prékopa, On logarithmic concave measures and functions, *Acta Sci. Math.* **34** (1973), 335–343.

[23] A. J. Radcliffe and E. Veomett, Vertex isoperimetric inequalities for a family of graphs on $\mathbb{Z}^k$. *Electron. J. Combin.* **19** (2) (2012), P45.

[24] I. Z. Ruzsa, Sum of sets in several dimensions, *Combinatorica* **14** (1994), 485–490.

[25] I. Z. Ruzsa, Sets of sums and commutative graphs, *Studia Sci. Math. Hungar.* **30** (1995), 127–148.

[26] R. Schneider, *Convex bodies: The Brunn-Minkowski theory*, 2nd expanded ed. Encyclopedia of Mathematics and its Applications **151**, Cambridge University Press, Cambridge, 2014.

[27] B. A. Slomka, A remark on discrete Brunn-Minkowski type inequalities via transportation of measure, *Unpublished*, `arXiv:2008.00738`.

[28] D. Wang and P. Wang, Discrete Isoperimetric Problems, *SIAM J. Appl. Math.*, **32** (4) (1977), 860–870.

Totally Greedy Sequences Generated by a Class of Second-Order Linear Recurrences With Constant Coefficients

Hebert Pérez-Rosés*[1]

[1]Dept. of Computer Science and Mathematics, Univ. Rovira i Virgili, Tarragona, Spain

Abstract

In the change-making problem the goal is to represent a certain amount of money with the least possible number of coins, chosen from a given set of denominations. The greedy algorithm picks the coin of largest possible denomination first. This strategy does not always produce the least number of coins, except when the set of denominations is endowed with certain properties, in which case it is called a greedy set. If the set of denominations is an infinite sequence, we call it totally greedy if every prefix subset is greedy. This paper investigates some totally greedy sequences generated by second-order linear recurrences with constant coefficients. In particular it investigates sufficient conditions for the sequence to be totally greedy.

1 Greedy sets and totally greedy sequences

In the *change-making problem* we are given a set of coin denominations $S = \{s_1 = 1, s_2, \ldots, s_t\}$, with $s_1 < \ldots < s_t$, and a target amount $k \in \mathbb{N}_0$ (where $\mathbb{N}_0$ denotes the set of nonnegative integers). The goal is to represent k using as few coins as possible from the given denominations. Mathematically, we are looking for a *payment vector* $(a_1, \ldots, a_t)$, such that: 1. $a_i \in \mathbb{N}_0$ for all $i = 1, \ldots, t$, 2. $\sum_{i=1}^{t} a_i s_i = k$, and 3. $\sum_{i=1}^{t} a_i$ is minimal.

This problem has been extensively studied in recent years (see for instance [1, 2, 3, 11]), and it is related to other problems involving integers, such as the *Frobenius problem* and the *postage stamp problem* [10]. It is also a special case of the well known *knapsack problem* [5]. Regarding its computational complexity, finding the optimal payment vector for a given k is NP-hard if the coins are large and represented in binary (or decimal) [4].

A simple approach for dealing with the problem is the *greedy algorithm*, which proceeds by first choosing the coin of the largest possible denomination, then the second largest, and so on. This idea is formalized in Algorithm 1:

Algorithm 1: GREEDY PAYMENT METHOD

Input : The set of denominations $S = 1, s_2, \ldots, s_t$, with $1 < s_2 < \ldots < s_t$, and a quantity $k \geq 0$.
Output: Payment vector $(a_1, a_2, \ldots, a_t)$.

```
for i:= t downto 1 do
    a_i := k div s_i;
    k := k mod s_i;
end
```

*Email: hebert.perez@urv.cat

Definition 1. *For a given set of denominations $S = 1, s_2, \ldots, s_t$, the* greedy payment vector *is the payment vector $(a_1, a_2, \ldots, a_t)$ produced by Algorithm 1, and* $\textsc{GreedyCost}_S(k) = \sum_{i=1}^{t} a_i$.

It is easy to verify that the greedy payment vector is not necessarily optimal (i.e. $\textsc{GreedyCost}_S(k)$ is not always minimal among all possible payment vectors) but there do exist some specific sets of denominations S for which the greedy payment vector is always optimal:

Definition 2. *If a set S of denominations* always *produces an optimal payment vector for* any *given amount k, then S is called* orderly, canonical, *or* greedy.

Greedy sets can be used, for instance, to construct circulant networks with efficient routing algorithms [8]. There is a polynomial-time algorithm that determines whether a given set of denominations is greedy [7, 10], as well as necessary and/or sufficient conditions for special families of denomination sets [1, 2, 3, 11]. The current paper continues along that path.

Note that a set S consisting of one or two denominations is always greedy. For sets of cardinal 3 we have the following [1]:

Proposition 3. *The set $S = \{1, a, b\}$ (with $a < b$) is greedy if, and only if, $b - a$ belongs to the set*

$$\mathfrak{D}(a) = \{a-1, a\} \cup \{2a-2, 2a-1, 2a\} \cup \ldots \{ma-m, \ldots ma\} \cup \ldots = \bigcup_{m=1}^{\infty} \bigcup_{s=0}^{m} \{ma - s\}$$

□

The *one-point theorem* provides a powerful necessary and sufficient condition (Theorem 2.1 [1]):

Theorem 4. *Suppose that $S = \{1, s_2, \ldots, s_t\}$ is a greedy set of denominations, and $s_{t+1} > s_t$. Now let $m = \left\lceil \frac{s_{t+1}}{s_t} \right\rceil$. Then $\hat{S} = \{1, s_2, \ldots, s_t, s_{t+1}\}$ is greedy if, and only if,* $\textsc{GreedyCost}_S(ms_t - s_{t+1}) < m$.
□

Notice that

$$(m-1)s_t + 1 \leq s_{t+1} \leq ms_t,$$

by the definition of m. A straightforward consequence of the one-point theorem is the following

Corollary 5. *[Lemma 7.4 of [1]] Suppose that $S = \{1, s_2, \ldots, s_t\}$ is a greedy set, and $s_{t+1} = us_t$, for some $u \in \mathbb{N}$. Then $\hat{S} = \{1, s_2, \ldots, s_t, s_{t+1}\}$ is also greedy.* □

Definition 6. *A set $S = \{1, s_2, \ldots, s_t\}$ is* totally greedy[1] *if every prefix subset $\{1, s_2, \ldots, s_k\}$, with $k \leq t$ is greedy.*

Obviously, a totally greedy set is also greedy, but the converse is not true in general. Take, for instance, the greedy set $\{1, 2, 5, 6, 10\}$, whose prefix subset $\{1, 2, 5, 6\}$ is not greedy.

Definition 6 can be extended to infinite sequences in a straightforward way:

Definition 7. *Let $S = \{s_n\}_{n=1}^{\infty}$ be an integer sequence, with $s_1 = 1$ and $s_i < s_{i+1}$ for all $i \in \mathbb{N}$. We say that S is* totally greedy *(or simply, greedy) if every prefix subset $\{1, s_2, \ldots, s_k\}$ is greedy.*

Totally greedy sequences are briefly mentioned in [3], where some sufficient conditions are also given, that allow to construct greedy sequences from recurrences, although the conditions are a bit cumbersome (see Corollary 2.12 of [3]). Here we provide a simpler set of sufficient conditions that produce greedy sequences from second-order homogeneous recurrences.

[1] Also called normal, or totally orderly.

2 Sequences of the form $G_{n+2} = pG_{n+1} + qG_n$

We will consider sequences $\{G_n\}_{n=1}^{\infty}$ generated by the recurrence

$$G_n = \begin{cases} 1 \text{ if } n = 1, \\ a \text{ if } n = 2, \\ pG_{n-1} + qG_{n-2}, \text{ if } n > 2, \end{cases} \tag{1}$$

where a, p, q are positive integers, with $a > 1$, and some additional restrictions that we will see later on.

The (shifted) Fibonacci sequence $\{F_n\}_{n=1}^{\infty}$, defined by $F_0 = 0$, $F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$, is a special case of Equation 1, taking $a = p = q = 1$. Similarly, the Lucas numbers (Sequence A000032 of [6]) and the Pell numbers (Sequence A000129 of [6]) are also special cases of Equation 1.

The characteristic polynomial associated with Equation 1 is $x^2 - px - q$, and its roots are

$$\lambda = \frac{1}{2}\left(p + \sqrt{p^2 + 4q}\right), \qquad \mu = \frac{1}{2}\left(p - \sqrt{p^2 + 4q}\right), \tag{2}$$

with $\mu + \lambda = p$ and $\mu\lambda = -q$. Since the roots λ and μ are real and distinct, the general term of $\{G_n\}_{n=1}^{\infty}$ is

$$G_{n+1} = c_1\lambda^n + c_2\mu^n, \tag{3}$$

where $c_1 = \dfrac{a - \mu}{\lambda - \mu}$ and $c_2 = \dfrac{\lambda - a}{\lambda - \mu}$.

It is quite easy to see that $\{G_n\}_{n=1}^{\infty}$ is monotonically increasing, $|\lambda| > |\mu|$, and $\lambda > 1$. Moreover, it can be easily shown that $\lambda > p$ and $\mu < 0$. Now, if $q \leq p$ we can bound the roots λ and μ with more precision.

Lemma 8. *If $\{G_n\}_{n=1}^{\infty}$ is a sequence defined by Equation 1, with $q \leq p$, and λ and μ are the roots of the characteristic polynomial, as defined in Equation 2, then*

$$-1 < \mu < 0 \quad \text{and} \quad p < \lambda < p + 1.$$

Proof: Straightforward. □

Note that as a consequence of the above results, c_1 is always positive, while c_2 can be positive or negative, depending on a. From now on, sequences that obey Equation 1, with $q \leq p$, will also be called *type-1-sequences*, and they will be the main focus of this section.

Now, in order to apply Theorem 4 we have to investigate the ratio

$$\frac{G_{n+1}}{G_n} = \frac{c_1\lambda^n + c_2\mu^n}{c_1\lambda^{n-1} + c_2\mu^{n-1}}, \tag{4}$$

where $\{G_n\}_{n=1}^{\infty}$ is a type-1-sequence.

Dividing the numerator and the denominator by λ^{n-1} we get

$$\frac{G_{n+1}}{G_n} = \frac{c_1\lambda + c_2\mu\left(\frac{\mu}{\lambda}\right)^{n-1}}{c_1 + c_2\left(\frac{\mu}{\lambda}\right)^{n-1}}. \tag{5}$$

Since $\left|\dfrac{\mu}{\lambda}\right| < 1$, $\left(\dfrac{\mu}{\lambda}\right)^{n-1} \longrightarrow 0$, and

$$\lim_{n\to\infty} \frac{G_{n+1}}{G_n} = \lambda \in (p, p+1). \tag{6}$$

It will also be useful (and instructive) to investigate how the different subsequences of $\left\{\dfrac{G_{n+1}}{G_n}\right\}$ approach the limit value of λ.

Lemma 9. *Let $\{G_n\}_{n=1}^{\infty}$ be a type-1-sequence. Then*

1. *If $a < \lambda$ (respectively $a > \lambda$) the subsequence $\left\{\frac{G_{2k+2}}{G_{2k+1}}\right\}_{k=0}^{\infty}$ is monotonically increasing (respectively decreasing).*
2. *If $a < \lambda$ (respectively $a > \lambda$) the subsequence $\left\{\frac{G_{2k+1}}{G_{2k}}\right\}_{k=1}^{\infty}$ is monotonically decreasing (respectively increasing).*

Proof: One way of proving the monotonicity of the subsequence $\left\{\frac{G_{2k+2}}{G_{2k+1}}\right\}$ is by investigating the difference

$$\frac{G_{2k+2}}{G_{2k+1}} - \frac{G_{2k+4}}{G_{2k+3}} = \frac{G_{2k+2}G_{2k+3} - G_{2k+1}G_{2k+4}}{G_{2k+1}G_{2k+3}} \tag{7}$$

in the first case, and the difference

$$\frac{G_{2k+1}}{G_{2k}} - \frac{G_{2k+3}}{G_{2k+2}} = \frac{G_{2k+1}G_{2k+2} - G_{2k}G_{2k+3}}{G_{2k}G_{2k+2}} \tag{8}$$

in the second case, i.e. in the subsequence $\left\{\frac{G_{2k+1}}{G_{2k}}\right\}$. Since both denominators are positive, we will investigate the sign of the numerators

$$G_{2k+2}G_{2k+3} - G_{2k+1}G_{2k+4} = c_1c_2\lambda^{2k}\mu^{2k}\left(\lambda\mu^2 + \lambda^2\mu - \mu^3 - \lambda^3\right) \tag{9}$$

and

$$G_{2k+1}G_{2k+2} - G_{2k}G_{2k+3} = c_1c_2\lambda^{2k-1}\mu^{2k-1}\left(\lambda\mu^2 + \lambda^2\mu - \mu^3 - \lambda^3\right), \tag{10}$$

respectively.

In the first case, the sign of the expression (9) depends solely on c_2, since c_1, λ^{2k}, and μ^{2k} are all positive, while $\left(\lambda\mu^2 + \lambda^2\mu - \mu^3 - \lambda^3\right) = -p\left(p^2 + 4q\right)$ is negative. If $a < \lambda$, then $c_2 > 0$, and (9) is negative, which means that $\left\{\frac{G_{2k+2}}{G_{2k+1}}\right\}$ is increasing. On the other hand, if $a > \lambda$, then $c_2 < 0$, and (9) is positive, which means that $\left\{\frac{G_{2k+2}}{G_{2k+1}}\right\}$ is decreasing.

In the second case, the sign of the expression (10) again depends solely on c_2, since c_1 and λ^{2k-1} are positive, while μ^{2k-1} and $\left(\lambda\mu^2 + \lambda^2\mu - \mu^3 - \lambda^3\right)$ are negative. The rest is similar.

□

Corollary 10. *Let $\{G_n\}_{n=1}^{\infty}$ be a type-1-sequence. Then there exists an integer $2 \le K_0 \le 3$ such that for all $n \ge K_0$ we have*

$$\frac{G_{n+1}}{G_n} \in (p, p+1) \tag{11}$$

Proof: We just have to check that $2 \le K_0 \le 3$. For all $n \ge 3$ we have

$$\frac{G_{n+1}}{G_n} = \frac{pG_n + qG_{n-1}}{G_n} = p + \frac{qG_{n-1}}{pG_{n-1} + qG_{n-2}} \in (p, p+1),$$

since $q \le p$ and $qG_{n-2} > 0$. Hence, $K_0 \le 3$.

Now, if additionally $a > q$, then $\frac{G_3}{G_2} = p + \frac{q}{a} \in (p, p+1)$, hence $K_0 = 2$.

□

Let's denote the prefix set $\{1, G_2, \ldots, G_k\}$ of $\{G_n\}_{n=1}^{\infty}$ by $G^{(k)}$. We know that $G^{(2)} = \{1, a\}$ is always greedy, and we will now investigate when $G^{(3)}$ is greedy:

Lemma 11. *Let $\{G_n\}_{n=1}^{\infty}$ be a type-1-sequence, then $G^{(3)} = \{1, a, pa+q\}$ is (totally) greedy if, and only if, $2 \leq a \leq p+q$.*

Proof: By Proposition 3, the set $\{1, a, pa+q\}$ is greedy if and only if $pa+q-a$ belongs to the set

$$\mathfrak{D}(a) = \{a-1, a\} \cup \{2a-2, 2a-1, 2a\} \cup \ldots \{ma-m, \ldots ma\} \cup \ldots.$$

If $a > p+q$ then $pa+q-a \notin \mathfrak{D}(a)$, so $G^{(3)}$ is not greedy. Hence $2 \leq a \leq p+q$. Let us now check that this condition is sufficient.

We may split the condition $2 \leq a \leq p+q$ into two cases:

1. $a < q$, and
2. $q \leq a \leq p+q$.

In the second case it is easy to see that $pa+q-a \in \mathfrak{D}(a)$, hence $G^{(3)}$ is greedy. In the first case let $m' = \left\lceil \frac{q}{a} \right\rceil > 1$.

$$\begin{aligned} pa+q-a &= pa+q-a+(m'-1)a-(m'-1)a \\ &= (p+m'-1)a-(m'a-q). \end{aligned}$$

Thus, $pa+q-a \in \mathfrak{D}(a)$ if, and only if, $0 \leq m'a-q \leq p+m'-1$. We already know that $m'a-q \geq 0$ by the definition of m'. As for the other inequality, we have

$$m'a-q < 2q-q = q \leq p < p+m'-1.$$

□

Now we are in the position to prove our main result:

Theorem 12. *Let $\{G_n\}_{n=1}^{\infty}$ be a type-1-sequence with $2 \leq a \leq p+q$. Then $\{G_n\}_{n=1}^{\infty}$ is totally greedy.*

Proof: The theorem is proved by induction. Lemma 11 guarantees that $G^{(3)}$ is greedy; that would be the base case. Now, let's suppose that $G^{(k)}$ is totally greedy for some arbitrary $k \geq 3$. We will prove that $G^{(k+1)}$ is also greedy (and hence totally greedy).

By Lemma 8 and Corollary 10 we know that $p < \dfrac{G_{k+1}}{G_k} < p+1$, so $m = \left\lceil \dfrac{G_{k+1}}{G_k} \right\rceil = p+1$. Now,

$$\begin{aligned} (p+1)G_k - G_{k+1} &= (p+1)G_k - (pG_k + qG_{k-1}) \\ &= G_k - qG_{k-1} = (pG_{k-1} + qG_{k-2}) - qG_{k-1} \\ &= (p-q)G_{k-1} + qG_{k-2}. \end{aligned}$$

To conclude the proof, note that $\textsc{GreedyCost}_{G^{(k)}}\left((p-q)G_{k-1} + qG_{k-2}\right) = p-q+q = p < p+1 = m$.

□

We can now apply Theorem 12 to some specific sequences, such as the (shifted) Fibonacci numbers $\{F_n\}_{n=1}^{\infty} = \{1, 2, 3, 5, 8, 13, \ldots\}$, and the (shifted) Pell numbers $\{P_n\}_{n=1}^{\infty} = \{1, 2, 5, 12, 29, 70 \ldots\}$.

A full version of this paper, including these and other results, can be found in [9].

Acknowledgements

The author has been partially supported by Grant 2021 SGR 00115 from the Government of Catalonia, by the project ACITECH PID2021-124928NB-I00, funded by MCIN/AEI/ 10.13039/501100011033/FEDER, EU, and by the project HERMES, funded by INCIBE and by the European Union NextGeneration EU/PRTR.

References

[1] Adamaszek, A. and M. Adamaszek: Combinatorics of the change-making problem. *European Journal of Combinatorics* **31** (2010), 47–63.

[2] Cai, X.: Canonical Coin Systems for Change-Making Problems. *Procs. of the 9th IEEE Int. Conf. on Hybrid Intelligent Systems* (2009), 499–504.

[3] Cowen, L.J., R. Cowen and A. Steinberg: Totally Greedy Coin Sets and Greedy Obstructions. *The Electronic Journal of Combinatorics* **15** (2008), #R90.

[4] Lueker, G.S.: Two NP-complete problems in nonnegative integer programming. *Tech. Rep. 178* (1975), Computer Science Lab., Princeton University.

[5] Magazine, M.J., G.L. Nemhauser and L.E. Trotter, Jr.: When the Greedy Solution Solves a Class of Knapsack Problems. *Operations Research* **23**(2) (1975), 207–217.

[6] *OEIS: The On-Line Encyclopedia of Integer Sequences.* `http://oeis.org/classic/index.html`.

[7] Pearson, D.: A polynomial-time algorithm for the change-making problem. *Operations Research Letters* **33** (2005), 231–234.

[8] Pérez-Rosés, H. M. Bras and J.M. Serradilla-Merinero: Greedy routing in circulant networks. *Graphs and Combinatorics* **38**(86) (2022). DOI: `https://doi.org/10.1007/s00373-022-02489-9`.

[9] Pérez-Rosés, H. "Totally Greedy Sequences Defined by Second-Order Linear Recurrences With Constant Coefficients". *arXiv:2405.16609* (2024).

[10] Shallit, J.: What This Country Needs is an 18c Piece. *The Mathematical Intelligencer* **25**(2) (2003), 20–23.

[11] Suzuki, Y. and R. Miyashiro: "Characterization of canonical systems with six types of coins for the change-making problem". *Theoretical Computer Science* **955** (2023), DOI: `https://doi.org/10.1016/j.tcs.2023.113822`.

An algebraic approach to the Weighted Sum Method in Multi-objective Integer Programming

J. M. Jiménez-Cobano*[1], H. Jiménez-Tafur†[2], and J.M. Ucha-Enríquez ‡[3]

[1]Instituto de Matemáticas de la Universidad de Sevilla, Spain
[2]Dpto. de Matemáticas. Universidad Pedagógica Nacional, Bogotá, Colombia
[1]Dpto. Matemática Aplicada I, Universidad de Sevilla, Spain

Abstract

In this work we present how to use test sets of Linear Integer Programming Problems to apply the classical *Weighted Sum Method* in bi-objective optimization. Although this method does not compute in general the complete set of non-dominated solutions, is one of the most widely used due to its simplicity.

The interest of using test sets computed with Gröbner bases is that these combinatorial tools compute exactly which weights should be considered to obtain the complete set of supported non-dominated solutions. Our approach can be extended to some problems in Multi-objective Non-Linear Integer Programming as well.

1 Preliminaries

1.1 Multi-objective optimization

Most real-life decision-making activities require more than one objective to be considered. These objectives can be conflicting, and thus some trade-offs are needed. As a result, a set of *Pareto-optimal solutions*, rather than a single solution, must be found.

A general multi-objective optimization problem can be written as

$$\begin{array}{ll} \min & f_1(\mathbf{x}), \ldots, f_r(\mathbf{x}) \\ s.t. & g_j(\mathbf{x}) \leq 0,\ j = 1, \ldots, J \\ & h_k(\mathbf{x}) = 0,\ k = 1, \ldots, K \\ & \mathbf{x} \in \mathbb{R}^d \end{array} \tag{1}$$

The space of the vectors of decision variables $\mathbf{x}$ is called the *search space*. The space formed by all the possible values of objective functions is called the *objective space*. Since in general there is no feasible point that minimises all the cost functions, we are interested in the *efficient points*: those feasible points $\mathbf{x}^\star$ such that there is no feasible $\mathbf{x}$ with $f_i(\mathbf{x}) \leq f_i(\mathbf{x}^\star)$ with at least one strict inequality for $i = 1, \ldots, r$. If $\mathbf{x}^\star$ is an efficient point, $(f_1(\mathbf{x}^\star), \ldots, f_r(\mathbf{x}^\star))$ is a *non-dominated point* in the objective space. The set of all non-dominated points is usually called the *Pareto front*.

The Weighted Sum Method (cf. [4]) combines all the multi-objective functions into a single objective function $w_1 f_1 + \cdots + w_r f_r$ with $\sum_{i=1}^{r} w_i = 1$ to express the preferences of the decision maker. So the aim of this approach is to describe (as accurate as possible) the set of of solutions of the following family of single objective problems:

*Email: josjimcob@alum.us.es
†Email: hjimenezt@pedagogica.edu.co
‡Email: ucha@us.es

$$\begin{array}{ll} \min & w_1 f_1(\mathbf{x}) + \cdots + w_r f_r(\mathbf{x}) \\ s.t. & g_j(\mathbf{x}) \leq 0,\, j = 1, \ldots, J \\ & h_k(\mathbf{x}) = 0,\, k = 1, \ldots, K \\ & \mathbf{x} \in \mathbb{R}^d \end{array} \tag{2}$$

It is well known that this method only produce the complete set of non-dominated solutions if the Pareto front is convex and it is not always clear how to select properly the w_i. The solutions obtained by this method are called *supported points.*

1.2 Bi-objective linear integer case

In this work we treat the bi-objective linear case for which objectives and constraints are linear functions and in which the variables are integer, that is

$$\begin{array}{ll} \min & \mathbf{c}_1^t\mathbf{x}, \mathbf{c}_2^t\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0}, \end{array} \tag{3}$$

for $\mathbf{b} \in \mathbb{Z}^m, A \in \mathbb{Z}^{m\times n}$. We present a combinatorial description of the classical Weighted Sum method for our problem considering the family

$$\begin{array}{ll} \min & w_1\mathbf{c}_1^t\mathbf{x} + w_2\mathbf{c}_2^t\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0}, \end{array} \tag{4}$$

for $w_1 + w_2 = 1$, using *test sets* computed via Gröbner bases. We will show how test sets provides the *exact* values of w_i that has to be considered to not drop any supported point.

In [8] an algebraic approach also based in test sets is proposed to apply the *ϵ-constraint method*, another classical approach that solves a family of several problems of only one objective to manage the multi-objective case. Tests sets in that case shows exactly which single objective problems are required to be solved to obtain all non-dominated solutions without redundant calculations.

1.3 Tests sets in linear integer programming

Given a linear integer programming problem with a single objective function

$$\begin{array}{ll} \min & \mathbf{c}^t\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \mathbb{Z}^d_{\geq 0} \end{array} \tag{5}$$

a fundamental tool is the *test set* associated to $\mathbf{c}$ and A:

Definition 1. *A test set of Problem 5 is a set $T \subset \ker(A) \subset \mathbb{Z}^d$ such that: 1) for any feasible solution $\mathbf{x}$ of Problem 5 that is not optimal, there exists $\mathbf{t} \in T$ such that $\mathbf{x} - \mathbf{t}$ is feasible and $\mathbf{c}^t(\mathbf{x} - \mathbf{t}) < \mathbf{c}^t\mathbf{x}$, and 2) given the optimal solution $\mathbf{x}^\star$ of Problem 5, $\mathbf{x} - \mathbf{t}$ is not feasible for any $\mathbf{t} \in T$.*

Test sets produced a natural way of solving Problem 5: starting from a feasible point, subtract elements of the test set as long as it is possible. Test sets can be obtained computing a Gröbner basis of the ideal associated to Problem 5 (cf. [3]) with respect to a suitable monomial ordering that takes into account the cost function $\mathbf{c}$ and codifying the exponents of the polynomials in vectors (see [11]) in which positive components correspond to the term leader of the polynomial. Test sets can be computed for a fixed $\mathbf{b}$ or, usually, valid for any possible $\mathbf{b}$.(*see*[**?**]

Up to our knowledge, the best implementation to compute test sets is `4ti2` ([5]). Test sets been introduced in [12] to solve nonlinear integer problems, and in [2], [6] or [7] for real size cases of Portfolio Selection and Reliability Redundancy Allocation problems.

2 The weighted sum method with test sets

Using test sets computed with Gröbner bases it is possible to compute exactly which values of w_1, w_2 produce new potential efficient points. This is possible because the Gröbner bases behind the test sets have the following property: if the exponents with respect to an ordering $\prec_2$ of the polynomials of a given base for another ordering $\prec_1$ are the same, the bases with respect to both orderings are the same one (cf. [3]).

Example 2. *Let us consider an illustrative example to get the general idea of our procedure. Given the bi-objective assignment problem*

$$\begin{array}{ll} \min & \mathbf{c}_1^t\mathbf{x}, \mathbf{c}_2^t\mathbf{x} \\ \text{s.t.} & \sum_{j=1}^3 x_{ij} = 1, 1 \leq i \leq 3 \\ & \sum_{i=1}^3 x_{ij} = 1, 1 \leq j \leq 3, \\ & x_{ij} \in \{0,1\}, \end{array} \tag{6}$$

with costs $\mathbf{c}_1 = (12, 12, 8, 15, 9, 1, 16, 4, 3)$ *and* $\mathbf{c}_2 = (6, 4, 11, 10, 19, 18, 16, 10, 17)$*, a test set*[1] *for the problem with objective* $\mathbf{c}_1$*, that is* $(1-w)\mathbf{c}_1 + w\mathbf{c}_2$ *with* $w = 0$*, is*

$$\begin{array}{rl} T = \{ & (-1,0,1,0,0,0,1,0,-1), (-1,0,1,1,0,-1,0,0,0), (-1,1,0,0,0,0,1,-1,0), \\ & (-1,1,0,1,-1,0,0,0,0), (0,-1,1,0,1,-1,0,0,0), (0,0,0,-1,1,0,1,-1,0), \\ & (0,0,0,0,1,-1,0,-1,1), (0,0,0,1,0,-1,-1,0,1), (0,1,-1,0,0,0,0,-1,1)\} \end{array}$$

and the optimum solution of the problem is $P_0 = (1,0,0,0,0,1,0,1,0)$*. The combination of costs* $(1-w)\mathbf{c}_1 + w\mathbf{c}_2$ *is*

$$(-6w+12, -8w+12, 3w+8, -5w+15, 10w+9, 17w+1, 16, 6w+4, 14w+3)$$

The cost of the first element $(1,0,1,0,0,0,1,0,1)$ *for this combination is*

$$-(-6w+12) + (3w+8) + 16 - (14w+3) = -5w+9,$$

so for every $w \in [0,1]$ *the exponent (corresponding to the leading term of the polynomials in the Gröbner basis) does not change: the cost is always positive, that is, the cost of positive components always surpass the cost of negative components and the exponent of the element does not change. On the contrary, if the element* $(0,0,0,1,0,1,1,0,1)$ *is considered, its cost is*

$$(-5w+15) - (17w+1) - (16) + (14w+3) = -8w+1.$$

For any $w \in [0, 1/8)$ *the exponent of this element does not change, but for* $w = 1/8$ *(and with respect to a monomial ordering that uses namely* $\mathbf{c}_2$ *to break ties) the exponent does change. Checking which are the* w *for which the exponent changes for every element in* T*, and considering the smallest one* w_0*, we can assure that*

- T *will be the test set of Problem 6 for* $w \in [0, w_0)$*. In this case* w_0 *is precisely* $1/8$*.*
- *A computation of the test set to solve Problem 6 for* $w = 1/8$ *will provide a new test set (and, eventually, a new optimum solution).*

The general procedure, given some $\mathbf{c}_1, \mathbf{c}_2$ the matrix A of the constraints and a feasible point P_{feas} (that implies the value of $\mathbf{b}$) is Algorithm 2. First a test set T corresponding to cost $\mathbf{c}_1$ and the associated optimum solution are computed. Then consider all the $w \in (0,1]$ that produce changes of exponent in the elements of T, and take the smallest one w_0 for which the test set of the minimisation problem with cost $(1-w)\mathbf{c}_1 + w\mathbf{c}_2$ is different to T. Repeat this process until w_0 turns out to be 1.

[1] We have implemented an ordering that takes into account $\mathbf{c}_1$ first, and break ties with $\mathbf{c}_2$. In `4ti2` this option is possible introducing matrices of costs, with each row corresponding to a different cost.

Algorithm 1 Weighted Sum with Test Sets

0: **input:** $\mathbf{c}_1, \mathbf{c}_2, A, \mathbf{b}, P_{\text{feas}}$ of Problem 3
0: **output:** Set of all supported points
0: $c_1 := \mathbf{c}_1$
0: SupPoints := $\emptyset$
0: **while** $c_1 \neq \mathbf{c}_2$ **do**
0: T = TestSet(c_1, A) $\{T$ does not depend on $\mathbf{b}\}$
0: P := Solve(P_{feas}, T)
0: SupPoints := SupPoints $\cup\{P\}$. $\{P$ can be superfluous$\}$
0: $w_0 = \min_{\mathbf{t}\in T}\{w \in (0,1] | w$ changes exponent of $\mathbf{t}\}$ $\{w_0$ can be equal to 1$\}$
0: $c_1 := (1 - w_0)c_1 + w_0\mathbf{c}_2$
0: **end while**
0: T = TestSet$(\mathbf{c}_2, A)$
0: P := Solve(P_{feas}, T)
0: SupPoints := SupPoints $\cup\{P\}$. {Optimum for $\mathbf{c}_2$ computed, just in case}
0: **return** SupPoints =0

The algorithm is correct because the number of different Gröbner basis for a given ideal is finite (cf. [9]).

Our method describes exactly which w_i are necessary to be selected because they produce a different test set, so potentially a new optimal point. Nevertheless, two different test sets can lead to the same optimal point.

Example 3. *In Problem 6, when the test set is computed for $w = 1/8$ a new test set is obtained:*

$$\begin{aligned} T' = \{ \quad & (-1,0,1,0,0,0,1,0,-1), (-1,0,1,1,0,-1,0,0,0), (-1,1,0,0,0,0,1,-1,0), \\ & (-1,1,0,1,-1,0,0,0,0), (0,-1,1,0,1,-1,0,0,0), (0,0,0,-1,0,1,1,0,-1), \\ & (0,0,0,-1,1,0,1,-1,0), (0,0,0,0,1,-1,0,-1,1), (0,1,-1,0,0,0,0,-1,1)\} \end{aligned}$$

However, the optimum point is the same one.

3 Applications to bi-objective non-linear integer programming

As it is explained in [12] test sets can be exploided to solve problems of type

$$\begin{array}{rl} \min & \mathbf{c}^t\mathbf{x} \\ \text{s.t.} & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \in \Omega \\ & \mathbf{x} \in \mathbb{Z}^n_{\geq 0}, \end{array} \tag{7}$$

for Ω described with non-linear (computable) conditions. The strategy is to calculate the linear optimum for the problem without the non-linear constraints and walking back (adding elements of the test set that worsens the values of the cost function) until points in Ω are reached. The best point obtained into Ω is the optimum of Problem 7.

Instead of $\mathbf{c}$ a family of costs $(1-w)\mathbf{c}_1 + w\mathbf{c}_2$ can be handled: with Algorithm 2 we can achieve the values of w that produce different linear optima for the whole family. Only walking back from them is required to obtain all supported points for the bi-objective counterpart of Problem 7. We present how to apply this framework to a thoroughly studied example in the literature.

In [1] a method to treat a family of three-objective redundancy allocation problem is presented. The functions to be optimised are the cost, weight and reliability of the system. We propose an alternative way to handle the case example (section 3) of three subsystems. We instead solve the bi-objective problem with respect to cost and weight and add the reliability f_R as an extra constraint (for which

we ask a convenient value ρ), as in [12]. Additionally, we have rearranged the weights (coefficients of f_2) to obtain more supported solutions.

We consider the resulting problem of the form

$$\begin{array}{ll} \min & f_1, f_2 \\ \text{s.t.} & 1 \leq \sum_{i=1}^{3} \sum_{j=1}^{m_i} x_{ij} \leq 7 \\ & f_R(\mathbf{x}) \leq \rho \\ & x_{ij} \in \mathbb{Z}_{\geq 0}^{14} \end{array} \tag{8}$$

with

$$\begin{aligned} f_1 = \; & 4x_{11} + 6x_{12} + 7x_{13} + 8x_{14} + 9x_{15} + \\ & + 3x_{21} + 4x_{22} + 5x_{23} + 7x_{24} + \\ & + 2x_{13} + 4x_{32} + 4x_{33} + 6x_{34} + 8x_{35}, \\ f_2 = \; & 9x_{11} + 6x_{12} + 6x_{13} + 3x_{14} + 2x_{15} + \\ & + 12x_{21} + 3x_{22} + 2x_{23} + 2x_{24} + \\ & + 10x_{13} + 6x_{32} + 4x_{33} + 3x_{34} + 2x_{35} \end{aligned}$$

and $m_1 = m_3 = 5, m_2 = 4$.

Applying Algorithn 2 to this problem produces the consecutive values of $w_0 = 1/3, 1/2, 1/2$ and $3/4$, and the following list of 6 optimal points of the linear part:

$$\begin{array}{c} (0,0,0,0,1,0,0,1,0,0,0,0,0,1), (0,0,0,0,1,0,0,1,0,0,0,1,0,0), \\ (0,0,0,1,0,0,1,0,0,0,0,1,0,0), (1,0,0,0,0,0,1,0,0,0,0,1,0,0), \\ (1,0,0,0,0,0,1,0,0,1,0,0,0,0), (1,0,0,0,0,1,0,0,0,1,0,0,0,0) \end{array}$$

For each of these optimal points we solve the non-linear corresponding problems with the strategy of walkback (as it is explained in [6]) for the values of ρ for which we are interested in, and obtain a suitable subset of the Pareto front of the non-linear Problem 8. For $\rho = 0.99$ the complete set of supported non-dominated points obtained is

$$\begin{array}{c} \{(0,0,0,0,5,0,1,3,0,0,0,0,0,5), \\ (1,1,0,0,0,0,3,0,0,2,0,0,0,0), \\ (2,0,0,0,0,1,1,0,0,2,0,0,0,0)\} \end{array}$$

We consider that the approach of solving this problem for different values of ρ is more convenient that the one proposed in [1] in which 6112 non-dominated points are reported. The size of this set is unmanageable for a decision maker.

4 Conclusions

We have presented an algebraic description of the weighted sum method to compute the supported non-dominated solutions of a bi-objective linear integer programming problem. A generalization to any number of objective functions is a work in progress and requires a complete understanding of how to calculate the subset of the *Gröbner fan* of the ideal corresponding to the the combinations of the costs with any number of parameters.

In addition we have presented how to extend the algebraic weighted sum method to some multi-objective non-linear integer problems, specifically to a widely studied example of redundancy allocation problem.

References

[1] D. Cao, A. Murat, R.B. Chinnam, Efficient exact optimization of multi-objective redundancy allocation problems in series-parallel systems, *Reliability Engineering and System Safety* **111** (2013), 154–163.

[2] F. J. Castro, M. J. Gago, M. I. Hartillo, J. Puerto, J. M. Ucha, An algebraic approach to integer portfolio problems. *European J. Oper. Res.* **210** (2011), no. 3, 647–659.

[3] D. A. Cox, J. Little, D. O'Shea, *Using algebraic geometry.* Graduate texts in mathematics: 185 (2nd). Springer, New York, 2005.

[4] M. Ehrgott, *Multicriteria optimization.* (2nd). Berlin: Springer, 2005.

[5] 4ti2 team. *4ti2—a software package for algebraic, geometric and combinatorial problems on linear spaces.* (2015) Available at `www.4ti2.de`.

[6] M. J. Gago, M. I. Hartillo, J. Puerto, J. M. Ucha J. M. Exact cost minimization of a series-parallel reliable system with multiple component choices using an algebraic method. *Comput. Oper. Res.* **40** (2011), no. 11, 2752–2759.

[7] Hartillo-Hermoso, M.I., Jiménez-Cobano, J.M., Ucha-Enríquez, J.M.: Finding multiple solutions in nonlinear integer programming with algebraic test-sets. In: Computer algebra in scientific computing 2018, Lille (France). LNCS, vol. 11077, pp. 230–237. Springer, Heidelberg (2016).

[8] M. I. Hartillo-Hermoso, H. Jiménez-Tafur, Haydee, J. M. Ucha-Enríquez, An exact algebraic ϵ-constraint method for bi-objective linear integer programming based on test sets. *European J. Oper. Res.* **282** (2020), no. 2, 453–463.

[9] T. Mora, L. Robbiano, The Gröbner fan of an ideal. *Journal of Symbolic Computation*, **6**(2–3)(1988), 183–208

[10] A. Schrijver, *Theory of linear and integer programming.* Wiley-Interscience Series in Discrete Mathematics. John Wiley Sons, Ltd., Chichester. A Wiley-Interscience Publication, 1986.

[11] R. Thomas, A geometric Buchberger algorithm for integer programming. *Mathematics of Operations Research*, **20(4)**(1985), 864–884.

[12] S.R. Tayur, R.R. Thomas, N.R. Natraj, An algebraic geometry algorithm for scheduling in presence of setups and correlated demands. *Mathematics Program*, **69** (1995), 369–401.

Sidorenko-type inequalities for Trees Discrete Mathematics Days 2024*

Natalie Behague[†1], Gabriel Crudele[‡1], Jonathan A. Noel[§1], and Lina Maria Simbaqueba[¶1]

[1]Department of Mathematics and Statistics, University of Victoria, Canada.

Abstract

Given two graphs H and G, the homomorphism density $t(H,G)$ represents the probability that a random mapping from $V(H)$ to $V(G)$ is a homomorphism. Sidorenko Conjecture states that for any bipartite graph H, $t(H,G)$ is greater or equal than $t(K_2,G)^{e(H)}$ for every graph G.

Introducing a binary relation $H \succcurlyeq T$ if and only if $t(H,G)^{e(T)} \geq t(T,G)^{e(H)}$ for all graphs G, we establish a partial order on the set of non-empty connected graphs. Employing a technique by Kopparty and Rossman [10], which involves the use of entropy to define a linear program, we derive several necessary and sufficient conditions for two trees T, F to satisfy $T \succcurlyeq F$. Furthermore, we show how important results and open problems in extremal graph theory can be reframed using this binary relation.

1 Introduction

One of the main objectives of extremal combinatorics is to study certain substructures in a large combinatorial object to understand the influence of local pattern frequencies on a global structure. This topic links many active areas of research, including the study of quasirandomness pioneered by Rödl [14], Thomason [17] and Chung, Graham and Wilson [4], the theory of combinatorial limits developed by Lovász and his collaborators, see [12], and the area of property testing in computer science spearheaded by Goldreich, Goldwasser and Ron [8].

A homomorphism from a graph H to a graph G is a function $f : V(H) \to V(G)$ such that $f(u)f(v) \in E(G)$ whenever $uv \in E(H)$. We denote by $Hom(H,G)$ the set of all possible homomorphisms between H and G. Let us denote $hom(H,G) = |Hom(H,G)|$. The homomorphism density, $t(H,G)$, is the probability that a random function $f : V(H) \to V(G)$ is a homomorphism.

$$t(H,G) = \frac{hom(H,G)}{v(G)^{v(H)}}.$$

Our focus is on proving inequalities for homomorphism densities of the following form:

$$t(F_2,G) \geq t(F_1,G)^{\alpha} \tag{1}$$

where F_1 and F_2 are fixed graphs, $\alpha > 0$ and the inequality in (1) holds for every graph G. Inequalities of this form are known as Sidorenko-type inequalities and several problems in extremal combinatorics

*The full version of this work can be found in [13].

[†]Email: nbehague@uvic.ca Supported by a PIMS Postdoctoral Fellowship.

[‡]Email: gabrielcrudele1@gmail.com.

[§]Email: noelj@uvic.ca. Research supported by NSERC Discovery Grant RGPIN-2021-02460 and NSERC Early Career Supplement DGECR-2021-00024 and a Start-Up Grant from the University of Victoria.

[¶]Email: lmsimbaquebam@uvic.ca Departmento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia. Research supported by a Mitacs Globalink Research Internship.

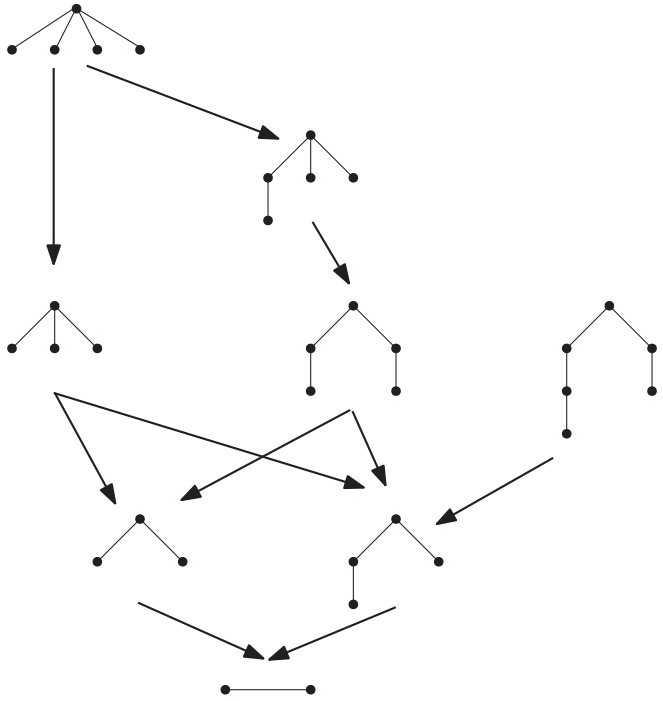


Figure 1: A section of the poset for trees.

can be expressed in terms of these inequalities. For example, the well-known Sidorenko Conjecture [16] states that $t(H,G) \geq t(K_2,G)^{|E(H)|}$ for every bipartite graph H. A systematic study of Sidorenko-type inequalities for graph homomorphisms via the method of tropicalization was recently initiated in [2, 3]. In [10], Kopparty and Rossman introduced a powerful method for using the information theoretic notion of entropy together with linear programming to prove Sidorenko-type inequalities. This approach is akin to the entropy-based approach which has seen recent success in the study of Sidorenko's Conjecture [5, 6] and other related problems [9, 11]. It was also used by Blekherman and Raymond [1] to give an illuminating alternative proof of the result of Sağlam [15] that

$$t(P_{k+2},G) \geq t(P_k,G)^{\frac{k+1}{k-1}} \tag{2}$$

for all $k \geq 2$ where, for all $\ell \geq 1$, P_ℓ denotes the path with ℓ vertices and $\ell - 1$ edges. This inequality was first conjectured by Erdős and Simonovits [7]. For a recent generalization of this result, see [2, Theorem 1.3].

Given two non-empty graphs H and T, we write $H \succcurlyeq T$ to mean that $t(H,G)^{e(T)} \geq t(T,G)^{e(H)}$ for every graph G. This binary relation is a partial order on the set of non-empty connected graphs. In Figure 1 we show the poset of some small trees.

2 The linear program.

Following the method introduced by Kopparty and Rossman, we reduce the problem of proving that $H \succcurlyeq T$ for forests H and T to solving a linear program. We obtained the full structure of the partial order on all pairs of trees with at most 8 vertices. Also, we characterize trees H such that $H \succcurlyeq S_k$ and $H \succcurlyeq P_4$, where S_k is the star on k vertices and P_4 is the path on 4 vertices.

Let $LP(H,T)$ be the following linear program. Let $\{w(\varphi) : \varphi \in Hom(H,T)\}$ be the variables.

$$\begin{aligned}
\text{maximize} \quad & \sum_{e\in E(T)} \sum_{\varphi\in Hom(H,T)} \mu_\varphi(e)\cdot w(\varphi) & \\
\text{subject to} \quad & \sum_{\varphi\in Hom(H,T)} \mu_\varphi(e)\cdot w(\varphi) \le 1 & \forall e \in E(T),\\
& \sum_{\varphi\in Hom(H,T)} \mu_\varphi(v)\cdot w(\varphi) \le 1 & \forall v \in V(T),\\
& w(\varphi) \ge 0 & \forall \varphi \in \mathrm{Hom}(H,T).
\end{aligned}$$

Where $\mu_\varphi(v) = |\varphi^{-1}(v)|$ for each $v \in V(T)$ and $\mu_\varphi(e) = |\varphi^{-1}(e)|$ for each $e \in E(T)$.

Lemma 1. *If H and T are forests such that the value of $LP(H,T)$ is equal to $e(T)$, then $H \succcurlyeq T$.*

We also define the dual of the linear program. Let $DLP(H,T)$ be the dual of $LP(H,T)$ with variables $\{y(m) : m \in V(T) \cup E(T)\}$ defined as follows:

$$\begin{aligned}
\text{minimize} \quad & \sum_{v\in V(T)} y(v) + \sum_{e\in E(T)} y(e) & \\
\text{subject to} \quad & \sum_{v\in V(T)} \mu_\varphi(v)\cdot y(v) + \sum_{e\in E(T)} \mu_\varphi(e)\cdot y(e) \ge e(H) & \forall \varphi \in \mathrm{Hom}(H,T),\\
& y(v) \ge 0 & \forall v \in V(T)\\
& y(e) \ge 0 & \forall e \in E(T).
\end{aligned}$$

Lemma 2. *If H and T are non-empty graphs such that the value of $DLP(H,T)$ is less than $e(T)$, then $H \not\succcurlyeq T$.*

3 Main results.

Given two trees H and T, the following theorems give sufficient or necessary conditions for $H \succcurlyeq T$. We let $\sigma(H)$ be the minimum of $|A|, |B|$ in the bipartition (A,B) for the tree.

Theorem 3. *If $H \succcurlyeq T$, then*

$$\frac{e(H)}{\sigma(H)} \ge \frac{e(T)}{\sigma(T)}.$$

The last theorem holds for any H and T bipartite graphs. For the sufficient condition, we say that a *fractional orientation* of a graph T is a function $f : V(T)\times V(T) \to [0,\infty)$ such that $f(u,v)+f(v,u) = 1$ for any edge $uv \in E(T)$ and $f(u,v) = 0$ if $uv \notin E(T)$.

The *out-degree* and *in-degree* of a vertex $v \in V(T)$ are $d_f^+(v) := \sum_{u\in V(T)} f(v,u)$ and $d_f^-(v) := \sum_{u\in V(T)} f(u,v)$, respectively.

Theorem 4. *If there exists a fractional orientation of T such that, for all $v \in V(T)$,*

$$\frac{d_f^-(v)\cdot(v(H)-\sigma(H)) + d_f^+(v)\cdot\sigma(H)}{e(H)} \le 1, \tag{3}$$

then $H \succcurlyeq T$.

Using Theorem 3 and 4, we get the characterization for stars.

Corollary 5. *Let $k \ge 3$ and let H be a non-empty tree. Then $H \succcurlyeq S_k$ if and only if $e(H) \ge (k-1)\sigma(H)$.*

Finally, the following gives a characterization for P_4.

Theorem 6. *Let H be a tree. Then $H \succcurlyeq P_4$ if and only if H has at least four vertices.*

Nevertheless, it is not easy to generalize the result of Theorem 6 to a more general case.

Theorem 7. *Let H be a k-vertex near-star with ℓ leaves. If $\frac{k+1}{2} \leq \ell \leq k-3$, then $H \not\succcurlyeq P_k$.*

We believe that the following weaker generalization may hold. This statement, if true, would support the rough intuition that path-like graphs are near the bottom of the partial order restricted to trees.

Conjecture 8. *For any $k \geq 1$, there exists $n_0(k)$ such that if H is a tree with at least $n_0(k)$ vertices, then $H \succcurlyeq P_{2k}$.*

References

[1] G. Blekherman and A. Raymond. Proof of the Erdős–Simonovits conjecture on walks. E-print arXiv:2009.10845v1, 2020.

[2] G. Blekherman and A. Raymond. A path forward: Tropicalization in extremal combinatorics. E-print arXiv:2108.06377v2, 2022.

[3] G. Blekherman, A. Raymond, M. Singh, and R. R. Thomas. Tropicalization of graph profiles. E-print arXiv:2004.05207v2, 2022.

[4] F. R. K. Chung, R. L. Graham, and R. M. Wilson. Quasi-random graphs. *Combinatorica*, 9(4):345–362, 1989.

[5] D. Conlon, J. H. Kim, C. Lee, and J. Lee. Some advances on Sidorenko's conjecture. *J. Lond. Math. Soc. (2)*, 98(3):593–608, 2018.

[6] D. Conlon and J. Lee. Finite reflection groups and graph norms. *Adv. Math.*, 315:130–165, 2017.

[7] P. Erdős and M. Simonovits. Compactness results in extremal graph theory. *Combinatorica*, 2(3):275–288, 1982.

[8] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, 1998.

[9] A. Grzesik, J. Lee, B. Lidický, and J. Volec. On tripartite common graphs. E-print arXiv:2012.02057v1, 2020.

[10] S. Kopparty and B. Rossman. The homomorphism domination exponent. *European J. Combin.*, 32(7):1097–1114, 2011.

[11] J. Lee. On some graph densities in locally dense graphs. *Random Structures Algorithms*, 58(2):322–344, 2021.

[12] L. Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[13] J.A. Noel N. Behague, G. Crudele and L. Simbaqueba. Sidorenko-type inequalities for pairs of trees. *preprint* `arXiv:2305.16542`, 2023.

[14] V. Rödl. On universality of graphs with uniformly distributed edges. *Discrete Math.*, 59(1-2):125–134, 1986.

[15] M. Sağlam. Near log-convexity of measured heat in (discrete) time and consequences. In *59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018*, pages 967–978. IEEE Computer Soc., Los Alamitos, CA, 2018.

[16] A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.*, 9(2):201–204, 1993.

[17] A. Thomason. Pseudorandom graphs. In *Random graphs '85 (Poznań, 1985)*, volume 144 of *North-Holland Math. Stud.*, pages 307–331. North-Holland, Amsterdam, 1987.

A note on generalized crowns in linear r-graphs*

Lin-Peng Zhang[†1,2], Hajo Broersma[‡2], and Ligong Wang[§1]

[1]School of Mathematics and Statistics, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, P.R. China.

[2]Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, P.O. Box 217, 7500 AE Enschede, the Netherlands.

Abstract

An r-graph H is called linear if any two edges of H intersect in at most one vertex. Let F and H be two linear r-graphs. If H contains no copy of F, then H is called F-free. The linear Turán number of F, denoted by $ex_r^{lin}(n, F)$, is the maximum number of edges in any F-free n-vertex linear r-graph. The crown $C_{1,3}$ is a linear 3-graph which is obtained from three pairwise disjoint edges by adding one edge that intersects all three of them in one vertex. In 2022, Gyárfás, Ruszinkó and Sárközy initiated the study of $ex_3^{lin}(n, F)$ for different choices of an acyclic 3-graph F. They established lower and upper bounds for $ex_3^{lin}(n, C_{1,3})$. In this paper, we generalize the notion of a crown to linear r-graphs for $r \geq 3$, and also generalize the above results to linear r-graphs.

1 Introduction

The result presented here is motivated by a number of very recent papers on linear Turán numbers. We extend a result on crown-free linear 3-graphs to linear r-graphs for $r \geq 3$. Throughout, we let r be an integer with $r \geq 3$.

Let $H = (V, E)$ be an r-graph consisting of a set of vertices $V = V(H)$ and a collection $E = E(H)$ of r-element subsets of V called edges. If any two edges in H intersect in at most one vertex, then H is said to be linear. Let F be a linear r-graph. Then H is called F-free if it contains no copy of F as its subhypergraph. The linear Turán number of F, denoted by $ex_r^{lin}(n, F)$, is the maximum number of edges in any F-free linear r-graph on n vertices. More generally, for two linear r-graphs F_1 and F_2, H is called $\{F_1, F_2\}$-free if it contains no copy of F_1 or F_2 as its subhypergraph. The linear Turán number of $\{F_1, F_2\}$, denoted by $ex_r^{lin}(n, \{F_1, F_2\})$, is the maximum number of edges in any $\{F_1, F_2\}$-free linear r-graph on n vertices.

A linear 3-graph is acyclic if it can be constructed in the following way. We start with one edge. Then at each step we add a new edge intersecting the union of the vertices of the previous edges in at most one vertex. In 2022, Gyárfás, Ruszinkó and Sárközy [5] initiated the study of $ex_3^{lin}(n, F)$ for different choices of an acyclic 3-graph F. In [5], they determined the linear Turán numbers of linear 3-graphs with at most 4 edges, except the crown, for which they gave lower and upper bounds (Theorem 1 below). Here the crown is a linear 3-graph which is obtained from three pairwise disjoint edges on 3

*The full version of this work can be found in [10] and will be published elsewhere. This research is supported by the National Natural Science Foundation of China (No. 12271439) and China Scholarship Council (No. 202206290003).

†Email: lpzhangmath@163.com. Research of L.-P.. Zhang. supported by the National Natural Science Foundation of China (No. 12271439) and China Scholarship Council (No. 202206290003).

‡Email: h.j.broersma@utwente.nl.

§Email: lgwangmath@163.com. Research of L. Wang. supported by the National Natural Science Foundation of China (No. 12271439).

vertices by adding one edge that intersects all three of them in one vertex. In [5], the authors used E_4 to denote a crown, but here we adopt the notation $C_{1,3}$ from the more recent paper [9].

Since the publication of [5], there have appeared several results involving the linear Turán number of some acyclic linear hypergraphs [6, 7, 8]. In the remainder, we focus on results involving $C_{1,3}$, as our aim is to present a natural generalization of these results to linear r-graphs.

In [5], Gyárfás, Ruszinkó and Sárközy obtained the following result.

Theorem 1 ([5])**.**

$$6\left\lfloor \frac{n-3}{4} \right\rfloor + \varepsilon \leq ex_3^{lin}(n, C_{1,3}) \leq 2n,$$

where $\varepsilon = 0$ *if* $n-3 \equiv 0,1 \pmod 4$, $\varepsilon = 1$ *if* $n-3 \equiv 2 \pmod 4$, *and* $\varepsilon = 3$ *if* $n-3 \equiv 3 \pmod 4$.

Indeed, for the lower bound in Theorem 1, the authors of [5] gave the following construction for obtaining a class of extremal linear $C_{1,3}$-free 3-graphs. We recall this construction for later reference. Start with the graph mK_4 consisting of m disjoint copies of the complete graph on four vertices. The graph mK_4 admits a one-factorization, *i.e.*, a decomposition of the edge set into three edge-disjoint perfect matchings. Each of these matchings corresponds to $2m$ vertex-disjoint pairs of edges. Add one new vertex for each of the matchings and form $2m$ triples by adding this vertex to each of the $2m$ pairs. Now ignore the edges of the mK_4. This construction consists of $n = 4m+3$ vertices and $6m$ triples, and it is easy to check that the corresponding 3-graph is linear and $C_{1,3}$-free. Thus for $n = 4m+3$, this construction provides an extremal 3-graph with $6\left\lfloor \frac{n-3}{4} \right\rfloor + \varepsilon$ edges, where ε is defined as in the above theorem. The construction can be adjusted to obtain extremal 3-graphs for the other residue classes modulo 4.

In a later paper [2], Carbonero, Fletcher, Guo, Gyárfás, Wang, and Yan proved that every linear 3-graph with minimum degree 4 contains a crown. The same group of authors conjectured in [1] that $ex_3^{lin}(n, C_{1,3}) \sim \frac{3n}{2}$, and proposed some ideas to obtain the exact bounds. After that, Fletcher [4] improved the upper bound to $ex_3^{lin}(n, C_{1,3}) \leq \frac{5n}{3}$.

Very recently, Tang, Wu, Zhang and Zheng [9] established the following result.

Theorem 2 ([9])**.** *Let G be any C_{13}-free linear 3-graph on n vertices. Then $|E(G)| \leq \frac{3(n-s)}{2}$, where s denotes the number of vertices in G with degree at least 6.*

The above result shows that the lower bound in Theorem 1 is essentially tight. Furthermore, the above result, combined with the results in [5], essentially completes the determination of the linear Turán numbers for all linear 3-graphs with at most 4 edges.

2 Crown-free linear r-graphs

In the remainder, we focus on the following natural generalization of the notion of a crown to linear r-graphs. An r-crown $C_{1,r}$ is a linear r-graph on r^2 vertices and $r+1$ edges obtained from r pairwise disjoint edges on r vertices by adding one edge that intersects all of them in one vertex. In fact, for our purposes we need a second generalization of the crown to linear r-graphs. We let $C^*_{1,r}$ denote the following linear r-graph on $r^2 - r + 3$ vertices and $r+1$ edges. It consists of a set of $r-2$ edges $\{e_1, e_2, \ldots, e_{r-2}\}$ that intersect in exactly one vertex v, two additional disjoint edges e_{r-1} and e_r that are also disjoint from $\{e_1, e_2, \ldots, e_{r-2}\}$, and one additional edge e intersecting each edge of $\{e_1, e_2, \ldots, e_r\}$ in exactly one vertex except for v. Note that both $C_{1,r}$ and $C^*_{1,r}$ are isomorphic to the crown in case $r = 3$.

In the following, we establish an upper bound on $ex_r^{lin}(n, \{C_{1,r}, C^*_{1,r}\})$, and a lower bound on $ex_r^{lin}(n, \{C_{1,r}, C^*_{1,r}\})$ when $r-1$ is a prime power.

In order to obtain a lower bound on $ex_r^{lin}(n, \{C_{1,r}, C^*_{1,r}\})$, we can use a similar construction as in the description following Theorem 1. We can construct a $\{C_{1,r}, C^*_{1,r}\}$-free linear r-graph on n vertices by using the notion of a transversal design.

Assume that n is a multiple of k for some integer $k \geq r-1$. A transversal design $T(n,k)$ is a linear k-graph on n vertices, in which the vertices are partitioned into k sets, each containing $\frac{n}{k}$ vertices, and where each pair of vertices from different sets belongs to exactly one edge on k vertices. Note that $T(n,k)$ is an $\frac{n}{k}$-regular k-partite linear k-graph. It can be found in [3] that such $T(n,k)$ exist for sufficiently large n when k divides n. In particular, $T(k^2,k)$ exists when k is a prime power.

Let $r-1$ be a prime power. Denote by $T'((r-1)^2, r-1)$ the linear $(r-1)$-graph obtained from $T((r-1)^2, r-1)$ by adding one edge for each set in the partition. Note that for $r=3$, $T'((r-1)^2, r-1)$ is a K_4. We next extend m disjoint copies of $T'((r-1)^2, r-1)$ to a $\{C_{1,r}, C^*_{1,r}\}$-free linear r-graph in the same way as we did for $r=3$ starting with mK_4. Consider a one-factorization of the linear $(r-1)$-graph $mT'((r-1)^2, r-1)$. Each of the r factors corresponds to $(r-1)m$ vertex-disjoint $(r-1)$-tuples. Add one new vertex for each of the factors and form $(r-1)m$ edges by adding this vertex to each of the $(r-1)m$ $(r-1)$-tuples. The resulting linear r-graph has $r(r-1)m$ edges and $(r-1)^2m+r$ vertices, and it is $\{C_{1,r}, C^*_{1,r}\}$-free. Let $n=(r-1)^2m+r$. Then the number of edges of the constructed r-graph is at least $r(r-1)\left\lfloor \frac{n-r}{(r-1)^2} \right\rfloor$, where $r-1$ is a prime power.

In order to obtain an upper bound on $ex_r^{lin}(n, \{C_{1,r}, C^*_{1,r}\})$, we generalize the result of Theorem 2 to linear r-graphs. We present our proof of the following theorem in the next section. In the final section, we complete the paper with a short discussion.

Theorem 3. *Let G be any $\{C_{1,r}, C^*_{1,r}\}$-free linear r-graph on n vertices, and let s denote the number of vertices with degree at least $(r-1)^2+2$. Then $|E(G)| \leq \frac{r(r-2)(n-s)}{r-1}$.*

3 Proof of Theorem 3

For the full proof see manuscript [10].

Before we present our proof, we need some additional notation, and we prove a key lemma. Let H be a linear r-graph, let $d_1 \geq d_2 \geq \ldots \geq d_r$ be positive integers, and let $e \in E(H)$. Then we use $D(e) \geq \{d_1, d_2, \ldots, d_r\}$ to denote that e can be written as $e = \{u_1, u_2, \ldots, u_r\}$ such that $d(u_i) \geq d_i$ for each $i \in [r] = \{1, 2, \ldots, r\}$. Here $d(v)$ denotes the degree, i.e., the number of edges containing the vertex v. We use the shorthand v-edge for an edge containing the vertex v.

Lemma 4. *Let G be a $\{C_{1,r}, C^*_{1,r}\}$-free linear r-graph, and let $e \in E(G)$ be such that $D(e) \geq \{(r-1)^2+1, (r-1)^2+1, (r-1)^2, \ldots, (r-1)^2\}$. Then*

$$S = \bigcup_{f \in E(G), f \cap e \neq \emptyset} f$$

contains exactly $(r-1)^3+r$ vertices, and all vertices in S have degree at most $(r-1)^2+1$. Moreover,

$$E_S = \{f : f \in E(G), f \cap S \neq \emptyset\}$$

contains at most $r(r-1)^2+1$ edges.

Proof. Without loss of generality, suppose $e = \{u_1, u_2, \ldots, u_r\}$ with $d(u_1) \geq d(u_2) \geq (r-1)^2+1$ and $d(u_i) \geq (r-1)^2$ for each $3 \leq i \leq r$. If $d(u_1) \geq (r-1)^2+2$, we can find a copy of $C_{1,r}$ in the following way. We start with the edge $e = \{u_1, u_2, \ldots, u_r\}$. We can find a u_r-edge $e_1 \neq e$ since $d(u_r) \geq (r-1)^2$. By considering i from $r-1$ to 2 one by one, we can find a u_i-edge e_{r-i+1} that does not share a vertex with any edge in $\{e_1, e_2, \ldots, e_{r-i}\}$. Finally, we can choose a u_1-edge e_r that does not share a vertex with $e_1, e_2, \ldots, e_{r-1}$. Hence, we have found a copy of $C_{1,r}$, a contradiction.

Therefore, we have $d(u_1) = d(u_2) = (r-1)^2+1$. For $p \in \{u_1, u_2, \ldots, u_r\}$, we use $G(p)$ to denote the set of all vertices outside e that lie on a common edge with p. Firstly, we have the following claim. (Due to page limitations, we omit the proofs for the following claims.)

Claim 3.1. $G(u_1) = G(u_2)$.

Similarly, we must have $G(u_i) \subset G(u_2)$ for each $3 \leq i \leq r$. Suppose to the contrary that there exists some $3 \leq i \leq r$ such that there is a u_i-edge $e_i \neq e$ containing some vertex not in $G(u_2)$. Then there are at most $r-2$ u_2-edges other than e intersecting e_i, so there are at least $(r-2)(r-1)+1$ u_2-edges that are disjoint from e_i. By the edge conditions that $d(u_1) \geq (r-1)^2+1$ and $d(u_s) \geq (r-1)^2$ for each $3 \leq s \leq r$, for each s satisfying the conditions $1 \leq s \leq r, s \neq 2$ and $s \neq i$ we can choose a u_s-edge e_s that is disjoint from $\{e_1, e_3, \ldots, e_{s-1}\}$, and then choose a u_2-edge e_2 that is disjoint from $\{e_1, e_3, \ldots, e_r\}$. So $\{e, e_1, e_2, \ldots, e_r\}$ forms a $C_{1,r}$, a contradiction.

Thus $S \setminus \{u_1, u_2, \ldots, u_r\} = G(u_2) = G(u_1) \supset G(u_i)$ for each $3 \leq i \leq r$. Denote by F the edge set each edge of which is disjoint from $\{u_1, u_2, \ldots, u_r\}$ and contains at least one vertex of S. It suffices to show that F must be empty.

For this purpose, we first construct $r-1$ auxiliary bipartite graphs as follows. Fix an h with $2 \leq h \leq r$, and let $H_h = (V_{H_h} = X_{H_h} \cup Y_{H_h}, E_{H_h})$, where $X_{H_h} = \{e_i | u_h \in e_i, e_i \neq e\}$, $Y_{H_h} = \{e_j | u_1 \in e_j, e_j \neq e\}$ and $E_{H_h} = \{\{e_i, e_j\} | e_i \cap e_j \neq \emptyset\}$. Then H_2 is an $(r-1)$-regular bipartite graph with partition classes of exactly $(r-1)^2$ vertices. For $3 \leq h \leq r$, H_h is a bipartite graph with one class of exactly $(r-1)^2$ vertices and the other class having at least $(r-1)^2 - 1$ vertices. Next, we prove two claims on the structure of these bipartite graphs.

Claim 3.2. *If G is $C_{1,r}$-free, then H_2 must contain a $K_{r-1,r-1}$.*

Claim 3.3. *If G is $C_{1,r}$-free, then H_h must contain a $K_{r-2,r-1}$ for each $2 \leq h \leq r$. Furthermore, the partition classes on $r-1$ vertices in these $K_{r-2,r-1}$'s are mutually disjoint.*

Let $\{e_1, e_2, \ldots, e_{(r-1)^2}\}$ denote the ordered sequence of all u_1-edges except for e. Without loss of generality, we assume that H_h contains the $(h-1)$-th $r-1$ u_1-edges of this sequence for $2 \leq h \leq r$. That means H_h contains $e_{(h-2)(r-1)+1}, e_{(h-2)(r-1)+2}, \ldots, e_{(h-1)(r-1)}$ for each $2 \leq h \leq r$. Denote by U_{h-1} the set of vertices in the $(h-1)$-th $r-1$ u_1-edges of the sequence for $2 \leq h \leq r$. We have another claim.

Claim 3.4. *Fix $2 \leq i \leq r$. Each u_i-edge contains only vertices of one vertex set from $\{U_1, U_2, \ldots, U_{r-1}\}$.*

Before we continue with the proof of Lemma 4, we note that the above analysis implies the following about the structure of H_i.

Remarks 3.1. *H_2 is the disjoint union of $r-1$ complete bipartite graphs $K_{r-1,r-1}$. Since $d(u_h) \geq (r-1)^2$ for each $3 \leq h \leq r$, H_h is either the disjoint union of $r-1$ complete bipartite graphs $K_{r-1,r-1}$ or the disjoint union of $r-2$ complete bipartite graphs $K_{r-1,r-1}$ and one complete bipartite graph $K_{r-2,r-1}$.*

As a consequence of Remarks 3.1, for each $1 \leq i \leq r-1$ there exist $r-1$ u_2-edges whose vertices except for u_2 are in U_i. Fix h with $3 \leq h \leq r$. There exists at most one s with $1 \leq s \leq r-1$ such that there exist $r-2$ u_h-edges whose vertices except for u_h are in U_s. For each $1 \leq i \neq s \leq r-1$, there exist $r-1$ u_h-edges whose vertices except for u_h are in U_i.

Now we are ready to prove the statement about F. If F is not an empty set, we let f be an edge of F. There must exist an s with $1 \leq s \leq r-1$ such that $|f \cap U_s| \geq 1$. Let $v \in f \cap U_s$. We choose a u_1-edge g containing v. By Remarks 3.1, there exist $r-2$ u_t-edges $g_1, g_2, \ldots, g_{r-2}$ with the property that each of them is disjoint from f and each of them intersects g. And there must exist another u_1-edge g' whose vertices except for u_1 are in U_t for some $1 \leq t \neq s \leq r-1$ such that g' is disjoint from f. Now the edges $f, g, g', g_1, g_2, \ldots, g_{r-2}$ constitute a $C^*_{1,r}$, a contradiction. This completes the proof of Lemma 4. □

Now we are ready to prove Theorem 3. Suppose to the contrary that G is a smallest (in terms of the number of vertices n) $\{C_{1,r}, C^*_{1,r}\}$-free linear r-graph such that G has more than $\frac{r(r-2)(n-s)}{r-1}$ edges. For each $v \in V(G)$, we define $I(v) = 1$ if $d(v) \leq (r-1)^2 + 1$, and $I(v) = 0$ otherwise.

We adopt the following useful observation from [9].

$$\sum_{e\in E(G)}\sum_{v\in V(G),v\in e}\frac{I(v)}{d(v)}=\sum_{v\in V(G)}\sum_{e\in E(G),v\in e}\frac{I(v)}{d(v)}=\sum_{v\in V(G)}I(v)=n-s.$$

Since $|E(G)|>\frac{r(r-2)(n-s)}{r-1}$, there must exist an edge $e=\{u_1,u_2,\ldots,u_r\}$ such that

$$\sum_{1\le i\le r}\frac{I(u_i)}{d(u_i)}<\frac{r-1}{r(r-2)}=\frac{r-1}{(r-1)^2-1}. \tag{1}$$

Without loss of generality, we assume $d(u_1)\ge d(u_2)\ge\ldots\ge d(u_r)$. Note that $d(u_r)\ge r-1$ and $d(u_2)\ge(r-1)^2$, as otherwise (1) would be violated. We can also deduce that $d(u_i)\ge(r-i)(r-1)+2$ for all $3\le i\le r-1$, as otherwise (1) would be violated. If $d(u_1)\ge(r-1)^2+2$, then we can easily find a $C_{1,r}$ in the following way. We start with the edge $e=(u_1,u_2,\ldots,u_r)$. We can find a u_r-edge $e_1\neq e$ since $d(u_r)\ge 2$. By considering i from $r-1$ to 2 one by one, we can find a u_i-edge e_{r-i+1} that does not share a vertex with any edge in $\{e_1,e_2,\ldots,e_{r-i}\}$. Finally, we can choose a u_1-edge e_r that does not share a vertex with $\{e_1,e_2,\ldots,e_{r-1}\}$, a contradiction. Therefore, we have $d(u_1)\le(r-1)^2+1$. By (1), we have $d(u_1)=d(u_2)=(r-1)^2+1$ and $d(u_i)\ge(r-1)^2$ for each $3\le i\le r$. Thus, $D(e)\ge\{(r-1)^2+1,(r-1)^2+1,(r-1)^2,\ldots,(r-1)^2\}$.

Now we define S and E_S as in Lemma 4. Let $G-S$ be the linear r-graph obtained by deleting the vertices of S and the edges of E_S. By Lemma 4, $G-S$ has $n'=n-((r-1)^3+r)$ vertices and at least $|E(G)|-(r(r-1)^2+1)$ edges. Furthermore, the number of vertices in $G-S$ of degree at least $(r-1)^2+2$ is exactly s. Therefore, we have

$$|E(G-S)|\ge|E(G)|-(r(r-1)^2+1)>\frac{r(r-2)(n-s)}{r-1}-(r(r-1)^2+1)>\frac{r(r-2)(n'-s)}{r-1},$$

which contradicts the assumption that G is a smallest counterexample to Theorem 3.

This completes the proof.

References

[1] A. Carbonero, W. Fletcher, J. Guo, A. Gyárfás, R. Wang, S. Yan, Crowns in linear 3-graphs, 2017, *arXiv:2017.14713*.

[2] A. Carbonero, W. Fletcher, J. Guo, A. Gyárfás, R. Wang, S. Yan, Crowns in linear 3-graphs of minimum degree 4, *The Electronic Journal of Combinatorics* **29**(4):#P4.17 (2022).

[3] C.J. Colbourn, J.H. Dinitz, *Handbook of Combinatorial Designs*, Second Edition, CRC Press, Boca Raton, Fl., (2007).

[4] W. Fletcher, Improved upper bound on the linear Turán number of the crown, 2021, *arXiv:2109.02729v1*.

[5] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, Linear Turán numbers of acyclic triple systems, *European Journal of Combinatorics* **99** (2022).

[6] A. Gyárfás, G.N. Sárközy, Turán and Ramsey numbers in linear triple systems, *Discrete Mathematics* **344** (3) (2021) 112258.

[7] A. Gyárfás, G.N. Sárközy, The linear Turán number of small triple systems or why is the wicket interesting? *Discrete Mathematics* **345** (11)(2022) 113025.

[8] G.N. Sárközy, Turán and Ramsey numbers in linear triple systems II, *Discrete Mathematics* **346** (1) (2023) 113182.

[9] C. Tang, H. Wu, S. Zhang, Z. Zheng, On the Turán number of the linear 3-graph C_{13}, *The Electronic Journal of Combinatorics* **29**(3):#P3.46 (2022).

[10] L.-P. Zhang, H. Broersma, L. Wang, A note on generalized crowns in linear r-graphs, 2024, `arXiv:2401.12339v1`.

A Kneser-type theorem for restricted sumsets*

Mario Huicochea †1

[1]CONACYT/UAZ, Zacatecas, Mexico

Abstract

Let $G = (G, +, 0_G)$ be a commutative group, A and B be nonempty finite subsets of G and $H = \{c \in G : \ c + A + B = A + B\}$. Kneser's Theorem is a fundamental result in Additive Number Theory and it establishes that $|A + B| \geq |A + H| + |B + H| - |H|$. For any subset S of $A \times B$, write $A \overset{S}{+} B = \{a+b : \ (a, b) \in S\}$. For any $c \in G$, set $r_{A,B}(c) = |\{(a, b) \in A \times B : \ a+b = c\}|$. An important problem in Additive Number Theory is to find a Kneser-type theorem for the restricted sumsets $A \overset{S}{+} B$. In particular, more than 20 years ago V. Lev proved that if $\{c \in A+B : \ r_{A,B}(c) \geq k\} \subseteq A \overset{S}{+} B$, for all $a \in A$ (resp $b \in B$) there is at most one $b' \in B$ (resp. $a' \in A$) such that $(a, b') \notin S$ (resp. $(a', b) \notin S$), and $A \overset{S}{+} B \neq A + B$, then

$$\left| A \overset{S}{+} B \right| > \left(1 - \frac{|A||B|}{(|A| + |B|)^2} \right) (|A| + |B|) - k - 1.$$

In the same paper, Lev proposed as a problem to improve $1 - \frac{|A||B|}{(|A|+|B|)^2}$ to something of the form $1 - w$ with $w \to 0$ whenever $\frac{|(A\times B)\setminus S|}{|A||B|} \to 0$. Lev's problem has been solved for some particular groups and some specific subsets S of $A \times B$. However, it remains open for arbitrary groups and arbitrary large subsets S of $A \times B$. Here, as a consequence of the main result of this paper, it is shown that if we take $-2k - s + 2$ instead of $-k - 1$ in the lower bound of $\left| A \overset{S}{+} B \right|$, then indeed we can take as the coefficient of $|A| + |B|$ something of the form $1 - w$ with $w \to 0$ whenever $\frac{|(A\times B)\setminus S|}{|A||B|} \to 0$.

1 Introduction

In this paper $\mathbb{R}, \mathbb{Z}, \mathbb{Z}^+, \mathbb{Z}_0^+$ denote the set of real numbers, integers, positive integers and nonnegative integers, respectively. Let $G = (G, +, 0_G)$ be a commutative group, H be a subgroup of G, A and B be subsets of G, $c \in G$ and $k \in \mathbb{Z}^+$. Write

$$\begin{aligned} A + B &:= \{a + b : \ a \in A, \, b \in B\} \\ A + c &:= A + \{c\} \\ -A &:= \{-a : \ a \in A\} \\ r_{A,B}(c) &:= |\{(a, b) \in A \times B : \ a + b = c\}| \\ A \overset{k}{+} B &:= \{d \in G : \ r_{A,B}(d) \geq k\} \\ \mathrm{Stab}(A) &:= \{b \in G : \ A + b = A\}. \end{aligned}$$

We will denote by $\pi_H : G \to G/H$ the canonical projection. To avoid confusion, $\pi_H(c)$ (resp. $\pi_H(A)$) will be an element (resp. subset) in the quotient group G/H, while $c + H$ (resp. $A + H$) will denote a subset of G.

*The full version of this work can be found in [6].

†Email: dym@cimat.mx Research of CONAHCYT/UAZ

One of the most important problems in Additive Number Theory is to find a sharp lower bound for a sumset in terms of the size of the sets and the technical properties demanded for these sets. A fundamental result in this direction is Kneser's Theorem which can be stated as follows.

Theorem 1. *Let G be a commutative group and A and B be nonempty finite subsets of G. Write $H = \mathrm{Stab}(A+B)$. Then*

$$|A+B| \geq |A+H| + |B+H| - |H|.$$

Proof. See [17, Thm.5.5]. □

An easy consequence of Kneser's Theorem is the next result.

Corollary 2. *Let G be a commutative group and A and B be nonempty finite subsets of G such that $|A+B| < |A| + |B| - 1$. Write $H = \mathrm{Stab}(A+B)$. Then*

$$|\pi_H(A+B)| = |\pi_H(A)| + |\pi_H(B)| - 1.$$

Proof. See [2, Ch.6]. □

There are a number of proofs, generalizations and applications of Kneser's Theorem; see [2, 12, 17].

Let G be a commutative group, A and B be nonempty subsets of G and S be a subset of $A \times B$. The restricted sumset of A and B by S is

$$A \overset{S}{+} B := \{a+b : \ (a,b) \in S\}.$$

Let $s \geq 0$ and $k \in \mathbb{Z}^+$. We say that $A \overset{S}{+} B$ is *s-regular* if for all $a \in A$ and $b \in B$, we have that $|\{b' \in B : \ (a,b') \notin S\}|, |\{a' \in A : \ (a',b) \notin S\}| \leq s$. We say that $A \overset{S}{+} B$ is *(k,s)-regular* if $A \overset{S}{+} B$ is s-regular and $A \overset{k}{+} B \subseteq A \overset{S}{+} B$. There are several problems where instead of considering the sum of each pair of elements in $A \times B$, we want to take just some of them. In particular, the cases where $S = \{(a,b) \in A \times B : \ a \neq b\}$ or $S = \{(a,b) \in A \times B : \ r_{A,B}(a+b) \geq n\}$ for a given $n \in \mathbb{Z}^+$ have been widely studied; see for example [1, 7, 8, 10, 13, 14, 17, 18]. Also for arbitrary large subsets S of $A \times B$, a number of results can be found nowadays; see [9, 11, 15, 16, 17]. An important problem in this area is to try to generalize Kneser's Theorem for restricted sumsets. If $A \overset{S}{+} B = A + B$, then Kneser's Theorem can be used to find a lower bound for $\left|A \overset{S}{+} B\right|$ in terms of $|A|, |B|$ and $|\mathrm{Stab}(A+B)|$. Thus it remains to study how can we bound $\left|A \overset{S}{+} B\right|$ below when $A \overset{S}{+} B \neq A + B$. An important step in this direction was given by V. Lev with the next theorem (which is stated in [8] with slightly different notation).

Theorem 3. *Let $k \in \mathbb{Z}^+$, G be a commutative group, A and B be nonempty finite subsets of G and S be a subset of $A \times B$ such that $A \overset{S}{+} B$ is $(k,1)$-regular. Write $w = \frac{|A||B|}{(|A|+|B|)^2}$. If $A \overset{S}{+} B \neq A + B$, then*

$$\left|A \overset{S}{+} B\right| > (1-w)(|A| + |B|) - k - 1.$$

Proof. See [8, Thm.4]. □

With the notation as in Theorem 3, notice that $w \leq \frac{1}{4}$ and the equality is achieved when $|A| = |B|$. Thus the coefficient $1 - w$ can be as small as $\frac{3}{4}$. In [8, Sec.4], Lev proposed as a problem to improve $1 - w$ in Theorem 3. There are already partial results.

- In the case $G = \mathbb{Z}/p\mathbb{Z}$, S. Guo and Z. W. Sun gave in [4] a lower bound for $\left|A\overset{S}{+}B\right|$ when $S = \{(a,b) \in A \times B : \ a - b \notin C\}$ for a subset C of $\mathbb{Z}/p\mathbb{Z}$.
- When $G = \mathbb{Z}$, Lev gave in [9] a lower bound for $\left|A\overset{S}{+}A\right|$. Later P. Mazur in [11] and X. Shao and W. Xu in [16] found nontrivial lower bounds for $\left|A\overset{S}{+}B\right|$ and inverse results in this direction.
- In the case G is torsion free or elementary abelian, H. Pan and Sun provided a nontrivial lower bound $\left|A\overset{S}{+}B\right|$ when $S = \{(a,b) \in A \times B : \ a - b \notin C\}$ for a subset C of G.
- For arbitrary finite commutative groups G, Lev in [7] and Guo in [3] gave lower bounds for $\left|A\overset{S}{+}B\right|$ when $S = \{(a,b) \in A \times B : \ a \neq b\}$. Later this was generalized by Y. O. Hamidoune, S. C. López and A. Plagne in [5].

More information about this topic can be found in Lev's nice survey [10]. Lev's problem remains open for arbitrary groups and large subsets S of $A \times B$, and this problem is the main motivation of this paper. Instead of considering just $(k,1)$-regular restricted sumsets, we will work with (k,s)-regular restricted sumsets.

2 Main results

To state the main result of this paper, we need two definitions. Let G be a commutative group, A and B be nonempty finite subsets of G and $m \in \{1, 2, \ldots, \min\{|A|, |B|\}\}$.

⋆ We say that (A, B, m) is a *Pollard triple* if

$$\sum_{k=1}^{m} \left|A\overset{k}{+}B\right| \geq m\,|A| + m\,|B| - 2m^2 + 3m - 2.$$

⋆ We say that (A, B, m) is a *Kneser triple* if there is a subset A' of A and a subset B' of B satisfying

$$\left|A \setminus A'\right| + \left|B \setminus B'\right| \leq m - 1$$

and

$$A'\overset{m}{+}B' = A' + B' = A\overset{m}{+}B.$$

For $m \in \{1, 2, \ldots, \min\{|A|, |B|\}\}$, a result of D. Grynkiewicz establishes that (A, B, m) is either Pollard or Kneser.

Theorem 4. *Let $s \geq 1$, $u \in [0, 1)$, G be a commutative group, A and B be nonempty finite subsets of G, $k \in \{2, 3, \ldots, \min\{|A|, |B|\}\}$ and S be a subset of $A \times B$ such that $|S| \geq (1-u)|A||B|$ and $A\overset{S}{+}B$ is (k,s)-regular. Assume that $A\overset{S}{+}B \neq A + B$.*

i) If $k \leq \sqrt{\frac{u|A||B|}{2}}$ and $\left(A, B, \left\lceil\sqrt{\frac{u|A||B|}{2}}\right\rceil\right)$ is a Pollard triple, then

$$\left|A\overset{S}{+}B\right| \geq |A| + |B| - \sqrt{8u|A||B|} - 2.$$

ii) If $k \leq \sqrt{\frac{u|A||B|}{2}}$ *and* $\left(A, B, \left\lceil \sqrt{\frac{u|A||B|}{2}} \right\rceil\right)$ *is a Kneser triple, then*

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - \sqrt{\frac{u|A||B|}{2}} - s.$$

iii) If $k > \sqrt{\frac{u|A||B|}{2}}$ *and* (A, B, k) *is a Pollard triple, then*

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - \frac{u|A||B|}{k} - 2k.$$

iv) If $k > \sqrt{\frac{u|A||B|}{2}}$ *and* (A, B, k) *is a Kneser triple, then*

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - k - s + 1.$$

If $k \leq \sqrt{\frac{u|A||B|}{2}}$, then i) and ii) in Theorem 4 lead to

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - \sqrt{8u|A||B|} - s - 2. \tag{1}$$

If $k > \sqrt{\frac{u|A||B|}{2}}$, then iii) and iv) in Theorem 4 imply

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - \frac{u|A||B|}{k} - 2k - s + 3 \geq |A| + |B| - \sqrt{2u|A||B|} - 2k - s + 3. \tag{2}$$

Using that $|A| + |B| \geq 2\sqrt{|A||B|}$, we get from (1) and (2) the next corollary.

Corollary 5. *Let* $s \geq 1$, $u \in [0, 1)$, G *be a commutative group,* A *and* B *be nonempty finite subsets of* G, $k \in \{2, 3, \ldots, \min\{|A|, |B|\}\}$ *and* S *be a subset of* $A \times B$ *such that* $|S| \geq (1-u)|A||B|$ *and* $A \overset{S}{+} B$ *is* (k, s)*-regular. Assume that* $A \overset{S}{+} B \neq A + B$. *Then*

$$\left|A \overset{S}{+} B\right| \geq \left(1 - \sqrt{2u}\right)(|A| + |B|) - 2k - s + 2.$$

Corollary 5 is a nontrivial step in the solution of Lev's problem, i.e. to solve the problem, it would be enough to have $-k - 1$ instead of $-2k - s + 2$ in the lower bound of $\left|A \overset{S}{+} B\right|$.

We sketch the proof of Theorem 4.

i) For $m \in \{1, 2, \ldots, \min\{|A|, |B|\}\}$, a result of D. Grynkiewicz, see [2, Thm.12.1], establishes that either the triple (A, B, m) is a Pollard triple or it is a Kneser triple.

ii) Assume that (A, B, m) is a Pollard triple. It is proven that

$$u|A||B| + m\left|A \overset{S}{+} B\right| \geq m\,|A| + m\,|B| - 2m^2,$$

which implies the claim of the theorem in this case. This crucial lemma is proven using some auxiliary subsets, partitions and elementary combinatorial arguments.

iii) Assume that (A, B, m) is a Kneser triple. Then there is a subset A' of A and a subset B' of B satisfying that

$$\left|A \setminus A'\right| + \left|B \setminus B'\right| \leq m - 1$$

and

$$A' \overset{m}{+} B' = A' + B' = A \overset{m}{+} B.$$

It is shown that

$$\left|A \overset{S}{+} B\right| \geq |A| + |B| - m - 1 - s.$$

This lemma is proven using partitions, projections and Kneser's Theorem.

References

[1] B. Bajnok, On the minimum size of restricted sumsets in cyclic groups, *Acta Math. Hungar.* **148** (2016), 228-256.

[2] D. J. Grynkiewicz, *Structural Additive Theory*, Developments in Mathematics 30 Springer, 2013.

[3] S. Guo, Restricted sumsets in a finite Abelian group, *Discrete Math.* **309** (2009), 6530-6534.

[4] S. Guo, Z. W. Sun, A variant of Tao's method with application to restricted sumsets, *J. Number Theory* **129** (2009), 434-438.

[5] Y. O. Hamidoune, S. C. López, A. Plagne, Large restricted sumsets in general Abelian groups, *European J. Combin.* **34** (2013), 1348-1364.

[6] M. Huicochea, A Kneser-type theorem for restricted sumsets, *SIAM J. Discrete Math.* **37** (2023), 83-93.

[7] V. Lev, Restricted set addition in groups, I. The classical setting, *J. London Math. Soc.* **62** (2000), 27-40.

[8] V. Lev, Restricted set addition in groups, II. A generalization of the Erdös-Heilbronn conjecture, *Electron. J. Combin.* **7** (2000) Paper 4, 1-10.

[9] V. Lev, Restricted set addition in groups, III. Integer sumsets with generic restrictions, *Periodica Math. Hungarica* **42** (2001), 89-98.

[10] V. Lev, Restricted set addition in abelian groups: results and conjectures, *J. Théor. Nombres Bordeaux* **17** (2005), 181-193.

[11] P. Mazur, A structure theorem for sets of small popular doubling, *Acta Arith.* **171** (2015), 221-239.

[12] M.B. Nathanson, *Additive Number Theory. Inverse Problems and the Geometry of Sumsets*, Grad. Texts in Math., vol. 165, Springer-Verlag, New York, 1996.

[13] H. Pan and Z. Wei, Restricted sumsets and a conjecture of Lev, *Israel J. Math.* **154** (2006), 21-28.

[14] F. Petrov, Restricted product sets under unique representability, *Mosc. J. Comb. Number Theory* **7** (2017), 73-78.

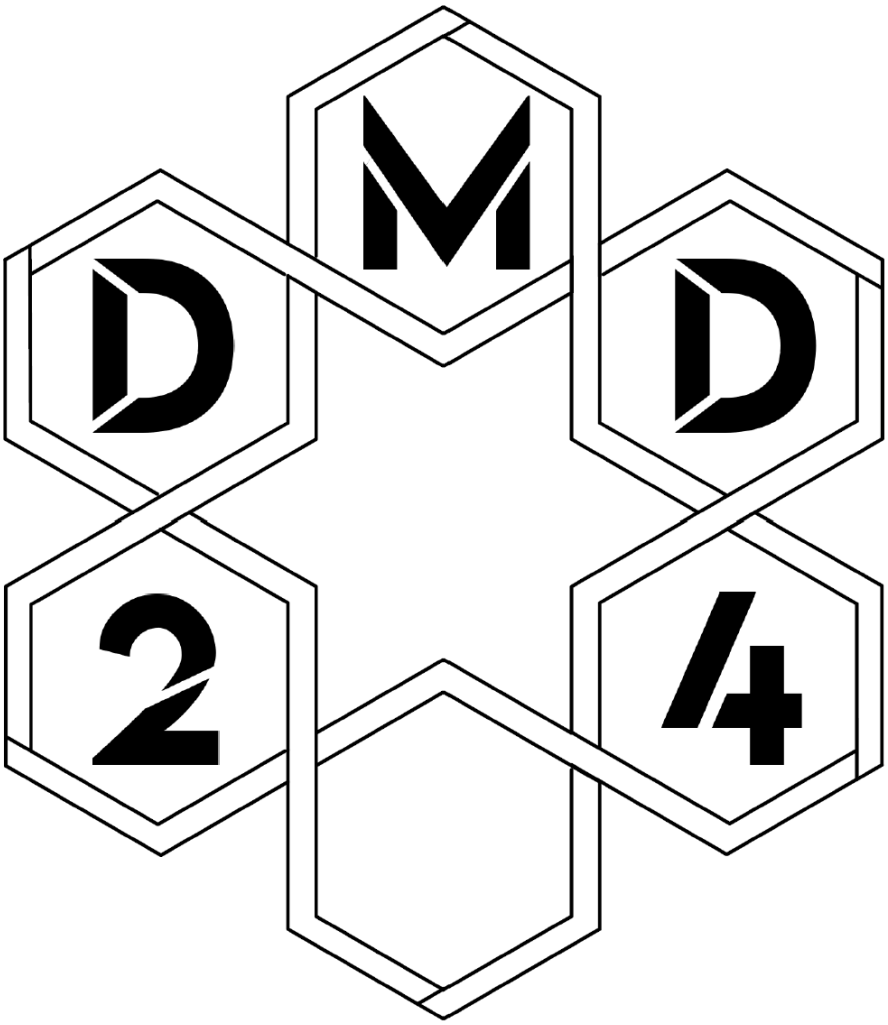
[15] X. Shao, On an almost all version of the Balog-Szemerédi-Gowers theorem, *Discrete Anal.* (2019) Paper 12, 1-18.

[16] X. Shao and W. Xu, A robust version of Freiman's 3k-4 theorem and applications, *Math. Proc. of the Cambridge Philos. Soc.* **166** (2019), 567-581.

[17] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Studies in Advanced Mathematics 105 Cambridge University Press, 2006.

[18] V. Vu and P. Wood, The inverse Erdös-Heilbronn problem, *Electron. J. Combin.* **16** (2009) Paper 100, 1-8.

Discrete Mathematics Days 2024



Universidad de Alcalá

Universidad de Alcalá
DEPARTAMENTO DE FÍSICA Y MATEMÁTICAS

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
Departament de Matemàtiques

DAM network

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN