

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Mathematics

2023

Challenges and Results in Extremal Combinatorics

Alweiss, Ryan

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp01x633f429v>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

CHALLENGES AND RESULTS IN EXTREMAL COMBINATORICS

Ryan Alweiss

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY

THE DEPARTMENT OF MATHEMATICS

Adviser: Noga Alon

May 2023

© Copyright by Ryan Alweiss, 2023. All rights reserved.

1 Abstract

In this thesis, we address several questions in extremal, probabilistic, and additive combinatorics, with applications to theoretical computer science.

We start by addressing the *sunflower lemma* of Erdős and Rado, and some related problems about set systems. A sunflower with r petals is a collection of r sets so that the intersection of each pair is equal to the intersection of all of them. Erdős and Rado proved the sunflower lemma: for any fixed r , any family of sets of size w with at least about w^w sets must contain a sunflower with r petals. The sunflower conjecture states that the bound on the number of sets can be improved to c^w for some constant c . In work with Lovett, Wu, and Zhang, we improve the bound to approximately $(\log w)^w$.

Next, we address a couple of problems about querying in graphs. We answer a question of Alon, Mossel, and Pemantle about the corruption detection model on graphs in the noisy setting. In work with He, Ben Hamida, and Moreira, we make progress on the *subgraph query problem* introduced by Ferber, Krivelevich, Sudakov, and Vieira.

We then discuss a question related to information theory, nailing down the “product dimension” $Q(s, r)$ of the vertex disjoint union of r cliques, each of size s .

Then, in joint work with Liu and Sawhney, we study discrepancy minimization for vectors in $\mathbb{R}^n$. Our main result is the analysis of a new simple random process in multiple dimensions, which we call a “self-balancing walk”, through a coupling argument. As a corollary, we obtain linear time algorithms for logarithmic bounds for the Komlós conjecture.

The final part of the thesis concerns a few questions in additive combinatorics. First, we solve a conjecture of Gowers from 2009 on clique differences in chains, ruling out any Sperner-type proof of the polynomial density Hales-Jewett Theorem for alphabets of size 2. In particular, this rules out Gowers’ proof strategy for polynomial density Hales-Jewett. Finally, we show primitive recursive bounds for the partition regularity of $\{x, x + y, xy\}$.

Contents

1	Abstract	3
I	Introduction	7
2	Acknowledgments	8
II	Improved Bounds for the Sunflower Lemma	8
3	Introductory Remarks	9
3.1	Robust Sunflowers	10
3.2	Proof Overview	13
3.3	Chapter Organization.	14
4	Proof of Theorem 3.9	14
4.1	A Reduction Step	17
4.2	A Final Step	20
4.3	Putting Everything Together	22
5	A Lower Bound for Robust Sunflowers	24
6	Subsequent Works and Corollaries	26
6.1	The Erdős-Szemerédi Sunflower Conjecture.	26
6.2	Intersecting Set Systems.	27
7	Rainbow Sunflowers	27

III	Blowing up Sunflowers	32
8	Introductory Remarks	32
8.1	Background and Main Result	32
8.2	Proof Outline	33
8.3	Blowing up Pairs of Sets	33
8.4	Sunflower Blowups	34
9	Proofs	35
9.1	Proof of Lemma 9.3	36
9.2	Proof of Lemma 9.4	37
9.3	Proof of Theorem 8.2	38
9.4	Proof of Corollary 8.4	38
9.5	Proof of Theorem 8.5	39
9.6	Proof of Corollary 8.7	39
IV	Noisy Corruption Detection	40
10	Introductory Remarks	40
11	Lower Bound for Simple Majority	41
12	Upper Bound for Robust Majority	44
13	Discussions and Open Problems	46
V	The Subgraph Query Problem	48
14	Introductory Remarks	48

15 Upper bounds	51
15.1 An Illustrative Example	51
15.2 The General Upper Bound	54
16 Lower bounds	58
16.1 Preliminaries	58
16.2 Proof of Theorem 14.2	62
17 Cliques in $G(n, 1/2)$	67
18 Discussions and Open Problems	71
 VI On The Product Dimension of Clique Factors	 73
19 Introductory Remarks	73
20 Preliminaries	76
21 Covering families	78
21.1 The Upper Bound	78
21.2 Prime s	79
21.3 General s	83
22 Concluding remarks and Open Problems	84
 VII Discrepancy Minimization via a Self-Balancing Walk	 91
23 Introductory Remarks	91
23.1 Online Algorithms	92
23.2 Algorithm Motivation and Description	93

23.3	Consequences of Theorem 23.1	94
23.4	Previous Approaches for Algorithmic Discrepancy Minimization	95
23.5	Overview of Analysis of the Self-Balancing Walk	96
23.6	Preliminaries and Conventions	96
23.7	Concurrent and Independent Work	96
24	Analysis	97
24.1	Properties of Spreading	97
24.2	Proof of Theorem 23.1	99
VIII	On a Conjecture of Gowers on Clique Differences	105
25	Introductory Remarks	105
26	Proof of Theorem 25.3	106
IX	Monochromatic Sums and Products of Polynomials	108
27	Introduction	108
28	The 3-color case	111
29	Partition Regularity of Sums and Products in T	112
30	Discussions and Open Problems	114
X	Bibliography	118

Part I

Introduction

2 Acknowledgments

I would first like to acknowledge my advisor, Noga Alon, for being the greatest advisor ever.

Thanks to my coauthors Noga Alon, Shachar Lovett, Yang Liu, Mehtaab Sawhney, Kewen Wu, Jiapeng Zhang.

I would also like to thank Jacob Fox, Peter Frankl, Timothy Gowers, Meghal Gupta, Mohammadmahdi Jahanara, Vishesh Jain, Gil Kalai, Noah Kravitz, Andrey Kupavskii, Alexander Moreira, Joel Moreira, Huy Pham, Ashwin Sah, Aaron Sidford, Matthew Wales, Yuval Wigderson, Yufei Zhao, and several anonymous reviewers for conferences and journals for their feedback on and discussion of various parts of this thesis.

Furthermore, I would like to thank all my friends and colleagues at Princeton, MIT, and elsewhere.

And last, but certainly not least, thanks to my mother Gayle, my father Gary, and my brother Max. I love you guys.

Part II

Improved Bounds for the Sunflower

Lemma

This chapter is based on joint work with Shachar Lovett, Kewen Wu, and Jiapeng Zhang, and has been published in the *Annals of Mathematics* [9]. It also won Best Paper Award at the ACM Symposium on the Theory of Computing (STOC) conference.

3 Introductory Remarks

Let X be a finite set. A set system $\mathcal{F}$ on X is a collection of subsets of X . We call $\mathcal{F}$ a w -set system if each set in $\mathcal{F}$ has size at most w .

Definition 3.1 (Sunflower). *A collection of sets $S_1, \dots, S_r$ is an r -sunflower if*

$$S_i \cap S_j = S_1 \cap \dots \cap S_r, \quad \forall i \neq j.$$

We call $K = S_1 \cap \dots \cap S_r$ the kernel and $S_1 \setminus K, \dots, S_r \setminus K$ the petals of the sunflower.

Erdős and Rado [36] proved that large enough set systems must contain a sunflower. Erdős and Rado originally called sunflowers Δ -systems, but the term “sunflower” was coined by Deza and Frankl [34] and is now more widely used.

Lemma 3.2 (Sunflower lemma [36]). *Let $r \geq 3$ and $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| \geq w! \cdot (r-1)^w$. Then $\mathcal{F}$ contains an r -sunflower.*

Erdős and Rado conjectured in the same paper that the bound in Lemma 3.2 can be drastically improved.

Conjecture 3.3 (Sunflower conjecture [36]). *Let $r \geq 3$. There exists $c = c(r)$ such that any w -set system $\mathcal{F}$ of size $|\mathcal{F}| \geq c^w$ contains an r -sunflower.*

The bound in Lemma 3.2 is of the form $w^{w(1+o(1))}$ where the $o(1)$ term depends on r . Despite nearly 60 years of research, the best known bounds were still of the form $w^{w(1+o(1))}$, even for $r = 3$. Kostochka [57] proved that any w -set system of size $|\mathcal{F}| \geq cw! \cdot (\log \log \log w / \log \log w)^w$ must contain a 3-sunflower for some absolute constant c .

In this chapter, we vastly improve the known bounds. We prove that any w -set system of size $(\log w)^{w(1+o(1))}$ must contain a sunflower. More precisely, we obtain the following:

Theorem 3.4 (Main theorem, sunflowers). *Let $r \geq 3$. For some constant C , any w -set system $\mathcal{F}$ of size $|\mathcal{F}| \geq (Cr^3 \log w \log \log w)^w$ contains an r -sunflower.*

We will implicitly assume throughout the chapter that $\log \log w > 0$. Formally, to handle the case of $w = 2$, we interpret $\log$ as logarithm in base 1.9 throughout this chapter.

3.1 Robust Sunflowers

We consider a “robust” generalization of sunflowers, the study of which was initiated by Rossman [76], who was motivated by questions in complexity theory. Later, it was studied by Li, Lovett and Zhang [61] in the context of the sunflower conjecture.

First, we define a more “robust” version of the property of having disjoint sets. Given a finite set X , we denote by $\mathcal{U}(X, p)$ the distribution over subsets $R \subset X$, where each element $x \in X$ is included in R independently with probability p (this is sometimes referred to as a “ p -biased distribution”).

Definition 3.5 (Satisfying set system). *Let $0 < \alpha, \beta < 1$. A set system $\mathcal{F}$ on X is (α, β) -satisfying if*

$$\Pr_{R \sim \mathcal{U}(X, \alpha)} [\exists S \in \mathcal{F}, S \subset R] > 1 - \beta.$$

The explanation for the name “satisfying” is that if the set system is interpreted as a disjunctive normal form (DNF) formula, then this condition is that the formula has more than a $1 - \beta$ probability of being satisfied on α -biased inputs; see [65]. As mentioned before, the property of being satisfying is a robust analogue of the property of having disjoint sets.

Lemma 3.6 ([65, Claim 14]). *If $\mathcal{F}$ is a $(1/r, 1/r)$ -satisfying set system and $\emptyset \notin \mathcal{F}$, then $\mathcal{F}$ contains r pairwise disjoint sets.*

Proof. Let $\mathcal{F}$ be a set system on X . Consider a random coloring of elements of X with r colors, so that each $x \in X$ has an equal probability of being colored with any of the r colors, independently of the other elements. For $1 \leq i \leq r$, let Y_i denote the subset of X colored with color i and let $\mathcal{E}_i$ denote the event that $\mathcal{F}$ contains an i -monochromatic set, namely,

$$\mathcal{E}_i = [\exists S \in \mathcal{F}, S \subset Y_i].$$

Note that $Y_i \sim \mathcal{U}(X, 1/r)$, and since we assume $\mathcal{F}$ is $(1/r, 1/r)$ -satisfying, we have

$$\Pr[\mathcal{E}_i] > 1 - 1/r.$$

By the union bound, with positive probability all of $\mathcal{E}_1, \dots, \mathcal{E}_r$ hold. In this case, $\mathcal{F}$ contains a set which is i -monochromatic for each $i = 1, \dots, r$. Such sets must be pairwise disjoint. $\square$

Given a set system $\mathcal{F}$ on X and a set $T \subset X$, the *link* of $\mathcal{F}$ at T is

$$\mathcal{F}_T = \{S \setminus T : S \in \mathcal{F}, T \subset S\}.$$

We now formally define a *robust sunflower* (which was called a *quasi-sunflower* in [76] and an *approximate sunflower* in [65]).

Definition 3.7 (Robust sunflower). *Let $0 < \alpha, \beta < 1$, $\mathcal{F}$ be a set system, and let $K =$*

$\bigcap_{S \in \mathcal{F}} S$ be the common intersection of all sets in $\mathcal{F}$. $\mathcal{F}$ is an (α, β) -robust sunflower if (i) $K \notin \mathcal{F}$, and (ii) $\mathcal{F}_K$ is (α, β) -satisfying. We call K the kernel.

The kernel is not allowed to appear in the set system for technical reasons; we do not want a set system consisting of a single set to be a robust sunflower. In order to circumvent this technical obstacle, when relevant we will ensure that the set systems we work with are w -uniform for some w ; i.e., all sets in the system will be of size w . Any w -set system can be turned into a w -uniform system by adding $w - |S|$ distinct dummy elements to the sets S of the system with $|S| < w$ so that no dummy element appears in more than one set.

Robust sunflowers are a generalization of satisfying set systems. In particular, an (α, β) -satisfying w -uniform system $\mathcal{F}$ will be an (α, β) -robust sunflower as long as $\mathcal{F}$ contains at least two distinct sets, because if $R \supset S$ for some $S \in \mathcal{F}$, then clearly $R \supset S \setminus K$.

Lemma 3.8 ([65, Corollary 15]). *Any $(1/r, 1/r)$ -robust sunflower contains an r -sunflower.*

Proof. Let $\mathcal{F}$ be a $(1/r, 1/r)$ -robust sunflower, and let $K = \bigcap_{S \in \mathcal{F}} S$ be the common intersection of the sets in $\mathcal{F}$. Note that by assumption, $\mathcal{F}_K$ does not contain the empty set as an element. Lemma 3.6 gives that $\mathcal{F}_K$ contains r pairwise disjoint sets $S_1, \dots, S_r$. Thus $S_1 \cup K, \dots, S_r \cup K$ is an r -sunflower in $\mathcal{F}$. $\square$

The proof of Theorem 3.4 follows from the following stronger theorem, by setting $\alpha = \beta = 1/r$ and applying Lemma 3.8. The theorem verifies a conjecture raised in [65], and answers a question of [76].

Theorem 3.9 (Main theorem, robust sunflowers). *Let $0 < \alpha, \beta < 1$. For some constant C , any w -uniform set system $\mathcal{F}$ of size $|\mathcal{F}| \geq \left(\frac{C}{\alpha^2} \cdot \left(\log w \log \log w + \left(\log \frac{1}{\beta} \right)^2 \right) \right)^w$ contains an (α, β) -robust sunflower.*

For fixed α, β , the bound of $(\log w)^{w(1+o(1))}$ for robust sunflowers in Theorem 3.9 is sharp; it cannot be improved beyond $(\log w)^{w(1-o(1))}$. We give an example demonstrating this in Lemma 5.1.

3.2 Proof Overview

We now outline our proof of Theorem 3.9. First, we define *spread* set systems.

Definition 3.10 (Spreadness, [65]). *We say that a w -set system $\mathcal{F}$ is κ -spread if $|\mathcal{F}| \geq \kappa^w$ and $|\mathcal{F}_T| \leq \kappa^{-|T|}|\mathcal{F}|$ for all non-empty T , where $|\mathcal{F}_T|$ is the size of the link at T .*

The paper [65] calls these “regular set systems”, but we use the more descriptive term “spread”. Furthermore, we will ultimately need to slightly generalize the above definition so that it depends on more than one parameter; see Definition 4.1 for more details.

Say $\mathcal{F}$ is a w -set system of size $|\mathcal{F}| \geq \kappa^w$ on a ground set X . Then either $\mathcal{F}$ is κ -spread, or there is a link $\mathcal{F}_T$ of size $|\mathcal{F}_T| \geq \kappa^{w-|T|}$. In the latter “structured” case, we can simply pass to the link and apply induction, much like in the original proof of Erdős and Rado [36].

Thus, it suffices to consider the “pseudorandom” case of w -set systems which are κ -spread. In [61, 65], it was conjectured that for some absolute C , a κ -spread $\mathcal{F}$ is necessarily $(1/3, 1/3)$ -satisfying if $\kappa \geq (\log w)^C$. We show that there is some $\kappa = (\log w)^{1+o(1)}$ which is sufficient (see Theorem 4.5), completing the proof of Theorem 3.9. We also show that this value of κ is tight, up to the $o(1)$ in the exponent.

We next outline how we obtain the bound on κ . We will prove that a κ -spread w -set system is (α, β) -satisfying for appropriate κ, w, α, β through a series of reductions. Let $\mathcal{F}$ be a w -set system which is κ -spread. Sample $W \sim \mathcal{U}(X, p)$ for some $p = O(1/\log w)$. We show that with high probability over the choice of W , for almost all sets $S \in \mathcal{F}$, there exists a set $S' \in \mathcal{F}$ such that: (i) $S' \setminus W \subset S \setminus W$; and (ii) $|S' \setminus W| \leq w'$, for some w' which we will take to be $w(1 - \varepsilon)$ for a small ε . We throw out sets S that do not satisfy this property (which we call the *bad* sets for W), and replace any S that does with $S' \setminus W$ (we call these the *good* sets for W). This yields a w' -set system $\mathcal{F}'$ which has almost as many sets as the original $\mathcal{F}$, and therefore will have almost the same spreadness (since the best κ for which $|\mathcal{F}| \geq \kappa^T |\mathcal{F}_T|$

and for which $|\mathcal{F}'| \geq \kappa^T |\mathcal{F}_T|$ will be similar).¹ If we sample $W' \sim \mathcal{U}(X, p')$ and find some $S' \setminus W \in \mathcal{F}'$ such that $S' \setminus W \subset W'$, then $S' \subset W \cup W'$. Since $W \cup W' \sim \mathcal{U}(x, q)$ for $q = p + p' - pp'$, the satisfyingness of $\mathcal{F}$ will follow from the satisfyingness of $\mathcal{F}'$, although with different values of w, κ, α, β . We iterate, creating $W \cup W' \cup W'' \cup \dots$, which will be a random set. In particular, applying this “reduction step” (Lemma 4.6) iteratively $t = \log w$ times with the correct parameters yields a set system with much smaller sets, and then we can apply Janson’s inequality (see Lemma 4.10) to finish. In order to obtain our nearly optimal bounds, we must use a slight generalization of spreadness; see Definition 4.1. The parameters change throughout the algorithm; we will specify exactly how later.

To conclude, let us comment on how we prove the key reduction step (Lemma 4.6). The main idea is to use an encoding argument, inspired by Razborov’s proof of Håstad’s switching lemma [52, 75]. We show that pairs (W, S) for which S is bad for W can be efficiently encoded, crucially relying on the spreadness condition. This allows to show that for a random W it is very unlikely that there will be many bad sets. This ensures that the w' -set system $\mathcal{F}' = \{S' \setminus W : S' \in \mathcal{F}, |S' \setminus W| \leq w'\}$ will have almost as many sets as $\mathcal{F}$.

3.3 Chapter Organization.

We present the proof of Theorem 3.9 in Section 4. We prove a matching lower bound for robust sunflowers in Section 5. We discuss some applications in Section 6. We introduce several conjectures aimed at proving the sunflower conjecture in Section 7.

4 Proof of Theorem 3.9

We proceed to prove Theorem 3.9. Let $\mathcal{F}$ be a set system, and let $\sigma : \mathcal{F} \mapsto \mathbb{Q}_{\geq 0}$ be a weight function that assigns nonnegative rational weights to sets in $\mathcal{F}$ which are not all 0.

¹In the language of DNFs, we take a random restriction and approximate the result by a smaller width DNF whose clauses come from removing some variables from clauses of the original DNF.

A weight function is necessary because in the reduction multiple sets S can be replaced by the same S' , and we must weight such S' accordingly. For simplicity, we do not permit irrational weights. We call the pair $(\mathcal{F}, \sigma)$ a *weighted set system*. For a subset $\mathcal{F}' \subset \mathcal{F}$ we write $\sigma(\mathcal{F}') = \sum_{S \in \mathcal{F}'} \sigma(S)$ for the sum of the weights of the sets in $\mathcal{F}'$.

A *weight profile* is a vector $\mathbf{s} = (s_0; s_1, \dots, s_k)$ where $s_0 \geq s_1 \geq \dots \geq s_k \geq 0$ and $s_0 > 0$. The s_i need not be rational. We assume implicitly that $s_i = 0$ for all $i > k$.

Definition 4.1 (Spread weighted set system). *Let $\mathbf{s} = (s_0; s_1, \dots, s_w)$ be a weight profile. A weighted set system $(\mathcal{F}, \sigma)$ is $\mathbf{s}$ -spread if*

- (i) $\sigma(\mathcal{F}) \geq s_0$;
- (ii) $\sigma(\mathcal{F}_T) \leq s_{|T|}$ for any link $\mathcal{F}_T$ with non-empty T .

In particular, $\mathcal{F}$ is a w -set system.

Definition 4.2 (Spread set system). *Let $\mathbf{s}$ be a weight profile. A set system $\mathcal{F}$ is $\mathbf{s}$ -spread if there exists a weight function $\sigma : \mathcal{F} \mapsto \mathbb{Q}_{\geq 0}$ such that $(\mathcal{F}, \sigma)$ is $\mathbf{s}$ -spread.*

This generalizes Definition 3.10; we call a set system κ -spread if it is $(1; \kappa^{-1}, \dots, \kappa^{-w})$ -spread for the uniform weight function.

We note that one may always normalize a weight profile to have $s_0 = 1$. However, keeping s_0 as a free parameter helps to simplify some of the arguments later.

The main idea is to show that set systems which are $\mathbf{s}$ -spread, for $\mathbf{s}$ appropriately chosen, are “random looking” and in particular must be (α, β) -satisfying. This motivates the following definition.

Definition 4.3 (Satisfying weight profile). *Let $0 < \alpha, \beta < 1$. A weight profile $\mathbf{s}$ is (α, β) -satisfying if any $\mathbf{s}$ -spread set system is (α, β) -satisfying.*

The following lemma concerning satisfying weight profiles, which appears in [65], underlies our proof of Theorem 3.9. We repeat the proof here for completeness.

Lemma 4.4. *Let $0 < \alpha, \beta < 1$ and $w \geq 2$. Let $\kappa = \kappa(w) > 1$ be a (not necessarily strictly) increasing function of w such that the weight profile $(1; \kappa^{-1}, \dots, \kappa^{-w})$ is (α, β) -satisfying. Then any w -uniform set system $\mathcal{F}$ of size $|\mathcal{F}| > \kappa^w$ must contain an (α, β) -robust sunflower.*

Proof. Assume the contrary, and let $\mathcal{F}$ be a w -uniform set system on X of size $|\mathcal{F}| > \kappa^w$ without an (α, β) -robust sunflower. If the uniform distribution $\sigma(S) = 1/|\mathcal{F}|$ for $S \in \mathcal{F}$ is $(1; \kappa^{-1}, \dots, \kappa^{-w})$ -spread, then $\mathcal{F}$ is (α, β) -satisfying and thus is an (α, β) -robust sunflower.

Thus, there exists some nonempty $K \subset X$ so that $|\mathcal{F}_K| > \kappa^{-|K|}|\mathcal{F}|$. Pick a maximal such K . Note that we cannot have $|K| = w$, as otherwise $|\mathcal{F}_K| = 1 < \kappa^{-w}|\mathcal{F}|$, and so we must have that $1 \leq |K| \leq w - 1$. Let $\mathcal{F}' = \mathcal{F}_K$. Note that $|\mathcal{F}'| > \kappa^{w-|K|}$; and for any non-empty set T disjoint from K , $|\mathcal{F}'_T| = |\mathcal{F}_{K \cup T}| \leq \kappa^{-|K|-|T|}|\mathcal{F}| \leq \kappa^{-|T|}|\mathcal{F}'|$, by the maximality of $|K|$. Let $\sigma'(S) = 1/|\mathcal{F}'|$ for $S \in \mathcal{F}'$. Then $(\mathcal{F}', \sigma')$ is $(1; \kappa^{-1}, \dots, \kappa^{-\ell})$ -spread for $\ell = w - |K|$, where $\kappa = \kappa(w) \geq \kappa(\ell)$. Thus $\mathcal{F}'$ is (α, β) -satisfying, and hence $\{S \cup K : S \in \mathcal{F}'\}$ is an (α, β) -robust sunflower contained in $\mathcal{F}$. $\square$

Lemma 4.4 motivates the following definition. For $0 < \alpha, \beta < 1$ and $w \geq 2$, let $\kappa(w, \alpha, \beta)$ be the least κ such that $(1; \kappa^{-1}, \dots, \kappa^{-w})$ is (α, β) -satisfying. Theorem 3.9 follows by combining Lemma 4.4 with the following theorem, which bounds $\kappa(w, \alpha, \beta)$.

Theorem 4.5. $\kappa(w, \alpha, \beta) = O\left(\frac{1}{\alpha^2} \cdot \left(\log w \log \log w + \left(\log \frac{1}{\beta}\right)^2\right)\right)$.

To prove Theorem 4.5 we must show that, if κ is the right-hand side of the above expression the following holds. Let $\mathcal{F}$ be a weighted w -set system with total weight 1 and where the weight of any link of a nonempty set T is at most $\kappa^{-|T|}$. Then $\mathcal{F}$ is (α, β) -satisfying, i.e. the chance that some set from $\mathcal{F}$ is contained in a subset of the ground set X sampled an α -biased distribution is at least $1 - \beta$. We prove Theorem 4.5 in the remainder of this section.

4.1 A Reduction Step

Let $\mathcal{F}$ be a w -set system on X , and fix $w' \leq w$. The main goal in this section is to reduce our problem for $\mathcal{F}$ to the same problem for some w' -set system $\mathcal{F}'$. We prove the following lemma in this section.

Lemma 4.6. *Let $\mathbf{s} = (s_0; s_1, \dots, s_w)$ be a weight profile, $w' \leq w$, $\delta > 0$ and define $\mathbf{s}' = ((1 - \delta)s_0; s_1, \dots, s_{w'})$. Assume $\mathbf{s}'$ is (α', β') -satisfying. Then for any $p > 0$, $\mathbf{s}$ is (α, β) -satisfying for*

$$\alpha = p + (1 - p)\alpha', \quad \beta = \beta' + \frac{(4/p)^w s_{w'}}{\delta s_0}.$$

In order to prove this lemma, we need the notions of good and bad sets with respect to some $W \subset X$.

Definition 4.7. *Let $W \subset X$. Given a set $S \in \mathcal{F}$, the pair (W, S) is said to be good if there exists a set $S' \in \mathcal{F}$ (possibly with $S' = S$) such that*

$$(i) \ S' \setminus W \subset S \setminus W;$$

$$(ii) \ |S' \setminus W| \leq w'.$$

If no such S' exists, we say that (W, S) is bad.

Note that if W contains a set in $\mathcal{F}$ (i.e., $S' \subset W$ for some $S' \in \mathcal{F}$) then all pairs (W, S) are good.

Lemma 4.8. *Let $(\mathcal{F}, \sigma)$ be an $\mathbf{s} = (s_0; s_1, \dots, s_w)$ -spread weighted set system on X . Let $W \subset X$ be a uniformly random subset of size $|W| = p|X|$ and $\mathcal{B}(W) = \{S \in \mathcal{F} : (W, S) \text{ is bad}\}$. Then $\mathbb{E}_W[\sigma(\mathcal{B}(W))] \leq (4/p)^w s_{w'}$.*

Proof. First, we simplify the setting a bit. We may assume by scaling σ and $\mathbf{s}$ by the same factor that $\sigma(S) = N_S$ is an integer for each $S \in \mathcal{F}$. Let $N = \sum N_S \geq s_0$. By scaling

back down by N , we can then identify $(\mathcal{F}, \sigma)$ with the uniform distribution on the multi-set system $\mathcal{F}' = \{S_1, \dots, S_N\}$, where each set $S \in \mathcal{F}$ is repeated N_S times in $\mathcal{F}'$. Thus

$$\sigma(\mathcal{B}(W)) = |\{i : S_i \in \mathcal{F}' \text{ and } (W, S_i) \text{ is bad}\}|.$$

Assume that (W, S_i) is bad in $\mathcal{F}'$. In particular, this means that W does not contain any set in $\mathcal{F}$. We describe (W, S_i) with a small amount of information. Let $|X| = n$ and $|W| = pn$. We encode (W, S_i) as follows:

1. The first piece of information is $W \cup S_i$, a subset of X of size between pn and $pn + w$.

The number of options for this is at most:

$$\sum_{i=0}^w \binom{n}{pn+i} \leq \sum_{i=0}^w \left(\frac{1-p}{p}\right)^i \binom{n}{pn} \leq \left(\frac{1-p}{p} + 1\right)^w \binom{n}{pn} = p^{-w} \binom{n}{pn}.$$

2. Given $W \cup S_i$, let j be minimal such that $S_j \subset W \cup S_i$, so that j depends only on $W \cup S_i$. Given S_j , there are fewer than 2^w possibilities for $A = S_i \cap S_j$. As such, we will let A be the second piece of information.

3. Because $S_j \subset W \cup S_i$ and $A = S_i \cap S_j$, we have $S_j \setminus W \subset A$. Since (W, S_i) is bad, $|A| \geq |S_j \setminus W| > w'$. The number of the sets in $\mathcal{F}'$ which contain A is $|\mathcal{F}'_A| \leq s_{w'}$. The third piece of information will be which one of these is S_i .

4. Finally, once we have specified S_i , we will specify $S_i \cap W$, which is one of 2^w possible subsets of S_i .

From these four pieces of information one can uniquely reconstruct (W, S_i) . Thus the total number of bad pairs (W, S_i) is bounded by

$$p^{-w} \binom{n}{pn} \cdot 2^w \cdot s_{w'} \cdot 2^w = (4/p)^w s_{w'} \binom{n}{pn}.$$

Because the number of sets $W \subset X$ of size $|W| = p|X|$ is $\binom{n}{pn}$, the lemma follows by taking expectation over W . $\square$

The following is a corollary of Lemma 4.8, where we replace sampling $W \subset X$ of size $|W| = p|X|$ with sampling $W \sim \mathcal{U}(X, p)$.

Corollary 4.9. *Let $(\mathcal{F}, \sigma)$ be an $\mathbf{s} = (s_0; s_1, \dots, s_w)$ -spread weighted set system on X . Let $W \sim \mathcal{U}(X, p)$ and $\mathcal{B}(W) = \{S \in \mathcal{F} : (W, S) \text{ is bad}\}$. Then $\mathbb{E}_W[\sigma(\mathcal{B}(W))] \leq (4/p)^w s_w$.*

Proof. The proof is by a reduction to Lemma 4.8. Replace the base set X with a much larger set X' (without changing $\mathcal{F}$, so the new elements do not belong to any set in $\mathcal{F}$). Let $W' \subset X'$ be a uniform set of size $|W'| = p|X'|$, and let $W = W' \cap X$. Then as the size of X' goes to infinity, the distribution of W approaches $\mathcal{U}(X, p)$. $\square$

Proof of Lemma 4.6. Let $(\mathcal{F}, \sigma)$ be an $\mathbf{s} = (s_0; s_1, \dots, s_w)$ -spread weighted set system on X . Let $W \sim \mathcal{U}(X, p)$. Say that W is δ -bad if $\sigma(\mathcal{B}(W)) \geq \delta s_0$. By applying Corollary 4.9 and Markov's inequality, we obtain that

$$\Pr[W \text{ is } \delta\text{-bad}] \leq \frac{\mathbb{E}[\sigma(\mathcal{B}(W))]}{\delta s_0} \leq \frac{(4/p)^w s_w}{\delta s_0}.$$

Fix W which is not δ -bad. If for some $S \in \mathcal{F}$, the pair (W, S) is good, then there exists $\pi(S) = S' \in \mathcal{F}$ (possibly with $S' = S$) such that (i) $S' \setminus W \subset S \setminus W$ and (ii) $|S' \setminus W| \leq w'$.

Define a new weighted set system $(\mathcal{F}', \sigma')$ on $X' = X \setminus W$ as follows:

$$\mathcal{F}' = \{\pi(S) \setminus W : S \in \mathcal{F} \setminus \mathcal{B}(W)\}, \quad \sigma'(S' \setminus W) = \sigma(\pi^{-1}(S')).$$

We claim that $\mathcal{F}'$ is $\mathbf{s}' = ((1 - \delta)s_0; s_1, \dots, s_{w'})$ -spread. To see that, note that $\sigma'(\mathcal{F}') =$

$\sigma(\mathcal{F} \setminus \mathcal{B}(W)) \geq (1 - \delta)s_0$ and that for any set $T \subset X'$,

$$\sigma'(\mathcal{F}'_T) = \sum_{S' \supset T} \sigma'(S') = \sum_{S: \pi(S) \supset T} \sigma(S) \leq \sum_{S \supset T} \sigma(S) = \sigma(\mathcal{F}_T) \leq s_{|T|}.$$

Finally, all sets in $\mathcal{F}'$ have size at most w' . Hence, by the assumption $\mathbf{s}'$ is (α', β') -satisfying. Thus, if we choose $W' \sim \mathcal{U}(X', \alpha')$ then we obtain that with probability more than $1 - \beta'$, there exists $S^* \in \mathcal{F}'$ such that $S^* \subset W'$. Recall that $S^* = S \setminus W$ for some $S \in \mathcal{F}$. Thus $S \subset W \cup W' \sim \mathcal{U}(X, p + (1 - p)\alpha')$. The value of β accounts for the case where W is δ -bad. $\square$

4.2 A Final Step

Repeated applications of the reduction step yield a set system which is almost as spread as the original set system, but whose sets are much smaller. Thus the spreadness guarantee becomes good compared to the size of the sets in the system. In this case, we can just directly show that the set system is satisfying. A similar argument appears in [76].

Lemma 4.10. *Let $0 < \alpha, \beta < 1$, $w \geq 2$, and set $\kappa = \max\{4\log(1/\beta), 2\} \cdot w/\alpha$. Let $(\mathcal{F}, \sigma)$ be an $\mathbf{s} = (s_0; s_1, \dots, s_w)$ -spread weighted set system where $s_i < \kappa^{-i}s_0$. Then $\mathcal{F}$ is (α, β) -satisfying.*

Proof. Again, we can assume by scaling that $N_S = \sigma(S)$ for $S \in \mathcal{F}$ are all integers. Let $\mathcal{F}'$ be the multi-set system where each $S \in \mathcal{F}$ is repeated N_S times, so that $\sum N_S > \kappa^w$. We may also assume that all sets in $\mathcal{F}'$ have size exactly w , by adding different dummy elements to each set of size less than w . We take care to scale by a large enough factor so that a negligible amount of weight falls on each dummy element, and so the spreadness hypothesis is preserved. Let $N = \sum N_S \geq s_0$ and let $\mathcal{F}' = \{S_1, \dots, S_N\}$. If σ' is the uniform weighting on sets of $\mathcal{F}'$, then $(\mathcal{F}', \sigma')$ also satisfies the assumption of the lemma. Note also that for

any set $W \subset X$, if W contains a set of $\mathcal{F}'$ then it also contains a set of $\mathcal{F}$, so it suffices to verify the (α, β) -satisfyingness of $\mathcal{F}'$.

The proof is by Janson's inequality. Let $W \sim \mathcal{U}(X, \alpha)$ and let $\mathcal{Z}_i$ be the indicator variable for $S_i \subset W$. Denote $i \sim j$ if S_i, S_j intersect. Define

$$\mu = \sum_i \mathbb{E}[\mathcal{Z}_i], \quad \Delta = \sum_{i \sim j} \mathbb{E}[\mathcal{Z}_i \mathcal{Z}_j].$$

We have $\mu = N\alpha^w$. To compute Δ , let p_ℓ denote the fraction of the N^2 pairs (i, j) with $1 \leq i, j \leq N$ such that $|S_i \cap S_j| = \ell$. Then

$$\Delta = \sum_{\ell=1}^w p_\ell N^2 \alpha^{2w-\ell}.$$

To bound p_ℓ , note that for each $S_i \in \mathcal{F}$, and any $R \subset S_i$ of size $|R| = \ell$, the number of $S_j \in \mathcal{F}$ such that $R \subset S_j$ is $|\mathcal{F}_R| \leq N/\kappa^{|R|}$ by the spreadness assumption. Let $q = \max\{4\log(1/\beta), 2\}$. Thus we can bound

$$\Delta \leq \sum_{\ell=1}^w \binom{w}{\ell} \kappa^{-\ell} N^2 \alpha^{2w-\ell} \leq \sum_{\ell=1}^w \left(\frac{w}{\alpha\kappa}\right)^\ell \mu^2 = \mu^2 \sum_{\ell=1}^w \left(\frac{1}{q}\right)^\ell < \frac{2\mu^2}{q}.$$

Note that in addition $\Delta \geq \mu$, as we include the pairs (i, i) in the summation for Δ . Thus by Janson's inequality,

$$\Pr[\forall i, \mathcal{Z}_i = 0] \leq \exp\left\{-\frac{\mu^2}{2\Delta}\right\} < \exp\left\{-\frac{q}{4}\right\} \leq \beta.$$

□

4.3 Putting Everything Together

We prove Theorem 4.5 in this subsection, where our goal is to bound $\kappa(w, \alpha, \beta)$. We will apply Lemma 4.6 iteratively, until we reach a case where we can apply Lemma 4.10.

Let $w \geq 2$ be fixed throughout, and let $\kappa > 1$, to be optimized later. We first introduce some notation. For $0 < \Delta < 1$ and $\ell \geq 1$, let $\mathbf{s}(\Delta, \ell) = (1 - \Delta; \kappa^{-1}, \dots, \kappa^{-\ell})$ be a weight profile. Let $A(\Delta, \ell), B(\Delta, \ell)$ be such that any $\mathbf{s}(\Delta, \ell)$ -spread set system is $(A(\Delta, \ell), B(\Delta, \ell))$ -satisfying.

Lemma 4.6 applied to $w' \geq w''$ and p, δ shows that we may take

$$\begin{aligned} A(\Delta, w') &\leq A(\Delta + \delta, w'') + p, \\ B(\Delta, w') &\leq B(\Delta + \delta, w'') + \frac{(4/p)^{w'}}{\delta(1 - \Delta)\kappa^{w''}}. \end{aligned}$$

We apply this iteratively for some widths $w_0, \dots, w_r$. Set $w_0 = w$ and $w_{i+1} = \lceil (1 - \varepsilon)w_i \rceil$ for some small ε as long as $w_i > w^*$ for some w^* . In particular, we need $w^* \geq 1/\varepsilon$ to ensure $w_{i+1} < w_i$, and we will optimize ε and w^* later. The number of steps is thus $r \leq (K \log w)/\varepsilon$ for some constant $K > 0$. Let $p_1, \dots, p_r$ and $\delta_1, \dots, \delta_r$ be the values we use for p, δ at each iteration. To simplify the notation, let $\Delta_i = \delta_1 + \dots + \delta_i$ and $\Delta_0 = 0$. Furthermore, define

$$\gamma_i = \frac{(4/p_i)^{w_{i-1}}}{\kappa^{w_i}}.$$

Then for $i = 1, \dots, r$, we may take

$$\begin{aligned} A(\Delta_{i-1}, w_{i-1}) &\leq A(\Delta_i, w_i) + p_i, \\ B(\Delta_{i-1}, w_{i-1}) &\leq B(\Delta_i, w_i) + \frac{\gamma_i}{\delta_i(1 - \Delta_{i-1})}. \end{aligned}$$

Set $p_i = p = \frac{\alpha}{2r}$ and $\delta_i = \sqrt{\gamma_i}$, where $r \leq (K \log w)/\varepsilon$ is the number of steps. We

will select the parameters so that $\Delta_i \leq 1/2$ for all i . Thus we may take $A(1/2, w^*)$ and $B(1/2, w^*)$ such that

$$A(0, w) \leq A(\Delta_r, w_r) + \alpha/2 \leq A(1/2, w^*) + \alpha/2,$$

$$B(0, w) \leq B(\Delta_r, w_r) + 2\Delta_r \leq B(1/2, w^*) + 2\Delta_r.$$

Plugging in the values for δ_i , we compute the sum

$$\begin{aligned} \Delta_r &= \sum_{i=1}^r \delta_i \leq \sum_{i=1}^r \sqrt{\frac{(4/p)^{w_{i-1}}}{\kappa^{(1-\varepsilon)w_{i-1}}}} = \sum_{i=1}^r \left(\frac{4/p}{\kappa^{1-\varepsilon}} \right)^{w_{i-1}/2} \leq \sum_{k \geq w^*} \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{k/2} \\ &= \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{w^*/2} \sum_{k \geq 0} \left(\frac{8K \log w}{\varepsilon \alpha \kappa^{1-\varepsilon}} \right)^{k/2} \leq 2^{1-w^*}, \end{aligned}$$

assuming $\kappa^{1-\varepsilon} \geq (32K \log w)/\varepsilon \alpha$.

Next, we will apply Lemma 4.10 to bound $A(1/2, w^*) \leq \alpha/2$ and $B(1/2, w^*) \leq \beta/2$. We use the simple observation that $(1/2; \kappa^{-1}, \dots, \kappa^{-w^*})$ -spread set systems must also be $(1; (\kappa/2)^{-1}, \dots, (\kappa/2)^{-w^*})$ -spread, in which case we can apply Lemma 4.10 and obtain that we need

$$\kappa/2 \geq (2 + 4 \log(2/\beta)) \cdot 2w^*/\alpha.$$

Let us now put the bounds together. We still have the freedom to choose $\varepsilon > 0$ and $w^* \geq 1/\varepsilon$.

To obtain $A(0, w) \leq \alpha$, $B(0, w) \leq \beta$, we also need $\Delta_r \leq \beta/4 < 1/2$. Thus all the constraints are:

1. $w^* \geq 1/\varepsilon$;
2. $\kappa^{1-\varepsilon} \geq (32K \log w)/\varepsilon \alpha$ where $K > 0$ is a constant;
3. $\kappa \geq 4 \cdot (2 + 4 \log(2/\beta)) \cdot w^*/\alpha$;
4. $2^{1-w^*} \leq \beta/4$.

Set $\varepsilon = 1/\log \log w$ and $w^* = c \cdot \max\{\log(1/\beta), \log \log w\}$ for some large enough constant $c \geq 1$. This ensures that the first and last conditions hold.

Thus we obtain that the result holds whenever

$$\kappa = \Omega \left(\max \left\{ \left(\frac{1}{\alpha} \right)^{1+2/\log \log w} \log w \log \log w, \frac{1}{\alpha} \left(1 + \log \frac{1}{\beta} \right)^2, \frac{1}{\alpha} \left(1 + \log \frac{1}{\beta} \right) \log \log w \right\} \right).$$

In particular, it suffices to set $\kappa = O \left(\frac{1}{\alpha^2} \cdot \left(\log w \log \log w + \left(\log \frac{1}{\beta} \right)^2 \right) \right)$.

5 A Lower Bound for Robust Sunflowers

In this section, we construct an example of a w -set system of size $(\log w)^{w(1-o(1))}$ which does not contain a robust sunflower, showing that Theorem 3.9 is tight up to the $o(1)$ in exponent.

For concreteness we fix $\alpha = \beta = 1/2$, but the construction can be easily modified for any other constant values of α, β . We assume that w is sufficiently large.

Lemma 5.1. *There exists a w -set system of size $((\log w)/8)^{w-\sqrt{w}} = (\log w)^{w(1-o(1))}$ which does not contain a $(1/2, 1/2)$ -robust sunflower.*

Let $c \geq 1$ be determined later. Let $X_1, \dots, X_w$ be pairwise disjoint sets of size $m = \log(w/c)$, and let X be their union. Let $\widehat{\mathcal{F}} = X_1 \times \dots \times X_w$ be the w -set system consisting of all sets which contain exactly one element from each of the X_i . We first argue that $\widehat{\mathcal{F}}$ is not satisfying.

Claim 5.2. *For $c \geq 1$, $\widehat{\mathcal{F}}$ is not $(1/2, 1/2)$ -satisfying.*

Proof. Let $R \sim \mathcal{U}(X, 1/2)$. The probability that some X_i is disjoint from R is $1 - (1 - 2^{-m})^w = 1 - (1 - c/w)^w > 1/2$ if $c \geq 1$. Thus if $c \geq 1$, with probability more than $1/2$, R is disjoint from some X_i and thus does not contain any set $S \in \widehat{\mathcal{F}}$. $\square$

The family $\widehat{\mathcal{F}}$ does contain a $(1/2, 1/2)$ -robust sunflower. For example, if T contains exactly one element from each of $X_1, \dots, X_{w-1}$, then $\widehat{\mathcal{F}}_T$ is isomorphic to X_w , and in particular is certainly $(1/2, 1/2)$ -satisfying. However, it turns out that the robust sunflowers in $\widehat{\mathcal{F}}$ all contain large kernels. We will exploit this to construct a large subsystem of $\widehat{\mathcal{F}}$ which does not contain a robust sunflower. Let $\varepsilon > 0$ be determined later, and choose $\mathcal{F} \subset \widehat{\mathcal{F}}$ to be a subsystem that satisfies:

$$|S \cap S'| \leq (1 - \varepsilon)w, \quad \forall S, S' \in \mathcal{F}, S \neq S'.$$

For example, we can obtain $\mathcal{F}$ by a greedy procedure, each time choosing an element S in $\widehat{\mathcal{F}}$ and deleting all S' such that $|S \cap S'| > (1 - \varepsilon)w$. The number of such S' is at most $\binom{w}{\varepsilon w} m^{\varepsilon w} \leq 2^w m^{\varepsilon w}$. Hence we can obtain $\mathcal{F}$ of size $|\mathcal{F}| \geq 2^{-w} m^{(1-\varepsilon)w}$.

Claim 5.3. *For $c \geq 1/\varepsilon$, $\mathcal{F}$ does not contain a $(1/2, 1/2)$ -robust sunflower.*

Proof. Consider any set $K \subset X$. We need to prove that $\mathcal{F}$ does not contain a $(1/2, 1/2)$ -robust sunflower with kernel K . If it does, $\mathcal{F}_K$ must contain at least two sets S and S' , which implies that $|K \cap X_i| \leq |S \cap X_i| \leq 1$ for all i . Furthermore $K \subset S \cap S'$ and so $|K| \leq |S \cap S'| \leq (1 - \varepsilon)w$. However, for such K we claim even $\widehat{\mathcal{F}}_K$ is not $(1/2, 1/2)$ -satisfying.

To prove this, let $I = \{i : |K \cap X_i| = 0\}$, so that $|I| \geq \varepsilon w$. Let $R \sim \mathcal{U}(X, 1/2)$. The probability that there exists $i \in I$ such that R is disjoint from X_i is $1 - (1 - 2^{-m})^{|I|} \geq 1 - (1 - c/w)^{\varepsilon w}$ which is more than $1/2$ for $c \geq 1/\varepsilon$. $\square$

To conclude the proof of Lemma 5.1 we optimize the parameters. Set $c = 1/\varepsilon$. We have $|\mathcal{F}| \geq 2^{-w} (\log(\varepsilon w))^{(1-\varepsilon)w}$. Setting $\varepsilon = 1/\sqrt{w}$ gives $|\mathcal{F}| \geq ((\log w)/8)^{w-\sqrt{w}} = (\log w)^{(1-o(1))w}$.

6 Subsequent Works and Corollaries

After the current work was made available on the ArXiv, Rao [74] simplified the proof of Theorem 4.5 using information theoretic techniques, and obtained a slightly better dependence on α and β . Furthermore, following the initial release of this work, the technique developed in this chapter has been used by Frankston, Kahn, Narayanan, and Park [41] to resolve a conjecture of Talagrand in random graph theory and later by Park and Pham [68] to resolve the Kahn-Kalai Conjecture. Rao then used the refinements of [41] to further improve the bound in Theorem 4.5 to $(C/\alpha) \log(w/\beta)$ for some constant C , hence improving the bound in Theorem 3.9 to $((C/\alpha) \log(w/\beta))^w$ and the bound in Theorem 3.4 to $(Cr \log(wr))^w$. Tao [86] gave another information theoretic proof.

We present a couple additional applications of Theorem 4.5 below. To present sharper (and cleaner) bounds, we instead use the version proved by Rao [74], building on the work of Frankston, Kahn, Narayanan, and Park [41]. Using the notations of Theorem 4.5, he proved that $\kappa(w, \alpha, \beta) \leq (C/\alpha) \log(w/\beta)$ for some constant C .

6.1 The Erdős-Szemerédi Sunflower Conjecture.

In [37], the Erdős-Rado sunflower conjecture is shown to imply the Erdős-Szemerédi sunflower conjecture. Plugging in our new bounds into their argument gives a corresponding improvement.

Theorem 6.1. *For any $r \geq 3$ there exists $c = c(r)$ such that the following holds. Let $\mathcal{F}$ be a set system on X , with $|X| = n$ and $|\mathcal{F}| \geq 2^{n(1-c/\log n)}$. Then $\mathcal{F}$ contains an r -sunflower.*

The Erdős-Szemerédi sunflower conjecture says the bound on the size of $\mathcal{F}$ can be improved to $2^{n(1-\varepsilon)}$ for some $\varepsilon = \varepsilon(r)$. This is known for $r = 3$, but the proof is algebraic, and does not generalize for larger values of r .

6.2 Intersecting Set Systems.

A set system is *intersecting* if any pair of its sets intersects. We note the following useful corollary of Theorem 4.5:

Theorem 6.2. *If $\mathcal{F}$ is an intersecting w -uniform set system, and for all T , $|\mathcal{F}_T| \leq \kappa^{-|T|}|\mathcal{F}|$, then $\kappa = O(\log w)$.*

An example from [65] shows that for $\kappa = \Omega(\log w / \log \log w)$, there are intersecting κ -spread w -uniform set systems, so the bound in Theorem 6.2 is close to tight.

Proof. If $\mathcal{F}$ is intersecting then it is not $(1/2, 1/2)$ -satisfying (apply Lemma 3.6 for $r = 2$). Thus by the improvement of Theorem 4.5 from [74], it cannot be $(C \log w)$ -spread for a large enough constant C . $\square$

Note that Theorem 6.2 is fairly robust in the sense that the bounds on the links of T for large T are required. For instance, for a small constant ε a family $\mathcal{F}$ of 2^{n^ε} random subsets of $[n]$ of size $w = n^{1/2+\varepsilon}$ is intersecting with high probability. Also with high probability it is $(1; \kappa, \dots, \kappa^{-\ell})$ -spread for ℓ up to $n^{\varepsilon/2} = w^{\Omega(1)}$, where $\kappa = w^{\Omega(1)}$. This example shows that in Theorem 4.5 the spreadness condition on the large sets is necessary. Thus it seems difficult to find a purely algebraic or analytic proof of Theorem 4.5, as one would have to use the spreadness condition on large sets in a useful but non-combinatorial way.

7 Rainbow Sunflowers

In this section, we discuss a possible line of attack toward the sunflower conjecture (Conjecture 3.3). We propose the following more general conjecture.

Conjecture 7.1 (Rainbow sunflower conjecture). *There exists a constant C such that the following holds. Let $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| \geq C^w$ over a ground set X . Assume that*

the elements of X are colored with either red, green, or blue uniformly and independently at random.

Then with high probability, there exists distinct sets $S_i, S_j, S_k \in \mathcal{F}$ such that if $Y = S_i \cap S_j \cap S_k$, then all elements of $S_i \setminus Y$ are red, all elements of $S_j \setminus Y$ are green, and all elements of $S_k \setminus Y$ are blue.

Note that Conjecture 7.1 implies Conjecture 3.3 for $r = 3$. Furthermore, this chapter gives a proof of Conjecture 7.1 where C^w is replaced by $(\log w)^{w(1+o(1))}$, and Rao [74] improved the bound to $(C \log w)^w$ for some constant C .

Unlike in Theorem 3.9, we do not know if $(C \log w)^w$ is tight, and in fact we suspect this is not the case. Conjecture 7.1 is in some sense similar to Theorem 3.9; the latter differs in that one must commit to a subfamily before seeing the coloring, whereas in the former the sets S_i, S_j, S_k may depend on the coloring. We propose another variant of Conjecture 7.1.

Conjecture 7.2. *There exists a constant C such that the following holds. Let $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| \geq C^w$ over a ground set X . Assume that the elements of X are colored with either red or blue uniformly and independently. In addition let $S \in \mathcal{F}$ be sampled uniformly.*

Then with high probability, there exists distinct sets $S_i, S_j \in \mathcal{F}$ such that all elements of $S_i \setminus S$ are red, all elements of $S_j \setminus S$ are blue, and $S_i \cap S = S_j \cap S \subsetneq S$.

Conjecture 7.2 also implies Conjecture 3.3 for $r = 3$, but we do not know of any nontrivial bounds for it. It would be interesting to have even a $w^{o(w)}$ bound. We can however prove the following lemma, which provides some evidence that Conjecture 7.1 might be true.

Lemma 7.3. *There exists a constant C such that the following holds. Let $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| \geq C^w$ over a ground set X . Assume that the elements of X are colored with either red or blue uniformly and independently.*

Then with high probability, there exists distinct sets $S_i, S_j \in \mathcal{F}$ such that if $Y = S_i \cap S_j$, then all elements of $S_i \setminus Y$ are red and all elements of $S_j \setminus Y$ are blue (equivalently, all

elements of $S_i \setminus S_j$ are red and all elements of $S_j \setminus S_i$ are blue).

In order to establish this, we first prove a weaker result along the same lines.

Lemma 7.4. *For all $\delta > 0$ there exists $C = C(\delta)$ such that the following holds. Let $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| = m \geq C^w$ over a ground set X . Assume that a random δ -fraction of the elements of X are colored white. Then:*

1. *The expected number of sets $S_j \in \mathcal{F}$, for which there is no $i \neq j$ with $S_i \setminus S_j$ all white, is at most C^w .*
2. *With high probability, for almost all $S_j \in \mathcal{F}$ there are at least m/C^w many sets $S_i \in \mathcal{F}$ such that $S_i \setminus S_j$ is all white.*

Proof. We will show that the lemma holds for $C(\delta) = 4/\delta$. Let W denote the white subset of X . Call a pair (S_j, W) *bad* if there is no $i \neq j$ such that $S_i \setminus S_j \subset W$. If (S_j, W) is a bad pair, then S_j is the only set of $\mathcal{F}$ contained in $S_j \cup W$. Thus, $S_j \cup W$ uniquely determines S_j , and another w bits encoding $S_j \cap W$ determine the pair (S_j, W) . Hence, the number of bad pairs (S_j, W) is only a factor of $(2/\delta)^w$ more than the number of possible W . Thus, for a random W , in expectation only $(2/\delta)^w < C^w$ of the sets $S_j \in \mathcal{F}$ are bad.

For the second part of the lemma, redefine bad so that (S_j, W) is bad if there are less than $K = m/C^w$ values of $i \neq j$ such that $S_i \setminus S_j \subset W$. Then each bad pair (S_j, W) can be identified by $S_j \cup W$ and a positive integer up to K , so in expectation only $K(2/\delta)^w = m \cdot 2^{-w}$ of the sets $S_j \in \mathcal{F}$ are bad. Then by Markov's inequality, with probability at least $1 - 2^{-w/2}$, for a $1 - 2^{-w/2}$ fraction of $S_j \in \mathcal{F}$ there are at least m/C^w many sets $S_i \in \mathcal{F}$ such that $S_i \setminus S_j$ is all white. $\square$

By adding dummy elements to X , Lemma 7.4 also holds with each element of X independently colored white with probability δ . Now we prove Lemma 7.3.

Proof of Lemma 7.3. We take a random red-blue coloring of X , where we can treat either the red or the blue elements as the white elements in Lemma 7.4, and then apply Lemma 7.4 with $\delta = 1/2$.

Choose a random $S \in \mathcal{F}$ and fix it for the remainder of the proof. By the second part of Lemma 7.4, in a random coloring with high probability there are at least m/C^w many sets $S_i \in \mathcal{F}$ so that $S_i \setminus S$ is all red.

Call a set S_i *good* if there exists S_j so that $S_j \setminus S_i$ is blue and $S_j \cap S = S_i \cap S$. Call it *bad* otherwise (this “bad” is different from the one in Lemma 7.4). For each $Y \subset S$, consider the family $\mathcal{F}_{S,Y}$ of sets in $\mathcal{F}$ that intersect S exactly at Y . By the first part of Lemma 7.4, in expectation only C^w of the sets $S_i \in \mathcal{F}_{S,Y}$ will be bad. Taking the union over all possible Y , it follows in expectation that only $(2C)^w$ of the sets $S_i \in \mathcal{F}$ will be bad.

So with high probability over the random coloring two events hold: (i) at most $(4C)^w$ of the sets $S_i \in \mathcal{F}$ are bad; and (ii) at least m/C^w of the sets $S_i \in \mathcal{F}$ will have $S_i \setminus S$ all red. So, as long as $m/C^w > (4C)^w$, there is some good S_i so that $S_i \setminus S$ is all red. Then for each such S_i there will be some S_j so that $S_i \setminus S_j \subset S_i \setminus S$ is all red and $S_j \setminus S_i$ is all blue. This concludes the proof. $\square$

The following is another corollary of Lemma 7.3. It follows from Lemma 7.3 in exactly the same way that the Erdős-Szemerédi sunflower lemma follows from the Erdős-Rado sunflower lemma (see [37]), so we omit the proof.

Lemma 7.5. *There exists $\delta > 0$ such that the following holds. Let $\mathcal{F}$ be a family of subsets of $[n]$ of size $|\mathcal{F}| > (2 - \delta)^n$. Color each element of $[n]$ either red or blue uniformly and independently at random.*

Then with high probability, there exists distinct $S_i, S_j \in \mathcal{F}$ such that $S_i \setminus S_j$ is all red and $S_j \setminus S_i$ is all blue.

In particular, this implies that if R is a uniform random subset of $[n]$, then with high

probability there exist some $S_i, S_j \in \mathcal{F}$ are such that $S_i \oplus R \subset S_j \oplus R$, where $\oplus$ denotes the exclusive or function.

Finally, we note that a statement similar to Conjecture 7.1 is equivalent to Conjecture 3.3.

Conjecture 7.6 (Monochromatic sunflower conjecture). *For all $\delta > 0$ there exists a constant $K = K(r, \delta)$ such that the following holds. Let $\mathcal{F}$ be a w -set system of size $|\mathcal{F}| \geq K^w$ on a ground set X . Color a random δ -fraction of the elements of X with red.*

Then with high probability, $\mathcal{F}$ contains an r -sunflower whose petals are all red.

Theorem 7.7. *Conjecture 7.6 is equivalent to Conjecture 3.3.*

Proof. Clearly, Conjecture 7.6 implies Conjecture 3.3, so now we prove the other direction. Assuming Conjecture 3.3, let $C = C(r)$ be such that any w -set system of size at least C^w contains an r -sunflower. By the second part of Lemma 7.4, if $K = K(r, \delta)$ is large enough, and $S \in \mathcal{F}$ is uniformly sampled, then with high probability there are at least $(2C)^w$ sets S_i with $S_i \setminus S$ all red. Thus there exists $Y \subset S$ such that at least C^w many sets S_i have $S_i \setminus S$ all red and $S_i \cap S = Y$. In particular, there are r sets $S_1, \dots, S_r$ forming an r -sunflower so that for $1 \leq i \leq r$ we have $S_i \setminus S$ red and $S_i \cap S = Y$. Because Y is contained in the kernel of the sunflower, all of the petals are red. $\square$

The above argument also shows that the optimal bounds in Conjecture 7.6 and Conjecture 3.3 are off by an exponential factor in w . For $r = 3$, we expect a similar relationship between Conjecture 7.1 and Conjecture 3.3.

Part III

Blowing up Sunflowers

This chapter is based on work [8] that has been submitted but not yet published.

8 Introductory Remarks

8.1 Background and Main Result

In the previous chapter, we presented the improvement on the best known bounds for the Erdős-Rado sunflower lemma [36]. Central to this result was an “encoding” argument, in which the ground set X of a set system $\mathcal{F}$ is colored in some way. Most pairs $(S, \mathcal{C})$ of a set $S \in \mathcal{F}$ and a coloring $\mathcal{C}$ of the ground set are then shown to have some property, by the construction of an explicit injection of the “bad” pairs that do not have the property into some small collection. In this chapter, we use this encoding idea to prove a rather surprising fact about set systems. A w -set system is a family of sets, each of size at most w . Throughout this part, let $\mathcal{F}$ be a w -set system on a ground set X , and let $k \geq 2$ be some fixed integer.

Definition 8.1. *Call a k -tuple of sets $(S_1, \dots, S_k)$ from $\mathcal{F}$ n -inflatable if for each $1 \leq i \leq k$ there exists a subfamily $\mathcal{F}_i \ni S_i$ of $\mathcal{F}$ so that $|\mathcal{F}_i| \geq |\mathcal{F}|/n$, and so that for any $T_i \in \mathcal{F}_i$ and $T_j \in \mathcal{F}_j$ with $i \neq j$, we have $T_i \cap T_j = S_i \cap S_j$.*

We show, that in a large set system almost all k -tuples are inflatable, where the corresponding families $\mathcal{F}_i$ are very large.

Theorem 8.2. *Let $k \geq 2$ be fixed, and let $\mathcal{F}$ be a w -set system on some ground set X . For all choices of n , all but at most $\binom{k+w-1}{k-1} \frac{k2^{w(k-1)}}{n} |\mathcal{F}|^k$ of the $|\mathcal{F}|^k$ -tuples $(S_1, \dots, S_k) \in \mathcal{F}^k$*

are n -inflatable.

It would be quite interesting to reproduce any of the results proved here in some other way, and perhaps this would shed some light on the sunflower conjecture.

8.2 Proof Outline

We prove Theorem 8.2 by introducing the concept of “mimicking” sets, which allows us to capture inflatability through looking at colorings. Given a k -tuple $(S_1, \dots, S_k)$ from $\mathcal{F}$ and a coloring $\mathcal{C}$ on X , we say that T “mimics” S_j if it satisfies certain natural coloring conditions, and its intersections with the S_i are the same as those of S_j . If enough sets mimic each S_j , we will have the desired inflatability property. We can use an encoding argument to count the number of “bad” k -tuples $(S_1, \dots, S_k)$ so that some S_j does not have too many sets T mimicking it.

8.3 Blowing up Pairs of Sets

For $k = 2$, Theorem 8.2 yields the following.

Theorem 8.3 (Theorem 8.2, $k = 2$). *If $\mathcal{F}$ is a w -set system, then all but at most $\frac{2^w(2w+2)}{n}|\mathcal{F}|^2$ pairs (S, T) of sets from $\mathcal{F}$ are n -inflatable.*

We make note of a particularly interesting consequence. First, recall the definition of a link from [9]. Given a set system $\mathcal{F}$ on X and a set $U \subset X$, the *link* of $\mathcal{F}$ at U is $\mathcal{F}_U = \{S \setminus U : S \in \mathcal{F}, U \subset S\}$.

We are now ready to state another corollary of Theorem 8.3.

Corollary 8.4. *If S, T are chosen randomly from $\mathcal{F}$, then with probability at least $1 - \frac{2^w(2w+2)}{n}$, the link $\mathcal{F}_U$, where $U = S \cap T$, contains two cross-wise disjoint subfamilies $\mathcal{F}_1, \mathcal{F}_2$ each of size at least $|\mathcal{F}|/n$.*

One could for instance take $n = 3^w$ so that the link of the intersection of almost any (as w goes to infinity) pair of sets contains two cross-wise disjoint subfamilies of size $|\mathcal{F}|/3^w$. Another consequence of Corollary 8.4 is that in any induced subgraph of a Kneser graph $KG(m, w)$, a typical edge is contained in a large complete bipartite subgraph. We state this explicitly as a theorem.

Theorem 8.5. *Let $G = KG(m, w)$ be a Kneser graph, and let G' be an induced subgraph. Then for any n , and for all but at most $\frac{2^w(2w+2)}{n}|V(G')|^2$ edges e of G' , there is some complete bipartite subgraph of G' with $|V(G')|/n$ vertices on each side that contains e .*

In particular, if we have a w -set system $\mathcal{F}$ with more than $\varepsilon|F|^2$ pairs of disjoint sets, then $n = \frac{1}{\varepsilon}2^w(2w+2)$ in Theorem 8.5 yields that there are some disjoint subfamilies $\mathcal{F}_1$ and $\mathcal{F}_2$ of size at least $\frac{\varepsilon|\mathcal{F}|}{2^w(2w+2)}$. In other words, any set system with many pairs of disjoint sets has a large pair of disjoint subfamilies.

8.4 Sunflower Blowups

We will also use Theorem 8.2 to extend the sunflower lemma of Erdős and Rado [36]. As mentioned in [9], Erdős and Rado originally called sunflowers Δ -systems, but Deza and Frankl [34] coined the name “sunflower”.

Definition 8.6. *For a fixed $k \geq 2$, a k -petal sunflower is a k -tuple of sets $(S_1, \dots, S_k)$ so that for any $1 \leq i < j \leq k$, $S_i \cap S_j = S_1 \cap \dots \cap S_k$. The $S_i \setminus (S_1 \cap \dots \cap S_k)$ are called the petals, and $S_1 \cap \dots \cap S_k$ is called the kernel.*

The case $k = 2$ is trivial, because any pair of sets form a 2-petal sunflower, so typically we assume $k \geq 3$.

Corollary 8.7. *Let $f_k(w)$ be such that any w -set system with $f_k(w)$ sets contains a k -petal sunflower. Then, there exists some constant D_k depending on k so that if $\mathcal{F}$ is a w -set*

system on a ground set X , then almost all (as w goes to infinity) of its k -petal sunflowers $(S_1, \dots, S_k)$ are n -inflatable with parameter $n = f_k(w)^k D_k^w$.

The sunflower conjecture ([36]) states that we may take $f_k(w) = (O_k(1))^w$. By using the modifications of the original argument of the author, Lovett, Wu, and Zhang [9] due to Frankston, Kahn, Narayanan, and Park [41], Rao [74] proved that we may take $f_k(w) = (O_k(\log w))^w$. Thus, if the sunflower conjecture is true, there is some constant C_k depending on k so that if $\mathcal{F}$ is a w -set system, then for a typical k -petal sunflower $(S_1, \dots, S_k)$, there exists subfamilies $\mathcal{F}_1, \dots, \mathcal{F}_k$ of $\mathcal{F}$ of size at least $|\mathcal{F}|/C_k^w$ so that any $S_1 \in \mathcal{F}_1, \dots, S_k \in \mathcal{F}_k$ form a sunflower with kernel $S_1 \cap \dots \cap S_k$. Unconditionally, there is some constant C_k depending on k such that a typical k -petal sunflower can be blown up to yield families $\mathcal{F}_1, \dots, \mathcal{F}_k$ of $\mathcal{F}$ of size at least $|\mathcal{F}|/(C_k \log(w)^k)^w$ so that any $S_1 \in \mathcal{F}_1, \dots, S_k \in \mathcal{F}_k$ form a k -petal sunflower. This greatly improves the known bounds in Theorem 1 of [43], which has found several (see [46], [45]) applications.

9 Proofs

We assume $w \geq 2$; $w = 1$ is a trivial case. We can assume all sets are of size exactly w by adding dummy elements to sets that are too small, since this does not affect inflatability of any k -tuple. We may also assume the ground set X has size divisible by k , since we may add dummy elements to X to make this the case. Thus, we may randomly partition X into k equal-sized parts X_i for $1 \leq i \leq k$. We say that the elements of X_i have *color* i and we refer to the partition of X into the X_i as a *coloring* $\mathcal{C}$ of the ground set X . We call a coloring *balanced* if there are the same number of elements of each of the k colors.

Definition 9.1. Fix a w -set system $\mathcal{F}$ on a ground set X , a balanced coloring $\mathcal{C}$ of X , and a k -tuple $(S_1, \dots, S_k)$ in $\mathcal{F}^k$. We say that $T \in \mathcal{F}$ mimics S_j with respect to $(S_1, \dots, S_k)$ if for all $i \neq j$, $S_i \cap T = S_i \cap S_j$, and $T \setminus S_j \subset X_j$.

The last condition, that all elements of $T \setminus S_j$ are colored with the color j , at first glance appears to be rather restrictive. It turns out that this is surprisingly not the case, in the sense that for a random k -tuple, many sets will mimic each S_j .

Definition 9.2. For a fixed $1 \leq j \leq k$, say that $(S_1, \dots, S_k)$ is n - j -bad (under a balanced coloring $\mathcal{C}$) if there are fewer than $|\mathcal{F}|/n$ sets $T \in \mathcal{F}$ such that T mimics S_j with respect to $(S_1, \dots, S_k)$. Say $(S_1, \dots, S_k)$ is n -bad under $\mathcal{C}$ if it is n - j -bad under $\mathcal{C}$ for some $1 \leq j \leq k$. A k -tuple is n -good under $\mathcal{C}$ if it is not n -bad under $\mathcal{C}$.

The following lemma shows the relationship between the good and inflatable properties:

Lemma 9.3. If $(S_1, \dots, S_k)$ is n -good under some coloring $\mathcal{C}$, then it is n -inflatable.

Therefore, it suffices to prove the following lemma:

Lemma 9.4. Let $\mathcal{F}$ be a set system on a ground set X . For any $1 \leq j \leq k$, if $\mathcal{C}$ is a uniform balanced coloring of X and $(S_1, \dots, S_k)$ is sampled uniformly from $\mathcal{F}^k$, then $(S_1, \dots, S_k)$ is n -bad with probability at most $\binom{k+w-1}{k-1} \frac{k2^{w(k-1)}}{n}$.

The proof of Lemma 9.4 in Section 9.2 follows an encoding argument similar in spirit to that of [9]. For each fixed $1 \leq j \leq n$, this allows to explicitly bound the number of pairs of a balanced coloring $\mathcal{C}$ and a k -tuple $(S_1, \dots, S_k)$ so that $(S_1, \dots, S_k)$ is n - j -bad under $\mathcal{C}$.

9.1 Proof of Lemma 9.3

Take $\mathcal{F}_j$ to be the family of T_j which mimic S_j under the coloring $\mathcal{C}$. Note $|\mathcal{F}_j| \geq |\mathcal{F}|/n$ for each $1 \leq j \leq k$ by assumption. It suffices to show that if $T_i \in \mathcal{F}_i$ and $T_j \in \mathcal{F}_j$ for $i \neq j$, then we have $T_i \cap T_j = S_i \cap S_j$. Clearly $T_i \supset T_i \cap S_j = S_i \cap S_j$, and similarly $T_j \supset T_j \cap S_i = S_i \cap S_j$. Hence, $T_i \cap T_j \supset S_i \cap S_j$. Now, we prove $T_i \cap T_j \subset S_i$. If $x \in T_i \cap T_j$ is not in S_i , then since $T_i \setminus S_i \subset X_i$, we have $x \in X_i$. Because X_i and X_j are disjoint, $x \notin X_j \supset T_j \setminus S_j$, and hence $x \notin T_j \setminus S_j$. Since $x \in T_j$, we have $x \in S_j$. But then $x \in T_i \cap S_j = S_i \cap S_j$, a contradiction. Hence, $T_i \cap T_j \subset S_i$, and similarly $T_i \cap T_j \subset S_j$, so $T_i \cap T_j = S_i \cap S_j$.

9.2 Proof of Lemma 9.4

For a fixed $1 \leq j \leq k$, we will bound the number of pairs of balanced colorings $\mathcal{C}$ and k -tuples $(S_1, \dots, S_k) \in \mathcal{F}^k$ so that $(S_1, \dots, S_k)$ is n - j -bad under $\mathcal{C}$. We will do this by an “encoding” argument; we can recover this pair from the following information.

1. The first piece of information will be $\mathcal{C}'$, which is the coloring obtained by taking $\mathcal{C}$ and then recoloring the elements of S_j with the color j . We claim that the number of possibilities for this is at most $\binom{|\mathcal{F}|}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k} \binom{k+w-1}{k-1}$. This is because there are at most $\binom{k+w-1}{k-1}$ possibilities for the number of times that each of the k colors appears as one of the w elements of S_j , and by the log-convexity of the factorial function, at most $\binom{|\mathcal{F}|}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k}$ possibilities for $\mathcal{C}'$ once the number of elements of each color is fixed.
2. The second piece of information will be S_i for each $i \neq j$. There are $|\mathcal{F}|^{k-1}$ possibilities for this.
3. The third piece of information will be $S_i \cap S_j$ for each $i \neq j$. There are at most $2^{w(k-1)}$ possibilities for this, as for each of the $k-1$ sets S_i with $i \neq j$ we specify one of 2^w possible subsets for $S_i \cap S_j$.
4. The final piece of information specifies which of the mutually mimicking sets T_j is S_j . The sets T_j so that $T_j \cap S_i = S_j \cap S_i$ for each $i \neq j$ and so that all elements T_j have color j in $\mathcal{C}'$ must have $T_j \setminus S_j \subset X_j$. Hence, they must mimic S_j with respect to $(S_1, \dots, S_k)$, and so there are at most $|\mathcal{F}|/n$ possibilities for them by assumption. We can thus identify S_j by a positive integer at most $|\mathcal{F}|/n$.

Thus, of the $\binom{|\mathcal{F}|}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k} |\mathcal{F}|^k$ possible pairs of $\mathcal{C}$ and $(S_1, \dots, S_k)$, the number

so that $(S_1, \dots, S_k)$ is n - j -bad under $\mathcal{C}$ is at most

$$\begin{aligned} & \binom{\mathcal{F}}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k} |\mathcal{F}|^{k-1} \binom{k+w-1}{w-1} 2^{w(k-1)} |\mathcal{F}|/n \\ &= \left(\binom{\mathcal{F}}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k} |\mathcal{F}|^k \right) \left(2^{w(k-1)} \binom{k+w-1}{w-1} \frac{1}{n} \right). \end{aligned}$$

By a simple union bound over k choices of $1 \leq j \leq k$, the number of possible pairs of $\mathcal{C}$ and $(S_1, \dots, S_k)$ so that $(S_1, \dots, S_k)$ is n -bad under $\mathcal{C}$ is at most

$$\left(\binom{\mathcal{F}}{|\mathcal{F}|/k, \dots, |\mathcal{F}|/k} |\mathcal{F}|^k \right) \left(\binom{k+w-1}{w-1} k 2^{w(k-1)} \frac{1}{n} \right).$$

This completes the proof of Lemma 9.4.

9.3 Proof of Theorem 8.2

Using Lemma 9.4, there exists a specific coloring $\mathcal{C}$ under which at most $\binom{k+w-1}{k-1} k 2^{w(k-1)} |\mathcal{F}|^k / n$ of the $|\mathcal{F}|^k$ k -tuples $(S_1, \dots, S_k)$ are n -bad. In combination with Lemma 9.3, this immediately implies Theorem 8.2.

9.4 Proof of Corollary 8.4

If (S, T) is n -inflatable, then there exists subfamilies $\mathcal{F}(S) \ni S$ and $\mathcal{F}(T) \ni T$ of $\mathcal{F}$ of size at least $|\mathcal{F}|/n$ such that $S' \cap T' = S \cap T = U$ for all $S' \in \mathcal{F}(S), T' \in \mathcal{F}(T)$. Then for all $S' \in \mathcal{F}(S), T' \in \mathcal{F}(T)$, we have $U \subset S', U \subset T'$, and $(S' \setminus U) \cap (T' \setminus U) = (S' \cap T') \setminus U = U \setminus U = \emptyset$. Thus if we take $\mathcal{F}_1 = \{S' \setminus U : S' \in \mathcal{F}(S)\}$ and $\mathcal{F}_2 = \{T' \setminus U : T' \in \mathcal{F}(T)\}$, we are done.

9.5 Proof of Theorem 8.5

Let $\mathcal{F} = V(G')$. Then by Theorem 8.3, all but at most $\frac{2^w(2w+2)}{n}|V(G')|^2$ pairs of sets of $\mathcal{F}$ are inflatable. If (S, T) is n -inflatable and S, T are disjoint, then this means exactly that the edge e between S and T in G' is contained in a $K_{|V(G')|/n, |V(G')|/n}$ in G' .

9.6 Proof of Corollary 8.7

By Theorem 8.2, there exists a constant $C = C_k$ so that all but at most $\frac{C_k^w}{n}|\mathcal{F}|^k$ of the k -tuples of sets from $\mathcal{F}$ are n -inflatable. Now, if any family of $f_k(w)$ sets contains a k -petal sunflower, then by a simple counting argument, a w -set system $\mathcal{F}$ so that $|\mathcal{F}| \geq f_k(w)$ will have at least $\left(\frac{|\mathcal{F}|}{f(w)}\right)^k$ k -petal sunflowers. If $D_k = 2C_k$, then if $n = f(w)^k D_k^w$, almost all (as w goes to infinity) k -petal sunflowers are inflatable.

Part IV

Noisy Corruption Detection

This chapter is based on my work [6] published in *Information Processing Letters*.

10 Introductory Remarks

Alon, Mossel, and Pemantle [4] considered the following one-player game. A finite simple connected graph G on n vertices is given, and each vertex is either “truthful” or “corrupt”. The player may query a vertex u about whether its neighbor v is truthful or corrupt. Truthful vertices always accurately report the status of their neighbors, but corrupt vertices will report adversarial answers. The setting in particular that they consider is where the graph G is a constant-degree expander.

Definition 10.1. ([4]) *A δ -good expander for a constant $\delta > 0$ is a graph G with $|V(G)| = n$ so that all $U \subset V(G)$ with $|U| \leq 2\delta n$ have more than $|U|$ neighbors outside U , and for every $U, W \subset V(G)$ with $|U| \geq \delta n$ and $|W| \geq n/4$, there is an edge between some vertex in U and some vertex in W .*

The main result of Alon, Mossel, and Pemantle is the following.

Theorem 10.2. *Let G be a δ -good expander, and let $V(G) = T \sqcup B$ be a partition of $V(G)$ into truthful vertices T and corrupt vertices B . If $|T| > |B|$, then after asking every vertex about all its neighbors, the player can find $T' \subset T$ and $B' \subset B$ so that $|T' \cup B'| > (1 - \delta)n$.*

Interestingly, finding T' and B' efficiently depends on how large T is compared to B .

Theorem 10.3. *In the above setting, if $|T| > (1/2 + \delta)n$, there is an algorithm that finds such T', B' in linear time.*

However, if we only assume $|T| > |B|$ one cannot expect a polynomial time algorithm unless $P = NP$.

Even in this “noiseless” case where truthful vertices always report accurately, the player cannot exactly identify T and B . For example, say x and y have no neighbors in T and correctly report that all their neighbors are in B . Then, the player cannot distinguish the case $x \in B, y \in T$ from $x \in T, y \in B$. The authors then ask about the corresponding results in the noisy case, where now a truthful vertex answers correctly and independently with probability $1 - \epsilon$ for some fixed $\epsilon > 0$ which we consider to be constant. In particular, if a vertex u is asked about a neighbor v multiple times, it may answer differently each time. Thus, the appropriate parameter to consider is the number of queries needed so that with high probability, the player can identify T', B' so that the symmetric differences $|T \Delta T'|$ and $|B \Delta B'|$ are small. In other words, with high probability the player can identify the status of almost all vertices.

In this part we answer these questions; $\Theta(n \log(n))$ queries are required if $|T| - |B| \leq n^{1/2-\beta}$ (for any constant $\beta > 0$) but $O(n)$ suffice if $|T| = (1/2 + \delta)n$ for some $\delta > 0$. Throughout this part we will consider δ, ϵ , and the degree of the graph G to be constants. We write $o(1)$ for quantities that go to 0 as $n \rightarrow \infty$.

11 Lower Bound for Simple Majority

Let G be a δ -good expander with $V(G) = T \sqcup B$ and $|T| - |B| = 1$, so that a simple majority of the $n = |V(G)|$ vertices are truthful. We begin this section with a simple upper bound for this case, which was communicated to the author by Noga Alon. We will show it is tight.

Theorem 11.1. *It suffices to ask $O(n \log n)$ queries to find a $1 - 2\delta$ fraction of the truthful vertices with high probability.*

Proof. The player can perform the following algorithm. For each $(u, v) \in E(G)$ the player

asks u about the status of v a total of $c \log(n)$ times for some $c = c(\epsilon)$ and then takes the majority of the answers given, breaking ties arbitrarily. If c is chosen appropriately, then for each pair $(u, v) \in V(G)$ so that u is truthful, with probability more than $1 - \frac{1}{n^2}$ most of the responses that u gives about v are correct. Hence the probability that this holds for every pair $(u, v) \in E(G)$ with u truthful is $1 - O(\frac{1}{n}) = 1 - o(1)$. Then for all (u, v) with u truthful, most of u 's opinions about v will be correct, so this reduces to precisely Theorem 10.2, the deterministic case considered in [4]. Hence $O(n \log n)$ queries suffice. $\square$

It turns out that $\Omega(n \log n)$ queries are in fact necessary, not just for expanders but for any constant-degree graph G , and as long as the number of truthful and corrupt vertices differ by at most $n^{1/2-\beta}$ for a constant $\beta > 0$.

Theorem 11.2. *If G is a constant-degree graph, $|V(G)| = n$, $|T| - |B| \leq n^{1/2-\beta}$ for an arbitrary fixed $\beta > 0$, and the player only makes $o(n \log n)$ queries, then there exists a strategy for the corrupt vertices so that any strategy the player uses to find T' with $|T \Delta T'| \leq \frac{n}{10}$ will succeed with probability at most $\frac{1}{2} + o(1)$.*

Note this implies that querying each edge once in each direction is not a viable strategy.

Proof. We will assume T is uniformly chosen over all sets of some size t (where t is known to the player) so that $t - (n - t) \leq n^{1/2-\beta}$. We describe an adversarial strategy for the corrupt vertices. Corrupt vertices will give the wrong answer with probability $1 - \epsilon$ randomly on each query, just as truthful vertices give the right answer with probability $1 - \epsilon$ randomly on each query.

This strategy can be restated as follows. Imagine that whenever the player queries a corrupt vertex u , with probability $1 - 2\epsilon$ it lies, but with probability 2ϵ a fair coin is flipped, so that the response of u in this case is random noise. Furthermore, imagine that whenever the player queries a truthful vertex u , with probability $1 - 2\epsilon$ u tells the truth, but with probability 2ϵ a fair coin is flipped, so again the response is random noise.

Now, call a vertex v *obscured* if a coin is flipped every time v is queried or queried about. Only random noise will enter and leave an obscured vertex. If $c = c(\epsilon)$ is sufficiently small, a vertex that is involved in at most $c \log(n)$ queries has a probability at least $n^{-\beta/2}$ of being obscured. If the total number of queries is at most $\frac{cn \log(n)}{20}$, then at most $\frac{n}{10}$ vertices are involved in more than $c \log(n)$ queries. So then in expectation at least $\frac{1}{2}n^{1-\beta/2}$ vertices of G will be obscured, so with high probability at least $n^{1-\beta}$ vertices of G will be obscured.

After all queries are answered, the player is then allowed to learn which query answers were chosen by a coin flip. As those answers provide no information for the player, we can assume without loss of generality that they are thrown out. All remaining answers are guaranteed to be false if given by corrupt vertices and true if given by truthful vertices. We now have a subgraph G' of G some of whose edges are labelled with true and some with false. All obscured vertices O of G have become isolated. Let $U = V(G) \setminus O$.

The following lemma implies that after we remove the obscured vertices, the chance that most remaining vertices are truthful is very close to $\frac{1}{2}$.

Lemma 11.3. *Given a uniformly random partition of $V(G)$ of size n into two sets of fixed size T, B so that $0 \leq |T| - |B| \leq n^{1/2-\beta}$, if O is a set of size at least $n^{1-\beta}$ vertices of G , $\Pr[|T \setminus O| \geq |B \setminus O|] \leq \frac{1}{2} + o(1)$.*

Proof. First, by conditioning on $|O|$, note that it suffices to consider any fixed $|O| \geq n^{1-\beta}$.

Now we will consider a random partition of $V(G)$ into sets T_1, T_2, B where $|T_1| = |B|$ and $T_1 \cup T_2 = T$. Then $|T_2| \leq n^{1/2-\beta}$. However, since at least $n^{1-\beta} - n^{1/2-\beta} \geq \frac{1}{2}n^{1-\beta}$ vertices of $T_1 \cup B$ are picked, and these are equally likely to come from T_1 and B , a simple computation (or the local central limit theorem) shows that $\Pr[||T_1 \cap O| - |B \cap O|| \leq n^{1/2-\beta}] = o(1)$ so that

$$\Pr[|T_1 \cap O| - |B \cap O| > n^{1/2-\beta}] \geq \frac{1}{2} - o(1)$$

and the lemma follows. □

Now for every partition (X, Y) of $V(G) \setminus O$ so that $(X, Y) = (T \cap U, B \cap U)$ and $(X, Y) = (B \cap U, T \cap U)$ are consistent (i.e. no $x \in X$ calls $y \in Y$ truthful in G' or vica versa, no $x_1 \in X$ calls $x_2 \in X$ corrupt in G' , and similar for $y_1, y_2 \in Y$) such that $|X| \geq |Y|$, we have by symmetry that

$$\begin{aligned} & \Pr[(X = T \cap U) \wedge (Y = B \cap U) \mid |X| = |T \cap U|, |Y| = |B \cap U|] \\ &= \Pr[(Y = T \cap U) \wedge (X = B \cap U) \mid |Y| = |T \cap U|, |X| = |B \cap U|] \end{aligned}$$

assuming that the latter events have nonzero probability. Because $|T| \geq |B|$ and $|X| \geq |Y|$,

$$\Pr[(|X| = |T \cap U|) \wedge (|Y| = |B \cap U|)] \geq \Pr[(|Y| = |T \cap U|) \wedge (|X| = |B \cap U|)].$$

Therefore, by Bayes, we have that

$$\Pr[(X = T \cap U) \wedge (Y = B \cap U)] \geq \Pr[(Y = T \cap U) \wedge (X = B \cap U)].$$

This means that if an oracle gives the player X and Y so that one of them is $T \cap U$ and the other is $B \cap U$, then the strategy that gives him the greatest chance of success (assuming $|O| = o(n)$, which happens with high probability) is to guess that T' is whichever of X, Y is larger. However, this strategy will still succeed only a $\frac{1}{2} + o(1)$ fraction of the time, because with probability $\frac{1}{2} - o(1)$ we have $|B \cap U| > |T \cap U|$ by Lemma 11.3. $\square$

12 Upper Bound for Robust Majority

Say now that $|T| = (\frac{1}{2} + \delta)n$ for some fixed $\delta > 0$. We will assume now that G is a δ -*excellent expander*, a graph so that any two disjoint subsets of $V(G)$ of size δn have an

edge between them. Note that spectral expanders with appropriate parameters satisfy this stronger hypothesis, but it is still weaker than spectral expansion.

Theorem 12.1. *On a constant-degree, δ -excellent expander G , there exists C depending on the maximum degree of the graph and on δ so that with high probability, in at most Cn queries we can find T', B' with $|T\Delta T'|, |B\Delta B'| \leq 6\delta n$.*

Proof. Query each edge of the graph G in both directions c times, for some c depending on δ and the maximum degree, i.e. for every $(u, v) \in E(G)$ ask u about the status of v c times and v about the status of u c times, and take the majority responses each time. In doing so, we can amplify the error probability ϵ down to some function ε of δ and ϵ .

Thus we can assume a $1 - \varepsilon$ accuracy rate, and that each edge of the graph is queried once in both directions. If ε is chosen appropriately, then with high probability the number of edges $(u, v) \in E(G)$ so that u is truthful but reports incorrectly about v is less than δn .

Let $T = T_1 \sqcup T_2$ be a split of T into two sets of $(\frac{1}{4} + \frac{1}{2}\delta)n$ vertices. By assumption, there are fewer than δn edges $(u, v) \in E(G)$ with $u \in T_1, v \in T_2$ so that at least one of u, v calls the other corrupt.

Now say there are disjoint $X, Y \subset V(G)$ with $|X| = |Y| = (\frac{1}{4} + \frac{1}{2}\delta)n$ so that there are fewer than δn pairs (x, y) with $x \in X, y \in Y, (x, y) \in E(G)$ so that one of x, y calls the other corrupt. Then $X \cup Y$ contains at least $(\frac{1}{2} + \delta)n + (\frac{1}{2} + \delta)n - n = 2\delta n$ truthful vertices. Assume without loss of generality that X does not have fewer truthful vertices than Y , so that X contains a set X' of at least δn truthful vertices.

Now, assume that Y contains a set Y' of at least $3\delta n$ corrupt vertices. Then by the expansion property, only at most δn vertices of Y' do not have a neighbor in X' , so that at least $2\delta n$ vertices of Y' have a neighbor in X' . Hence there are at least $2\delta n$ pairs $(x', y') \in E(G)$ with $x' \in X', y' \in Y'$. For at least δn of these, x' reports correctly and calls y' corrupt, a contradiction. Hence Y has fewer than $3\delta n$ corrupt vertices. By assumption, $|X \cap T| \geq |Y \cap T|$, so $|X\Delta T| \leq |Y\Delta T|$ and $|(X \cup Y)\Delta T| \leq 2|Y\Delta T| \leq 6\delta n$.

Thus, our algorithm will look for two disjoint sets of vertices X, Y with $|X| = |Y| = (\frac{1}{4} + \frac{1}{2}\delta)n$ so that fewer than δn pairs $x \in X, y \in Y$ are such that $(x, y) \in E(G)$ and one of x, y calls the other corrupt. With high probability such sets X, Y exist, and in fact arbitrary X, Y with $X \cup Y = T$ suffice. For such X, Y we have that $|(X \cup Y)\Delta T| \leq 6\delta n$. $\square$

13 Discussions and Open Problems

There are several natural questions here. The first is to discern how many queries are needed when $|T| - |B|$ is sublinear but too big for our argument to work, say $n^{1/2}$, although the lower bound for the $|T| - |B| \leq n^{1/2-\beta}$ case will give something here as well. It would be interesting to see in what regime we require only $o(n \log n)$ queries. Is there a transition at $|T| - |B| = n^\gamma$ for some $\gamma < 1$?

The other natural question is to figure out whether or not the algorithm for the $|T| = (\frac{1}{2} + \delta)n$ case can be made efficient. The algorithm we present here takes exponential time but it is not at all clear that this is necessary, and in fact we conjecture that it is not. It may also be interesting to figure out whether the assumptions on the expansion in the upper bound can be relaxed, perhaps to the original δ -good expansion assumption of [4]. We conjecture that it can be. Noga Alon points out that there is a trivial algorithm when the graph is a good spectral expander. In that case if we simply query all edges (after enough amplification), and for any u we take the majority reports of its neighbors, we will correctly determine the status of almost all vertices with high probability. But this crucially uses that for most vertices u , the neighborhood $N(u)$ has a robust majority of truthful vertices. This follows from an appropriate spectral expansion, but is a stronger assumption than the one we make in this part.

Furthermore, our proof of the lower bound inspires the following question. What if the task is instead to find two sets of vertices, one of which has small symmetric difference with

the set of all truthful vertices, and the other of which has small symmetric difference with the set of all corrupt vertices, but without specifying which is which? This task may be easier, although we conjecture it is not unless the corrupt vertices almost always lie. Noga Alon also suggests a generalization of this, a “list-coloring” version of the problem, i.e. what if one wants to find a small number of candidates for T so that some candidate for T is close to T ? We feel the player should be able to learn some information about T .

Finally, it is worth mentioning that in their paper [4], Alon, Mossel, and Pemantle considered the directed setting. Our results generalize in a fairly straightforward way to the directed setting, albeit again with analogous slightly stronger guarantees for expansion necessary in the “robust majority” case.

Part V

The Subgraph Query Problem

The results of this part are based on joint work [10] with Chady Ben Hamida, Xiaoyu He, and Alexander Moreira.

14 Introductory Remarks

The *subgraph query problem*, introduced by Ferber, Krivelevich, Sudakov, and Vieira [39], has been the subject of recent attention in extremal combinatorics and theoretical computer science. The problem is to determine the smallest number of adaptive queries of the form “is $(u, v) \in E(G)$?” that need to be made to an Erdős-Rényi random graph $G \sim G(n, p)$ to find a copy of a given subgraph H with probability at least $\frac{1}{2}$.

Several variants of the problem appear in the literature. Ferber, Krivelevich, Sudakov, and Vieira [39, 40] first studied the subgraph query problem when H is a long path or cycle of order comparable to n , exhibiting asymptotically optimal algorithms for finding long paths and cycles. For example, as long as $p \geq \frac{\log n + \log \log n + \omega(1)}{n}$ is above the threshold for Hamiltonicity in $G(n, p)$, they showed that a Hamiltonian cycle can be found by the time one reveals $(1 + o(1))n$ edges. Here and henceforth we write $\log$ for the natural logarithm and $\lg$ for the base-2 logarithm.

In connection with the online Ramsey number, Conlon, Fox, Grinshpun, and He [32] studied the case where $H = K_m$ is a fixed complete graph, $p \rightarrow 0$, and the number of vertices n is allowed to be arbitrarily large. They defined the function $f(H, p)$ to be the number of queries needed to find a copy of H in the countably infinite random graph $G(\infty, p)$ with

probability $\frac{1}{2}$, and proved that

$$p^{-(2-\sqrt{2})m+O(1)} \leq f(K_m, p) \leq p^{-\frac{2}{3}m-O(1)}. \quad (1)$$

In this part, we study the behavior of $f(H, p)$ as $p \rightarrow 0$ for an arbitrary fixed graph H . We will use the phrases “build H in T time” and “find H in T queries” interchangeably for the statement $f(H, p) \leq T$.

Recall that a graph H is d -degenerate if every subgraph of H contains a vertex of degree at most d , and the *degeneracy* of H is the least d for which H is d -degenerate. Equivalently, H is d -degenerate if and only if there is an acyclic orientation of H with maximum outdegree at most d . Degeneracy is the natural notion of sparsity in graph Ramsey theory (see e.g. the recent proof of the Burr-Erdős conjecture by Lee [59]).

In the subgraph query problem, a d -degenerate graph can be built by adding one vertex at a time so that each new vertex has degree at most d at the time it is built. Since a common neighbor of d given vertices can be found in $O(p^{-d})$ queries, this shows that $f(H, p) = O_H(p^{-d})$ whenever H is d -degenerate.

Our first main result is that this trivial bound is never tight when $d \geq 2$. Define the *depth* Δ of a graph H with degeneracy d to be the smallest Δ for which there exists an acyclic orientation of H with maximum outdegree at most d and longest directed path of length at most Δ (we use the convention that the length of the path with $n + 1$ vertices is n). Let $\log_t(x)$ denote the t -fold iterated logarithm of x .

Theorem 14.1. *If H is a graph with degeneracy $d \geq 2$ and depth $\Delta \geq 1$, then*

$$f(H, p) = O_H\left(\frac{p^{-d} \log_{\Delta+1}(p^{-1})}{\log_{\Delta}(p^{-1})}\right).$$

Roughly speaking, one of the main innovations is to exploit the observation that in a

random graph $G(n, 1/n)$, the degrees of vertices are approximately Poisson with mean 1. Thus, the maximum degree is $\Omega(\log n / \log \log n)$ despite the fact that the average degree is constant. Repeatedly finding these vertices of exceptionally large degree allows us to find H slightly faster.

We will also show that the behavior in Theorem 14.1 can be correct up to the polylogarithmic factor. Let the trifold graph be the graph obtained from the triangle K_3 by adding a common neighbor to each pair of vertices (see Figure 15.1 in Section 15).

Theorem 14.2. *If H is the trifold graph and $\ell = \frac{\log(1/p)}{2 \log \log(1/p)}$, then*

$$f(H, p) = \Omega(p^{-2}/\ell^4).$$

Note that the trifold is 2-degenerate, so Theorems 14.1 and 14.2 together prove that $f(H, p) = o(p^{-2})$ and $f(H, p) = \Omega(p^{-2+\varepsilon})$ for every $\varepsilon > 0$. This is the first example of a graph for which it is known that $f(H, p)$ does not grow like a power of p^{-1} .

The question of querying for subgraphs in random graphs was also studied by Feige, Gamarnik, Neeman, Racz, and Tetali [38], and by Racz and Schiffer in the related planted clique model [72]. Feige, Gamarnik, Neeman, Racz, and Tetali restricted their attention to the balanced random graph $G(n, \frac{1}{2})$ and asked for the minimum number of queries needed to find a clique of order $(2 - o(1)) \lg n$ (which approaches the clique number) with probability at least $1/2$. For $\delta < 2$, define $\alpha_\star(\delta)$ to be the supremum over $\alpha \leq 2$ for which a clique of order $\alpha \lg n$ can be found with probability at least $1/2$ in at most n^δ queries for all n sufficiently large. They showed under the additional assumption that only a bounded number of rounds of adaptiveness are used that $\alpha_\star(\delta) < 2$ for all $\delta < 2$, and asked if this could be proved unconditionally.

Our last theorem answers this question affirmatively. We are grateful to Huy Pham [69] for communicating to us the main idea of the proof.

Theorem 14.3. *For all $2/3 < \delta < 2$,*

$$\alpha_\star(\delta) \leq 1 + \sqrt{1 - \frac{(2 - \delta)^2}{2}} < 2.$$

The proof is an adaptation of the lower bound proof for (1) by [32] to take the size of the vertex set into account. The exact value of $\alpha_\star(\delta)$ remains open for all δ , and the best known lower bound is $\alpha_\star(\delta) \geq 1 + \frac{\delta}{2}$ when $1 \leq \delta < 2$ (see [38, Lemma 6]).

Organization. In Section 15, we describe a new algorithm for finding any d -degenerate graph and prove that it achieves the runtime described in Theorem 14.1. In Section 16, we give a new argument for proving lower bounds on $f(H, p)$, proving Theorem 14.2. In Section 17 we give a short proof of Theorem 14.3, using a variation of the methods in [32]. Finally, Section 18 highlights a few of the many open questions that remain about $f(H, p)$.

We will write $b = p^{-1}$ for the expected number of queries needed to find a single edge in $G(\infty, p)$. No attempt will be made to optimise the implicit constants in any of our proofs. We use $A \lesssim B$ to mean $A = O(B)$. For the sake of clarity of presentation, we systematically omit floor and ceiling signs whenever they are not crucial.

15 Upper bounds

15.1 An Illustrative Example

As mentioned in the introductory remarks, there is a straightforward algorithm for finding any d -degenerate graph H in $O_H(b^d)$ time. In this section we prove Theorem 14.1 by providing a new algorithm that beats the trivial algorithm by an iterated logarithmic factor.

We begin by illustrating how the algorithm works with a specific 2-degenerate graph.

Definition 15.1. *The trifold is the graph on 6 vertices and 9 edges obtained from the*

triangle K_3 by adding a common neighbor to each pair of vertices.

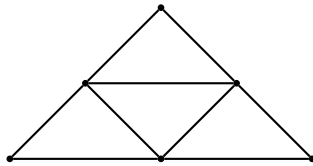


Figure 1: The triforme graph.

The key step in building the triforme quickly is to build a large book.

Definition 15.2. *The book $B_{d,t}$ is the graph on $d+t$ vertices obtained by removing the edges of a clique K_t from a complete graph K_{d+t} . The t vertices of the removed clique are called the pages of the book and the remaining d vertices are called its spine.*

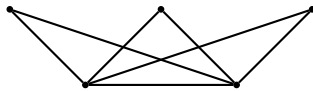


Figure 2: The book $B_{2,3}$.

Note that when d and t are fixed positive integers, $B_{d,t}$ is d -degenerate and thus we have $f(B_{d,t}, p) = O_t(b^d)$. The key observation is that this can be improved substantially even if we allow t to grow slowly as p tends to 0.

Lemma 15.3. *If $d \geq 2$ and $\ell = \frac{\log b}{2 \log \log b}$, then $f(B_{d,\ell}, p) = O(b^d \ell^{-1/2})$.*

Proof. We will exhibit an algorithm which finds $B_{d,\ell}$ in $G(\infty, p)$ with constant probability (w.c.p.) in $O(b^d \ell^{-1/2})$ time. The algorithm has three steps.

First, we find w.c.p. $d-1$ vertices $v_1, \dots, v_{d-1}$ of the spine forming a clique in $O(b^{d-2})$ time, which is possible because K_{d-1} is $(d-2)$ -degenerate. Assume this step succeeds.

Next, we build a large pool S of common neighbors of $v_1, \dots, v_{d-1}$, which will serve as candidates for the remaining vertex v_d of the spine and for the pages of the book. In $d-1 = O(1)$ queries we can check a single new vertex u to see if it is a common neighbor

of $v_1, \dots, v_{d-1}$, and u has a probability p^{d-1} of being such a common neighbor. We check a total of $4b^d \ell^{-1/2}$ possible u , and each common neighbor successfully found is added to S . Since the outcomes of all queries are independent, $|S|$ is distributed like the binomial random variable $\text{Bin}(4b^d \ell^{-1/2}, p^{d-1})$ with mean $4b \ell^{-1/2}$, so w.c.p. $|S| \geq 2b \ell^{-1/2}$.

For the last step, assuming the previous two steps succeed, we will find a star $K_{1,\ell}$ contained in S . Along with vertices $v_1, \dots, v_{d-1}$ already chosen, this forms the desired book.

To find this star, remove vertices from S until it has size exactly $2b \ell^{-1/2}$, and then query all pairs of vertices in S in $O(b^2 \ell^{-1})$ time. The induced subgraph on S is just an Erdős-Rényi random graph $G(2b \ell^{-1/2}, p)$. It suffices to show that w.c.p. there exists a vertex of degree at least ℓ therein. This fact is a consequence of the observation that the degrees are approximately Poisson.

To give a quick proof of this fact, divide S into two sets S_1, S_2 of size $r = b \ell^{-1/2}$, let $u_1, \dots, u_r$ be the vertices of S_1 , and let X_i be the number of neighbors of u_i in S_2 . Then $\{X_i\}_{i=1}^r$ are r i.i.d. random variables distributed like $\text{Bin}(r, p)$, so

$$\mathbb{P}[X_i \geq \ell] \geq \binom{r}{\ell} p^\ell (1-p)^{r-\ell} \geq \frac{(r-\ell)^\ell p^\ell (1-p)^r}{\ell!}.$$

As $p \rightarrow 0$, we can bound $r - \ell > b \ell^{-1/2}/2$, $(1-p)^r \rightarrow 1$, and $\ell! < \ell^\ell$. Thus,

$$\mathbb{P}[X_i \geq \ell] \geq \Omega\left(\frac{1}{2^\ell \ell^{3\ell/2}}\right).$$

When $\ell = \frac{\log b}{2 \log \log b}$, this fraction is certainly $\Omega(b^{-\frac{4}{5}})$. In particular, since there are $r = b^{1-o(1)}$ independent random variables X_i , w.c.p. some X_i is at least ℓ , as desired.

Letting the vertex of degree ℓ be the last vertex v_d of the book's spine and its ℓ neighbors in S be the pages of the book, we have found a copy of $B_{d,\ell}$ w.c.p. in $O(b^d)$ total queries, as desired. $\square$

We are now ready to prove a stronger version of Theorem 14.1 when H is the triforce.

Theorem 15.4. *If H is the triforce graph and $\ell = \frac{\log b}{2 \log \log b}$, then $f(H, p) = O(b^2 \ell^{-\frac{1}{2}})$.*

Proof. We exhibit an algorithm for finding H w.c.p. in $O(b^2 \ell^{-\frac{1}{2}})$ time.

Using Lemma 15.3 with $d = 2$, build a copy of $B_{2,\ell}$. Let x and y be the two vertices of its spine and let Z be its pages. In $O(b^2 \ell^{-\frac{1}{2}})$ time we can w.c.p. find two sets of vertices S_x, S_y , each of size $b \ell^{-\frac{1}{2}}$, so that everything in S_x is adjacent to x and everything in S_y is adjacent to y . Now, we will query all pairs between S_x and Z as well as all pairs between S_y and Z . This takes only $O(b \ell^{1/2})$ time, which is negligible.

We claim that w.c.p. there exist $x' \in S_x$, $z \in Z$, and $y' \in S_y$ so that $x' \sim z$ and $z \sim y'$. This follows because w.c.p., $\Theta(\ell^{1/2})$ vertices of Z have a neighbor in S_x , and among these vertices w.c.p. at least one has a neighbor in S_y . Now, let z' be any common neighbor of x and y in Z other than z . It follows that the six vertices x, y, z, x', y', z' form a triforce, and we have found it in $O(b^2 \ell^{-\frac{1}{2}})$ queries w.c.p., as desired. $\square$

15.2 The General Upper Bound

Roughly speaking, the main trick in the proofs of Lemma 15.3 and Theorem 15.4 is that we can find vertices of much larger than average degree in a random graph with constant average degree. We will iterate this trick many times to prove the general statement in Theorem 14.1.

We will construct an arbitrary d -degenerate graph H recursively. If the vertices of H are ordered $v_1, \dots, v_n$ in the degeneracy order, the algorithm will maintain a “cloud” of candidates C_i for the image of vertex v_i which shrinks as the algorithm progresses. On step i , the algorithm chooses v_i from C_i and then shrinks the clouds corresponding to neighbors of v_i to stay consistent with this choice.

Proof of Theorem 14.1. Let H be a graph on n vertices with degeneracy $d \geq 2$ and depth

Δ . Order its vertices $v_1, \dots, v_n$ so that each v_i has at most d neighbors v_j with $j < i$, and the longest left-to-right path $v_{i_0}, \dots, v_{i_r}$ with $i_0 < \dots < i_r$ has length $r = \Delta$. Let Δ_i be the length (in edges) of the longest left-to-right-path ending at v_i , so that $\Delta_i \leq \Delta$ for all i . Finally, define

$$L(x) = \frac{\log x}{3n \log \log x}$$

and $\ell_i = L^{(\Delta_i)}(b)$ is obtained from b by iterating L Δ_i times.

We describe an algorithm for finding an injection ϕ from H to $G(\infty, p)$ in a series of rounds, assuming p is sufficiently small. There are many points at which the algorithm may fail. However, each round succeeds with probability $\Omega_H(1)$ conditional on the success of all previous rounds, and there are $n = O_H(1)$ rounds, so the entire algorithm succeeds with $\Omega_H(1)$ probability. The algorithm can then be repeated a number of times depending only on H until its success probability reaches $\frac{1}{2}$; this only changes the implicit constant in $f(H, p)$.

We begin by setting aside n disjoint sets (“clouds”) $C_1, \dots, C_n$ which will change throughout the algorithm. We initialize these to $C_1^{(0)}, \dots, C_n^{(0)}$ of order $|C_i^{(0)}| = b^d / \ell_i$, so that $C_i^{(0)}$ is the set of candidates for $\phi(v_i)$. We proceed in n rounds, so that $C_j^{(k)}$ will refer to the state of cloud C_j after round k . After the k th round we will have nonempty disjoint sets $C_j^{(k)}$, and we always have $C_j^{(k-1)} \supset C_j^{(k)}$. In the round k a number of queries are made to decide the value of $\phi(v_k) \in C_k^{(k-1)}$. Thus $C_k^{(k)}$ is the singleton $\{\phi(v_k)\}$ and C_k remains a singleton until the end. For each j with $j > k$ and $v_j \sim v_k$, the set $C_j^{(k-1)}$ is updated to a subset $C_j^{(k)}$ consisting of all elements of $C_j^{(k-1)}$ adjacent to $\phi(v_k)$. We say that a vertex v_j is *living* after round k if $j > k$, and *dead* otherwise.

Two properties are maintained. The first is that after round k , for any $i \leq k$ and $j \leq n$ and any $u_i \in C_i^{(k)}$ and $u_j \in C_j^{(k)}$, $u_i \sim u_j$ if $v_i \sim v_j$. In other words, the adjacency relations are correct within the dead vertices and between the dead vertices and the clouds $C_i^{(k)}$ for the living ones.

The second property is that the size of the set $C_j^{(k)}$ must be

$$c_j^{(k)} = \begin{cases} \frac{b^{d-m}}{\ell_j} & \text{if } v_j \text{ is living and has } m < d \text{ dead left-neighbors,} \\ \ell_j & \text{if } v_j \text{ is living and has exactly } d \text{ dead left-neighbors,} \\ 1 & \text{if } v_j \text{ is dead.} \end{cases}$$

The queries on round k are made to guarantee two properties. First, on round k , vertices are thrown out of $C_k^{(k-1)}$ until it has exactly ℓ_k vertices (this is possible because $c_k^{(k-1)} \geq \ell_k$ when b is sufficiently large). Then, consider the j so that $j > k$ and $v_j \sim v_k$. If such a v_j has exactly $d - 1$ dead left-neighbors, then j is called *active* on round k , and otherwise j is called *inactive*. For each active j , all pairs in $C_k^{(k-1)} \times C_j^{(k-1)}$ are queried.

Each round is divided into an *active portion*, which happens first, and then an *inactive portion*.

The active portion of round k succeeds if a candidate $u_k \in C_k^{(k-1)}$ is found to have at least $c_j^{(k)}$ neighbors in $C_j^{(k-1)}$ for all the active j . One such candidate u_k is picked for $\phi(v_k)$ and $C_j^{(k)}$ is chosen to be exactly $c_j^{(k)}$ neighbors of u_k in $C_j^{(k-1)}$.

Then, for all inactive j , all pairs $\{u_k\} \times C_j^{(k-1)}$ are queried. The inactive portion of round k succeeds if after these queries a total of $c_j^{(k)}$ neighbors are found for u_k in $C_j^{(k-1)}$ for each of the inactive j 's as well. We say that the round succeeds if both the active and inactive portions succeed. The algorithm only continues past round k if round k succeeds.

By induction on k , the algorithm maintains all the required properties and produces a valid injection $\phi : H \rightarrow G(\infty, p)$ if it succeeds on every round. It remains to show that the probability of success on each round is $\Omega_H(1)$.

For each $u \in C_k^{(k-1)}$ and $j > k$ for which $v_j \sim v_k$, let $d_j(u)$ be number of neighbors u has in $C_j^{(k-1)}$. Note that $d_j(u)$ is distributed like $\text{Bin}(c_j^{(k-1)}, p)$, since each vertex of $C_j^{(k-1)}$ is adjacent to u independently with probability p .

Suppose j is active in round k , so that $c_j^{(k-1)} = b/\ell_j$. This time, we get

$$\mathbb{P}[d_j(u) \geq c_j^{(k)}] = \mathbb{P}[\text{Bin}(b/\ell_j, p) \geq \ell_j] \geq \binom{b/\ell_j}{\ell_j} p^{\ell_j} (1-p)^{b/\ell_j - \ell_j}.$$

Using the facts that $1-x \geq e^{-2x}$ for all $x \in [0, \frac{1}{2}]$ and that $\ell_j \rightarrow \infty$ as $p \rightarrow 0^+$, we see that $(1-p)^{b/\ell_j - \ell_j} \geq e^{-2/\ell_j} \rightarrow 1$. Also, $\binom{a}{b} \geq (a/b)^b$ for all $a \geq b \geq 1$, and so

$$\mathbb{P}[d_j(u) \geq c_j^{(k)}] \geq \binom{b/\ell_j}{\ell_j} p^{\ell_j} (1-p)^{b/\ell_j - \ell_j} = \Omega(\ell_j^{-2\ell_j}).$$

There are at most n total j , so taking a product over all active j , we arrive at a lower bound

$$\mathbb{P}[d_j(u) \geq c_j^{(k)} \text{ for all active } j] \geq \Omega_H(\ell_j^{-2n\ell_j}).$$

Since each $u \in C_k^{(k-1)}$ is individually a successful candidate for $\phi(v_k)$ with this probability, and these are ℓ_k independent events, it follows that

$$\mathbb{P}[\text{Active portion round } k \text{ succeeds}] \geq \Omega_H(\min(1, \ell_k \ell_j^{-2n\ell_j})).$$

Finally, we observe that for every j active in round k , $\Delta_j \geq \Delta_k + 1$ since every left-to-right path ending at k extends to a longer one ending at j . Thus, $\ell_j \leq L(\ell_k)$, and the function L was chosen so that $L(x)^{2nL(x)} \leq x$ for x sufficiently large. It follows that the active portion of round k succeeds with probability $\Omega_H(1)$, as desired.

Now we look at the inactive j in round k . Then $c_j^{(k)} = pc_j^{(k-1)}$, so

$$\mathbb{P}[d_j(u_k) \geq c_j^{(k)}] = \mathbb{P}[\text{Bin}(c_j^{(k-1)}, p) \geq pc_j^{(k-1)}] = \Omega(1).$$

Thus, conditional on the success of the active portion of round k , the inactive portion succeeds with probability $\Omega_H(1)$ as well.

We have now shown that the algorithm, iterated $O_H(1)$ times, succeeds with probability $\frac{1}{2}$. It remains to bound the total number of queries made. In the active portion of each round, queries are only made between sets $C_k^{(k-1)}$ and $C_j^{(k-1)}$ if j is relevant, which implies that $c_j^{(k)} = b/\ell_j$. Also, elements of $C_k^{(k-1)}$ were thrown out until it had size exactly $\ell_k = O(b)$, so the number of queries made in the active portion of any round is at most $O(b^2/L^{(\Delta)}(b)) = O(b^d/L^{(\Delta)}(b))$.

In the inactive portion of each round, queries are made between a single vertex u_k and sets $C_j^{(k-1)}$ of size at most b^d/ℓ_j each. Thus, the number of queries made in the inactive portion of any round is also $O(b^d/L^{(\Delta)}(b))$.

Since there are at most $n = O_H(1)$ rounds and at most n choices of j involved in each round, we find that

$$f(H, p) = O_H\left(\frac{b^d}{L^{(\Delta)}(b)}\right) = O_H\left(\frac{p^{-d} \log_{\Delta+1}(p^{-1})}{\log_{\Delta}(p^{-1})}\right).$$

as desired. □

16 Lower bounds

16.1 Preliminaries

In this section, we will prove lower bounds for $f(H, p)$. Because N queries necessarily involve at most $2N$ vertices, it suffices to prove lower bounds for finding a copy of H in $G(2N, p)$ rather than in $G(\infty, p)$. Following [32], we will lower bound the number of queries it takes to build a copy of H by showing that the expected number of copies of H we can build in some amount of time is not too large.

Definition 16.1. *If H is a graph without isolated vertices, define $t(H, p, N)$ to be the maximum (over all querying strategies) expected number of copies of H that can be found in*

$G(\infty, p)$ in N queries. Since we are working on $G(2N, p)$, if $H = H' \cup \{v_1, \dots, v_t\}$ has t isolated vertices, define $t(H, p, N) = (2N)^t \cdot t(H', p, N)$.

If we show that we cannot build so many copies of H (in expectation) in some time, this gives us a lower bound on how long it takes to build a single copy of H .

Lemma 16.2. *If $N \geq f(H, p)$, then*

$$f(H, p) \cdot t(H, p, N) \geq N/4.$$

Thus, upper bounds on $t(H, p, N)$ will yield lower bounds on $f(H, p)$. The proof of Lemma 16.2 is straightforward, but we include it for completeness.

Proof. By definition, there exists a strategy which finds H with probability $\frac{1}{2}$ in $f(H, p)$ queries. Given N queries, we can iterate this strategy $\lfloor N/f(H, p) \rfloor$ independent times on disjoint vertex sets. By linearity of expectation, this implies

$$t(H, p, N) \geq \left\lfloor \frac{N}{f(H, p)} \right\rfloor \cdot \frac{1}{2} \geq \frac{N}{4f(H, p)}. \quad \square$$

Thus, it suffices to produce upper bounds on $t(H, p, N)$. Fortunately, we can recursively bound $t(H, p, N)$ in terms of $t(H', p, N)$ for some subgraphs H' . The following bounds are proved in [32].

Lemma 16.3 ([32]). *If H is any graph, $p \in (0, 1)$, and $N \geq p^{-1-\varepsilon}$ for some $\varepsilon > 0$, then the following inequalities hold:*

$$t(H, p, N) \leq \min_{e \in E(H)} t(H \setminus e, p, N) \tag{2}$$

$$t(H, p, N) \lesssim p \max_{e \in E(H)} t(H \setminus e, p, N) \tag{3}$$

$$t(H, p, N) \lesssim pN \min_{e \in E(H)} t(H \setminus \{u, v\}, p, N), \tag{4}$$

where u, v are the vertices of e in (4). In the latter two inequalities, the implicit constants are allowed to depend only on H .

In general, the bounds of Lemma 16.3 are not tight. In certain cases, we will improve these bounds using the crucial observation that large enough sets of vertices in a random graph have few common neighbors. For any vertex subset $U \subseteq V(G)$ of a graph G , write $d(U)$ for the number of common neighbors of every vertex in U .

Lemma 16.4. *Let $\ell = \frac{\log b}{2 \log \log b}$, let $k, n \geq 2$ be absolute constants, and let $G = G(2N, p)$.*

1. *If $p^k N \lesssim 1$, then there exists $C > 0$ so that*

$$\mathbb{P}[\max_{|U|=k} d(U) > C\ell] < p^n,$$

where the maximum is taken over all k -subsets U of $V(G)$.

2. *If $p^k N = (1/N)^{\Omega(1)}$, then there exists $C > 0$ so that*

$$\mathbb{P}[\max_{|U|=k} d(U) > C] < p^n.$$

Proof. For an arbitrary set of k vertices U , note that $d(U) \sim \text{Bin}(2N - k, p^k)$ as there are $2N - k$ other vertices of $G(2N, p)$ and each vertex has a p^k chance of being adjacent to all members of U . Hence, we find that

$$\mathbb{P}[d(U) \geq t] = \mathbb{P}[\text{Bin}(2N - k, p^k) \geq t] \leq \binom{2N - k}{t} p^{tk} \leq \left(\frac{2Ne}{t} p^k\right)^t. \quad (5)$$

If $p^k N \lesssim 1$, then (5) implies $\mathbb{P}[d(U) \geq t] \leq (O(1/t))^t$. Next we take a union bound over at most $(2N)^k$ choices of U , which shows that, if we take $t = C\ell$ for a sufficiently large C depending on n ,

$$\mathbb{P}[\max_{|U|=k} d(U) \geq t] \leq (O(1/t))^t \cdot N^k < p^n.$$

This proves the first part of the lemma.

If $p^k N = (1/N)^{\Omega(1)}$, then (5) implies the stronger bound $\mathbb{P}[d(U) \geq t] = N^{-\Omega(t)}$. For any fixed n , if C is a large enough constant, the probability that $\max_{|U|=k} d(U) > C$ will be below $N^{-\Omega(C)} N^k \leq p^n$ by the union bound. $\square$

The power of Lemma 16.4 is that the final graph that we find after N queries is a subgraph of $G(2N, p)$, so we can bound the number of common numbers of any vertex set U of constant size without even seeing the graph. It allows us to prove new upper bounds on $t(H, p, b^d)$.

Lemma 16.5. *Let $\ell = \frac{\log b}{2 \log \log b}$, let H be a graph, let $v \in V(H)$, and let $H' = H \setminus \{v\}$.*

1. *If $d(v) = d$, then*

$$t(H, p, b^d) \lesssim \ell t(H', p, b^d).$$

2. *If $d(v) > d$, then*

$$t(H, p, b^d) \lesssim t(H', p, b^d).$$

Proof. Let $k = d(v)$, and let the neighbors of v in H be $v_1, \dots, v_k$. Fix a query strategy that maximizes $t(H, p, b^d)$.

For any subset $U = \{u_1, \dots, u_k\}$ of k vertices of the final graph $G \subset G(2b^d, p)$ that is found, let $H'(U)$ be the number of maps $H' \rightarrow G$ so that for all $1 \leq i \leq k$, v_i maps to u_i . Then we have that

$$t(H, p, b^d) = \mathbb{E} \left[\sum_{|U|=k} d(U) H'(U) \right] \leq \mathbb{E} \left[\left(\max_{|U|=k} d(U) \right) \left(\sum_{|U|=k} H'(U) \right) \right], \quad (6)$$

where the sum is taken over all k -sets of vertices U .

Assume $d(v) = d$. By the first part of Lemma 16.4, there is a large constant $C = C(H)$ so that

$$\mathbb{P} \left[\max_{|U|=k} d(U) > C\ell \right] < p^{d|V(H')|+1}.$$

We will break up the expectation in (6) depending on the size of $\max_{|U|=k} d(U)$. If $\max_{|U|=k} d(U) \leq C\ell$, the contribution to right side of (6) is $O(\ell t(H', p, b^d))$. Now, it holds that $\max_{|U|=k} d(U) > C\ell$ with probability at most $p^{d|V(H')|+1}$, so the contribution from these terms is bounded by $p^{d|V(H')|+1}(2b^d)^{|V(H')|} = o(1)$.

Likewise, when $d(v) > d$, by the second part of Lemma 16.4 there is a C so that the case of $\max_U d(U) \leq C$ contributes $O(t(H', p, b^d))$ to the right side of (6), and the case of $\max_{|U|=k} d(U) > C$ contributes $o(1)$. $\square$

16.2 Proof of Theorem 14.2

We now begin the proof of Theorem 14.2. The main idea is to apply Lemma 16.5 to obtain new upper bounds on $t(H, p, N)$ for various subgraphs H of the triforce, and then combine these with Lemma 16.3 to bound $t(H, p, N)$ for the triforce itself.

We describe the subgraphs of the triforce to which we will apply Lemma 16.5. Any copy of the triforce must arise from a copy of one of the graphs formed by deleting two edges from the triforce. There are 8 such graphs up to isomorphism, which we denote by H_i for $1 \leq i \leq 8$ (see Figure 3).

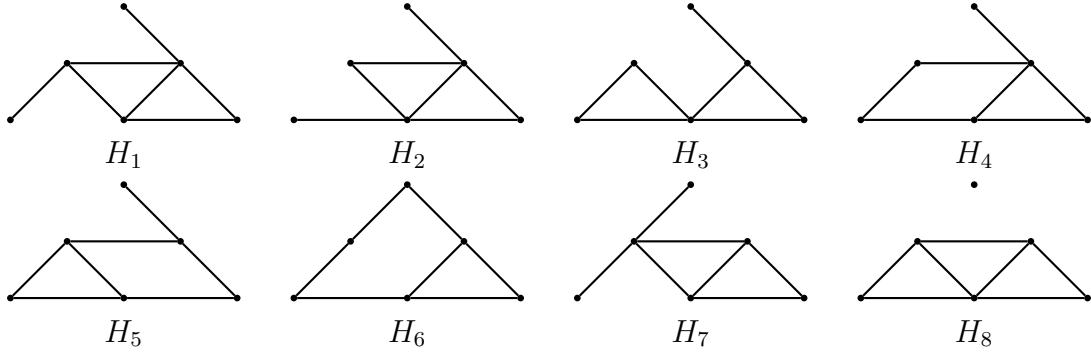


Figure 3: The eight subgraphs (up to isomorphism) with 7 edges of the triforce.

We will prove that the first 6 of these graphs are hard to construct quickly, although it turns out that there is no need to analyze H_1 , H_2 , or H_7 . The last subgraph H_8 is more

difficult to handle, and we will bound copies of it using a different analysis.

Lemma 16.6. *For all H_i such that $1 \leq i \leq 6$, $t(H_i, p, b^2) \lesssim b^2 \ell^2$.*

Proof. For each graph H_i with $1 \leq i \leq 6$, it is possible to remove two vertices of degree at least two to arrive at the path P_3 on four vertices. Thus, we may apply the first part of Lemma 16.5 twice to show that

$$t(H_i, p, b^2) \lesssim \ell^2 t(P_3, p, b^2),$$

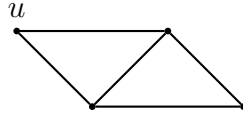
for all $1 \leq i \leq 6$. Lastly, note that $t(P_3, p, b^2) \lesssim bt(K_2, p, b^2) \lesssim b^2$ by applying (4) from Lemma 16.3.

Hence, for H_i with $1 \leq i \leq 6$, $t(H_i, p, b^2) \lesssim \ell^2 t(P_3, p, b^2) \lesssim \ell^2 b^2$ as desired. $\square$

We must now deal with H_8 , on which the reductions of Lemma 16.5 and Lemma 16.3 are not sufficient to provide the bounds that we want.

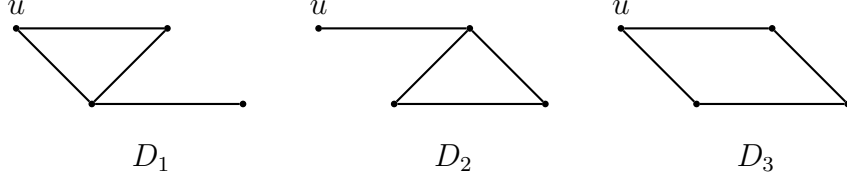
We need one last definition. Given a graph H with a distinguished vertex u , let $t^u(H, p, N)$ be the maximum expected number of copies of H we can build in time N so that u maps to the same vertex in each copy. It is important to emphasize that the image of u is not determined ahead of time, and we may pick it adaptively based on the queries made so far.

Lemma 16.7. *Let D be the diamond graph depicted below:*



Then $t^u(D, p, b^2) \lesssim \ell^3$.

Proof. As usual, we fix a query strategy maximizing $t^u(D, p, b^2)$ and let $G \subset G(2b^2, p)$ be the final graph built. Consider the following 3 subgraphs of D which we call D_1 , D_2 , and D_3 respectively.



Every copy of D must arise from adding an edge to a graph isomorphic to one of the D_i . For each $u' \in V(G)$, let $X_i(u')$ be the random variable counting the number of copies of D so that u maps to u' , and the last edge built in D is the edge missing from D_i . For each $1 \leq i \leq 3$, the number of copies of D we can build so that u maps to the same vertex in each copy, and so that D arises from some copy of D_i , is bounded by $\max_{u'} X_i(u')$. Thus,

$$t^u(D, p, b^2) \leq \sum_{i=1}^3 \mathbb{E}[\max_{u'} X_i(u')]. \quad (7)$$

Also, define the random variable $X_i(u', j)$ to be the number of copies of D_i with u mapping to u' that turn into a copy of D after query j . In particular this number is 0 if the query j finds a non-edge. We have that $X_i(u') = \sum_j X_i(u', j)$.

By the first part of Lemma 16.4, in the random graph $G(2b^2, p)$ any two vertices have $O(\ell)$ common neighbors with overwhelmingly high probability. We can assume this is the case here as the contribution to the expectation $t^u(D, p, b^2)$ from other cases is $o(1)$. In particular, this means that each new edge built can turn at most $O(\ell^2)$ copies of D_1 , D_2 or D_3 into D . For example, if an edge (u', v') is built in G , then the number of copies of D_2 that can be completed into D is exactly the number of ways to choose a common neighbor w' of u' and v' , and then a common neighbor of v' and w' . As we assumed that codegrees are all $O(\ell)$, there are only $O(\ell^2)$ total choices for this copy of D_2 .

This means we may assume that $X_i(u', j)$ is stochastically dominated (up to a constant) by $\ell^2 \text{Bin}(1, p)$. As the results of all queries are independent, it follows that $X_i(u')$ is stochas-

tically dominated by a constant times $\ell^2 \text{Bin}(b, p)$. Now it is a short computation that

$$\mathbb{P}[\text{Bin}(b, p) > 100\ell] < p^5.$$

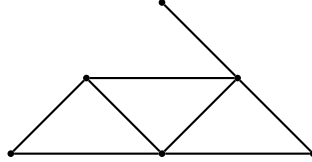
In particular, there exists a $C > 0$ such that $\mathbb{P}[X_i(u') > C\ell^3] < p^5$ for all $1 \leq i \leq 3$ and all $u' \in V(G)$. Also, the maximum possible number of diamonds with a given vertex u' is $(2b)^3$, so

$$t^u(D, p, b^2) \leq 3C\ell^3 + \mathbb{P}[X_i(u') > C\ell^3 \text{ for some } i, u'] \cdot (2b)^3 = O(\ell^3),$$

by the union bound over all $O(b^2)$ choices of $1 \leq i \leq 3$ and $u' \in V(G)$, as desired. $\square$

Finally, we deal with the graph H_8 . This graph behaves differently from the other ones, in that it is not the case that $t(H_8, p, b^2) = O(b^{2+o(1)})$ (in fact one can build $\Omega(b^3)$ copies of H_8 due to the isolated vertex). We will need to add one edge to H_8 and analyze the resulting graph instead.

Lemma 16.8. *Let H^* be the following graph:*



Then $t(H^*, p, b) = O(b\ell^3)$.

Proof. Let u be the vertex adjacent to the leaf of H^* . For any vertex v of our final graph $G \subset G(2b^2, p)$, let $f(v)$ be the number of copies of the diamond D so that u maps to v . As before, we may assume that all degrees are $O(b)$ as the contribution to $t(H^*, p, b^2)$ is trivial otherwise. Furthermore, we may assume that any two vertices have $O(\ell)$ common neighbors, as again the contribution is trivial from other cases by the first part of Lemma 16.4. Given any v , there are $f(v)$ choices for a copy of D including it, $d(v)$ choices for the leaf off of v ,

and $O(\ell)$ choices for the remaining vertex of H^* , as it is the common neighbor of v and of its degree 4 neighbor in H^* . Thus, we obtain:

$$\begin{aligned} t(H^*, p, b^2) &\lesssim \ell \mathbb{E} \left[\sum_v f(v) d(v) \right] \lesssim \ell \mathbb{E} \left[\left(\max_v f(v) \right) \left(\sum_v d(v) \right) \right] \\ &\lesssim b \ell \mathbb{E} \left[\max_v f(v) \right] \lesssim b \ell t^u(D, p, b^2) \lesssim b \ell^4, \end{aligned}$$

where the last inequality follows from Lemma 16.7. $\square$

Putting all of the bounds together completes the proof of Theorem 14.2.

Proof of Theorem 14.2. We will apply (3) twice to the trifold H . Applying it once, we find

$$t(H, p, b^2) \lesssim p \max_{e \in E(H)} t(H \setminus e, p, b^2), \quad (8)$$

and there are only two nonisomorphic subgraphs of H of the form $H \setminus e$. One of them is H^* , for which we have $t(H^*, p, b^2) = O(b \ell^4)$ by Proposition 16.8.

If H' is the other subgraph of the form $H \setminus e$, where an inner edge is deleted from the trifold, then we apply (3) again to find

$$t(H', p, b^2) \lesssim p \max_{3 \leq i \leq 6} t(H_i, p, b^2), \quad (9)$$

since all the graphs $H' \setminus e$ are isomorphic to one of H_3, H_4, H_5, H_6 . We have by Proposition 16.6 that $t(H_i, p, b^2) = O(b^2 \ell^3)$.

It follows from (9) that $t(H', p, b^2) = O(b \ell^3)$. Together with the fact that $t(H^*, p, b^2) = O(b \ell^4)$ and (8), this proves that $t(H, p, b^2) = O(\ell^4)$. The theorem follows by one application of Lemma 16.2 with $N = b^2$. $\square$

17 Cliques in $G(n, 1/2)$

In this section, we prove Theorem 14.3 using an idea of Huy Pham [69]. The argument is a modification of the proof of Theorem 1 in [32] when the number of vertices in $G(n, p)$ is bounded beforehand.

Let $G = G(n, \frac{1}{2})$. For each vertex subset $U \in V(G)$, let $e_t(U)$ be the number of queries made between pairs of vertices in U after query t . We will study the weight function

$$w(U, t) = \begin{cases} 2^{-\binom{|U|}{2} + e_t(U)} & \text{if all queries so far in } U \text{ succeeded} \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $w(U, t)$ is exactly the probability that $G[U]$ is a clique conditional on the information revealed after query t . The standard method of conditional expectation proceeds by studying the evolution of the function

$$w_k(t) = \sum_{|U|=k} w(U, t),$$

which is a martingale, and has the property that a k -clique is found after query t only if $w_k(t) \geq 1$. Our modification studies instead a restricted version of this sum. Namely, define $m(U, t)$ to be the size of the maximum matching in the known edges of U after query t . Then,

$$w_{k,m}(t) = \sum_{|U|=k, m(U,t) \geq m} w(U, t).$$

Restricting to only sets with large maximum matchings has the function of radically reducing the number of terms in the sum $w_{k,m}(t)$. We pay for it in that $w_{k,m}(t)$ is no longer a martingale and its expectation is harder to study. Nevertheless, it remains true that if a k -clique is found after query t then $w_{k,m}(t) \geq 1$ for every $m \leq k/2$.

Lemma 17.1. *For any $0 \leq m \leq k/2$ and any fixed querying strategy that uses $t \leq \binom{n}{2}$ queries,*

$$\mathbb{E}[w_{k,m}(t)] \leq t2^{-(2k-2m-1)} \cdot \mathbb{E}[w_{k-2,m-1}(t)].$$

Proof. For every set $U \in \binom{[n]}{k}$, we say that U is m -critical at query s if s is the smallest number for which $m(U, s) \geq m$. In particular, U does not contribute to $w_{k,m}(t)$ until $t = s$, after which it contributes $w(U, t)$, which is a martingale. This means that if

$$w_{k,m}^*(s) = \sum'_{|U|=k} w(U, s),$$

where the sum is restricted to only sets U which are m -critical at query s , then

$$\mathbb{E}[w_{k,m}(t) - w_{k,m}(t-1)] = \mathbb{E}[w_{k,m}^*(t)],$$

and so

$$\mathbb{E}[w_{k,m}(t)] = \sum_{s \leq t} \mathbb{E}[w_{k,m}^*(s)]. \quad (10)$$

Next, we will show

$$w_{k,m}^*(s) \leq 2^{-(2k-2m-2)} w_{k-2,m-1}(s). \quad (11)$$

To see this, note that every U that appears on the left side must contain the edge (u, v) built after query s , since $m(U, s) > m(U, s-1)$. Furthermore, $U' = U \setminus \{u, v\}$ is a set with $k-2$ vertices and an $m-1$ matching. Finally, every edge in U but not U' is incident to (u, v) . It is easy to check that if (u, v) is an edge that lies in every m -matching of U , then at most $2m-2$ other edges are incident to (u, v) . Thus, there are at least $2(k-2) - (2m-2) = 2k-2m-2$ unqueried pairs in U but not in U' , and

$$w(U, t) \leq 2^{-(2k-2m-2)} w(U', t).$$

Summing over all m -critical sets U , we get the desired inequality (11). Taking expectations of both sides,

$$\mathbb{E}[w_{k,m}^*(s)] \leq 2^{-(2k-2m-1)} \mathbb{E}[w_{k-2,m-1}(s)].$$

Note that we gained another factor of $\frac{1}{2}$ here because there is a $\frac{1}{2}$ chance that the query (u, v) fails and $w_{k,m}^*(s) = 0$. Plugging this into (10), we get

$$\mathbb{E}[w_{k,m}(t)] \leq 2^{-(2k-2m-1)} \sum_{s \leq t} \mathbb{E}[w_{k-2,m-1}(s)].$$

The expectations on the right side are nondecreasing as a function of s , so we can bound this by

$$\mathbb{E}[w_{k,m}(t)] \leq 2^{-(2k-2m-1)} \sum_{s \leq t} \mathbb{E}[w_{k-2,m-1}(s)] \leq t 2^{-(2k-2m-1)} \cdot \mathbb{E}[w_{k-2,m-1}(t)]$$

as desired. □

Now we may iterate Lemma 17.1 until $m = 0$ to prove the following general bound.

Lemma 17.2. *For any $0 \leq m \leq k/2$ and any fixed querying strategy that uses $t \leq \binom{n}{2}$ queries,*

$$\mathbb{E}[w_{k,m}(t)] \leq t^m n^{k-2m} 2^{-\binom{k}{2} + m(m-1)}.$$

Proof. We induct on m . The base case $m = 0$ is just the unrestricted weight function

$$\mathbb{E}[w_{k,0}(t)] = \mathbb{E}[w_k(t)] = w_k(0) = \binom{n}{k} 2^{-\binom{k}{2}} \leq n^k 2^{-\binom{k}{2}},$$

for all k , as desired. Assuming the statement is true for some $m \geq 0$ and all $k \geq 2m$, Lemma 17.1 provides the inductive step for $m + 1$ and all $k \geq 2m + 2$. □

It remains to prove Theorem 14.3 using Lemma 17.2.

Proof of Theorem 14.3. Recall that $\mathbb{E}[w_{k,m}(t)]$ is an upper bound on the probability one can find a k -clique in t queries. By Lemma 17.2, we see that whenever n, k, t are such that there exists $m \leq \frac{k}{2}$ for which

$$\mathbb{E}[w_{k,m}(t)] \leq t^m n^{k-2m} 2^{-\binom{k}{2} + m(m-1)} < \frac{1}{2},$$

then it is impossible to find a k -clique in t queries in $G(n, \frac{1}{2})$ with probability at least $\frac{1}{2}$. It is cleaner to compute the base-2 logarithm of this quantity. Taking $t = n^\delta$ and $k = \alpha \lg n$ and writing $\ell = \lg n$ as a shorthand, we get

$$\begin{aligned} \lg \left(t^m n^{k-2m} 2^{-\binom{k}{2} + m(m-1)} \right) &= (\alpha \ell - m(2 - \delta))\ell - \binom{\alpha \ell}{2} + m(m-1) \\ &\leq \left(\alpha - \frac{\alpha^2}{2} \right) \ell^2 - (2 - \delta)m\ell + m^2 + O(\ell). \end{aligned}$$

If $m = c\ell$ where $c \leq \alpha/2$, then

$$\left(\alpha - \frac{\alpha^2}{2} \right) \ell^2 - (2 - \delta)m\ell + m^2 = \left(\alpha - \frac{\alpha^2}{2} - (2 - \delta)c + c^2 \right) \ell^2$$

is minimized at $c = \frac{2-\delta}{2}$. Assuming that $\alpha \geq 2 - \delta$, we find that for this choice of c ,

$$\lg(\mathbb{E}[w_{k,m}(t)]) \leq \left(\alpha - \frac{\alpha^2}{2} - \frac{(2 - \delta)^2}{4} \right) \ell^2 + O(\ell).$$

In particular, this shows that whenever $\alpha \geq 2 - \delta$ satisfies

$$\alpha - \frac{\alpha^2}{2} - \frac{(2 - \delta)^2}{4} < 0,$$

then for sufficiently large n , it is impossible to find a clique with $\alpha \lg n$ vertices in n^δ queries.

Thus, $\alpha_*(\delta)$ is bounded above by the (larger) solution to the above quadratic, which is

$$\alpha_+ = 1 + \sqrt{1 - \frac{(2 - \delta)^2}{2}} > 2 - \delta,$$

as desired. □

18 Discussions and Open Problems

The immediate question that arises from our work is to classify the graphs H for which $f(H, p) = b^{d-o(1)}$. The natural first step is the case $d = 2$. To this end, we first establish a large family of 2-degenerate graphs H for which $f(H, p) = O(b^{2-\varepsilon})$ for some $\varepsilon > 0$.

Definition 18.1. *We call a graph H $(1, 1)$ -degenerate if H can be vertex-partitioned into induced subgraphs $T_1, \dots, T_n$ which are trees, such that for all $k \in \{1, \dots, n\}$ and all $v \in T_k$,*

$$|N(v) \cap \bigcup_{i=1}^{k-1} T_i| \leq 1.$$

It is easy to see that if H is $(1, 1)$ -degenerate, then H is 2-degenerate. One can show by induction on the number of trees n that if H is $(1, 1)$ -degenerate, then $f(H, p) = O(b^{2-\varepsilon})$ for some $\varepsilon > 0$. We prove this, and conjecture that the converse is true.

Theorem 18.2. *If H is $(1, 1)$ -degenerate, then $f(H, p) = O(b^{2-\varepsilon})$ for some $\varepsilon = \varepsilon(H) > 0$.*

Proof. We induct on the number of trees n in the $(1, 1)$ -degenerate partition $T_1, \dots, T_n$ of H . The explicit value of $\varepsilon(H)$ we pick is $\varepsilon(H) = \max\{|V(T_1)|, \dots, |V(T_n)|\}^{-1}$.

If $n = 1$, then H is a tree so that $f(H, p) = O(b)$, as desired. Now, say $n \geq 2$ and let $V(H) = V(H') \cup V(T_n)$. For $\varepsilon' = \varepsilon(H')$, we can build a copy of H' with probability at least $3/4$ in some time $O(b^{2-\varepsilon'})$ by the inductive hypothesis. Let $\varepsilon = \min\{\varepsilon', |V(T_n)|^{-1}\}$. Now, for each $v \in V(T_n)$ there exists at most one $u \in V(H')$ so that v and u are adjacent. If there

does not exist such a u , we let C_v be a set of $b^{1-\varepsilon}$ previously unexplored vertices, and if there is, let C_v be a set of $b^{1-\varepsilon}$ neighbors of u . We can find such a set C_v in at most $O(b^{2-\varepsilon})$ queries (with constant probability).

Now, query all edges between C_v and $C_{v'}$ for every pair of vertices $v, v' \in V(T_n)$. This takes time $O(b^{2-2\varepsilon})$. We claim that there will exist a copy of T_n with $v \in C_v$ for all $v \in V(T_n)$, because the expected number of copies of T_n is on the order of $b^{(1-\varepsilon)|V(T_n)|} b^{|V(T_n)|-1} = \Theta(1)$, and it is a standard result that the threshold for containment of a tree is the same as the expectation threshold. This gives that we have at least one such copy of T_n with probability at least $3/4$. This copy of T_n extends the original copy of H' to our desired copy of H , all with probability at least $1/2$ in time $O(b^{2-\varepsilon})$. $\square$

Conjecture 18.3. *If H is a 2-degenerate graph that is not $(1, 1)$ -degenerate, then $f(H, p) = b^{2-o(1)}$.*

In the case $d = 2$, we were able to construct a particular 2-degenerate graph H for which $f(H, p) \geq b^{d-o(1)}$. The existence of such graphs when $d \geq 3$ remains open.

Conjecture 18.4. *For all integers $d \geq 2$, there exists a d -degenerate graph H for which $f(H, p) = b^{d-o(1)}$.*

There is a natural random process for constructing d -degenerate graphs on n vertices. Namely, starting with a K_d , $n - d$ vertices are added one at a time, and each new vertex is given d neighbors uniformly at random among the previous ones. If n is sufficiently large, it is plausible that the random d -degenerate graph constructed in this manner should satisfy $f(H, p) = b^{d-o(1)}$ asymptotically almost surely.

Part VI

On The Product Dimension of Clique Factors

This chapter is based on joint work with Noga Alon [3], published in the *European Journal of Combinatorics*.

19 Introductory Remarks

The product dimension of a graph $G = (V, E)$ is the minimum possible cardinality d of a collection of proper vertex colorings of G such that every pair of nonadjacent vertices have the same color in at least one of the colorings (and so that any two distinct vertices are colored differently in some coloring). Equivalently, this is the minimum d so that one can assign to every vertex v a vector in Z^d , so that two vertices are adjacent if and only if the corresponding vectors differ in all coordinates (and so that no two distinct vertices are assigned the same vector). If G does not contain two distinct non-adjacent vertices with the same neighborhoods, as will be the case in this chapter, we can take the parenthetical distinctness conditions for granted. The product dimension is also the minimum number of complete graphs so that G is an induced subgraph of their tensor product, where the tensor product of graphs $H_1, \dots, H_d$ is the graph whose vertex set is the cartesian product of the vertex sets of the graphs H_i , and two vertices $(u_1, u_2, \dots, u_d)$ and $(v_1, v_2, \dots, v_d)$ are adjacent iff u_i is adjacent (in H_i) to v_i for all $1 \leq i \leq d$. Yet another equivalent definition is the minimum number of subgraphs of the complement $\overline{G}$ of G so that each subgraph is a vertex disjoint union of cliques, and every edge of $\overline{G}$ belongs to at least one of the subgraphs (and

also every pair of distinct vertices are not adjacent in some subgraph).

For positive integers $s, r \geq 2$ let $K_s(r)$ denote the graph consisting of r pairwise vertex disjoint copies of the complete graph K_s . Any two non-adjacent vertices of this graph have different neighborhoods. Let $Q(s, r)$ denote the product dimension of this graph. Lovász, Nešetřil and Pultr [63] (see also [1]) proved that $Q(2, r) = \lceil \log_2(2r) \rceil$. The proof of the upper bound is simple. If $q = \lceil \log_2(2r) \rceil$ then $2^q \geq 2r$. Hence one can assign distinct binary vectors of length q to the $2r$ vertices of $K_2(r)$ so that the vectors assigned to each pair of adjacent vertices are antipodal, i.e. they differ in all coordinates. It is easy to check that two vertices are adjacent if and only if the corresponding vectors differ in all coordinates, showing that $Q(2, r) \leq q$.

The lower bound is proved in [63] by a linear algebra argument, and the proof given in [1] applies exterior algebra. There is yet another (similar) short proof that proceeds by assigning to each vertex of $K_s(r)$ a multilinear polynomial in $x_1, x_2, \dots, x_q$ that depends on the coloring used, and by showing that these polynomials are linearly independent. As mentioned in the abstract, Lovász, Nešetřil and Pultr [63] raised the problem of estimating $Q(s, r)$ for larger values of s . More recently, Kleinberg and Weinberg considered the same problem, motivated by the investigation of prophet inequalities for intersection of matroids [56]. In this chapter, we determine the asymptotic behavior of $Q(s, r)$ for any fixed $s \geq 2$ and large r .

Theorem 19.1. *For every fixed s , $Q(s, r) = (1 + o(1)) \log_2 r$, where the $o(1)$ -term tends to 0 as r tends to infinity.*

The main tool in the proof is the method of Gargano, Körner and Vaccaro in their work on Sperner capacities [48]. For completeness, and since we are interested in the behavior of the $o(1)$ -term in the theorem above, we describe a variant of the method as needed here, in a combinatorial way that avoids any application of information theoretic techniques. The proof is based on what we call here Z_s -covering families of vectors.

Let Z_s denote the ring of integers modulo s . For a subset $A \subset Z_s$ and a vector $v = (v_1, v_2, \dots, v_q) \in Z_s^q$, we say that v is A -covering if for every $a \in A$ there is an $1 \leq i \leq q$ so that $v_i = a$. The vector v is covering if it is Z_s -covering. A family $\mathcal{F} \subset Z_s^q$ is A -covering if for every ordered pair of distinct vectors $u, v \in \mathcal{F}$, the difference $u - v$ is A -covering. $\mathcal{F}$ is covering if it is Z_s -covering. Therefore, a family $\mathcal{F}$ of vectors in Z_s^q is covering if every element of Z_s appears in at least one coordinate of the difference between any two distinct vectors in the family. Let $R(s, q)$ denote the maximum possible cardinality of a covering family of vectors in Z_s^q . The following simple statement describes the connection between $Q(s, r)$ and $R(s, q)$.

Lemma 19.2. *If $R(s, q) \geq r$ then $Q(s, r) \leq q$.*

Note that by definition $R(s, q) = 1$ for all $q < s$. Our main result about $R(s, q)$ is the following.

Theorem 19.3. *1. For every $q \geq s \geq 2$, $R(s, q) \leq 2^{q-1}$. Equality holds for $s = 2$.
2. For every fixed s , $R(s, q) \geq (2 - o(1))^q$, where the $o(1)$ -term tends to 0 as $q \rightarrow \infty$.*

The rest of this chapter is organized as follows. Section 2 contains the proof of Proposition 19.2 and that of a simple combinatorial lemma. In Section 3 we present the proof of Theorem 19.3 and note that in view of Proposition 19.2 it implies Theorem 19.1. The proof supplies better estimates for prime values of s , and we thus first present the proof for this special case (which suffices to deduce the assertion of Theorem 19.1 for every s) and then describe briefly the proof for general s . The final Section 4 contains some concluding remarks and open problems, including some (modest) estimates for $R(s, q)$ when q is not much larger than s .

To simplify the presentation we omit, throughout the chapter, all floor and ceiling signs whenever these are not crucial. Thus, we ignore all divisibility issues.

20 Preliminaries

We first prove Proposition 19.2. If $R(s, q) \geq r$, then there exists a matrix of elements in Z_s with r rows and q columns so that the difference of any two rows is covering. We will use the q columns of this matrix to find q graphs, each being a disjoint union of cliques, that cover the complement of $K_s(r)$. This complement is a complete multipartite graph with r parts of size s . Label the vertices of this graph with elements of Z_s so that in each part all labels are used exactly once. We will associate each row of the matrix to a part and each column to a vertex disjoint union of cliques. For a column $(a_1, \dots, a_r)^T$, consider the following graph. For each $0 \leq k < s$ take the $k + a_i$ th vertex (taken modulo s) from the i th part of size s and take the union of the s cliques obtained as k ranges between 0 and $s - 1$. This is clearly a union of s vertex disjoint cliques. Now, suppose we have some two vertices of the graph we are trying to cover in different parts, say the $\ell + m$ th vertex of part i and the ℓ th vertex of part j for some $1 \leq i < j \leq q$ and some $\ell, m \in Z_s$. Then the difference of the i th and j th rows contains an m in some column $(a_1, \dots, a_r)^T$, so that $a_i - a_j \equiv m$ in Z_s , and then the disjoint union of cliques corresponding to this column will cover our desired edge.

One can also phrase the proof using the proper coloring definition of $Q(s, r)$. Say we are given a matrix of r vectors $\{v_1, \dots, v_r\}$ over Z_s^q which are a Z_s -covering family. We can associate the vector $v_i + (j, \dots, j)$ to the j th vertex in the i th clique of $K_s(r)$. These vectors are all distinct for different vertices, because if $v_i + (j, \dots, j) = v_{i'} + (j', \dots, j')$, then $v_i - v_{i'}$ is not covering, so $i = i'$ and $j = j'$. Now, we define q colorings of $K_s(r)$ so that if a vertex x is associated to $(c_1, \dots, c_q)$, it is colored with c_k in the k th coloring. These colorings are proper, because if x and y are the j th and j' th vertex in the i th clique for $j \neq j'$, then their associated vectors will have a difference $(j - j', \dots, j - j') \neq \widehat{0}$. Now, say we are given two distinct non-adjacent vertices, say the j th vertex in the i th clique and the j' th vertex in the i' th clique, where $i \neq i'$. Then $v_i + (j, \dots, j)$ and $v_{i'} + (j', \dots, j')$ will share a coordinate;

$v_i - v_{i'}$ is covering and thus will be $j' - j$ in some coordinate.

Next, we need the following simple lemma.

Lemma 20.1. *Let H be a bipartite graph with classes of vertices A_1, A_2 where $|A_1| = n_1$, $|A_2| = n_2$, each vertex of A_1 has degree d_1 , and each vertex of A_2 has degree d_2 . Furthermore suppose that $d_2 \geq \log(2n_2)$. Then there is a union of vertex-disjoint stars with centers in A_1 , each star having at least $\frac{d_1}{4\log(2n_2)}$ leaves, such that all vertices of A_2 are leaves.*

Proof. Define a random subset S of A_1 by choosing each vertex of A_1 to be in S with probability $p = \frac{\log(2n_2)}{d_2}$ uniformly and independently. We claim that with positive probability, each vertex of A_2 has between 1 and $4\log(2n_2)$ neighbors in S . The proof is a simple union bound; a fixed vertex $v \in A_2$ has probability

$$(1 - p)^{d_2} < e^{-pd_2} = \frac{1}{2n_2}$$

of having no neighbors in S , and probability at most

$$\binom{d_2}{4\log(2n_2)} p^{4\log(2n_2)} \leq \left(\frac{ped_2}{4\log(2n_2)} \right)^{4\log(2n_2)} = \left(\frac{e}{4} \right)^{4\log(2n_2)} < \frac{1}{2n_2}$$

of having more than $4\log(2n_2)$ neighbors, proving the claim. Fix an S with this property.

We now finish the proof of the lemma by an application of Hall's theorem. For all $S' \subseteq S$, let $N(S')$ denote the set of all neighbors of S' and let $e(S', A_2)$ denote the number of edges from S' to A_2 . Then $|N(S')| \geq \frac{e(S', A_2)}{4\log(2n_2)} = \frac{d_1}{4\log(2n_2)} |S'|$. Hence every subset of S expands by a factor of at least $\frac{d_1}{4\log(2n_2)}$. Thus by Hall's theorem, there is a union of disjoint stars whose centers are exactly the vertices of S , each having at least $\frac{d_1}{4\log(2n_2)}$ leaves. Every remaining vertex of A_2 is adjacent to some vertex in S , so we can simply add it to an existing star. $\square$

21 Covering families

21.1 The Upper Bound

The following proposition implies the assertion of Theorem 19.3, part 1.

Lemma 21.1. *Fix $s \geq 2$, and let $\mathcal{F} \subset Z_s^q$ be a $\{0,1\}$ -covering family of vectors. Then $|\mathcal{F}| \leq 2^{q-1}$. For $s = 2$ equality holds.*

Proof. Put $m = |\mathcal{F}|$. Let p be a prime divisor of s and consider the vectors in $\mathcal{F}$ as vectors in Z_p^q by reducing their coordinates modulo p . Note that these vectors form a $\{0,1\}$ -covering family over Z_p , and so are distinct. Let $v_i = (v_{i1}, v_{i2}, \dots, v_{iq})$, $(1 \leq i \leq m)$ be the vectors in $\mathcal{F}$ (considered as elements of Z_p^q).

For each $1 \leq i \leq m$ define two polynomials P_i, Q_i in the variables $x_1, x_2, \dots, x_q$ over Z_p as follows.

$$P_i(x_1, \dots, x_q) = \prod_{j=1}^q (x_j - v_{ij}), \quad Q_i(x_1, \dots, x_q) = \prod_{j=1}^q (x_j - v_{ij} - 1).$$

It is not difficult to check that for every i , $Q_i(v_i) \neq 0$ and $P_i(v_i) = 0$. In addition, for every $1 \leq i \neq i' \leq m$, $P_{i'}(v_i) = 0$ (as there is a coordinate j for which $v_{ij} - v_{i'j} = 0$) and $Q_{i'}(v_i) = 0$ (as there is a j so that $v_{ij} - v_{i'j} = 1$).

Similar reasoning gives that for the vectors $v_i + J$, where J is the all 1-vector of length q , $P_i(v_i + J) \neq 0$, $Q_i(v_i + J) = 0$, and for every $i' \neq i$, $P_{i'}(v_i + J) = Q_{i'}(v_i + J) = 0$. Therefore, for each member of the collection of $2m$ polynomials $\{P_i, Q_i : 1 \leq i \leq m\}$ there is an assignment of values of the variables in which this member is nonzero and all others vanish. This easily implies that the set of $2m$ polynomials P_i, Q_i is linearly independent in Z_p , and as each of its members lies in the space of multilinear polynomials with the m variables x_j , the number, $2m$, of these polynomials is at most the dimension of this space

which is 2^q . It follows that $|\mathcal{F}| = m \leq 2^{q-1}$, as needed. For $s = 2$ the family of all binary vectors in which the first coordinate is 1 is $\{0, 1\}$ -covering, showing that $R(2, q) \geq 2^{q-1}$ and completing the proof. $\square$

21.2 Prime s

For prime $s \geq 3$, we will prove that $R(s, q) \geq (2 - o(1))^q$ where the $o(1)$ -term tends to 0 as $q \rightarrow \infty$. The crux of the proof is a Markov chain argument from [48], which we will iterate $O(\log s)$ times.

A *balanced word* over Z_s^q is a word containing the letters 1 through $s-1$ an equal number of times. A *special balanced word* is a balanced word such that the first $\frac{q}{(s-1)/2}$ letters are 1 and 2 in some order, the next $\frac{q}{(s-1)/2}$ letters are 3 and 4 in some order, and so on. Construct a bipartite graph between the set A_2 of balanced words w over Z_s^q and the set A_1 of permutations π on q elements defined as follows: w and π are adjacent if and only if $\pi(w)$ is a special balanced word. By symmetry all $n_1 = q!$ vertices in A_1 have the same degree d_1 , and all vertices in A_2 have the same degree d_2 . We have

$$n_2 = |A_2| = \binom{q}{q/(s-1), \dots, q/(s-1)} \leq (s-1)^q$$

and d_1 is the total number of special balanced words, so

$$d_1 = \binom{2q/(s-1)}{q/(s-1)}^{(s-1)/2} > \frac{2^q}{q^{s/2}}.$$

Furthermore, $d_2 = \frac{n_1}{n_2} d_1 = \left(\left(\frac{q}{s-1}\right)!\right)^{s-1} d_1 \geq \log(2n_2)$ so by Lemma 20.1 there exists a way to map balanced words to some set T of permutations π of q elements, so that each balanced word is associated to exactly one permutation, and each permutation in T is associated to at least $\frac{d_1}{4 \log(2n_2)} > \frac{2^q/q^{s/2}}{4 \log(2(s-1)^q)} > \frac{2^q}{q^s}$ balanced words. Thus, we can partition the balanced

words into sets $S_1, S_2, \dots$ so that for each S_i we have $|S_i| > \frac{2^q}{q^s}$ and for all i there exists π_i so that $\pi_i(s_i)$ is a special balanced word for all $s_i \in S_i$.

Given all of the special balanced words of length q , any two of them have a difference vector which covers $\{\pm 1\}$. The idea of the proof will be to amplify this set $\{-1, 1\}$, first to $\{-\alpha, -1, 1, \alpha\}$ for a primitive root α modulo s , and after r steps to $\{\pm \alpha^b\}$ for $0 \leq b < 2^r$. At each stage, the number of vectors will be $(2 - o(1))^L$ where L is the length of the vectors. Thus after $O(\log s)$ steps we will have a set of vectors that is Z_s^* covering. We can then add an extra coordinate of 0 to all of the vectors to make them Z_s -covering.

We describe the first step of this iteration in detail. Fix a primitive root α modulo s , which will be constant throughout the steps. Also fix $n = \frac{100}{\epsilon}(\log(s))^2$, which will again be constant throughout the steps. Initially we set $q = q_0 = 100(s^2 - 1)$, ensuring it is divisible by $s - 1$. We will construct words of length $q_1 = qn$ by stringing together balanced words of length q in a specific way. If $x_i \in S_j$, then force $x_{i+1} \in \alpha S_j$; here we mean that if we take x_{i+1} and multiply its letters by α^{-1} pointwise, the result will be in S_j . Consider all vectors of length qn constructed according to this rule, by concatenating n balanced words of length q in this way. There are more than $n_2(\frac{2^q}{q^s})^{n-1}$ such words, because at each stage other than the first we must pick x_{i+1} so that $x_i \in \alpha S_j$ for some j , and thus there are more than $\frac{2^q}{q^s}$ choices for x_{i+1} . We will make x_1, x_n identical over all such words; this costs us a factor of n_2^2 and thus we now have more than $\frac{1}{n_2}(\frac{2^q}{q^s})^{n-1}$ such words. Using that $n_2 < s^q$, we can find more than $\frac{2^{q(n-1)}}{q^{s(n-1)}s^q} = (2 - o(1))^{qn}$ words of length qn of the form $x_1 x_2 \dots x_n$ so that x_1, x_n are fixed, where $o(1)$ is as ϵ goes to 0 and n goes to infinity. We now note that for any two different words of this form $x_1 \dots x_n$ and $x'_1 \dots x'_n$, with $x_1 = x'_1, x_n = x'_n$, there is some minimal $i \geq 2$ so that $x_i \neq x'_i$. But then there is some k so that $x_{i-1} = x'_{i-1} \in S_k$, and that means $x_i, x'_i \in \alpha S_k$. Because $x_i \neq x'_i$, we have $\alpha^{-1}x_i \neq \alpha^{-1}x'_i$, and both $\alpha^{-1}x_i$ and $\alpha^{-1}x'_i$ are in S_k . It follows that $\alpha^{-1}x_i - \alpha^{-1}x'_i$ covers $\{\pm 1\}$ and so $x_i - x'_i$ covers $\{\pm \alpha\}$. Similarly, there is some maximal $i < n$ so that $x_i \neq x'_i$. Then $x_{i+1} = x'_{i+1} \in \alpha S_j$ for some j , so $x_i \in S_j$,

$x'_i \in S_j$. As $x_i \neq x'_i$, it follows that $x_i - x'_i$ must cover $\{\pm 1\}$ as coordinates. When $s = 5$, setting $\alpha = 2$ and applying this construction already gives a covering family for $\mathbb{Z}_s^* = \mathbb{Z}_5^*$ with $(2 - o_N(1))^N$ vectors of length N . In the general case we iterate this argument to find $(2 - o_N(1))^N$ vectors of length N , so that after the r th iteration the vectors we get cover $\{\pm \alpha^b\}$ for all $0 \leq b < 2^r$. We describe how to do this inductively.

After r iterations, we find for some $q = q_r$ (that depends on s) a family of $M^q = (2 - o_q(1))^q$ (balanced) vectors over $\mathbb{Z}_s^q$ which covers $\pm \alpha^b$ for all $0 \leq b < 2^r$. We call these vectors $y_1, \dots, y_{M^q}$ and let $Y = \{y_1, \dots, y_{M^q}\}$. We repeat the above argument. The y_i play the role of the special balanced words. Again we construct a bipartite graph. On one side there is the set A_2 of all balanced words of length q and on the other side there is A_1 , permutations of q elements. We have π is adjacent to a (balanced) vector y if and only if $\pi(y) \in Y$. Again $n_2 = |A_2| \leq (s-1)^q$ but now $d_1 = M^q$. It can easily be verified that $d_2 \geq \log(2n_2)$. Thus by Lemma 20.1 the balanced words of length q will be split into sets $S_1, S_2, \dots$ so that all S_i satisfy $|S_i| \geq \frac{d_1}{4 \log(2n_2)} \geq \frac{M^q}{4 + 4q \log(s-1)} = (2 - o_q(1))^q$, and furthermore for each S_i , there exists a permutation π_i so that for all $s_i \in S_i$, $\pi_i(s_i) \in Y$. Now we will again construct a Markov chain. Consider all words of length $qn = q_{r+1}$ consisting of n balanced words of length q of the form $x_1 x_2 \dots x_n$, so that if $x_i \in S_j$, then $x_{i+1} \in \alpha^{2^r} S_j$. If we have two such words $x_1 x_2 \dots x_n$ and $x'_1 x'_2 \dots x'_n$ so that $x_1 = x'_1$, $x_n = x'_n$, let $i > 1$ be minimal so that $x_i \neq x'_i$. Then if $x_{i-1} = x'_{i-1} \in S_j$, $x_i, x'_i \in \alpha^{2^r} S_j$. This means that $x_i - x'_i$ covers $\{\pm \alpha^b\}$ modulo s for $2^r \leq b < 2^{r+1}$. Now, if we let $x_i, i < n$, be maximal so that $x_i \neq x'_i$, then $x_{i+1} = x'_{i+1} \in \alpha^{2^r} S_j$ for some j and so x_i, x'_i are not equal but are both in S_j . This means $x_i - x'_i$ covers $\{\pm \alpha^b\}$ modulo s for $0 \leq b < 2^r$. So indeed the family covers $\{\pm \alpha^b\}$ modulo s for $0 \leq b < 2^{r+1}$, as long as x_1 and x_n are fixed over the family. We can always find

$$\frac{\min_i |S_i|^{n-1}}{\binom{q}{q/(s-1), \dots, q/(s-1)}} \geq \frac{(2 - o(1))^{qn}}{(s-1)^q} = (2 - o(1))^{qn}$$

such words if n is large enough, where the $o(1)$ terms tend to 0 as $q \rightarrow \infty$ and n is sufficiently large. We will soon see that our choice of $n = \frac{100}{\epsilon}(\log(s))^2$ is sufficient. Iterating the argument $\log_2(s)$ times allows us to find $(2 - o(1))^q$ vectors of some length q which cover Z_s^* . Adding a single coordinate where all vectors are 0 gives us vectors of length $q + 1$ that cover Z_s , without changing the asymptotic analysis.

With care, we can extract quantitative bounds. We assume $\epsilon > 0$ is a fixed constant and show that we only require $q = \left(\frac{O(1)}{\epsilon}(\log(s))^2\right)^{\log(s)}$ to have a $(2 - \epsilon)^q$ size covering system over Z_s^q for large s . Say that after r iterations, we have $M^q = M_r^q$ vectors of length $q = q_r$ for some $M \leq 2$ (which is nearly 2). Then $\min_i S_i \geq \frac{M^q}{5q \log(s)}$ and thus we find at least

$$\begin{aligned} \frac{\min_i |S_i|^{n-1}}{\binom{q}{q/(s-1), \dots, q/(s-1)}} &\geq \frac{M^{qn}}{M^q (5q \log(s))^{n s^q}} \geq \frac{M^{qn}}{(2s)^q (5q \log(s))^n} \\ &= \left(\frac{M}{(2s)^{1/n} q^{1/q} (5 \log(s))^{1/q}} \right)^{qn} = M_{r+1}^{q_{r+1}} \end{aligned}$$

vectors of length $qn = q_{r+1}$. Recall that $n = \frac{100}{\epsilon}(\log(s))^2$ and $q_0 = (1 + o(1))100s^2$, where $o(1)$ is as s goes to infinity, so when we iterate the Markov chain argument $\log_2(s)$ times, we lose a factor of at most

$$\begin{aligned} \frac{M_0}{M_{\log_2(s)}} &\leq (2s)^{2 \log(s)/n} \left(\prod_{j=0}^{\infty} (q_0 n^j)^{1/q_0 n^j} \right) (5 \log(s))^{1/q_0} (5 \log(s))^{2 \log(s)/n} \\ &\leq (10s \log(s))^{2 \log(s)/n} 2^{10 \log(s)/s^2} (5 \log(s))^{1/100s^2}. \end{aligned}$$

In the last inequality here the bound on the second term holds because the relevant infinite product is at most $(q_0^{1/q_0})^{100} < 2^{10 \log s / s^2}$. For large s this is easily seen to be smaller than, say, $1 + \epsilon/4$. Since for large s , $M_0 > 2 - \epsilon/2$, for $s > s_0(\epsilon)$ we get $M_{\log_2(s)} > 2 - \epsilon$. Thus at the end we have $q = 100s^2 \left(\frac{O(1)}{\epsilon}(\log(s))^2\right)^{\log_2(s)}$ and at least $(2 - \epsilon)^q$ vectors of length q which cover Z_s . One can easily modify the argument to work for any larger q , or simply

use the super-multiplicative property of $R(s, q)$ (see the beginning of Section 4) to conclude, taking $\epsilon = 1/\log s$, that for every large s and for all $q > s^{(3+o(1)) \log \log s}$, $R(s, q) > (2 - \frac{1}{\log s})^q$. This completes the proof of Theorem 19.3, part 2, for prime s .

21.3 General s

Given an arbitrary fixed integer $s > 2$, we now show how to find $(2 - o(1))^q$ vectors in Z_s^q which form a Z_s -covering family; this shows $R(s, q) \geq (2 - o(1))^q$ even for composite s . The general strategy is similar to our strategy in the previous section, except now we work over Z , and the set we are covering does not grow so quickly. As before, we are not concerned with the vectors covering 0, because we can simply add an extra coordinate to deal with it. Since the argument is very similar to the one described in the previous subsection, we only provide a brief description omitting some of the formal details.

We prove the stronger statement that for any fixed s , we can find a $(2 - o_q(1))^q$ size family over Z that covers $[-s, s]$, i.e. the difference of any two vectors contains all integers between $-s$ and s as coordinates. Let S be the least common multiple of the first s positive integers. We will assume without loss of generality that $2S \mid q$.

At the first step of our iteration, we consider vectors over Z^q with an equal number of each element of $[2S]$ as coordinates so that the first $\frac{q}{S}$ coordinates are 1s and 2s in some order, the next $\frac{q}{S}$ are an equal number of 3s and 4s, and so on. We can find $(2 - o_q(1))^q$ of these and they cover $\{\pm 1\}$. Furthermore, they have the property that for any ordered pair of these vectors, there is a coordinate in which the first has an even integer $2k$ and the second has the odd integer $2k - 1$ for some $1 \leq k \leq S$. This property is crucial and maintained throughout our iterations.

Now we describe the m th step of our iteration, for $2 \leq m \leq s$. Define a bijection $f = f_m$ on $[2S]$ so that for all integers $1 \leq k \leq S$, $f(2k) = f(2k - 1) + m$. We can do this for instance by setting $f(1) = 1, f(2) = m + 1, f(3) = 2, f(4) = m + 2, \dots, f(2m - 1) = m, f(2m) = 2m$

and then set $f(x) = f(x - 2m) + 2m$ for $x > 2m$ as long as $x \leq 2S$.

We then apply the same Markov chain argument as before, defining sets S_i and constructing words $x_1x_2\cdots$ starting and ending at the same vectors. Now, however, when x_i is in some set S_j , instead of demanding $x_{i+1} \in \alpha S_j$ we require $x_{i+1} \in f(S_j)$, meaning that if we apply f^{-1} to each coordinate of x_{i+1} , the result will be in S_j . Looking at the first place where two vectors differ gives us a difference of $\pm m$. Looking at the last place where they differ shows that the crucial property is preserved, and also that the differences $\pm 1, \dots, \pm(m-1)$ are retained.

Thus after $s = O(1)$ iterations, this algorithm produces $(2 - o_q(1))^q$ vectors of length q which cover all of $[-s, s]$, after we add an extra coordinate to deal with covering 0. This completes the proof of Theorem 19.3.

22 Concluding remarks and Open Problems

A natural open problem is to study the functions $R(s, q)$ and $Q(s, r)$ in general. There are several simple properties that $R(s, q)$ satisfies. We know that $R(s, q)$ is (weakly) increasing in q , because to create r covering vectors for Z_s of length $q' > q$, we can take r vectors for Z_s of length q and pad them with $q' - q$ zeroes at the end. Furthermore, we know that $R(s, q)$ is super-multiplicative, i.e. $R(s, q_1 + q_2) \geq R(s, q_1)R(s, q_2)$, because if $m = R(s, q_1)$, $n = R(s, q_2)$ then we can find vectors $v_1, \dots, v_m$ of length q_1 and $w_1, \dots, w_n$ of length q_2 that form covering families. The mn vectors v_iw_j obtained by concatenating v_i and w_j are clearly a covering family of length $q_1 + q_2$.

For $q < s$, we have $R(s, q) = 1$, because of course we can take a single vector in Z_s^q , but if we take two then their difference can only cover a set of size q and cannot cover Z_s .

The next natural question is studying the value of $R(s, s)$.

Lemma 22.1. *$R(s, s) \leq s$, and $R(s, s) \geq p$ where p is the smallest prime factor of s . When*

$p = 2$ this is tight, that is, if s is even, then $R(s, s) = 2$.

Proof. For the lower bound, for each $0 \leq a < p$ we have a vector $(0, a, 2a, \dots, (s-1)a)$ reduced modulo s . This is a covering system for Z_s , since all positive integers smaller than p are relatively prime to s .

For the upper bound, assume there was a covering family in Z_s^s with $s+1$ vectors, so that the difference of any two of these vectors has all values of Z_s exactly once. By the pigeonhole principle, there exist two vectors v and w so that the difference of their first and second coordinates is the same. But then $v - w$ has the same value in its first and second coordinate, a contradiction.

When s is even, $R(s, s) \geq 2$ as 2 is the least prime factor of s . If $R(s, s) \geq 3$ then there exist 3 vectors in Z_s^s , $(a_1, \dots, a_s), (b_1, \dots, b_s), (c_1, \dots, c_s)$, which are covering. But then $\sum(a_i - b_i) \equiv \sum_{j=0}^{s-1} j \equiv (s-1)\frac{s}{2} \equiv \frac{s}{2}$ modulo s , and similarly $\sum(b_i - c_i) \equiv \sum(c_i - a_i) \equiv \frac{s}{2}$ modulo s . Hence, these three sums are all $\frac{s}{2}$ modulo s , so they must add up to $\frac{3s}{2} \equiv \frac{s}{2}$ modulo s . But they add up to 0, so this is a contradiction. $\square$

Nonetheless, the problem of determining $R(s, s)$ for all s remains open, and the lower bound in the last proposition is not tight. A computer search gives that $R(15, 15) \geq 4$ ([62]).

For q which is only a little bigger than s , we can prove a reasonable upper bound, using essentially the same observation applied in the proof that $R(s, s) \leq s$.

Lemma 22.2. *If $2(q-s)^2 < s-1$, then $R(s, q) \leq s + 2(q-s) + 2(q-s)^2$.*

Proof. Assume a contradiction for some $q > s$. By the definition of $R(s, q)$, for any $r \leq R(s, q)$ there exists a matrix of r rows and q columns so that the difference of any two rows contains all values modulo s ; in particular there exists such a matrix for $r = s + 1 + 2(q-s) + 2(q-s)^2$. We will double count the number of pairs of rows and columns r_i, r_j, c_k, c_ℓ with the following property: the 2×2 submatrix formed by r_i, r_j, c_k, c_ℓ has sums of opposite corners equal.

For any two rows r_i, r_j , we examine their difference. This is a vector over Z_s with q entries containing all elements modulo s , so at most $\binom{q-s+1}{2}$ pairs of its entries can be equal. This means that for the pair of rows r_i, r_j , there exist at most $\binom{q-s+1}{2}$ pairs of columns c_k, c_ℓ that satisfy our property. So the total number of such pairs as $1 \leq i, j \leq r$, $1 \leq k, \ell \leq q$ range is at most $\binom{r}{2} \binom{q-s+1}{2}$. However, if we instead fix c_k, c_ℓ we see that the difference of these two columns is a vector of length r over Z_s which has at least $r-s$ pairs of equal elements. This means that there are at least $\binom{q}{2} (r-s)$ r_i, r_j, c_k, c_ℓ with our property. Hence

$$\binom{q}{2} (r-s) \leq \binom{r}{2} \binom{q-s+1}{2}$$

and so $2q(q-1)(r-s) \leq r(r-1)(q-s+1)(q-s)$. By our assumption, $2q > r$ and so $2(q-1) \geq r-1$. Hence $2q(q-1)(r-s) \leq r(r-1)(q-s+1)(q-s) \leq 4q(q-1)(q-s+1)(q-s)$. Thus $r-s \leq 2(q-s+1)(q-s) = 2(q-s) + 2(q-s)^2$ and so $r \leq s + 2(q-s) + 2(q-s)^2$, a contradiction. $\square$

Note that this argument gives a bound only when $q = s + O(\sqrt{s})$, and the bound is $s + O((q-s)^2)$. We believe that this is not tight.

There is a nontrivial (though weak) bound which holds in the regime $q = s + \omega(\sqrt{s})$ as long as $q < Cs$ for some fixed constant $C < \log_2 e$. The natural open problem here is to study cases when q is larger but still fairly small, for instance if $q = 1.5s$, $q = s \log(s)$, or $q = s^2$. Furthermore, it would be interesting to figure out how large q has to be so that $R(s, q) > (2-\epsilon)^s$ for a small positive constant ϵ ; the following argument shows that if ϵ is small enough, then $\frac{q}{s}$ must be bigger than an absolute constant which is above 1.44, but we believe this is far from tight.

Lemma 22.3. *If $1 \leq \frac{q}{s} < C < \log_2(e)$, then $R(s, q) \leq (2-\epsilon)^q$ for $\epsilon = \epsilon(C) > 0$.*

Proof. Say that $R(s, q) \geq r$ so there are r vectors in Z_s^q which are covering. Place them in a matrix with r rows and q columns. Now, for each column select uniformly an element of

Z_s , and add it to all the elements of that column, reducing modulo s . This does not change the covering property. After having done this, replace all entries of this matrix by their reduction modulo 2. Given two entries of a matrix in the same column and different rows, if initially they differed by k modulo s they now have some probability p_k of being equal that depends only on k , which is the probability that if $x \in Z_s$ is chosen randomly and uniformly then the reductions of $x, x + k$ modulo s have the same parity. One can easily check that $p_0 = 1$, $p_1 = 1/s$, $p_2 = (s - 2)/s$, and so on, so that $\prod_{i \in Z_s} p_i = e^{-s(1+o(1))}$ by Stirling's formula. Furthermore, if we take this product over all $i \in Z_s$ except for a subset of size $2cs$ for a small constant c , we have a bound of $e^{-s(1-\epsilon)}$, where ϵ is a small positive constant that tends to 0 with c .

Let $K = cs$. For any pair of rows, the probability their colors match in all but at most $2K$ places is at most $e^{-s(1-\epsilon')}$ for some ϵ' that tends to 0 with c , by a union bound over all the at most $2\binom{q}{2K}$ possible sets of places where they do not match. Thus if we define a graph on our r vectors where two are adjacent iff their values upon reduction modulo 2 match in all but at most $2K$ places, this graph will in expectation have edge density $e^{-s(1-\epsilon')}$ and thus for some fixed choice of the random shifts will have at most $r^2 e^{-s(1-\epsilon')}$ edges. Fixing these shifts one can remove less than $r^2 e^{-s(1-\epsilon')}$ vertices from this graph and get an independent set. But this set must be of size at most $(2 - \delta)^q$ for some $\delta = \delta(c) > 0$, because this size is bounded by the cardinality of a family of disjoint Hamming balls in $\{0, 1\}^q$, each of radius $K = cs$. Thus we have $r \leq r^2 e^{-s(1-\epsilon')} + (2 - \delta)^q$, and the same inequality holds for any $r' < r$ by applying the same reasoning to a set of r' of our vectors.

Now if $q < Cs$ for some fixed $C < \log_2 e$ then setting $q > q_0(\epsilon', C)$, $r' = 3(2 - \delta)^q$ violates the last inequality (with r replaced by r') since for this value of r' and large q , $(r')^2 e^{-s(1-\epsilon')} < r'/2$ and $(2 - \delta)^q < r'/2$ so their sum is smaller than r' . This establishes the assertion of the proposition. $\square$

On the opposite extreme, one may ask what happens if $s \geq 3$ is a small fixed positive

integer and q grows. We know by Proposition 21.1 that $R(2, q) = 2^{q-1}$, so it is natural to ask what happens when 2 is replaced by a larger positive integer.

Conjecture 22.4. *For any fixed $s \geq 3$, $R(s, q) = o(2^q)$, and furthermore $R(s, q) = \Theta(2^q/q^c)$ where $c = c(s) > 0$ is a constant that depends only on s .*

Note that the vectors over Z_3^q with a zero in the first coordinate and with exactly $\lfloor q/2 \rfloor$ ones and $\lceil q/2 \rceil$ zeroes are a covering system, so $R(3, q) = \Omega(2^q/\sqrt{q})$. When s is odd the best upper bound known for $R(s, q)$ is $O(2^q)$, as shown in Proposition 21.1. When s is even, if the vectors of a covering system for Z_s^q are reduced modulo 2, any two differ in at least $s/2$ places, so there are at most $O(2^q/q^{\lfloor \frac{s-1}{4} \rfloor})$ of them, as the Hamming balls of radius $\lfloor \frac{s-1}{4} \rfloor$ centered at these reduced vectors are pairwise disjoint. It would be interesting to establish the above conjecture and to find the relevant constants c if they indeed exist. In particular it seems plausible that $c = 1/2$ when $s = 3$, and if so then the lower bound for $R(3, q)$ is essentially optimal. We conjecture that this is indeed the case.

Conjecture 22.5. $R(3, q) = \Theta(2^q/\sqrt{q})$.

In [29] it is shown that $R(3, q) \leq (\frac{1}{2} + o(1))2^q$ when q is even and $R(3, q) \leq (\frac{1}{3} + o(1))2^q$ when q is odd, leaving this problem open. Note that we do not even know how to prove the weaker claim that a $\{-1, 0, 1\}$ covering system (or equivalently just a $\{\pm 1\}$ covering system) over Z must have size $o(2^q)$.

Our original motivation for studying the function $R(s, q)$ here is its connection to the product dimension $Q(s, r)$ of the disjoint union of r cliques, each of size s . This connection is described in Proposition 19.2. The results here suffice to determine the asymptotic behaviour of $Q(s, r)$ for every fixed s as r tends to infinity, but do not provide tight bounds when r is not much bigger than s . We conclude this short chapter with several simple comments about this range of the parameters. The first remark is that the product dimension $Q(s, r)$ is at least s

for every $r \geq 2$. To see this fix a vertex u of the first clique and observe that in every proper coloring of the graph $K_s(r)$ of r disjoint cliques, each of size s , there is at most one pair uv with a vertex v of the second clique so that u and v have the same color. As altogether there are s such pairs, and each one has to be monochromatic in at least one vertex coloring in a collection exhibiting an upper bound for the product dimension, the number of such colorings is at least s . Another comment is that $Q(s, r)$ is clearly monotone non-decreasing in both r and s . Therefore, for every $r, s \geq 2$, $Q(s, r) \geq Q(2, r) = \lceil \log_2(2r) \rceil$.

A *transversal design* $TD(r, s)$ of order s and block size r (with multiplicity $\lambda = 1$) is a set V of sr elements partitioned into r pairwise disjoint groups, each of size s , and a collection of blocks, each containing exactly one element of each group, so that every pair of elements from distinct groups is contained in exactly one block. A transversal design is *resolvable* if its blocks can be partitioned into parallel classes where the blocks in any parallel class partition the set V . There is a substantial amount of literature about transversal designs, see [31]. It is not difficult to check that $Q(s, r) = s$ if and only if a resolvable $TD(r, s)$ exists and hence the known results about resolvable transversal designs supply nearly precise information for the range $r \leq s$ (it is easy to see that such a design cannot exist for $r > s$). In particular, for every prime power s , $Q(s, s) = s$ and therefore by the obvious monotonicity, for any prime power s , $Q(s, r) = s$ for every $2 \leq r \leq s$, and if p is a prime power then for every $s, r \leq p$, $Q(s, r) \leq p$. (Note that Propositions 19.2 and 22.1 also imply that $Q(s, s) = s$ when s is a prime.)

It is not difficult to prove that for every s, r_1, r_2 ,

$$Q(s, r_1 r_2) \leq Q(s, r_1) + Q(s, r_2). \quad (12)$$

Indeed, given the graph $K_s(r_1 r_2)$ consisting of $r_1 r_2$ disjoint cliques, each of size s , we can split the cliques into r_1 disjoint groups, each consisting of r_2 cliques. Define $Q(s, r_1)$ proper

colorings in which the cliques in every group are colored the same, based on the system of colorings that shows that the product dimension of $K_s(r_1)$ is $Q(s, r_1)$. Add to these $Q(s, r_2)$ additional colorings, whose restrictions to the r_2 cliques in each group are exactly the colorings showing that the product dimension of $K_s(r_2)$ is $Q(s, r_2)$. The resulting $Q(s, r_1) + Q(s, r_2)$ colorings establish (12).

The above comments together with the results in the previous sections provide upper and lower bounds for $Q(s, r)$ for all s and r , but these bounds are quite far from each other when r is much bigger than s but much smaller than $2^{s^3 \log \log s}$. In particular, for $r = 2^s$ the bounds we have are

$$s \leq Q(s, 2^s) \leq (1 + o(1)) \frac{s^2}{\log s}.$$

It would be interesting to close this gap.

Part VII

Discrepancy Minimization via a Self-Balancing Walk

This chapter is based on joint work with Yang Liu and Mehtaab Sawhney [11]. It won Best Student Paper Award at the Symposium on the Theory of Computing (STOC) conference and is currently under review at the *Siam Journal on Computing* after being invited for a submission.

23 Introductory Remarks

Consider the *vector balancing problem*: given a set of vectors $v_1, v_2, \dots, v_t \in \mathbb{R}^n$, we choose signs $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t \in \{-1, 1\}$ so that the *discrepancy* $\|\sum_{i=1}^t \varepsilon_i v_i\|_\infty$ is as small as possible. This encompasses many problems in discrepancy theory, including the Komlós conjecture and set discrepancy minimization. The Komlós conjecture asks for the best bound $B(n, t)$ such that for any matrix $A \in \mathbb{R}^{n \times t}$ with columns with ℓ_2 -norm at most 1, there is a signing $\varepsilon \in \{-1, 1\}^t$ with $\|A\varepsilon\|_\infty \leq B(n, t)$. The Komlós conjecture states that one may take $B(n, t) = O(1)$. The best known bound is $B(n, t) = O(\sqrt{\log \min(n, t)})$ due to Banaszczyk [13, 14]. Similarly, the case where $v_1, v_2, \dots, v_t \in \{0, 1\}^n$ captures the problem of minimizing set discrepancy. Here, Spencer’s famous “six standard deviations suffice” [81] shows that there exists a signing $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t \in \{-1, 1\}$ so that $\|\sum_{i=1}^t \varepsilon_i v_i\|_\infty = O(\sqrt{t \max(1, \log(n/t))})$. In particular, if $t = n$, then discrepancy $6\sqrt{n}$ is achievable, and this is tight up to a constant factor.

While the original proofs of Banaszczyk and Spencer [13, 14, 81] were nonconstructive,

there has been significant interest in making them algorithmic. The first major results in this direction, which gave polynomial time algorithms for Spencer’s result [81], were achieved by Bansal [15] and Lovett-Meka [64]. Since then, there have been several other constructive discrepancy minimization algorithms [78, 35, 16, 18, 17, ?], including those matching Banaszczyk’s bound for the Komlós Conjecture [13, 14] due to Bansal, Dadush, Garg [16] and Bansal, Dadush, Garg, and Lovett [17]. However, these algorithms do not seem to extend to an online setting.

23.1 Online Algorithms

The *online discrepancy minimization* problem, proposed by Spencer [80], concerns assigning weights $\varepsilon_i \in \{-1, 1\}$ to vectors $v_1, v_2, \dots, v_t \in [-1, 1]^n$ which arrive one at a time, while trying to maintain a low ℓ_∞ norm of all the partial sums $w_i = \sum_{j=1}^i \varepsilon_j v_j$. Against adaptive adversaries, the best possible bound is $\Omega(\sqrt{t})$ as the adversary may choose the next vector v_{i+1} to be orthogonal to w_i . Similarly, a greedy algorithm where ε_{i+1} is chosen so that $\varepsilon_{i+1} \langle v_{i+1}, w_i \rangle \leq 0$, achieves a bound of $O(\sqrt{t})$.

However, it was open whether one could get any improvement over the trivial $O(\sqrt{t})$ bound in the *oblivious* version of this vector balancing problem. Here, an adversary fixes vectors $v_1, \dots, v_t \in [-1, 1]^n$ ahead of time, and the player may use randomness. No deterministic algorithm can do better than $\Omega(\sqrt{t})$ as this is essentially equivalent to the case of an adaptive adversary. Our main result provides a probabilistic algorithm that achieves an ℓ_∞ bound of $O(\max_{i \in [t]} \|v_i\|_2 \log(nt/\delta))$ with probability at least $1 - \delta$ for any $\delta > 0$. For v_i in $[-1, 1]^n$, our result, assuming no further structure on the v_i , gives a bound of $O(\sqrt{n} \log(nt/\delta))$.

The best previous results in this direction required that v_i were sampled in an iid manner from a distribution $\mathbf{p}$ which is known beforehand. In this direction there are a number of previous works each achieving different guarantees. In the most restrictive setting where

the vectors v_i were sampled uniformly from $[-1, 1]^n$, Bansal and Spencer [21] showed an ℓ_∞ guarantee of $O(\sqrt{n} \log t)$. In the more general setting where when the vectors v_i were sampled from $\mathbf{p}$ supported on $[-1, 1]^n$, Aru, Narayanan, Scott, and Venkatesan [12] achieved a bound of $O_n(\sqrt{\log t})$ (where the implicit dependence on n is super-exponential) and Bansal, Jiang, Singla, and Sinha [20] achieved an ℓ_∞ guarantee of $O(n^2 \log(nt))$.

23.2 Algorithm Motivation and Description

We now motivate and describe the self-balancing walk which is at the heart of the chapter. Let the input vectors be $v_1, \dots, v_t$, where $\|v_i\|_2 \leq 1$ for all $i \in [t]$, and let $\varepsilon_1, \dots, \varepsilon_t \in \{-1, 1\}$ be chosen later. The algorithm maintains the current partial sum $w_{i-1} = \varepsilon_1 v_1 + \dots + \varepsilon_{i-1} v_{i-1}$, and will decide the sign of v_i probabilistically depending on w_{i-1} and v_i . First, it is natural that the algorithm should pick $\varepsilon_i = -1, 1$ with probability $1/2$ each if w_{i-1} and v_i are orthogonal vectors. Additionally, the more correlated v_i and w_{i-1} are, the higher probability that the algorithm picks $\varepsilon_i = -1$.

We set the algorithm, which we call the *self-balancing walk*, to pick $\varepsilon_i = 1$ with probability $\frac{1}{2} - \frac{\langle w_{i-1}, v_i \rangle}{2c}$, where $c = 30 \log(nt/\delta)$ is a constant upper bound on $|\langle w_{i-1}, v_i \rangle|$ with high probability. This way, if w_{i-1} and v_i are orthogonal, the algorithm picks $\varepsilon_i = -1, 1$ with equal probability. Additionally, the bias is stronger further from the origin, and it is stronger if one of the signings reduces the norm by a larger amount. The bias is also linear in the dot product between w_{i-1} and v_i , which is a key property we will use in our analysis.

Our main result is that the self-balancing walk maintains low discrepancy with high probability.

Theorem 23.1. *For any vectors $v_1, v_2, \dots, v_t \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ for all $i \in [t]$, algorithm $\text{BALANCE}(v_1, \dots, v_t, \delta)$ maintains $\|w_i\|_\infty = O(\log(nt/\delta))$ for all $i \in [t]$ with probability $1 - \delta$.*

Our bound is tight up to a square root log factor, due to a $\Omega(\sqrt{\log t / \log \log t})$ lower

Algorithm 1: $\text{BALANCE}(v_1, \dots, v_t, \delta)$ – Given a sequence of input vectors $v_1, \dots, v_t$ it assigns the vectors ± 1 signs online to maintain low discrepancy with probability $1 - \delta$.

```

1  $w_0 \leftarrow 0$ .
2  $c \leftarrow 30 \log(nt/\delta)$ .
3 for  $1 \leq i \leq t$  do
4   if  $|\langle w_{i-1}, v_i \rangle| > c$  or  $\|w_{i-1}\|_\infty > c$  then
5     Fail. ▷ Algorithm terminates with failure.
6    $p_i \leftarrow \frac{1}{2} - \frac{\langle w_{i-1}, v_i \rangle}{2c}$ .
7    $\varepsilon_i \leftarrow 1$  with probability  $p_i$ , and  $\varepsilon_i \leftarrow -1$  with probability  $1 - p_i$ .
8    $w_i \leftarrow w_{i-1} + \varepsilon_i v_i$ .
```

bound in [20]. We conjecture that a variant of our algorithm can maintain an ℓ_∞ bound of $O(\sqrt{\log nt})$ instead of $O(\log nt)$, and leave this as an interesting open problem.

23.3 Consequences of Theorem 23.1

The self-balancing works against oblivious adversaries. Therefore, Theorem 23.1 implies tight bounds up to logarithmic factors in n and t for all of Questions 1-5 in [20]. We state Questions 4 and 5, which are about oblivious adversaries, as these generalize the stochastic and prophet models discussed in the other questions raised in [20].

- [20, §8, Question 4] Is there an online algorithm which maintains discrepancy $\text{poly}(n, \log t)$ on any sequence of vectors in $[-1, 1]^n$ chosen by an oblivious adversary?
- [20, §8, Question 5] Is there an online algorithm which maintains discrepancy $\text{poly}(s, \log n, \log t)$ on any sequence of s -sparse vectors in $[-1, 1]^n$ chosen by an oblivious adversary?

Theorem 23.1 directly implies a nearly tight bound of $O(\sqrt{n} \log(nt))$ for Question 4, and $O(\sqrt{s} \log(nt))$ for Question 5.

Our results extend to the applications discussed in [20, 19], including improved bounds to online geometric discrepancy problems, balancing against convex bodies and multicolor discrepancy. This is discussed in Section 6.

Finally we also obtain linear time algorithms for logarithmic bounds for the Komlós Conjecture. In what follows, $\text{nnz}(A)$ denotes the number of nonzero entries in the matrix A .

Theorem 23.2. *Given a matrix $A \in \mathbb{R}^{n \times t}$ with columns with ℓ_2 -norm at most 1, we can find with high probability in $O(\text{nnz}(A))$ time a vector $x \in \{-1, 1\}^t$ such that $\|Ax\|_\infty = O(\sqrt{\log t \cdot \log n})$.*

This requires a minor modification of our algorithm which we sketch at the end of Section 24. Previous constructive discrepancy minimization algorithms [15, 64, 16, 35, 17] involved expensive linear algebra or solving linear or semidefinite programs, although [16, 17] achieve stronger bounds of $\|Ax\|_\infty = O(\sqrt{\log t}), O(\sqrt{\log n})$ respectively. This result therefore can be seen as a stepping stone towards giving input-sparsity time algorithms for discrepancy problems, a direction mentioned by Dadush [87].

23.4 Previous Approaches for Algorithmic Discrepancy Minimization

Here we describe previous approaches to algorithm discrepancy minimization and the difficulties in extending previous methods to the online setting against oblivious adversaries. Previous approaches [15, 64, 35, 16, 17] either solve linear or semidefinite programs or perform random walks on the sign vector $(\varepsilon_1, \dots, \varepsilon_t)$, all of which require knowing all input vectors $v_1, \dots, v_t$ at the beginning.

The results of [21] and [20] work for the restricted online setting where all vectors come from a fixed distribution $\mathbf{p}$ and work by choosing the sign $\varepsilon_i = -1, 1$ to minimize a potential function of the current point such as $\Phi(w) = \sum_{i=1}^n \cosh(\lambda w_i)$. This approach has significant difficulties working in the setting of oblivious adversaries as algorithms minimizing potentials are deterministic, and a lower bound of $\Omega(t^{1/2})$ holds for any deterministic algorithm against oblivious adversaries.

23.5 Overview of Analysis of the Self-Balancing Walk

A natural approach to analyzing the self-balancing walk would be to show that some potential function or exponential moment is increasing slowly in expectation, as is done with several analyses of (sub)martingales. However, this mode of analysis encounters significant difficulties due to the fact that the partial sum w_{i-1} might be orthogonal to v_i . This prevents us from arguing that some potential function is pointwise a (sub)martingale.

We instead maintain a distributional guarantee that the distribution of w_i over executions of our algorithm is less “spread” out than an associated normal distribution. This allows us to transfer the tail bounds on normal distributions to the distribution of w_i .

23.6 Preliminaries and Conventions

For a vector v , we let $v(i)$ denote the i -th coordinate of v . For positive semidefinite matrices A, B we write $A \preceq B$ if $B - A$ is positive semidefinite. For a positive semidefinite matrix $M \in \mathbb{R}^{m \times m}$ we define $\mathcal{N}(0, M)$ as the normal distribution with covariance M , i.e. $\mathbb{E}_{x \sim \mathcal{N}(0, M)}[x_i x_j] = M_{ij}$. For a subset $S \subseteq \mathbb{R}^k$, we write 1_S for the indicator function of the set S .

23.7 Concurrent and Independent Work

In independent and concurrent work, Bansal, Jiang, Meka, Singla, Sinha [19] (building on the techniques of [20]) achieve similar guarantees to the present work (with worse poly-log factors) for the online Komlós problem restricted to the setting where vectors are sampled randomly from a fixed distribution $\mathbf{p}$ inside the unit sphere. However [19] uses potential based techniques as in [20] and thus their results do not extend to minimizing discrepancy in the (more general) oblivious adversary model which is the primary focus of this chapter.

24 Analysis

24.1 Properties of Spreading

We now define the key notion for the analysis – the notion of one random variable being a *spread* of another.

Definition 24.1. *We say that random variable Y on $\mathbb{R}^n$ is a spread of random variable X on $\mathbb{R}^n$ if there exists a coupling of X and Y such that $\mathbb{E}[Y|X] = X$.*

The definition above appears in mathematical economics literature under the name “mean-preserving spread” [77] and is closely related to “second-order stochastic dominance” [51, 50, 77]. The term “spread” here is used for the following reason. Consider the coupling between X and Y such that $\mathbb{E}[Y|X] = X$. Then the random variable $Y - X$ conditional on X has mean 0. Thus, if Y is a spread of X , one can obtain Y by first sampling X , and then adding a mean 0 random variable Z whose distribution may depend on the specific value of X sampled. Furthermore, since whether Y is a spread of the random variable X only depends on the respective distributions of Y and X , we will often refer to distributions as spreads of one another. Equivalently, X and Y can be coupled so that X, Y form a two-step martingale; the former perspective, however, is far more useful here.

Lemma 24.2. *Let distribution Y be a spread of X . For any convex function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$, we have that $\mathbb{E}_{x \sim X} \Phi(x) \leq \mathbb{E}_{y \sim Y} \Phi(y)$.*

Proof. Couple X and Y in the manner which demonstrates that Y is a spread of X . Then

$$\mathbb{E}_{y \sim Y} \Phi(y) = \mathbb{E}[\Phi(Y)|X] \geq \mathbb{E}\Phi(\mathbb{E}[Y|X]) = \mathbb{E}\Phi(X)$$

where we have used Jensen’s inequality and that $\mathbb{E}[Y|X] = X$. □

Spreading is transitive and preserved by linear transformations.

Lemma 24.3 (Spreading is transitive). *If Z is a spread of Y and Y is a spread of X , then Z is a spread of X .*

Lemma 24.4 (Linear transformations maintain spreading). *If Y is a spread of X , then for any linear transformation M on $\mathbb{R}^n$ we have that MY is a spread of MX .*

The following is a slightly more abstract property of spreading that we need for Theorem 23.1. Intuitively, it is a formal statement of the intuition that you can add up pairs of variables that spread other variables.

Lemma 24.5. *Consider random variables, X, Y, W, Z , where X, Y are dependent. Suppose that W is a spread of X , and Z spreads the conditional distribution of $Y - X$ given $X = x$ for any value x . Then $W + Z$ is a spread of Y .*

Remark. *It is implicit in the above definition that X and Y live on the same probability space.*

Proof. The proof produces the desired coupling between Y and $W + Z$ as follows.

- First sample X and then sample W using the coupling that demonstrates that W is a spread of X . Fix the value of X sampled, so $X = x$.
- Now, sample Y conditioned on $X = x$.
- Finally sample Z from the coupling which demonstrates that Z is a spread of $Y - X$.

We prove that $W + Z$ is a spread of Y . Note that we have

$$\begin{aligned}\mathbb{E}[W + Z|X, Y] &= \mathbb{E}[W|X, Y] + \mathbb{E}[Z|X, Y] \\ &= \mathbb{E}[W|X] + \mathbb{E}[Z|X, Y - X] \\ &= X + (Y - X) = Y\end{aligned}$$

where each equality follows by construction. Therefore, $\mathbb{E}[W + Z|Y] = Y$. The result follows. $\square$

Finally, we prove a relation in terms of spreading for normal distributions and bounded mean-zero random variables.

Lemma 24.6 (Spreading real variables by Gaussians). *Let X be a real-valued random variable with $\mathbb{E}[X] = 0$ and $|X| \leq C$ almost surely. Then $G = \mathcal{N}(0, \pi C^2/2)$ is a spread of X .*

Proof. We first prove that the Bernoulli distribution with equal weight on $\pm C$ is a spread of any such variable X . To see this, define Y conditional on X to be C with probability $(X + C)/(2C)$ and $-C$ with probability $(-X + C)/(2C)$. Then Y is clearly supported on only $\pm C$ and

$$\mathbb{E}[Y|X] = \frac{C(X + C)}{2C} + \frac{-C(-X + C)}{2C} = X,$$

so the Bernoulli distribution with equal weight on $\pm C$ is a spread of X . To see that G is a spread of the Bernoulli distribution with equal weight on $\pm C$, let Z be distributed as $|G|$ if $X = C$ and $-|G|$ otherwise. The result follows by noting that

$$\mathbb{E}[|G|] = \sqrt{\frac{\pi}{2}} C \mathbb{E}_{Z \sim \mathcal{N}(0,1)}[|Z|] = C$$

and Z is distributed as G . $\square$

24.2 Proof of Theorem 23.1

We now formalize the notion of the distribution at time i induced by our algorithm. For $i \in [t]$, this is defined to be the distribution of w_i , except with all mass where the algorithm failed (line 5 of our algorithm) moved to 0.

Definition 24.7 (Distribution induced by BALANCE). *We define the distribution induced by $\text{BALANCE}(v_1, \dots, v_t, \delta)$ recursively.*

- *At $i = 0$ we have all mass of the distribution at 0.*
- *Move all mass with $|\langle w_i, v_{i+1} \rangle| > c$ or $\|w_i\|_\infty > c$ to 0 – this mass will stay at 0 for the remainder of the process. We refer to such w_i as being corrupted.*
- *For the remaining mass $i \in [t]$, evolve the distribution according to lines 7, 8 of our algorithm.*

$\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t, \delta))$ will denote the distribution of the vector w_i after i time steps of the above procedure.

One key observation is that $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t, \delta))$ is symmetric about the origin at each stage. We will ultimately compare $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t, \delta))$ to $\mathcal{N}(0, M_i)$ for an appropriate set of covariance matrices M_i .

Definition 24.8. *Fix $L = 2\pi$ and $c \geq 1$. Define $M_0 = 0$. For $i \geq 1$ define M_i inductively as*

$$M_i = (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T) + L v_i v_i^T.$$

We will ultimately show by induction that the distribution $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t))$ is spread by $\mathcal{N}(0, M_i)$. For intuition why this might hold, the formula in Definition 24.8 is essentially saying that the distribution $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t))$ will undergo a linear transformation by $(I - c^{-1}v_i v_i^T)$ in expectation, and then accumulate a constant variance in direction v_i .

We now note that these covariance matrices M_i as defined in Definition 24.8 are pointwise upper bounded independent of time.

Lemma 24.9. *Let $M \succeq 0$, $c \geq 1$, $L \geq 0$. If $M \preceq LcI$ then for any vector $v \in \mathbb{R}^n$ with $\|v\|_2 \leq 1$*

$$M' = (I - c^{-1}vv^T)M(I - c^{-1}vv^T) + Lvv^T$$

satisfies $0 \preceq M' \preceq LcI$.

Proof. It is direct that $0 \preceq M'$. Note that

$$\begin{aligned} M' &\preceq (I - c^{-1}vv^T)LcI(I - c^{-1}vv^T) + Lvv^T \\ &= LcI - L(2 - c^{-1}\|v\|_2^2)vv^T + Lvv^T \preceq LcI \end{aligned}$$

as $2 - c^{-1}\|v\|_2^2 \geq 2 - c^{-1} \geq 1$. □

Applying Lemma 24.9 inductively gives us the following immediate corollary.

Corollary 24.10. *For all $i \in [t]$ we have that $0 \preceq M_i \preceq LcI$.*

The following lemma is the key step in our analysis, where we show that the distribution $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t))$ is spread by normal distributions with covariance matrices defined in Definition 24.8.

Lemma 24.11. *$\mathcal{N}(0, M_i)$ is a spread of $\mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t, \delta))$ for all times $i \in [t]$.*

Proof. We write $\mathcal{D}_i := \mathcal{D}_i(\text{BALANCE}(v_1, \dots, v_t, \delta))$ for simplicity throughout the proof. We can compute that if the algorithm does not fail (line 5 of our algorithm) then

$$\begin{aligned} \mathbb{E}[w_i | w_{i-1}] &= w_{i-1} + (2p_i - 1)v_i \\ &= w_{i-1} - c^{-1}v_i v_i^T w_{i-1} \\ &= (I - c^{-1}v_i v_i^T)w_{i-1}. \end{aligned}$$

Define w'_{i-1} to be w_{i-1} except when w_{i-1} became corrupted – in this case set w'_{i-1} to 0. Let $\mathcal{D}'_{i-1}$ be the distribution of w'_{i-1} . As $\mathcal{D}_{i-1}$ is symmetric, $\mathcal{D}_{i-1}$ is a spread of $\mathcal{D}'_{i-1}$. Finally we say w'_{i-1} is corrupted if the corresponding w_{i-1} is corrupted.

Define random variables

$$R(w'_{i-1}, v_i) = \begin{cases} 0 & \text{if } w'_{i-1} \text{ is corrupted,} \\ 1 + c^{-1}\langle w'_{i-1}, v_i \rangle & \text{with probability } p_i \\ -1 + c^{-1}\langle w'_{i-1}, v_i \rangle & \text{if } w'_{i-1} \text{ is not corrupted,} \\ -1 + c^{-1}\langle w'_{i-1}, v_i \rangle & \text{with probability } 1 - p_i \\ & \text{if } w'_{i-1} \text{ is not corrupted,} \end{cases}$$

where p_i is defined in line 6 of our algorithm. By definition, w_i is distributed as

$$(I - c^{-1}v_i v_i^T)w'_{i-1} + R(w'_{i-1}, v_i)v_i.$$

By induction and Lemma 24.4 we have that

$$\mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T))$$

is a spread of $(I - c^{-1}v_i v_i^T)\mathcal{D}_{i-1}$. By Lemma 24.3 therefore it follows that $\mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T))$ is a spread of $(I - c^{-1}v_i v_i^T)\mathcal{D}'_{i-1}$.

Note that by definition, $R(w_{i-1}, v_i)$ is mean 0. Furthermore $R(w_{i-1}, v_i)$ is supported on $[-2, 2]$ as $|\langle w_{i-1}, v_i \rangle| \leq c$ if w'_{i-1} is not corrupted. By Lemma 24.6, we have that $\mathcal{N}(0, 2\pi)$ is

a spread of $R(w'_{i-1}, v_i)$ for each w'_{i-1} . Thus by Lemma 24.5, we have

$$\begin{aligned} & \mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T)) + \mathcal{N}(0, 2\pi v_i v_i^T) \\ &= \mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T) + 2\pi v_i v_i^T) = \mathcal{N}(0, M_i) \end{aligned}$$

is a spread of $\mathcal{D}_i$, as desired.

To be precise, in the application of Lemma 24.5, we take $X = \mathbb{E}[w_i | w'_{i-1}] = (I - c^{-1}v_i v_i^T)w'_{i-1}$ which by Lemma 24.4 is spread by $W = \mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T))$, Y to be w_i , and note that $Y - X = R(w'_{i-1}, v_i)v_i$ is always spread by $Z = \mathcal{N}(0, 2\pi)v_i$. Therefore by Lemma 24.5 we have that Y or w_i is spread by $W + Z = \mathcal{N}(0, (I - c^{-1}v_i v_i^T)M_{i-1}(I - c^{-1}v_i v_i^T) + 2\pi v_i v_i^T) = \mathcal{N}(0, M_i)$ as desired. $\square$

We can get tail bounds on M_i because they are always bounded (Corollary 24.10).

Lemma 24.12. *For any vector $u \in \mathbb{R}^n$ with $\|u\|_2 \leq 1$ we have that*

$$\mathbb{E}_{x \sim \mathcal{N}(0, M_i)} [e^{\langle x, u \rangle^2 / (4Lc)}] \leq \sqrt{2}.$$

Proof. Note that $\langle x, u \rangle$ is distributed as $\mathcal{N}(0, u^T M_i u)$ and by Corollary 24.10 we have $u^T M_i u \leq Lc$. This implies that $\mathcal{N}(0, Lc)$ is a spread of $\mathcal{N}(0, u^T M_i u)$. The result then follows by noting that

$$\mathbb{E}_{x \sim \mathcal{N}(0, u^T M_i u)} [e^{x^2 / (4Lc)}] \leq \mathbb{E}_{x \sim \mathcal{N}(0, Lc)} [e^{x^2 / (4Lc)}] = \sqrt{2},$$

where we have used Lemma 24.2 on the convex function $e^{x^2 / (4Lc)}$. $\square$

We are now ready to complete the proof of Theorem 23.1 by combining the fact that $\mathcal{D}_i$ is spread by $\mathcal{N}(0, M_i)$ with the tail bounds in Lemma 24.12.

Proof of Theorem 23.1. It suffices to bound the total amount of additional corrupted mass in $\mathcal{D}_i$ compared to $\mathcal{D}_{i-1}$. To bound this note for any vector $u \in \mathbb{R}^n$ with $\|u\|_2 \leq 1$ that

$$\Pr_{w \sim \mathcal{D}_i} [|\langle w, u \rangle| > c] \leq e^{-c^2/(4Lc)} \mathbb{E}_{w \sim \mathcal{D}_i} \left[e^{\langle w, u \rangle^2/(4Lc)} \right] \quad (13)$$

$$\leq e^{-c/4L} \mathbb{E}_{w \sim \mathcal{N}(0, M_i)} \left[e^{\langle w, u \rangle^2/(4Lc)} \right] \leq \sqrt{2} e^{-c/4L} \leq (2nt)^{-1} \delta \quad (14)$$

by the choice of c . Here, we have used that $\mathcal{N}(0, M_i)$ is a spread of $\mathcal{D}_i$ through Lemma 24.11, Lemma 24.2 on the convex function $e^{\langle x, u \rangle^2/(4Lc)}$, and Lemma 24.12. Now, at step $i \in [t]$ union bound over the choices $u = v_i, e_1, e_2, \dots, e_n$ where e_j denotes the unit basis vector for coordinate j . $\square$

Proof of Theorem 23.2. We modify the self-balancing walk to instead use the parameter $c = 30 \log(t/\delta)$ for $\delta = \frac{1}{\text{poly}(t)}$ and note that we no longer need to maintain that $\|w_i\|_\infty \leq c$ at every step and instead only at the end. As in Equations (13) and (14) we have

$$\Pr_{w_{i-1} \sim \mathcal{D}_{i-1}} [|\langle w_{i-1}, v_i \rangle| > c] \leq \sqrt{2} e^{-c/4L} \leq \delta/t.$$

Therefore, a union bound over $i \in [t]$ shows that the algorithm does not fail for steps $i \in [t]$ with probability $1 - \delta$. Again, as in Equations (13) and (14) for any basis vector e_j we have that

$$\Pr_{w_t \sim \mathcal{D}_t} [|\langle w_t, e_j \rangle| > \sqrt{8cL \log n}] \leq \sqrt{2} e^{-8cL \log n / 4cL} \leq \frac{2}{n^2}.$$

Now union bounding over all $j \in [n]$ and using that $\sqrt{8cL \log n} = O(\sqrt{\log t \cdot \log n})$ gives the bound. $\square$

Part VIII

On a Conjecture of Gowers on Clique Differences

This part is based on work that has been submitted [7] but not published.

25 Introductory Remarks

The polynomial density Hales-Jewett conjecture is one of the central remaining open questions in ergodic Ramsey theory, and is a common generalization of the Bergelson-Leibman theorem [23] and the density Hales-Jewett theorem [47]. One special case was proposed by Gowers as a potential polymath project [84].

Conjecture 25.1 (Gowers, 2009, [84]). *For every $\delta > 0$, there exists n such that if A is any collection of at least $\delta 2^{\binom{n}{2}}$ graphs with vertex set $\{1, 2, \dots, n\}$, then A contains two distinct graphs G and H such that $G \subset H$ and $H \setminus G$ is a clique.*

In the same blog post, Gowers asks the following question in an attempt to prove Conjecture 25.1.

Question 25.2 (Gowers, 2009). *Is it true that for every $\delta > 0$ there exists a sequence of distinct graphs $G_1 \subset G_2 \subset \dots \subset G_r$ such that for every subset $A \subset \{1, 2, \dots, r\}$ of size at least δr there exist $i < j$ such that $i, j \in A$ and $G_j \setminus G_i$ is a clique?*

Given a sequence of distinct graphs $G_1 \subset G_2 \subset \dots \subset G_r$, we form the graph G on them by connecting G_i, G_j for $i < j$ when $G_j \setminus G_i$ is a clique. In this short chapter we settle Question 25.2 in the negative, and we show that G always has an independent set of size at least $\frac{r}{20}$.

Theorem 25.3. *For every sequence of distinct graphs $G_1 \subset G_2 \subset \dots \subset G_r$, there exists $A \subset \{1, 2, \dots, r\}$ of size at least $r/20$ so that for all $i, j \in A$, $G_j \setminus G_i$ is not a clique.*

26 Proof of Theorem 25.3

Lemma 26.1. *Given $a < b < c < d$, if $G_c \setminus G_a, G_d \setminus G_b$ are cliques, then $G_c \setminus G_b$ is also a clique.*

Proof. $G_c \setminus G_a$ and $G_d \setminus G_b$ are cliques, so their nonempty edge intersection $G_c \setminus G_b$ is also a clique. $\square$

Lemma 26.2. *For any y , it cannot be true that all of G_y, G_{y+1}, G_{y+2} have at least three right-neighbors and three left-neighbors in G .*

Proof. Assume this was the case. Then there is some $z \geq y + 3$ so that $G_z \setminus G_y$ is a clique. Also there is some $x < y$ so that $G_{y+1} \setminus G_x$ is a clique. By Lemma 26.1 on $(x, y, y + 1, z)$, then $G_{y+1} \setminus G_y$ is a clique. Similarly, $G_{y+2} \setminus G_{y+1}$ is a clique. Also there is some $u < y$ so that $G_{y+2} \setminus G_u$ is a clique, and then by Lemma 26.1 on $(u, y, y + 2, z)$, $G_{y+2} \setminus G_y$ is a clique, but it is the edge-disjoint union of the cliques $G_{y+2} \setminus G_{y+1}$ and $G_{y+1} \setminus G_y$, which is a contradiction. $\square$

Call a graph G_i *good* if it has either at most two right-neighbors in G or at most two left-neighbors in G . For any $G_1 \subset G_2 \subset \dots \subset G_r$, at least $\frac{r-2}{3}$ are good by Lemma 26.2 because we cannot have three bad in a row. So without loss of generality at least $\frac{r-2}{6}$ of the graphs have at most two right-neighbors. This means we can go through these graphs from left to right and greedily select an independent set of size at least $\frac{r-2}{18} \geq \frac{r}{20}$. We are done.

Following the initial proof, Noga Alon communicated an improvement of the constant in Theorem 25.3.

Lemma 26.3. *There are no $G_{3i-2}, G_{3i-1}, G_{3i}$ so that:*

- *There is some $j > 3i - 1$ so that $G_j \setminus G_{3i-2}$ is a clique.*
- *There is some $k > 3i - 1$ and some $\ell < 3i - 1$ so that $G_k \setminus G_{3i-1}$ and $G_{3i-1} \setminus G_\ell$ are cliques.*
- *There is some $m < 3i - 1$ so that $G_{3i} \setminus G_m$ is a clique.*

Proof. Assume this were the case. Then if $\ell = 3i - 2$, $G_{3i-1} \setminus G_{3i-2}$ is a clique. Else if $\ell < 3i - 2$, then by Lemma 26.1 on $(\ell, 3i - 2, 3i - 1, j)$, $G_{3i-1} \setminus G_{3i-2}$ is a clique anyway. Similarly, $G_{3i} \setminus G_{3i-1}$ is also a clique. Finally if $j = 3i$ or $m = 3i - 2$ then $G_{3i} \setminus G_{3i-2}$ is a clique. Else if $j > 3i$ and $m < 3i - 2$ then by Lemma 26.1 on $(m, 3i - 2, 3i, j)$, $G_{3i} \setminus G_{3i-2}$ is a clique anyway. This contradicts Lemma 26.2. $\square$

For each $i \geq 1$, either G_{3i-2} violates the first bulleted condition, G_{3i-1} the second, or G_{3i} the third. Select one of them correspondingly. Direct all edges from left to right. Now on this induced subgraph, every vertex G_{3i-2} has outdegree 0, every vertex G_{3i-1} has either indegree 0 (if there does not exist ℓ) or outdegree 0 (if there does not exist k), and every vertex G_{3i} has indegree 0. So every vertex has indegree or outdegree 0, and thus there is an independent set of size at least $\lceil [r/3]/2 \rceil \geq \frac{r-2}{6}$.

Part IX

Monochromatic Sums and Products of Polynomials

This part is based on a note that has been submitted but not yet published [5].

27 Introduction

A pattern $\mathcal{F}$ is *partition regular* over a set S , usually taken to be $\mathbb{N}$, if in any coloring of S with finitely many colors, there is some monochromatic copy of $\mathcal{F}$. Partition regularity is an important subfield of Ramsey theory, and encompasses results such as van der Waerden's theorem, which says that the pattern $\{x, x + d, \dots, x + nd\}$ is partition regular over $\mathbb{N}$, and Schur's theorem, that $\{y, x, x + y\}$ is partition regular over $\mathbb{N}$. In 1933, Rado [73] completely characterized partition regularity over $\mathbb{N}$ when $\mathcal{F}$ is taken to be the solution set to a linear system of equations.

Over polynomials, the situation is much less well understood. The pattern $\{x, y, xy\}$ is known to be partition regular over $\mathbb{N}$, because we can color powers of 2 and then find a pattern $\{2^x, 2^y, 2^{x+y}\}$ using Schur. It is natural to wonder whether there is a common generalization of this multiplicative form of Schur's theorem and the ordinary additive form. A central and longstanding open problem in partition regularity due to Hindman (see e.g. [22], [55], [66]) is to prove that the pattern $\{x, y, x + y, xy\}$ is partition regular over $\mathbb{N}$. In a 2016 breakthrough, Moreira [66] proved that the pattern $\{x, x + y, xy\}$ is partition regular over $\mathbb{N}$. Before 2016, it was not even known if $\{x + y, xy\}$ was partition regular. Moreira's

proof relies on the properties of $\mathbb{N}$, using extensively the following notion from dynamics.

Definition 27.1. *A subset S of $\mathbb{N}$ is syndetic if it has bounded gaps, i.e. there exists a finite set $F \subset \mathbb{Z}$ so that $S + F \supset \mathbb{N}$. A subset T of $\mathbb{N}$ is thick if it contains arbitrary long intervals. A piecewise syndetic subset U is the intersection of a thick set with a syndetic set.*

Moreira iterates the “piecewise syndetic” version of the van der Waerden theorem, in which the size of the set F blows up very quickly. Thus, he does not obtain primitive recursive bounds.

Following Moreira’s work, there has recently been exciting progress on Hindman’s conjecture, with Bowen and Sabok [28] proving it over the rationals. Over the integers, the problem is much more difficult. Very recently, Bowen [25] proved that any two-coloring of $\mathbb{N}$ has infinitely many monochromatic $\{x, y, x + y, xy\}$.

In this paper, we prove that $\{x, x + y, xy\}$ is partition regular in a more general setting. Let a *good* polynomial be a polynomial with zero constant term and non-negative integer coefficients. Let T be the semi-ring of nonzero good polynomials of a variable t .

Theorem 27.2. *For any set of good polynomials $P_1, \dots, P_k$, the pattern $\{x, x + P_1(y), \dots, x + P_k(y), xy\}$ is partition regular over T with primitive recursive bounds, and is thus also partition regular over $\mathbb{N}$ with primitive recursive bounds.*

Even the following special case was previously unsolved, though Green [49] has asked for “reasonable bounds”. The bounds we obtain are still quite large, as they require iterating the polynomial van der Waerden theorem of a semigroup, but because we only ever need to use it for a fixed number of colors, they are primitive recursive.

Corollary 27.3. *The pattern $\{x + y, xy\}$ is partition regular over $\mathbb{N}$ with primitive recursive bounds, i.e. there is some primitive recursive $N = N(n)$ so that when $[N] = \{1, 2, \dots, N\}$ is colored with n colors, there is a monochromatic $\{x + y, xy\}$.*

Moreira’s proof [66] works for polynomials over a finite field, but not for polynomials with rational or integer coefficients; he notes specifically that his Lemma 7.4 breaks down. Partition regularity over T (with primitive recursive bounds) is stronger than partition regularity over $\mathbb{N}$ (with primitive recursive bounds); we prove this in detail in Section 30.

An important feature of our proof is that it considers high degree polynomials, which we believe is necessary for settling Hindman’s conjecture. Previous work in this area implicitly works over a setting in which we conjecture that Hindman’s conjecture is actually false. This is discussed more in Section 30. Because our proof is the first to take advantage of higher degree polynomials, it relies in an intrinsic way on the polynomial van der Waerden theorem.

Whereas Moreira uses the linear van der Waerden theorem to prove $\{x, x + y, xy\}$ is partition regular, we prove a much better bound by instead using the polynomial van der Waerden theorem. Recall the polynomial van der Waerden theorem of Bergelson and Leibman [23], which states that for any set of good polynomials $P_1, \dots, P_k$, the pattern $x + P_1(y), \dots, x + P_k(y)$ is partition regular over $\mathbb{N}$. This holds over any semigroup [66] and there is a simple color-focusing proof [83] straightforwardly generalizing the ordinary color-focusing proof of van der Waerden theorem. A result of Shelah [79] provides primitive recursive bounds in the polynomial Hales-Jewett theorem of Bergelson and Leibman [60], and thus over any semigroup [66].

To some extent, working over the polynomials themselves can be viewed as analogous to extending results like van der Waerden Theorem to the Hales-Jewett Theorem; working over these more general (semi)-rings like T abstracts away the properties of $\mathbb{N}$, and captures the combinatorial structure. In particular, we avoid tools from dynamics such as the piecewise syndetic version of the van der Waerden theorem, and in fact our proof is “purely combinatorial” does not use the notions of syndetic, piecewise syndetic, or thickness at all. This again distinguishes our proof from the existing proofs in the literature.

We hope our new perspective will continue to find further applications and shine some

light on Hindman's conjecture. Based on the ideas here, in upcoming work with Bowen and Sabok we find several simple new proofs that there are infinitely many monochromatic $\{x, y, x + y, xy\}$ in any two-coloring of T and thus of $\mathbb{N}$, which was not known until 2022.

The structure of the paper is as follows. First, in Section 28, we describe a proof of the partition regularity of the pattern $\{x, x + y, xy\}$ in the 3-color case, to convey the general idea of our paper. Then, we discuss our proof of $\{x, x + y, xy\}$ for arbitrarily many colors in Section 29. Finally we conclude in Section 30 with some discussions and open problems.

28 The 3-color case

We will first describe the proof of the partition regularity of the pattern $\{a, a + b, ab\}$ for polynomials over 3 colors, red, blue, and green.

We will call a polynomial P of one or more variables *good* if it has zero constant term and all nonnegative integer coefficients. In particular, if x is in some semi-ring or ring R and P is a good polynomial, $P(x)$ will also be in R . Let T be the semi-ring of good polynomials of a single variable t ; for us, a semi-ring or ring does not necessarily have 1. Our goal is to prove $\{x, x + y, xy\}$ is partition regular over T . The *size* $s(P)$ of a good polynomial P is the largest c so that its degree and its coefficients are all at most c . The size of the 0 polynomial will be 0. Recall the polynomial van der Waerden theorem [23], which by the polynomial Hales-Jewett theorem holds over any semigroup [85].

Assume that there exists a coloring of the semi-ring T without the pattern $\mathcal{F} = \{a, a + b, ab\}$. By the polynomial van der Waerden theorem, for any c , there exist some x, d so that $x + P(d)$ is red for all good P of size at most c . Pick $c = c_0$ to be specified later.

Now, because we do not have the pattern $\{a, a + b, ab\}$, and for any good P of size less than c_0 , and we have the pattern $\{(x + P(d)), (x + P(d)) + d\}$, we know $d(x + P(d))$ cannot

be red for any good P of size less than c_0 . Hence, the terms $dx + dP(d)$ are all blue or green.

Again by the polynomial van der Waerden theorem over semigroups [85], if c_0 is large enough, we can find without loss of generality some f, e of the form $dP(d)$ (even of the form $P(d^2)$) so that $dx + f + Q(e)$ are all blue for good Q of size at most c_1 and that $x + \frac{f}{d} + P(d, \frac{e}{d})$ are red for good P of size at most c_1 , as we will have $x + \frac{f}{d} + P(d, \frac{e}{d}) = x + P'(d)$ for good P' of size at most c_0 . We rename $x + \frac{f}{d}$ as x , so that $x + P(d, \frac{e}{d})$ are all red, while now $dx + P(e)$ are all blue, again for all good P of size at most c_1 . In particular, c_0 will be a primitive recursive function of c_1 .

Now, because $\{x, x + de\}$ and $\{x + \frac{e}{d}, x + \frac{e}{d} + de\}$ are red, dex and $dex + e^2$ are not red. As $\{dx, dx + e\}$ and $\{dx + e, dx + e + e\}$ are blue, they also cannot be blue. So $dex, e(dx + e) = dex + e^2$ are both green. This means de^3x cannot be green. Because $\{x, x + de^3\}$ are red, de^3x cannot be red. But because $\{dx, dx + e^3\}$ are blue, de^3x cannot be blue. This gives a contradiction, and we are done.

29 Partition Regularity of Sums and Products in T

We now present the proof of the general case. The idea is the same as in the last section, although it is a little more technical to write out formally. We prove the following lemma. The c_i will shrink sufficiently quickly; what matters for us is that c_i is a primitive recursive function of c_{i+1} .

Lemma 29.1. *For any coloring C of T and any positive integers k and c_k , there exist some sequence of positive integers $c_0 \gg c_1 \gg c_2 \gg \dots \gg c_k$ and elements $b_1, \dots, b_k$ of T , so that the following holds.*

Let $a_1 = b_1$ and $a_i = a_{i-1}b_i$ for $2 \leq i \leq k$. In particular, $b_i = \frac{a_i}{a_{i-1}}$ for $i \geq 2$.

Then there exists some x_k so that for all good polynomials P of size at most c_i and all

$0 \leq i < k$,

$$\left(\prod_{j=1}^i a_j \right) x_k + P(a_{i+1}, b_{i+2}, \dots, b_k)$$

is the same color in the coloring C as $\left(\prod_{j=1}^i a_j \right) x_k$.

When $i = 0$, $\left(\prod_{j=1}^i a_j \right) x_0 = x_0$; when $i = k - 1$, $P(a_{i+1}, \dots) = P(a_k)$.

Proof. We induct on k . The base case $k = 0$ is trivial. Assume now that the lemma holds for some k . We assume that c_k is a large enough function of c_{k+1} , to be specified later. By the polynomial van der Waerden Theorem over semigroups [85], for any n there exists some f, e which are good polynomials of a_k^{k+1} so that for any good polynomial P with $s(P) \leq c_{k+1}$, all of $\left(\prod_{j=1}^k a_j \right) x_k + f + P(e)$ are the same color as $\left(\prod_{j=1}^k a_j \right) x_k + f$.

Now, we set $x_{k+1} = x_k + \frac{f}{\prod_{j=1}^k a_j}$ and $b_{k+1} = \frac{e}{a_k}$ so that $e = a_{k+1}$. For $0 \leq i < k$ and P of size at most c_{k+1} ,

$$\begin{aligned} (a_1 \cdots a_i) (x_{k+1}) + P(a_{i+1}, b_{i+2}, \dots, b_k, b_{k+1}) &= (a_1 \cdots a_i) \left(x_k + \frac{f}{\prod_{j=1}^k a_j} \right) + P(a_{i+1}, b_{i+2}, \dots, b_k, b_{k+1}) \\ &= (a_1 \cdots a_i) x_k + \frac{f}{a_{i+1} \cdots a_k} + P(a_{i+1}, b_{i+2}, \dots, b_k, b_{k+1}) \\ &= (a_1 \cdots a_i) x_k + P' \left(\frac{a_k}{a_{i+1}}, \dots, \frac{a_k}{a_{k-1}}, a_k \right) + P(a_{i+1}, b_{i+2}, \dots, b_k, b_{k+1}) \\ &= (a_1 \cdots a_i) x_k + P''(a_{i+1}, b_{i+2}, \dots, b_k, b_{k+1}) \end{aligned}$$

for some good polynomial P'' , recalling again that f is a polynomial of a_k^{k+1} . We can ensure that P'' has size less than c_k , by choosing c_k large enough.

For $i = k$,

$$(a_1 \cdots a_k) (x_{k+1}) + P(a_{k+1}) = (a_1 \cdots a_k) \left(x_k + \frac{f}{\prod_{j=1}^k a_j} \right) + P(e) = (a_1 \cdots a_k) x_k + f + P(e)$$

has the same color as $(a_1 \cdots a_k) x_{k+1} = (a_1 \cdots a_k) x_{k+1} + f$ for all $s(P) \leq c_{k+1}$. $\square$

This proof yields primitive recursive bounds. In particular, c_k will be a primitive recursive function of c_{k+1} , as due to Shelah [79], we have primitive recursive bounds in the polynomial Hales-Jewett theorem and thus in polynomial van der Waerden over a semigroup.

By passing to a subset of the $(\prod_{j=1}^i a_j)x$ with a fixed color, we can finish. Explicitly, if assume without loss of generality that $a_1 \cdots a_{\ell_1}$, and so on are the same color for $\ell_1 < \ell_2 < \cdots$, and set $c_1 = a_1 \cdots a_{\ell_1}$, $c_2 = a_{\ell_1+1} \cdots a_{\ell_2}$, and in general $c_n = a_{\ell_{n-1}+1} \cdots a_{\ell_n}$, then all $(\prod_{j=1}^i c_j)x + P(c_{j+1}, \dots, c_n)$ are the same color. In particular, $\{x, x + P_1(c_1), \dots, x + P_k(c_1), c_1x\}$ are the same color, for any finite family of good polynomials P_i . What we actually get is slightly more general, and very similar to Theorem 1.4 of [66].

30 Discussions and Open Problems

Some discussion about partition regularity over various rings and semi-rings is in order. There is an important way in which our proof differs from existing proofs in the literature. For $d \geq 1$, let P_d be the set of polynomials of a countable set of variables $t = x_1, x_2, \dots, x_n, \dots$ with nonnegative integer coefficients, constant term, and degree at most d in each variable. Let $P = P_\infty$. In particular, $P \supset T$. Note that for $d < \infty$, P_d is actually not a semi-ring, as for instance it contains x_1 but does not contain x_1^{d+1} .

Moreira's result [66] is stated over $\mathbb{N}$, and unlike our proof, it is more natural in that setting. However, as it only uses multilinear polynomials, it can be made to work over P_1 [25] using the methods of [26]. Previous work in the area, such as Bowen and Sabok's proof of the partition regularity of $\{x, y, x + y, xy\}$ over the rationals [28] and Bowen's proof of the 2 color case [27] can similarly be made to work over, say, P_5 . In other words, these proofs are not using high degree polynomials. We conjecture this is insufficient to settle the whole Hindman conjecture.

Conjecture 30.1. *If $1 \leq d < \infty$, then $\{x, y, x + y, xy\}$ is not partition regular over P_d .*

We do know that this conjecture is true for $d = 1$. It is a nice exercise to check that 2-coloring of P_1 where reducible polynomials are red and irreducible polynomials are blue avoids a monochromatic $\{x, y, x + y, xy\}$. We make a complimentary conjecture.

Conjecture 30.2. *For all $1 \leq i < j$, there exists some pattern $\mathcal{F}$ so that $\mathcal{F}$ is partition regular in P_j but not in P_i , but $\mathcal{F}$ does appear in P_i (i.e. the 1-coloring of $\mathcal{F}$ contains a monochromatic P_i).*

In particular, the previous conjecture says that we do *not* think the pattern $\{x, y, x + y, xy\}$ is an example of such a pattern. Clearly if $1 \leq i \leq j \leq \infty$, and F is partition regular over P_i , then it is partition regular over $P_j \supset P_i$.

Lemma 30.3. *A $\mathcal{F}$ is partition regular over $P = P_\infty$ if and only if it is partition regular over T . If a pattern is partition regular over T (with primitive recursive bounds), then it is partition regular over $\mathbb{N}$ (with primitive recursive bounds).*

Proof. Clearly if some pattern is partition regular over T , then it is partition regular over $P \supset T$. If a pattern $\mathcal{F}$ is not partition regular over T , then we can find a coloring c of T avoiding $\mathcal{F}$, and color any $P(x_1, \dots, x_n) \in P$ with the color of $P(t, t, \dots, t)$ in c , and so $\mathcal{F}$ will not be partition regular over P either.

If there is a coloring c of $\mathbb{N}$ avoiding $\mathcal{F}$, then we can color $p(t) \in T$ with the color of $p(2)$ in c . This coloring of T will avoid $\mathcal{F}$. Thus, partition regularity over T (with primitive recursive bounds) in turn implies partition regularity over $\mathbb{N}$ (with primitive recursive bounds). $\square$

Any nontrivial insight into the relationship between partition regularity over various settings would be interesting. We conjecture that the second part of Lemma 30.3 is not strict.

Conjecture 30.4. *A polynomial pattern $\mathcal{F}$ is partition regular over T if and only if it is partition regular over $\mathbb{N}$.*

In principle, one could ask about the problem for other semi-rings. Recall that an integral domain D has *characteristic 0* if for any positive integer k , there exist some $x \in D$ so that $kx := x + \cdots + x \neq 0$. The semi-ring T is the most important semi-ring to consider in the following sense.

Lemma 30.5. *If some family $\mathcal{F}$ of good polynomials $P_1(a_1, \dots, a_m), \dots, P_k(a_1, \dots, a_m)$ is partition regular over the semi-ring T , then it is partition regular over any integral domain R of characteristic 0.*

Proof. Assume $\mathcal{F}$ is not partition regular over R .

Say that there exists a sequence $t_1, \dots, t_n, \dots$ in R so that $P(t_n) \neq 0$ for all good polynomials P of size at most n . By a standard compactness argument, there is a limit t , and a coloring of all good polynomials of t avoiding $\mathcal{F}$.

If there exists a c so that for all x , there is some good polynomial P of size at most c , so that $P(x) = 0$, then taking the product of all of these polynomials, there exists some P so that $P(x) = 0$ identically for $x \in R$. If P is not a monomial and has degree d , then we may pick an appropriate positive integer k so that $Q(x) = P(kx) - k^d P(x)$ is also identically 0, but is a nonzero polynomial with integer coefficients of smaller degree. Q may not be good, but this is not a problem, all we require is that it has integer coefficients. We can reduce the degree until we find some kx^d which is identically 0 for $x \in R$. Because R is an integral domain, this means that $kx = 0$, a contradiction. $\square$

Partition regularity over T is also interesting for the following reason. If there is a coloring of T that avoids some polynomial pattern $\mathcal{F}$, we can assume without loss of generality that it is *polynomially syndetic*, i.e. there exists a c so that for any element x and any color K , there is some polynomial P of size at most c so that $P(x)$ is color K . This is again by

compactness; if there exists a color K and a sequence $x_1, \dots, x_n, \dots$ so that $P(x_n)$ is never color K when P is size at most c , then we can create a coloring of good polynomials of a limit x that avoids $\mathcal{F}$ and does not use color K , and reduce the number of colors used.

Part X

Bibliography

The following is a list of all the sources cited in my thesis, including my own papers [3, 5, 6, 7, 8, 9, 10, 11].

References

- [1] N. Alon, Covering graphs by the minimum number of equivalence relations, *Combinatorica* 6 (1986), 201–206.
- [2] N. Alon, On the capacity of digraphs, *European J. Combinatorics* 19 (1998), 1–5.
- [3] N. Alon and R. Alweiss, On the product dimension of clique factors, *European J. Combin.* 86 (2020), 103097.
- [4] N. Alon, E. Mossel, and R. Pemantle, Corruption Detection on Networks, *Theory of Computing*, 16(1), (2015), 1–23.
- [5] R. Alweiss, Monochromatic Sums and Products of Polynomials, <https://arxiv.org/abs/2211.00766>. Submitted.
- [6] R. Alweiss: Noisy corruption detection. *Inform. Process. Lett.*, 155, (2020), 105897.
- [7] R. Alweiss, On a Conjecture of Gowers on Clique Differences, 2020. <https://arxiv.org/abs/2011.04039>. Submitted.
- [8] R. Alweiss. Set system intersections can typically be blown up. 2020. <https://arxiv.org/abs/2003.11202>. Submitted.

- [9] R. Alweiss, S. Lovett, K. Wu, and J. Zhang, Improved bounds for the sunflower lemma, *Ann. of Math. (2)* 194, (2021), 795–815.
- [10] R. Alweiss, C. Ben Hamida, X. He, and A. Moreira, On the subgraph query problem, *Combin. Probab. Comput.* 30 (2020), 1–16.
- [11] R. Alweiss, Y.P. Liu, M. Sawhney, Discrepancy Minimization via a Self-Balancing Walk, *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, (2021), 14–20.
- [12] J. Aru, B. Narayanan, A. Scott, and R. Venkatesan, Balancing sums of random vectors, *Discrete Anal.*, (2018), Paper No. 4, 17.
- [13] W. Banaszczyk, Balancing vectors and gaussian measures of n-dimensional convex bodies, *Random Struct. Algorithms*, 12 (1998), 351–360.
- [14] W. Banaszczyk, On series of signed vectors and their rearrangements, *Random Struct. Algorithms*, 40 (2012), 301–316.
- [15] N. Bansal, Constructive algorithms for discrepancy minimization, in *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010*, IEEE Computer Society, (2010), 3–10.
- [16] N. Bansal, D. Dadush, and S. Garg, An algorithm for Koml’s conjecture matching Banaszczyk’s bound, in *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016*, IEEE Computer Society (2016), 788–799.
- [17] N. Bansal, D. Dadush, S. Garg, and S. Lovett, The gram-schmidt walk: a cure for the Banaszczyk blues, in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018*, ACM, (2018), 587–597.

- [18] N. Bansal and S. Garg, Algorithmic discrepancy beyond partial coloring, in Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, (2017), 914–926.
- [19] N. Bansal, H. Jiang, R. Meka, S. Singla, and M. Sinha, Online discrepancy minimization for stochastic arrivals, in Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, (2021), 2842–2861
- [20] N. Bansal, H. Jiang, S. Singla, and M. Sinha, Online vector balancing and geometric discrepancy, in Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, ACM, (2020), 1139–1152
- [21] N. Bansal and J. H. Spencer, On-line balancing of random inputs, *Random Struct. Algorithms*, 57(4), (2019), 879–891.
- [22] V. Bergelson. Ergodic Ramsey Theory-an update. Volume 228 of London Math. Soc. Lecture Note Ser, (1996), 1–61.
- [23] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden’s and Szemerédi’s theorems. *J. Amer. Math. Soc.*, 9(3) (1996), 725–753.
- [24] A. Blokhuis, On the Sperner capacity of the cyclic triangle, *J. Algebraic Combin.* 2(2) (1993), 123–124.
- [25] M. Bowen, personal communication.
- [26] M. Bowen, Composite Ramsey Theorems via Trees, available online at <https://arxiv.org/abs/2210.14311>
- [27] M. Bowen, Monochromatic Sums and Products in 2-colorings of $\mathbb{N}$, available online at <https://arxiv.org/abs/2205.12921>

- [28] M. Bowen and M. Sabok, Monochromatic Sums and Products in the Rationals, available online at <https://arxiv.org/abs/2210.12290>
- [29] A. Calderbank, P. Frankl, R. Graham, W. Li, and L. Schepp, The Sperner capacity of linear and nonlinear codes for the cyclic triangle, *J. Algebraic Combin.* 2(1) (1993), 31–48.
- [30] B.P. Cavalar, M. Kumar, B. Rossman, Monotone Circuit Lower Bounds from Robust Sunflowers, *Latin American Symposium on Theoretical Informatics*, (2020), 311–322.
- [31] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs*, Second Edition, Chapman and Hall/CRC, 2006.
- [32] D. Conlon, J. Fox, A. Grinshpun, and X. He, Online Ramsey numbers and the subgraph query problem, *Building Bridges II*, *Bolyai Soc. Math. Stud.*, 28 (2019).
- [33] S. Chow, S. Lindqvist, S. Prendiville, Rado’s criterion over squares and Higher Powers, *Journal of the European Math. Soc.*, 23(6), (2021), 1925–1997.
- [34] M. Deza and P. Frankl, Every large set of equidistant $(0, +1, -1)$ -vectors forms a sunflower, *Combinatorica*, 1(3), (1981), 225–231.
- [35] R. Eldan and M. Singh, Efficient algorithms for discrepancy minimization in convex sets, *Random Struct. Algorithms*, 53 (2018), 289–307.
- [36] P. Erdős and E. Rado, Intersection theorems for systems of sets, *Journal of the London Mathematical Society*, 35(1), (1960), 85–90.
- [37] P. Erdős and E. Szemerédi, Combinatorial properties of systems of sets, *Journal of Combinatorial Theory, Series A*, 24(3), (1978), 308–313.

- [38] U. Feige, D. Gamarnik, J. Neeman, M.Z. Rácz and P. Tetali, Finding cliques using few probes, *Random Struct. Algorithms* 56 (2020), 142–153.
- [39] A. Ferber, M. Krivelevich, B. Sudakov and P. Vieira, Finding Hamilton cycles in random graphs with few queries, *Random Struct. Algorithms* 49 (2016), 635–668.
- [40] A. Ferber, M. Krivelevich, B. Sudakov and P. Vieira, Finding paths in sparse random graphs requires many queries, *Random Struct. Algorithms* 50 (2017), 71–85.
- [41] K. Frankston, J. Kahn, B. Narayanan, and J. Park, Thresholds versus fractional expectation-thresholds, *Ann. of Math. (2)* 194 (2), (2021), 475–495.
- [42] A. Frieze and R. Kannan, A new approach to the planted clique problem, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science* 2 (2008), 187–198.
- [43] A. Furedi: On finite set-systems whose every intersection is a kernel of a star, *Discrete Mathematics* 47 (1983), 129–132.
- [44] Z. Füredi, On the Prague dimension of Kneser graphs, *Numbers, information and complexity* (2000), 125–128.
- [45] P. Frankl and Z. Furedi: Exact solution of some Turán-type problems, *Journal of Combinatorial Theory, Ser. A* 45 (1987), 226–262.
- [46] P. Frankl and Z. Furedi: Forbidding just one intersection, *Journal of Combinatorial Theory, Ser. A* 39 (1985), 160–176.
- [47] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett Theorem. *Journal d’Analyse Mathématique*, 57, (1991), 64–119.

- [48] L. Gargano, J. Körner and U. Vaccaro, Sperner capacities, *Graphs Combin.* 9(1), (1993), 31–46.
- [49] B. Green, Some Open Problems.
- [50] J. Hadar and W. R. Russell, Rules for ordering uncertain prospects, *The American Economic Review*, 59 (1969), 25–34.
- [51] G. Hanoch and H. Levy, The efficiency analysis of choices involving risk, *The Review of Economic Studies*, 36 (1969), 335–346.
- [52] J. Håstad, Computational limits of small-depth circuits, ISBN 0262081679, (1987).
- [53] D. Hefetz, M. Krivelevich, M. Stojakovic and T. Szabó, *Positional Games*, Birkhäuser, (2014).
- [54] N. Hindman, Partitions and sums and products. *Trans. Amer. Math. Soc.*, 247, (1979), 227–245.
- [55] N. Hindman, I. Leader, and D. Strauss. Open Problems in Partition Regularity. *Comb., Prob., Comput.*, 12(5-6), (2003), 571–583.
- [56] R. Kleinberg and M. Weinberg, Matroid prophet inequalities, in *Proceedings of the 44nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2012*, ACM, (2012), 123–136.
- [57] A.V. Kostochka, A bound of the cardinality of families not containing Δ -systems, *The Mathematics of Paul Erdős II*, (1997), 229–235.
- [58] M. Krivelevich, Positional games, *Proceedings of the International Congress of Mathematicians 3* (2014), 355–379.
- [59] C. Lee, Ramsey numbers of degenerate graphs. *Ann. of Math.* 185 (2017), 791–829.

- [60] V. Bergelson and A. Leibman. Set polynomial and polynomial extensions of the Hales-Jewett theorem. *Ann. of Math*, 150, (1999), 33–75.
- [61] X. Li, S. Lovett, and J. Zhang, Sunflowers and quasi-sunflowers from randomness extractors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*, (2018).
- [62] Y. Liu, Personal Communication.
- [63] L. Lovász, J. Nešetřil and A. Pultr, On the product dimension of graphs, *J. Combin. Theory Ser. B* 29, (1980), 47–67.
- [64] S. Lovett and R. Meka, Constructive discrepancy minimization by walking on the edges, *SIAM J. Comput.*, 44 (2015), 1573–1582.
- [65] S. Lovett, N. Solomon, and J. Zhang, From DNF Compression to Sunflower Theorems via Regularity, *34th Computational Complexity Conference, CCC 2019*, July 18-20, 2019.
- [66] J. Moreira, Monochromatic sums and products in $\mathbb{N}$, *Ann. Math. (3)* 185, (2017), 1069–1090.
- [67] J. Nešetřil and A. Pultr, A Dushnik-Miller type dimension of graphs and its complexity, *Fundamentals of Computation Theory, Proc. Conf. Poznan, Kórnik, 1977*, Springer Lect. Notes in Comp. Sci. 56, (1977), 482–493.
- [68] J. Park, H.T. Pham, A Proof of the Kahn-Kalai Conjecture, 2022. <https://arxiv.org/abs/2203.17207>.
- [69] H.T. Pham, personal communication.

- [70] S. Poljak, A. Pultr and V. Rödl, On the dimension of Kneser graphs, Algebraic Methods in Graph Theory, Proc. Colloq. in Szeged, Hungary, 1978, L. Lovász and V. T. Sós, Eds., Proc. Colloq. Math. Soc. J. Bolyai 25, (1981), 631–646.
- [71] D. H. J. Polymath, A new proof of the density Hales-Jewett theorem, Ann. Math., 175 (2012), 1283–1327.
- [72] M. Z. Rácz and B. Schiffer, Finding a planted clique by adaptive probing, preprint (2019), <https://arxiv.org/abs/1903.12050>.
- [73] R. Rado. Studien zur kombinatorik. Math. Zeit., 36, (1933), 242–280.
- [74] A. Rao, Coding for Sunflowers, Discrete Analysis, 2020, 8 pages.
- [75] A. Razborov, Bounded Arithmetic and Lower Bounds in Boolean Complexity, Feasible Mathematics II, Feasible Mathematics II, (1995), 384–386.
- [76] B. Rossman, The Monotone Complexity of k -Clique on Random Graphs, SIAM J. Comput., 43(1), (2014), 256–279.
- [77] M. Rothschild and J. E. Stiglitz, Increasing risk: I. A definition, Journal of Economic Theory, 2 (1970), 225–243.
- [78] T. Rothvoss, Constructive discrepancy minimization for convex sets, SIAM Journal on Computing, 46 (2017), 224–234.
- [79] S. Shelah, A partition theorem, Scientiae Math Japonicae, 2002, 413–438. Paper 679 at the Shelah Archive: <https://shelah.logic.at/files/199205/679.pdf>
- [80] J. Spencer, Balancing games, J. Comb. Theory, Ser. B, 23 (1977), 68–74.
- [81] J. Spencer, Six standard deviations suffice, Trans. Amer. Math. Soc., 289 (1985), 679–706.

- [82] J. Spencer, A. Srinivasan, and P. Tetali, The discrepancy of permutation families, in Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997), ACM, New York, (1997), 692–701.
- [83] M. Walters, Combinatorial proofs of the polynomial van der Waerden theorem and the polynomial Hales–Jewett theorem. *Journal of the London Math. Soc.*, 61(1), (2000), 1–12.
- [84] <https://archive.ph/wip/4I8uq>
- [85] <https://archive.ph/wip/DKpZ1>
- [86] <https://archive.ph/wip/UJJft>
- [87] <https://archive.ph/zrh9C>