PROBLEMS IN COMBINATORICS: HAMMING CUBES AND THRESHOLDS

By

JINYOUNG PARK

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Mathematics

Written under the direction of

Jeff Kahn

And approved by

New Brunswick, New Jersey

May, 2020

ABSTRACT OF THE DISSERTATION

PROBLEMS IN COMBINATORICS: HAMMING CUBES AND THRESHOLDS

by JINYOUNG PARK

Dissertation Director:

Jeff Kahn

This thesis consists of five parts, treated in Chapters 3-7 respectively. The material in Chapter 7 is separate, but the other topics are interconnected: Chapters 5 and 6 are both concerned with asymptotic enumeration on the cube (and share some ideas), while Chapters 3 and 4, though of independent interest, also provide crucial ingredients for Chapter 5.

Chapter 3 proves, answering questions of Y. Rabinovich, "stability" versions of upper bounds on maximal independent set counts in graphs under various restrictions. Roughly these say that being close to the maximum implies existence of a large induced matching or triangle matching (depending on assumptions).

The main result of Chapter 4 is an isoperimetric inequality for the Hamming cube Q_n, which can be written:

$$\int h_A^\beta d\mu \geq 2\mu(A)(1-\mu(A)).$$

Here μ is uniform measure on $V = \{0,1\}^n$ $(= V(Q_n))$; $\beta = \log_2(3/2)$; and, for $S \subseteq V$ and $x \in V$,

$$h_S(x) = \begin{cases} d_{V\setminus S}(x) & \text{if } x \in S, \\ 0 & \text{if } x \notin S \end{cases}$$

(where $d_T(x)$ is the number of neighbors of x in T).

This implies inequalities involving mixtures of edge and vertex boundaries, with related stability results, and suggests some more general possibilities.

Chapter 5 proves that the number of maximal independent sets in Q_n is asymptotically (with $N = 2^n$)

$$2n2^{N/4},$$

as was conjectured by Ilinca and Kahn in connection with a question of Duffus, Frankl and Rödl.

The value is a natural lower bound derived from a connection between maximal independent sets and induced matchings. The proof that it is also an upper bound draws on various tools, among them some of the results of Chapters 3 and 4, and isoperimetric ideas originating in work of Sapozhenko in the 1980's.

Chapter 6 proves that the number of (proper) 4-colorings of Q_n is asymptotically

$$6e2^N,$$

as was conjectured by Engbers and Galvin in 2012. The proof uses a combination of information theory (entropy) and, again, ideas related to the work of Sapozhenko mentioned above.

Chapter 7 proves a conjecture of Talagrand, a fractional version of the "expectation-threshold" conjecture of Kahn and Kalai. It is shown that for any increasing family $\mathcal{F}$ on a finite set X,

$$p_c(\mathcal{F}) = O(q_f(\mathcal{F}) \log \ell(\mathcal{F})),$$

where $p_c(\mathcal{F})$ and $q_f(\mathcal{F})$ are the threshold and "fractional expectation-threshold" of $\mathcal{F}$, and $\ell(\mathcal{F})$ is the maximum size of a minimal member of $\mathcal{F}$. This easily implies several heretofore difficult results and conjectures in probabilistic combinatorics, including thresholds for perfect hypergraph matchings (Johansson–Kahn–Vu), bounded degree spanning trees (Montgomery), and bounded degree graphs (new). The machinery developed also resolves (and vastly extends) the "axial" version of the random multi-dimensional assignment problem (earlier considered by Martin–Mézard–Rivoire and Frieze–Sorkin). The approach in this chapter builds on a recent breakthrough of Alweiss, Lovett, Wu and Zhang on the Erdős–Rado "Sunflower Conjecture."

Acknowledgements

I would like to express my deepest gratitude to my advisor, Jeff Kahn. He is the best advisor in every aspect, and I was extremely fortunate to have him as my advisor.

I am very thankful to the rest of my thesis committee members David Galvin, Bhargav Narayanan, and Shubhangi Saraf for their advice and continued support.

I would like to thank my husband, Hanbom, for his selfless support, and my daughter, Dahnbi, for her existence.

Most results of this thesis have already been published or under review for publications: the results in Chapters 3, 4, 5, 6, and 7 are contained in [38], [39], [40], [37], and [16], respectively.

Table of Contents

Chapter 1

Introduction

1.1 Maximal independent sets

Chapter 3 discusses stability aspects of maximal independent set counts. We recall two well-known bounds for $\text{mis}(G)$:

Theorem 1.1 (Moon-Moser [52]). *For any n-vertex graph G,*

$$\text{mis}(G) \leq 3^{n/3},$$

with equality iff G is the disjoint union of $n/3$ triangles.

Theorem 1.2 (Hujter-Tuza [29]). *For any n-vertex, triangle-free graph G,*

$$\text{mis}(G) \leq 2^{n/2},$$

with equality iff G is a perfect matching.

As usual M is an *induced matching* of G if it is an induced subgraph of G that is a matching. Similarly, T is an *induced triangle matching* of G if it is an induced subgraph of G that is a vertex-disjoint union of triangles.

Write $\text{itm}(G)$ for the number of triangles in a largest induced triangle matching in G, and $\text{im}(G)$ for the number of edges in a largest induced matching.

In what follows we will usually prefer to work with $\log \text{mis}$ ($\log = \log_2$), thought of as the number of bits needed to specify a maximal independent set. Note that $\text{itm}(G) \log 3$ and $\text{im}(G)$ are obvious lower bounds on $\log \text{mis}(G)$. We will be interested in questions suggested by Yuri Rabinovich [54] concerning "stability" aspects of upper bounds on mis, meaning, roughly: does large mis imply existence of a large induced triangle matching or large induced matching (as appropriate)? Formally, his conjectures were unquantified versions of the following three statements, whose proofs are the content of Chapter 3. (The questions were motivated by [55], which includes a proof of Theorem 1.4 for bipartite graphs.)

Theorem 1.3. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = \Omega(\varepsilon)$ such that for an n-vertex graph G, if $\mathrm{itm}(G) < (1-\varepsilon)\frac{n}{3}$ then $\log \mathrm{mis}(G) < (\frac{1}{3}\log 3 - \delta)n$.*

Theorem 1.4. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = \Omega(\varepsilon)$ such that for a triangle-free n-vertex graph G, if $\mathrm{im}(G) < (1-\varepsilon)\frac{n}{2}$ then $\log \mathrm{mis}(G) < (\frac{1}{2} - \delta)n$.*

Theorem 1.4 applies to bipartite graphs, of course. If G is bipartite with bipartition $X \cup Y$, then $\log \mathrm{mis}(G)$ is trivially at most $\min\{|X|, |Y|\}$ (since a maximal independent set is determined by its intersection with either of X, Y); so the statement is uninteresting unless G is close to balanced. But Rabinovich asked whether something analogous also holds for unbalanced (bipartite) G; more precisely, whether something along the following lines is true.

Theorem 1.5. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = 2^{-O(1/\varepsilon)}$ such that for a bipartite graph G on $X \cup Y$ with $|X| = n$ and $|Y| = 2n$, if $\mathrm{im}(G) < (1-\varepsilon)n$ then $\log \mathrm{mis}(G) < (1-\delta)n$.*

The proof of this is easily adapted to $|Y| = Bn$ (with δ then $\delta(\varepsilon, B)$), but to keep things simple we just state the result for $B = 2$.

Rabinovich suspected that, as in Theorems 1.3 and 1.4, $\delta(\varepsilon)$ should be linear in ε, but this is not true. In fact, Theorem 1.5 is tight (up to the implied constant); a construction to show this will be given in Section 3.2.2.

The proofs of Theorems 1.3 and 1.4 are similar, while that of Theorem 1.5 is related but somewhat trickier. The general approach has its roots in an idea for counting (ordinary) independent sets due to A.A. Sapozhenko [60], [61].

Strictly speaking we prove the theorems only for sufficiently large n, since we occasionally hide minor terms in $o(1)$'s. Of course combined with the characterizations of equality in Theorems 1.1 and 1.2 this does give the stated versions, though the δ's we produce may not be valid for small n. Since we are really interested in large n anyway, this approach seems preferable to carrying explicit error terms.

One reason to be interested in Theorem 1.4—or in what its proof actually gives; see Theorem 3.4—is its key role in a proof of Theorem 1.8. While Theorem 3.4 is one of the easier ingredients in the proof of Theorem 1.8, it is in some sense the basis for the whole; in particular, it was understanding the connection between induced matchings and stability that first suggested that the conjecture of [30], which had seemed out of reach, might in fact be manageable.

1.2 Hamming cubes

In Chapter 4 we discuss an isoperimetric inequality for the Hamming cube. Write Q_n for the n-dimensional Hamming cube and V for $V(Q_n)$. (A few basic definitions are given in Section 2.2.)

For $T \subseteq V$ let $d_T(x)$ be the number of neighbors of x in T $(x \in V)$ and define $h_S : V \to \mathbb{N}$ by

$$h_S(x) = \begin{cases} d_{V\setminus S}(x) & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases} \tag{1.1}$$

For $f : V \to \mathbb{N}$, a probability measure ν on V and $X \subseteq V$, we set

$$\int_X f d\nu = \sum_{x\in X} f(x)\nu(x).$$

We also use $\int$ for $\int_V$.

The main result of Chapter 4 is the following isoperimetric inequality. We use β for $\log_2(3/2)$ ($\approx .585$) and μ for uniform measure on V.

Theorem 1.6. *For any $A \subseteq V$,*

$$\int h_A^\beta d\mu \geq 2\mu(A)(1-\mu(A)). \tag{1.2}$$

The form of Theorem 1.6 is inspired by the following inequality of Talagrand [62].

Theorem 1.7. *For any $A \subseteq V$,*

$$\int \sqrt{h_A} d\mu \geq \sqrt{2}\mu(A)(1-\mu(A)).$$

Notice that Theorem 1.6 is tight in two ways: it holds with equality for subcubes of codimensions 1 and 2, and for subcubes of codimension 2 it does not hold for any smaller value of β. As far as we know the $\sqrt{2}$ in Theorem 1.7 could be replaced by 2 when $\mu(A) = 1/2$ (but of course not in general). The difference between 2 and $\sqrt{2}$ wouldn't have mattered in [62], but getting the right constant when $\mu(A)$ is close to 1/2 was crucial for applications, particularly Theorem 1.8, which was our original motivation.

Chapters 5 and 6 discuss two asymptotic enumeration problems on Hamming cubes. In Chapter 5 we prove the following statement, which was conjectured by Ilinca and Kahn [30] in connnection with a question of Duffus, Frankl and Rödl [7]. We use $\text{mis}(G)$ for the number of maximal independent sets (MIS's) of a graph G, and N for 2^n.

Theorem 1.8.

$$\text{mis}(Q_n) \sim 2n2^{N/4} \tag{1.3}$$

(where $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$).

(The original question from [7], answered in [30], just asked for the asymptotics of $\log \text{mis}(Q_n)$.)

In Chapter 6 we prove the following statement, which was conjectured by Engbers and Galvin [10]. We use $C_q(G)$ for the number of q-colorings of a graph G (where, here and throughout, *coloring* means *proper vertex* coloring).

Theorem 1.9. $C_4(Q_n) \sim 6e2^N$.

The general context for Theorems 1.9 and 1.8 is asymptotic enumeration in the spirit of, prototypically, Erdős, Kleitman and Rothschild [11], who showed that a.a.[1] triangle-free graphs are bipartite. Here we typically have some collection $\mathcal{C}$ (really, a *sequence* of collections $\mathcal{C}_n$) and the goal is to say that some natural, easily understood subcollection $\mathcal{C}_n'$ accounts for a.a. of $\mathcal{C}_n$.

Within this broad context Theorems 1.9 and 1.8 are closest to a short sequence of results beginning nearly forty years ago with the asymptotic solution of Dedekind's

[1] that is, all but a $o(1)$ fraction as $n \to \infty$

Problem by Korshunov [43] (and Sapozhenko [58]). Other results in the sequence give asymptotics for the number of independent sets in Q_n (again Korshunov and Sapozhenko [44] and Sapozhenko [59]) and the numbers of proper 3-colorings of Q_n due to Galvin [22]. (Similar ideas appear in work on certain statistical physics models, e.g. in [24, 53], to mention just the earliest and most recent instances.) The reader familiar with the earlier combinatorial results will note the striking purity of Theorem 1.8, which involves no terms akin to the powers of e in [44, 59, 22] and Theorem 1.9 or the far messier "extra" terms in the Dedekind asymptotic.

1.3 Thresholds

The most important contribution in Chapter 7 is the proof of a conjecture of Talagrand [67] that is a fractional version of the "expectation-threshold" conjecture of Kahn and Kalai [35]. For an increasing family $\mathcal{F}$ on a finite set X, we write (with definitions in Section 7.2) $p_c(\mathcal{F})$, $q_f(\mathcal{F})$ and $\ell(\mathcal{F})$ for the threshold, fractional expectation-threshold, and size of a largest minimal element of $\mathcal{F}$. In this language, our main result is the following.

Theorem 1.10. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q_f(\mathcal{F}) \log \ell(\mathcal{F}).$$

As observed in Section 7.1.1, $q_f(\mathcal{F})$ is a more or less trivial lower bound on $p_c(\mathcal{F})$, and Theorem 1.10 says this bound is never far from the truth. (Apart from the constant K, the upper bound is tight in many of the most interesting cases.)

Thresholds have been a—maybe *the*—central concern of the study of random discrete structures (random graphs and hypergraphs, for example) since its initiation by Erdős and Rényi [12], with much work around identifying thresholds for specific properties (see [3, 31]), though it was not observed until [4] that *every* increasing $\mathcal{F}$ admits a threshold (in the Erdős–Rényi sense; see Section 7.1.1). See also [18] for developments, since [17], on the very interesting question of *sharpness* of thresholds.

Our second main result is Theorem 7.6, which was motivated by work of Frieze and

Sorkin [20] on the "random multi-dimensional assignment problem." The statement, Theorem 7.6, is postponed until we have filled in some background in Section 7.1.3.

Chapter 2

Preliminaries for Chapters 3-6

For a graph G and disjoint $X, Y \subseteq V(G)$, $\nabla(X,Y)$ is the set of edges joining X and Y. We use "$\sim$" for adjacency, $N(X)$ for the set of neighbors of X (with $N(\{x\}) = N_x$), and $d_S(x) = |N_x \cap S|$. We also use $N^2(X)$ for $N(N(X))$. (We will only use these when G is bipartite with X contained in one side of the bipartition, so e.g. will not need to worry about whether $N(X)$ can include vertices of X.)

As usual, $G[S]$ is the subgraph of G induced by $S \subseteq V(G)$, and $[n]$ is shorthand for $\{1, \ldots, n\}$. We use log for $\log_2$. Following a common abuse, we pretend all large numbers are integers, to avoid cluttering the discussions with irrelevant floor and ceiling symbols.

Recall that a *composition* of m is a sequence $(a_1, \ldots, a_s)$ of positive integers with $\sum a_i = m$ (the a_i's are the *parts* of the composition), and that:

Proposition 2.1. *The number of compositions of m is 2^{m-1} and the number with at most $b \leq m/2$ parts is $\sum_{i<b} \binom{m-1}{i} < \exp_2[b \log(em/b)]$.*

2.1 Maximal independent set counts

For the proof of Theorem 1.4 we need the following upper bounds on mis for paths and cycles, given by Z. Füredi [21].

Proposition 2.2. *Let γ (≈ 1.325) be the unique real solution of the equation $1+\gamma = \gamma^3$.*

1. *For P_n, the path with n vertices,*

$$\mathrm{mis}(P_n) \leq 2\gamma^{n-2}.$$

2. For C_n, the cycle with n vertices,

$$\text{mis}(C_n) \le 3\gamma^{n-3}.$$

2.1.1 Algorithm

Here we isolate an algorithmic framework that will play key roles in the proofs of theorems in Chapter 3, and Lemmas 5.1 and 5.3. This is motivated by an idea for counting (ordinary) independent sets due to Sapozhenko [61], but the analyses that we use in Chapters 3 and 5 seem new.

Let G be given. For the algorithm we fix some order "$\prec$" on $V = V(G)$.

[**Algorithm**] Given I an MIS in G, let $X_0 = V(G)$ and repeat for $i = 1, 2, \ldots$:

(1) Let x_i be the first (in $\prec$) vertex of X_{i-1} among those with largest degree in X_{i-1}.

(2) If $x_i \in I$ then let $X_i = X_{i-1} \setminus (\{x_i\} \cup N(x_i))$; otherwise, let $X_i = X_{i-1} \setminus \{x_i\}$. Set $\xi_i = \mathbf{1}_{\{x_i \in I\}}$.

(3) STOP: the stopping rule will vary.

Let $X = X(I)$ be the final X_i and $H = H(I) = G[X]$. Notice that $\xi = \xi(I) = (\xi_1, \xi_2, \ldots)$ encodes a complete description of the run of the algorithm (so we may also write $H = H(\xi)$), including, in particular, the identities of the x_i's; also that

$$\xi(I) \text{ determines } X \text{ and } I \setminus X \tag{2.1}$$

and

$$I \cap X \text{ is an MIS of } H. \tag{2.2}$$

2.2 Hamming cubes

Recall that the n-dimensional hypercube Q_n has vertex set $\{0,1\}^n$, with two vertices adjacent iff they differ in exactly one coordinate. Thus Q_n is n-regular and bipartite with (unique) bipartition $\mathcal{E} \cup \mathcal{O}$, where $\mathcal{E}$ and $\mathcal{O}$ are the sets of even and odd vertices (the *parity* of x being the parity of the number of 1's in x).

We use A, B, C and W for subsets of V and E for $E(Q_n)$. For $x \in V$, x_i is (as usual) the ith coordinate of x, and x^i is the vertex obtained from x by flipping x_i. For any A,

$$A^i = \{x^i : x \in A\},$$

the *vertex-boundary* of A is

$$\partial A = \{x \notin A : x \sim y \text{ for some } y \in A\},$$

and the *edge-boundary* of A is

$$\nabla A = \{(x, y) : x \in A, y \notin A\}.$$

We also use

$$\nabla(A, B) = \{(x, y) : x \in A, y \in B\},$$

$$\nabla_i A = \{(x, x^i) : x \in A, x^i \notin A\},$$

$$\nabla_I A = \cup_{i \in I} \nabla A_i \quad (I \subseteq [n]),$$

and

$$\nabla_i(A, B) = \{(x, x^i) : x \in A, x^i \in B\}.$$

We say C is *a codimension k subcube* if there are $I \subseteq [n]$ of size k and $z \in \{0, 1\}^I$ such that

$$C = \{x \in V : x_i = z_i \text{ for all } i \in I\}.$$

2.2.1 k-components

For a graph G and positive integer k, say $u, v \in V = V(G)$ are *k-linked* if there is a path from u to v of length at most k, and $A \subset V$ is *k-linked* if for any $u, v \in A$, there are vertices $u = u_0, u_1, \ldots, u_l = v$ in A such that u_{i-1}, u_i are k-linked for each $i \in [l]$. Then for $A \subseteq V$ the *k-components* of A are its maximal k-linked subsets. (So we use "component" for a set of vertices rather than a subgraph.) In what follows we will only be interested in $k = 2$, and use i_A for the number of 2-components of A. Notice that

distinct 2-components of A have disjoint neighborhoods.

In particular, when the given graph is Q_n, we use the next lemma in bounding the numbers of certain types of 2-linked sets in Q_n. It follows from the fact (see e.g. [42, p. 396, Ex.11]) that the infinite Δ-branching rooted tree contains precisely

$$\frac{\binom{\Delta m}{m}}{(\Delta-1)m+1} \leq (e\Delta)^{m-1}$$

rooted subtrees with m vertices.

Lemma 2.3. *If G is a graph with maximum degree Δ, then the number of m-vertex subsets of $V(G)$ which contain a fixed vertex and induce a connected subgraph is at most $(e\Delta)^m$.*

Proposition 2.4 ([22], Lemma 1.6)**.** *For each fixed k, the number of k-linked subsets of $V(Q_n)$ of size x containing some specified vertex is at most $2^{O(x\log n)}$.*

Proposition 2.5. *For any $Y \subseteq \mathcal{E}$ and $x, b \in \mathbb{Z}^+$ with $b \leq |Y|/2$, the number of possibilities for an $X \subseteq \mathcal{E}$ with $|X| = x$, $i_X \leq b$ and each 2-component of X meeting Y is at most*

$$\binom{|Y|}{b} n^{O(x)} \tag{2.3}$$

(and similarly with $\mathcal{E}$ replaced by $\mathcal{O}$).

Proof. The number of possibilities for the (say ordered, though this overcounts) list of sizes, say $x_1, \ldots, x_t$, of the 2-components of X is at most the number of compositions of x, so at most 2^{x-1} (see Proposition 2.1). Given this list—so also i_X—the number of ways to choose "roots" in Y for the 2-components is at most $\binom{|Y|}{i_X} \leq \binom{|Y|}{b}$, and then Lemma 2.3 bounds the number of ways to complete the 2-components by $(en^2)^{\sum x_i} = n^{O(x)}$, which absorbs the initial 2^{x-1}. □

2.2.2 Isoperimetry

As is common in this area, we will need to know a little about isoperimetric behavior of small subsets of Q_n. See e.g. [44, Lemma 1.3] or [23, Claim 2.5] for the following proposition.

Proposition 2.6. *For $A \subseteq \mathcal{E}$ with $|A| \le N/4$,*

$$\frac{|N(A)| - |A|}{|N(A)|} = \Omega(1/\sqrt{n}).$$

Lemma 2.7. *For A a subset of $\mathcal{E}$ or $\mathcal{O}$ and $k = n^{o(1)}$,*

$$\text{if } a := |A| = n^k, \text{ then } |N(A)| > (1 - o(1))(an/k).$$

Proof. This is similar to [22, Lemma 6.1]—and a routine application of [45]—so we will be brief, referring to [22] for some elaboration.

It is of course enough to consider $A \subseteq \mathcal{E}$. By the main theorem of [45] (see [22, Lemma 1.10]) we may assume A is an even Hamming ball; that is,

$$B(v, l) \subseteq A \subseteq B(v, l + 2), \tag{2.4}$$

for some v and l with $l \equiv |v| \pmod 2$ (where $|v| = \sum v_i$ and, with ρ denoting distance, $B(v, r) = \{w \in \mathcal{E} : \rho(v, w) \le r\}$ is the even Hamming ball of radius r about v). We just discuss $v \in \mathcal{E}$, in which case we may assume $v = \underline{0}$.

Elementary calculations show that (assuming a is as in the lemma) the l in (2.4) is asymptotic to k (since for $l = n^{o(1)}$, $|\binom{[n]}{\le l} \cap \mathcal{E}| = n^{l-o(l)}$). It's then easy to see that each of $|N(B(\underline{0}, l))|, |N(B(\underline{0}, l + 2))|$ is asymptotic to an/k, and the lemma follows. □

2.2.3 Sapozhenko and Galvin

Finally we recall what we need from the aforementioned results of Sapozhenko and Galvin (adapted to present purposes).

For $A \subseteq V$, the *closure* of A is $[A] = \{x \in V : N(x) \subseteq N(A)\}$ and A is *closed* if $A = [A]$. Given A (always a subset of $\mathcal{E}$ or $\mathcal{O}$), we use G and B for $N(A)$ and $B(A)$ $(:= \{y : N(y) \subseteq A\})$.

Let

$$\mathcal{G}(g) = \{A \subseteq \mathcal{E} \text{ 2-linked } : |G| = g\} \tag{2.5}$$

and

$$\mathcal{H}(g, b) = \{A \subseteq \mathcal{E} \text{ 2-linked } : |G| = g, |B| = b\}.$$

The first of our lemmas here is from [57] but we refer to the more accessible [23, Lemma 3.1]:

Lemma 2.8. *For any* $\gamma < 2$ *and each* $g \in [n^4, \gamma^n]$,

$$|\mathcal{G}(g)| \leq 2^{g-\Omega(g/\log n)}.$$

Lemma 2.9 ([22], Lemma 7.1)**.** *For any* $\gamma < 2$ *and each* $g, b \leq \gamma^n$,

$$|\mathcal{H}(g,b)| < 2^n 2^{g-b-\Omega(g/\log n)}.$$

The next two lemmas are special cases of Lemmas 5.3-5.5 in [23].

Lemma 2.10. *For* g *as in Lemma 2.8 and* $\mathcal{G} = \mathcal{G}(g)$, *there are* $\mathcal{W} = \mathcal{W}(g) \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ *with*

$$|\mathcal{W}| = 2^{O(g\log^2 n/n)}$$

and $\varphi = \varphi_g : \mathcal{G} \to \mathcal{W}$ *such that for each* $A \in \mathcal{G}$, $(S, F) := \varphi(A)$ *satisfies:*

(a) $S \supseteq [A]$, $F \subseteq G$;

(b) $d_F(u) \geq n - n/\log n \ \forall u \in S$.

In the next lemma, we take $t = |G| - |[A]|$.

Lemma 2.11. *For* $g \in [n^4, N/4]$ *and* $\mathcal{G} = \mathcal{G}(a, g) := \{A \subseteq \mathcal{E} : A$ *is 2-linked and closed,* $|A| = a$ *and* $|G| = g\}$, *there are* $\mathcal{W} = \mathcal{W}(a, g) \subseteq 2^{\mathcal{E}} \times 2^{\mathcal{O}}$ *with*

$$|\mathcal{W}| = 2^{O(t\log^2 n/\sqrt{n})}$$

and $\varphi = \varphi_{a,g} : \mathcal{G} \to \mathcal{W}$ *such that for each* $A \in \mathcal{G}$, $(S, F) := \varphi(A)$ *satisfies:*

(a) $S \supseteq A \ (= [A]), F \subseteq G$;

(b) $d_F(u) \geq n - \sqrt{n}/\log n \quad \forall u \in S$;

(c) $|S| \leq |F| + O(t/(\sqrt{n}\log n))$.

2.3 Entropy

We next briefly recall relevant entropy background; see e.g. [50] for a less hurried introduction.

Let X, Y be discrete random variables. The binary entropy of X is

$$H(X) = \sum_x p(x) \log \frac{1}{p(x)},$$

where $p(x) = \mathbb{P}(X = x)$ (and, recall, log is $\log_2$). The *conditional entropy of X given Y* is

$$H(X|Y) = \sum_y p(y) \sum_x p(x|y) \log \frac{1}{p(x|y)} \tag{2.6}$$

(where $p(x|y) = \mathbb{P}(X = x|Y = y)$).

The next lemma lists a few basic properties.

Lemma 2.12.

(a) $H(X) \leq \log |Range(X)|$, *with equality iff X is uniform from its range;*

(b) $H(X, Y) = H(X) + H(Y|X)$*;*

(c) $H(X_1, \ldots, X_n|Y) \leq \sum H(X_i|Y)$ *(note $(X_1, \ldots, X_n)$ is a discrete random variable);*

(d) *if Z is determined by Y, then $H(X|Y) \leq H(X|Z)$.*

We also need the following version of *Shearer's Lemma* [5]. (This statement is more general than the original, but is easily extracted from the proof in [5].)

Lemma 2.13. *If $X = (X_1, , \ldots, , X_k)$ is a random vector and $\alpha : 2^{[k]} \to \mathbb{R}^+$ satisfies*

$$\sum_{A \ni i} \alpha_A = 1 \quad \forall i \in [k], \tag{2.7}$$

then

$$H(X) \leq \sum_{A \subseteq [k]} \alpha_A H(X_A), \tag{2.8}$$

where $X_A = (X_i : i \in A)$.

Finally, we will need the following standard fact (see e.g. Lemma 16.19 in [15]; this is also implied by Lemma 2.13 with α_A equal to 1 if $|A| = 1$ and zero otherwise).

Proposition 2.14. *For* $k \le \frac{1}{2}n$*,*

$$\sum_{i=0}^{k} \binom{n}{i} \le 2^{H(\frac{k}{n})n}.$$

Chapter 3

Stability for maximal independent sets

3.1 Proofs of Theorems 1.3 and 1.4

In this section, I is always a maximal independent set in G. We use [**Algorithm**] in Section 2.1.1 with the following stopping rule:

Terminate the process if $d_{X_i}(x) \le 2$ for all $x \in X_i$.

Recall that $X = X(I)$ is the final X_i produced by [**Algorithm**] and $H = H(I) = G[X]$.

3.1.1 Proof of Theorem 1.3

The argument for Theorem 1.3 goes roughly as follows. By (2.1) and (2.2) we find that

$$\text{mis}(G) \le \sum_{\xi} \text{mis}(H(\xi)) \tag{3.1}$$

(If we restrict the sum to *possible* ξ's—those corresponding to actual I's—then we have equality in (3.1).)

It turns out that running the algorithm for very long is "expensive" in the sense that the loss in $|X|$, and so in possibilities for $I \cap X$, outweighs what is contributed to (3.1) by possibilities for ξ; this limits the number of I's with $t(I)$ large. Similarly, the difference between the bounds in Theorems 1.1 and 1.2 says there are "few" I's for which the triangle-free part of H is large. (Note H, having maximum degree at most two, is a disjoint union of triangles and a triangle-free part, below called R.) But the part of mis(G) corresponding to I's for which both t and R are small must come mainly from counting choices for the restriction of I to the triangles of H, and these are limited by our assumption on itm(G).

To begin with, the following lemma bounds the number of I's with large $t(I)$.

Lemma 3.1. *Let* $\alpha = -\log(4 \cdot 3^{-4/3})$ (≈ 0.113). *For any* $x \in [0,1]$,

$$\log|\{I : t(I) \geq xn\}| \leq (\frac{1}{3}\log 3 - \alpha x + o(1))n. \tag{3.2}$$

Proof. For given t and s, consider I's for which $t(I) = t$ and $s(I) = s$. Note that for each such I, $|X| \leq n - (t+3s)$, so by Theorem 1.1 we have $\text{mis}(H) \leq 3^{(n-(t+3s))/3}$. Also, there are at most $\binom{t}{s}$ possibilities for $\xi(I)$, so by (2.1) and (2.2) we have

$$|\{I : t(I) = t, s(I) = s\}| \leq \binom{t}{s} 3^{(n-(t+3s))/3}, \tag{3.3}$$

so

$$|\{I : t(I) = t\}| \leq \sum_{s=0}^{t} \binom{t}{s} 3^{(n-(t+3s))/3} = 3^{n/3}\alpha_1^t,$$

where $\alpha_1 = 4 \cdot 3^{-4/3}$. Thus,

$$|\{I : t(I) \geq xn\}| \leq 3^{n/3}\alpha_1^{xn}/(1-\alpha_1),$$

yielding (3.2). □

Let $T = T(I)$ be the union of the triangles in X (so the unique maximal induced triangle matching in H), $R = R(I) = H[X \setminus V(T)]$, and $r = r(I) = |V(R)|$. Note that there are no edges between $V(T)$ and $V(R)$, since H has maximum degree at most 2, so

$$\text{mis}(H) = \text{mis}(T)\text{mis}(R). \tag{3.4}$$

Note also that R is triangle-free, so

$$\log \text{mis}(R) \leq r/2 \tag{3.5}$$

by Theorem 1.2. Now, the following lemma bounds the number of I's with large r.

Lemma 3.2. *Let* $\beta = -\log(2^{1/2}3^{-1/3})$ (≈ 0.028). *For any* $y \in [0,1]$,

$$\log|\{I : r(I) \geq yn\}| \leq (\frac{1}{3}\log 3 - \beta y + o(1))n. \tag{3.6}$$

Proof. By (3.4) and (3.5), we have

$$|\{I : r(I) = r, t(I) = t, s(I) = s\}| \leq \binom{t}{s} 3^{(n-(t+3s+r))/3} 2^{r/2},$$

so

$$\begin{aligned}|\{I : r(I) = r\}| &\leq \sum_{t=0}^{n}\sum_{s=0}^{t}\binom{t}{s}3^{(n-(t+3s+r))/3}2^{r/2}\\ &\leq 3^{n/3}\beta_1^r/(1-\alpha_1),\end{aligned}$$

where $\alpha_1 = 4\cdot 3^{-4/3}$ (as in Lemma 3.1) and $\beta_1 = 2^{1/2}3^{-1/3}$. Thus,

$$|\{I : r(I) \geq yn\}| \leq 3^{n/3}\beta_1^{yn}/((1-\alpha_1)(1-\beta_1)),$$

which gives (3.6). □

Lemma 3.3. *If* $\mathrm{itm}(G) < (1-\varepsilon)n/3$ *then for any* $x, y \in [0,1]$,

$$\log|\{I : t(I) < xn, r(I) < yn\}| \leq ((1-\varepsilon)\frac{1}{3}\log 3 + x + y/2 + o(1))n. \tag{3.7}$$

Proof. For any I, with $H = H(I)$ and $r = r(I)$, we have (using (3.4), (3.5) and $|V(T(I))| = 3\mathrm{itm}(H) < (1-\varepsilon)n$)

$$\mathrm{mis}(H) \leq 3^{(1-\varepsilon)n/3}2^{r/2}.$$

Therefore,

$$|\{I : t(I) = t, r(I) = r\}| \leq 2^t 3^{(1-\varepsilon)n/3}2^{r/2},$$

so

$$\begin{aligned}|\{I : t(I) < xn, r(I) < yn\}| &\leq \sum_{t<xn}\sum_{r<yn}3^{(1-\varepsilon)n/3}2^{r/2+t}\\ &\leq 3^{(1-\varepsilon)n/3}\cdot 2^{xn+1}\cdot(\sqrt{2}-1)^{-1}2^{(yn+1)/2},\end{aligned}$$

giving (3.7). □

Proof of Theorem 1.3. This is now just a matter of combining the above bounds. With $\delta_1 = \varepsilon\alpha/8$ and $\delta_2 = \varepsilon\beta/4$, Lemmas 3.1 - 3.3 give (respectively)

$$\log|\{I : t(I) \geq \delta_1 n/\alpha\}| \leq (\frac{1}{3}\log 3 - \delta_1 + o(1))n,$$

$$\log|\{I : r(I) \geq \delta_2 n/\beta\}| \leq (\frac{1}{3}\log 3 - \delta_2 + o(1))n$$

and (using $\frac{1}{3}\log 3 > 1/2$)

$$\log|\{I : t(I) < \delta_1 n/\alpha, r(I) < \delta_2 n/\beta\}| \le (\frac{1}{3}\log 3 - \varepsilon/4 + o(1))n.$$

Thus, with $\delta = \min\{\delta_1, \delta_2, \varepsilon/4\}$ $(= \Omega(\varepsilon))$, we have

$$\log \text{mis}(G) \le (\frac{1}{3}\log 3 - \delta + o(1))n.$$

□

3.1.2 Proof of Theorem 1.4

We first give the slightly stronger version of Theorem 1.4 mentioned in Chapter 1. For $I \subseteq V(G)$, write $m(I) = m_G(I)$ for the maximum size of an induced matching M satisfying

- each edge of M meets I and
- there are no edges joining $V(M)$ (the set of vertices covered by M) and $I \setminus V(M)$.

Given G we now write $\mathcal{I} = \mathcal{I}(G)$ for the collection of maximal independent sets of G and set

$$\mathcal{I}_\varepsilon = \mathcal{I}(G, \varepsilon) = \{I \in \mathcal{I}(G) : m(I) < (1-\varepsilon)n/2\}$$

and $\text{mis}(G, \varepsilon) = |\mathcal{I}_\varepsilon|$.

Theorem 3.4. *For any $\varepsilon > 0$ there is a $\delta = \Omega(\varepsilon)$ such that for any n-vertex, triangle free G,*

$$\log \text{mis}(G, \varepsilon) < (1-\delta)n/2.$$

(We have omitted the corresponding strengthening of Theorem 1.3.)

As mentioned earlier, the argument for Theorem 3.4 is similar to the one in Section 3.1.1, so we will try to be brief. We again start from the algorithm in Section 2.1.1, and continue to use the notation (X, H etc.) defined in the paragraph following the algorithm's description. (For most of this we just need $I \in \mathcal{I}$; the role of $\mathcal{I}_\varepsilon$ will appear in Lemma 3.8.)

Lemma 3.5. *Let* $\alpha = -\log(\frac{1}{\sqrt{2}} + \frac{1}{4})$ (≈ 0.063). *For any* $x \in [0,1]$,

$$\log |\{I \in \mathcal{I} : t(I) \geq xn\}| \leq (\frac{1}{2} - \alpha x + o(1))n. \tag{3.8}$$

Proof. Arguing as for (3.3) in Section 3.1.1, we obtain

$$|\{I \in \mathcal{I} : t(I) = t, s(I) = s\}| \leq \binom{t}{s} 2^{(n-(t+3s))/2}, \tag{3.9}$$

where we used $\text{mis}(H) \leq 2^{(n-(t+3s))/2}$, as given by Theorem 1.2 (since G is triangle-free). Thus

$$|\{I \in \mathcal{I} : t(I) = t\}| \leq \sum_{s=0}^{t} \binom{t}{s} 2^{(n-(t+3s))/2} = 2^{n/2}\alpha_1^t,$$

where $\alpha_1 = \frac{1}{\sqrt{2}} + \frac{1}{4}$, and

$$|\{I \in \mathcal{I} : t(I) \geq xn\}| \leq 2^{n/2}\alpha_1^{xn}/(1-\alpha_1),$$

yielding (3.8). □

Say an edge vw of H is *isolated* if $H[\{v,w\}]$ is a component of H. Let $M = M(I)$ be the set of isolated edges in H, $R = R(I) = H[X \setminus V(M)]$, and $r = r(I) = |V(R)|$. Notice that M satisfies the two •'s from the definition of $m(I)$ (the first by maximality of I, the second by the definition of M and the fact that there are no edges joining X and $I \setminus X$); so if $I \in \mathcal{I}_\varepsilon$ then $|M| < (1-\varepsilon)n/2$. Also, since there are no edges between $V(M)$ and $V(R)$,

$$\text{mis}(H) = \text{mis}(M)\text{mis}(R). \tag{3.10}$$

Note that R is triangle-free, so is a vertex-disjoint union of isolated vertices, cycles with at least 4 vertices, and paths with at least 3 vertices. Combining this with Proposition 2.2, we obtain an upper bound for $\text{mis}(R)$. (Recall that $\gamma \approx 1.325$ was defined in Proposition 2.2.)

Lemma 3.6. *With* R *and* r *as above,* $\text{mis}(R) \leq (3\gamma)^{r/4}$.

Proof. Let l_p (resp. l_c) be the number of vertices in the union of all paths (resp. cycles) in R. Clearly $l_p + l_c \leq r$, while the number of paths (resp. cycles) in R is at most $l_p/3$

(resp. $l_c/4$). Thus

$$\begin{aligned} \text{mis}(R) &\leq (2/\gamma^2)^{l_p/3}(3/\gamma^3)^{l_c/4}\gamma^r \\ &< (3/\gamma^3)^{r/4}\gamma^r = (3\gamma)^{r/4}, \end{aligned}$$

where the first inequality is given by Proposition 2.2 and the second follows from the fact that $(2\gamma^{-2})^{1/3} < (3\gamma^{-3})^{1/4}$. □

Lemma 3.7. *Let* $\beta = -\log(2^{-1/2}(3\gamma)^{1/4})$ (≈ 0.0023). *For any* $y \in [0,1]$,

$$\log|\{I \in \mathcal{I} : r(I) \geq yn\}| \leq (\frac{1}{2} - \beta y + o(1))n. \tag{3.11}$$

Proof. By (3.10) and Lemma 3.6,

$$|\{I \in \mathcal{I} : r(I) = r, t(I) = t, s(I) = s\}| \leq \binom{t}{s} 2^{(n-(t+3s+r))/2}(3\gamma)^{r/4},$$

and summing this over t and s gives

$$|\{I \in \mathcal{I} : r(I) = r\}| \leq 2^{n/2}\beta_1^r/(1-\alpha_1),$$

where $\alpha_1 = \frac{1}{\sqrt{2}} + \frac{1}{4}$ (as in Lemma 3.5) and $\beta_1 = 2^{-1/2}(3\gamma)^{1/4}$. Thus,

$$|\{I \in \mathcal{I} : r(I) \geq yn\}| \leq 2^{n/2}\beta_1^{yn}/((1-\alpha_1)(1-\beta_1)),$$

which gives (3.11). □

Lemma 3.8. *For any* $x, y \in [0,1]$,

$$\log|\{I \in \mathcal{I}_\varepsilon : t(I) < xn, r(I) < yn\}| \leq ((1-\varepsilon)/2 + x + (\log(3\gamma)/4)y + o(1))n. \tag{3.12}$$

Proof. As in the proof of Lemma 3.3 (now using $|M| < (1-\varepsilon)n/2$),

$$\text{mis}(H) \leq 2^{(1-\varepsilon)n/2}(3\gamma)^{r/4}$$

for any $I \in \mathcal{I}_\varepsilon$ with $r(I) = r$. Therefore,

$$|\{I \in \mathcal{I}_\varepsilon : t(I) = t, r(I) = r\}| \leq 2^t 2^{(1-\varepsilon)n/2}(3\gamma)^{r/4},$$

and summing over the relevant t's and r's gives

$$|\{I \in \mathcal{I}_\varepsilon : t(I) < xn, r(I) < yn\}| \leq 2^{(1-\varepsilon)n/2} \cdot 2^{xn+1} \cdot ((3\gamma)^{1/4} - 1)^{-1}(3\gamma)^{(yn+1)/4};$$

so we have (3.12). □

Proof of Theorem 3.4. With $\delta_1 = \varepsilon\alpha/8$ and $\delta_2 = \varepsilon\beta/(2\log(3\gamma))$, Lemmas 3.5, 3.7 and 3.8 give (respectively)

$$\log|\{I \in \mathcal{I} : t(I) \geq \delta_1 n/\alpha\}| \leq (\frac{1}{2} - \delta_1 + o(1))n,$$

$$\log|\{I \in \mathcal{I} : r(I) \geq \delta_2 n/\beta\}| \leq (\frac{1}{2} - \delta_2 + o(1))n,$$

and

$$\log|\{I \in \mathcal{I}_\varepsilon : t(I) < \delta_1 n/\alpha, r(I) < \delta_2 n/\beta\}| \leq (\frac{1}{2} - \varepsilon/4 + o(1))n.$$

Thus, with $\delta = \min\{\delta_1, \delta_2, \varepsilon/4\}$, we obtain

$$\log \text{mis}(G, \varepsilon) \leq (\frac{1}{2} - \delta + o(1))n.$$

□

3.2 Proof of Theorem 1.5

For a bipartite graph G on $X \cup Y$, say $X' \subseteq X$ is *irredundant* if $\forall x \in X'$, $N(x) \nsubseteq N(X' \setminus \{x\})$. (So for this discussion "irredundant" sets are always subsets of X.) Denote the number of irredundant sets in G by $\text{irr}(G)$.

Proposition 3.9. *For any G as above,* $\text{mis}(G) \leq \text{irr}(G)$.

Proof. This follows from the observation that for each maximal independent set I there is an irredundant set $J \subseteq I \cap X$ with $N(J) = N(I \cap X)$ $(= Y \setminus I)$; namely, this is true whenever $J \subseteq I \cap X$ is minimal with $N(J) = N(I \cap X)$. □

Thus the following statement implies Theorem 1.5.

Theorem 3.10. *For any $\varepsilon > 0$, there is a $\delta = \delta(\varepsilon) = 2^{-O(1/\varepsilon)}$ such that for a bipartite graph G on $X \cup Y$ with $|X| = n$ and $|Y| = 2n$, if* $\text{im}(G) < (1 - \varepsilon)n$ *then* $\log \text{irr}(G) < (1 - \delta)n$.

For the rest of this section, G is as in Theorem 3.10.

3.2.1 Proof

The algorithm we use for Theorem 3.10 is slightly different from the one in section 2.1.1. In what follows, I is always an irredundant set (thus $I \subseteq X$).

Algorithm Let $X_0 = X$, $Y_0 = Y$ and $M = M_\varepsilon = 12/\varepsilon$. Fix an order "$\prec$" on X. For a given I, repeat for $i = 1, 2, \ldots$:

1. Let x_i be the first vertex of X_{i-1} in $\prec$ among those with largest degree in X_{i-1}.

2. If $x_i \in I$ then set $Y_i = Y_{i-1} \setminus N(x_i)$; otherwise, set $Y_i = Y_{i-1}$. In either case, set $X_i = X_{i-1} \setminus \{x_i\}$.

3. Terminate the process if $d_{Y_i}(x) < M$ for all $x \in X_i$.

Let $X = X(I) = X_t$ and $Y^* = Y^*(I) = Y_t$ be the final X_i and Y_i, respectively. Set $t = t(I)$ and $H = H(I) = G[X \cup Y^*]$. As in section 2.1.1, define $\xi = \xi(I) = (\xi_1, \xi_2, \cdots, \xi_t)$ by $\xi_i := \mathbf{1}_{\{x_i \in I\}}$, and let $|\xi|$ be the length of ξ (so $|\xi(I)| = t(I)$). Finally, let $s = s(I) = |\mathrm{supp}(\xi)|$ and define $\psi = \psi(I) = I \cap X$.

Notice that I is determined by (ξ, ψ), namely (as earlier) $I \setminus X$ is determined by ξ (and $I \cap X = \psi$).

Consider a random (uniform) irredundant set $\mathbf{I}$. Our various parameters $(\xi, \psi, \ldots)$ are then random variables, which will be denoted by $\boldsymbol{\xi}$ and so on. Since each of $\mathbf{I}$ and $(\boldsymbol{\xi}, \boldsymbol{\psi})$ determines the other and $\boldsymbol{\xi}$ determines $\mathbf{t}$, we have (using Lemma 2.12)

$$\begin{aligned} H(\mathbf{I}) &= H(\boldsymbol{\xi}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}) \\ &= H(\mathbf{t}) + H(\boldsymbol{\xi}|\mathbf{t}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}) \\ &\leq \log n + H(\boldsymbol{\xi}|\mathbf{t}) + H(\boldsymbol{\psi}|\boldsymbol{\xi}). \end{aligned} \tag{3.13}$$

Notice that, by Lemma 2.12 (a),

$$H(\boldsymbol{\xi}|\mathbf{t} = t) \leq t$$

for any t and

$$H(\boldsymbol{\psi}|\boldsymbol{\xi} = \xi) \leq n - |\xi| \ (= n - t) \tag{3.14}$$

for any ξ. Thus the sum of the last two terms in (3.13) is at most

$$\sum_t \mathbb{P}(\mathbf{t}=t)\left[H(\boldsymbol{\xi}|\mathbf{t}=t)+\sum_{|\xi|=t}\mathbb{P}(\boldsymbol{\xi}=\xi|\mathbf{t}=t)H(\boldsymbol{\psi}|\boldsymbol{\xi}=\xi)\right]\leq n,$$

and we would like to somewhat improve these bounds. (Since we aim for $H(\mathbf{I}) < n-\Omega(n)$, the $\log n$ in (3.13) is irrelevant.) The next lemma, giving such a gain in (3.14) when t is small, is our main point.

Lemma 3.11. *For any ξ with $|\xi| = t < \varepsilon n/2$,*

$$H(\boldsymbol{\psi}|\boldsymbol{\xi}=\xi)\leq n-t-\vartheta n,$$

where $\vartheta = \vartheta(\varepsilon) = 2^{-O(1/\varepsilon)}$.

Proof. Given ξ as in the Lemma, set

$$\tilde{X}=\tilde{X}(\xi)=\{x\in X : N_{Y^*}(x)\subseteq N_{Y^*}(X\setminus\{x\})\}.$$

We have

$$(1-\varepsilon)n > \mathrm{im}(G)\geq \mathrm{im}(H)\geq n-t-|\tilde{X}|,$$

where the last inequality holds since for each $x\in X\setminus\tilde{X}$ there is some $y_x\in Y^*$ with $N_X(y_x)=\{x\}$, and $\{(x,y_x) : x\in X\setminus\tilde{X}\}$ is an induced matching of H of size $|X\setminus\tilde{X}| = n-t-|\tilde{X}|$. Thus

$$|\tilde{X}| > \varepsilon n - t > \varepsilon n/2. \tag{3.15}$$

For each $x\in\tilde{X}$ fix some $Z_x\subseteq X\setminus\{x\}$ such that

$$N_{Y^*}(x)\subseteq N_{Y^*}(Z_x), \tag{3.16}$$

$$|Z_x| < M \tag{3.17}$$

and

$$\forall z\in X \quad |\{x\in\tilde{X} : z\in Z_x\}| < 2M. \tag{3.18}$$

To see that we can do this: For each $y\in N_{Y^*}(\tilde{X})$ let Π_y be a partition of $N_X(y)$ into blocks of size 2 or 3. (Note $y\in N_{Y^*}(\tilde{X})$ implies $d_X(y)\geq 2$.) Then to form Z_x, for each $y\in N_{Y^*}(x)$ choose one $x'\neq x$ from the block of Π_y containing x and take $x'\in Z_x$.

Note that each $x \in X$ has degree less than M in H (see step 3 of the algorithm), so we have (3.17) and (3.18) (and (3.16) is clear).

Let $W_x = Z_x \cup \{x\}$ ($x \in \tilde{X}$), and $\psi_A = \psi \cap A$ for any $A \subseteq X$. Note that for each $x \in X$,

$$\begin{aligned} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) &\leq \log[2^{|W_x|} - 1] \\ &= |W_x| + \log(1 - 2^{-|W_x|}) \\ &< |W_x| - 2^{-M} \log e. \end{aligned} \tag{3.19}$$

(The first inequality follows from irredundancy: we cannot have $\psi_{W_x} = W_x$.)

Now aiming to use Lemma 2.13, form $\alpha : 2^X \to \mathbb{R}_{\geq 0}$ by assigning weight $1/(2M)$ to each W_x (thus assigning each set weight some multiple of $1/(2M)$, with the total weight of the sets containing any given x' at most 1 by (3.18)) and supplementing with weights on the singletons to get to (2.7). Then by Lemma 2.13,

$$\begin{aligned} H(\boldsymbol{\psi}|\boldsymbol{\xi} = \xi) &\leq \sum_{A \subseteq X} \alpha_A H(\boldsymbol{\psi}_A|\boldsymbol{\xi} = \xi) \\ &= \sum_{x \in \tilde{X}} \alpha_{W_x} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) + \sum_{x \in X} \alpha_{\{x\}} H(\boldsymbol{\psi}_{\{x\}}|\boldsymbol{\xi} = \xi). \end{aligned} \tag{3.20}$$

Now (3.19) and the fact that α assigns total weight $|\tilde{X}|/(2M)$ to the W_x's give

$$\sum_{x \in \tilde{X}} \alpha_{W_x} H(\boldsymbol{\psi}_{W_x}|\boldsymbol{\xi} = \xi) < \sum_{x \in \tilde{X}} \alpha_{W_x} |W_x| - |\tilde{X}|(2M2^M)^{-1} \log e,$$

while the second sum in (3.20) is at most $\sum_{x \in X} \alpha_{\{x\}}$ (since $H(\boldsymbol{\psi}_{\{x\}}|\boldsymbol{\xi} = \xi) \leq 1$). Thus the entire bound in (3.20) is at most

$$\begin{aligned} \sum_{x \in \tilde{X}} \alpha_{W_x} |W_x| + \sum_{x \in X} \alpha_{\{x\}} - |\tilde{X}|(2M2^M)^{-1} \log e &= |X| - |\tilde{X}|(2M2^M)^{-1} \log e \\ &< n - t - \vartheta n, \end{aligned}$$

where $\vartheta = (\varepsilon/2)(2M2^M)^{-1} \log e = 2^{-O(1/\varepsilon)}$ (see (3.15)) and we use

$$\sum_{x \in \tilde{X}} \alpha_{W_x} |W_x| + \sum_{x \in X} \alpha_{\{x\}} = \sum_{x \in X} \sum_{A \ni x} \alpha_A = |X|.$$

□

Corollary 3.12. *Let $\zeta = \mathbb{P}(\mathbf{t} < \varepsilon n/2)$. Then with ϑ as in Lemma 3.11,*

$$H(\boldsymbol{\psi}|\boldsymbol{\xi}) \leq n - \mathbb{E}\mathbf{t} - \zeta\vartheta n.$$

Proof. Using Lemma 3.11 and (3.14) we have

$$\begin{aligned} H(\boldsymbol{\psi}|\boldsymbol{\xi}) &= \sum_t \sum_{|\xi|=t} \mathbb{P}(\boldsymbol{\xi} = \xi) H(\boldsymbol{\psi}|\boldsymbol{\xi} = \xi) \\ &\leq \sum_{t<\varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(n - t - \vartheta n) + \sum_{t \geq \varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(n - t) \\ &= n - \mathbb{E}\mathbf{t} - \zeta\vartheta n. \end{aligned}$$

□

The gain for larger t is easier. Noting that

$$\mathbf{s} \leq s_0 := 2n/M = \varepsilon n/6,$$

setting $H(1/3) = 1 - \gamma$ and using Proposition 2.14, we have, for any $t \geq \varepsilon n/2$,

$$H(\boldsymbol{\xi}|\mathbf{t} = t) \leq \log \sum_{s \leq s_0} \binom{t}{s} \leq H(1/3)t = (1 - \gamma)t,$$

whence (recall $\zeta = \mathbb{P}(\mathbf{t} < \varepsilon n/2)$)

$$\begin{aligned} H(\boldsymbol{\xi}|\mathbf{t}) &\leq \sum_{t<\varepsilon n/2} \mathbb{P}(\mathbf{t} = t)t + \sum_{t \geq \varepsilon n/2} \mathbb{P}(\mathbf{t} = t)(1 - \gamma)t \\ &\leq \mathbb{E}\mathbf{t} - (1 - \zeta)\gamma\varepsilon n/2. \end{aligned}$$

Finally, combining this with (3.13) and Corollary 3.12 yields

$$\begin{aligned} H(\mathbf{I}) &\leq \log n + n - [\zeta\vartheta + (1 - \zeta)\gamma\varepsilon/2]n \\ &\leq \log n + n - \vartheta n \end{aligned}$$

(since the ϑ produced in Lemma 3.11 is much smaller than $\gamma\varepsilon/2$), proving Theorem 3.10.

3.2.2 Tightness

Define a bipartite graph B_m on $X \cup Y = [m] \cup [2m]$ (disjoint copies, of course) as follows.

1. If $x \in X$ and $x \leq m-1$, then $x \sim y$ iff $y = x$ or $y = m-1+x$.

2. If $x = m \in X$, then $x \sim y$ iff $m \leq y \leq 2m-2$.

It is easy to see that $\text{im}(B_m) = m-1$, and $\text{mis}(B_m) = 2^m - 1$.

Now, for $\varepsilon > 0$ and n with $1/\varepsilon$ and εn integers, let G be the union of εn disjoint copies of $B_{1/\varepsilon}$. Then G is bipartite on $[n] \cup [2n]$, $\text{im}(G) = (1-\varepsilon)n$, and $\text{mis}(G) = (2^{1/\varepsilon} - 1)^{\varepsilon n}$. So,

$$
\begin{aligned}
\log \text{mis}(G) &= \varepsilon n \log(2^{1/\varepsilon} - 1) \\
&= \varepsilon n (\frac{1}{\varepsilon} + \log(1 - 2^{-1/\varepsilon})) \\
&= n(1 - 2^{-1/\varepsilon}\varepsilon \log e + O(2^{-2/\varepsilon}))
\end{aligned}
$$

(where the implied constant does not depend on ε).

Chapter 4

An isoperimetric inequality for the Hamming cube and some consequences

We first introduce some applications of Theorem 1.6. We use V for $V(Q_n)$, and $x = a \pm b$ means $a - b \leq x \leq a + b$. Recall that $\beta = \log_2(3/2)(\approx .585)$.

4.1 Applications

4.1.1 First application: separating the cube

Isoperimetric inequalities beginning with Harper [26] (and for edge boundaries also Lindsey [47]) give lower bounds in terms of $|A|$ on the sizes of ∂A and ∇A; e.g.

$$|\nabla A| \geq |A| \log_2(2^n/|A|), \tag{4.1}$$

with equality iff A is a subcube. We are interested in hybrid versions of these. In what follows we assume (A, B, W) is a partition of V, with W thought of as small. The next conjecture is a simple illustration of what we have in mind, followed by something general.

Conjecture 4.1. *There is a fixed K such that if $\mu(A) = 1/2$, then*

$$|\nabla(A, B)| + K\sqrt{n}\ |W| \geq 2^{n-1}.$$

Results of Margulis [49] and Talagrand [62] (motivated by [49]) imply tradeoffs between $|\nabla A|$ and $|\partial A|$, but don't seem to help here. Theorem 1.6 implies a weaker version of Conjecture 4.1:

Corollary 4.2. *For A, B, W as in Conjecture 4.1,* $|\nabla(A, B)| + n^{\beta}|W| \geq 2^{n-1}$.

4.1.2 Second application: stability for "almost" isoperimetric subsets

A simple (though now suboptimal) "stability" statement for edge boundaries says:

Theorem 4.3. *For a fixed k, if $|A| = 2^{n-k}$ and $|\nabla A| < (1+\varepsilon)|A|\log_2(2^n/|A|)$, then there is a subcube C with $\mu(C\Delta A) = O(\varepsilon)$ (where the implied constant depends on k).*

This was proved for $k = 1$ by Friedgut, Kalai and Naor [19]; then for $k = 2, 3$ by Bollobás, Leader and Riordan, who conjectured the general statement (see [8]); and finally in full by Ellis [8]. These all based on Fourier analysis; e.g. at the heart of [8] is Talagrand's extension [63] of [36]. Even stronger, very recent results of Ellis, Keevash and Lifshitz [9] are more elementary but rather involved.

Notice that if A is (sufficiently) close to a codimension k subcube then there is an $I \subseteq [n]$ of size k with $\nabla A \approx \nabla_I A$. In fact the implication goes both ways; this follows (more or less) from Theorem 4.3, but is also easy without that machine:

Proposition 4.4. *Assume $|A| = (1 \pm \varepsilon)2^{n-k}$ and*

$$|\nabla A \setminus \nabla_I A| \le \varepsilon|A|,$$

where I is a k-subset of $[n]$. Then there is a (codimension k) subcube C with $|A\Delta C| = O(\varepsilon)|A|$ (where the implied constant depends on k).

The original motivation for Theorem 1.6 arose in connection with our efforts to prove Theorem 1.8 whose proof is completed in Chapter 5. What it needed from isoperimetry was a variant of Theorem 4.3—really, just of the original result of [19]—of the following type.

If (A, B, W) is a partition of V with $\mu(A), \mu(B) \approx 1/2$ (so W is "small") and
$$|\nabla(A, B)| \approx 2^{n-1}, \text{ then}$$
$$\nabla A \approx \nabla_i A \quad \text{for some } i.$$

Of course this depends on quantification; e.g. it can fail with $\mu(W)$ as small as $\Theta(n^{-1/2})$ (let W consist of strings of weight $\lfloor n/2 \rfloor$). Note also that here the full edge boundary of A need *not* be small, since there is no restriction (beyond $n|W|$) on $|\nabla(A, W)|$.

The following consequence of Theorem 1.6 is a (limited) statement of the desired type, the case $k = 1$ of which suffices for Theorem 1.8. (Recall $\beta = \log_2(3/2)$.)

Theorem 4.5. *For $k \in \{1, 2\}$ the following holds. Suppose (A, B, W) is a partition of V with $\mu(A) = (1 \pm \varepsilon)2^{-k}$, $\mu(W) \leq \varepsilon n^{-\beta}$ and*

$$|\nabla(A, B)| < (1 + \varepsilon)k2^{n-k}. \tag{4.2}$$

Then there is $I \subseteq [n]$ of size k such that

$$|\nabla_i A| = (1 - O(\varepsilon))2^{n-k} \quad \forall i \in I. \tag{4.3}$$

Furthermore, there is a codimension k subcube C such that

$$\mu(C \Delta A) = O(\varepsilon). \tag{4.4}$$

Conjecture 4.6. *The statement in Theorem 4.5 holds for all $k \in \mathbb{P}$, even with n^β replaced by $2^n/\partial(|A|)$.*

(The implied constant in (4.3) and (4.4) would necessarily depend on k.)

Note Theorem 4.5 implies an isoperimetric statement—similar to those in Section 4.1.1—of which it is a stability version; namely:

Corollary 4.7. *For $k \in \{1, 2\}$, the assumptions of Theorem 4.5 imply $|\nabla(A, B)| > (1 - O(\varepsilon))k2^{n-k}$.*

(And of course similarly for whatever one can establish in the direction of Conjecture 4.6.)

Finally, the next observation provides a general approach to proving something like the statement in Theorem 4.5 for other values of k. (Its proof is similar to the derivation of Theorem 4.5 from Theorem 1.6 and is omitted.)

Theorem 4.8. *Fix $k \in \mathbb{P}$ and suppose there are $f, g : [0, 1] \to \Re^+$ such that (i) g is continuous with $g(2^{-k}) = k2^{-k}$ and (ii) f is increasing and strictly concave, with $f(0) = 0$, $f(k) = k$ and*

$$\int f(h_A) d\mu \geq g(\mu(A)) \quad \forall A \subseteq V.$$

Then the conclusions of Theorem 4.5 hold (with implied constants depending on f and g) for A, B, W as in the theorem, except with the bound on w replaced by $w \leq \varepsilon/f(n)$.

(For the cases covered by Theorem 4.5, Theorem 1.6 gives the hypothesis of Theorem 4.8 with $f(x)$ equal to x^β when $k = 1$ and $(4/3)x^\beta$ when $k = 2$.)

Theorem 1.6 is proved in Section 4.2. Section 4.3 derives the case $k = 1$ of Theorem 4.5 and then indicates the small changes needed for $k = 2$, and in passing derives Corollary 4.2 (see following Corollary 4.13). The easy proof of Proposition 4.4 is given in Section 4.4.

4.2 Proof of Theorem 1.6

Lemma 4.9. *Let $X \subseteq V$ and let f be a real-valued function on V. If*

$$\frac{1}{\mu(X)} \int_X f^\beta d\mu = T^\beta, \tag{4.5}$$

then

$$\frac{1}{\mu(X)} \int_X (f+1)^\beta d\mu \geq (T+1)^\beta. \tag{4.6}$$

Proof. Set $g(x) = f^\beta(x)$ for $x \in X$. Then the l.h.s. of (4.5) is $\mathbb{E}g$ and the l.h.s. of (4.6) is $\mathbb{E}(g^{1/\beta}+1)^\beta$, where $\mathbb{E}$ refers to uniform measure on X. But $p(x) := (x^{1/\beta}+1)^\beta$ is easily seen to be convex; so, by Jensen's inequality,

$$\mathbb{E}(g^{1/\beta}+1)^\beta \geq ((\mathbb{E}g)^{1/\beta}+1)^\beta,$$

which implies (4.6). □

The proof of Theorem 1.6 proceeds by induction on n. (This is also true of Theorem 1.7, but beyond this the arguments seem to be different.) It is easy to see that the theorem holds for $n = 1$, so we suppose $n \geq 2$.

Given A, fix an $i \in [n]$. Let

$$V_0 = \{x \in V : x_i = 0\},$$

$$V_1 = \{x \in V : x_i = 1\},$$

$$A_0 = A \cap V_0,$$

and

$$A_1 = (A \cap V_1)^i = \{x^i : x \in A, x_i = 1\} \subseteq V_0.$$

Let μ' be uniform measure on V_0. For simplicity, write h_0 (h_1, h, resp.) for h_{A_0} (h_{A_1}, h_A, resp.), a function on V_0 (V_0, V, resp.).

Let $\mu'(A_0) = a_0$, $\mu'(A_1) = a_1$, and $\mu(A) = a = (a_0 + a_1)/2$. Then by induction hypothesis, for $i = 0, 1$,

$$\int h_i^\beta d\mu' \geq 2a_i(1 - a_i). \tag{4.7}$$

We may assume $a_0 \geq a_1$. Note that

$$h(x) = \begin{cases} h_0(x) + 1 & \text{if } x \in A_0 \setminus A_1, \\ h_0(x) & \text{if } x \in A_0 \cap A_1, \\ h_1(x^i) + 1 & \text{if } x^i \in A_1 \setminus A_0, \\ h_1(x^i) & \text{if } x^i \in A_0 \cap A_1; \end{cases} \tag{4.8}$$

so

$$\begin{aligned} \int h^\beta d\mu &= \int_{A_0} h^\beta d\mu + \int_{(A_1)^i} h^\beta d\mu \\ &= \int_{A_0} h^\beta d\mu + \int_{A_1 \setminus A_0} (h_1 + 1)^\beta d\mu + \int_{A_0 \cap A_1} h_1^\beta d\mu \\ &\geq \int_{A_0} h^\beta d\mu + \int_{A_1} h_1^\beta d\mu \\ &\geq \int_{A_0} h^\beta d\mu + a_1(1 - a_1) \end{aligned} \tag{4.9}$$

(the last inequality by (4.7)). Thus the theorem will follow if we show

$$\int_{A_0} h^\beta d\mu \geq 2a(1 - a) - a_1(1 - a_1) \ (= a_0 + a_1^2 - (a_0 + a_1)^2/2). \tag{4.10}$$

The rest of this section is devoted to the proof of (4.10). Let $Z = \mathrm{supp}(h_0) \setminus A_1$ and $X = \mathrm{supp}(h_0) \cap A_1$ (see Figure 6.1); thus

$$2\int_{A_0} h^\beta d\mu = \int_Z (h_0 + 1)^\beta d\mu' + \int_X h_0^\beta d\mu' + \int_{A_0 \setminus (A_1 \cup Z)} 1 d\mu'. \tag{4.11}$$

Observation 4.10. *We may assume* $A_1 \subseteq A_0$.

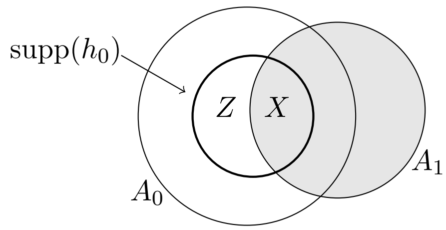

Figure 4.1:

Proof. If there is $x \in A_1 \setminus A_0$ then we can find $y \in A_0 \setminus A_1$ since $\mu'(A_0) \geq \mu'(A_1)$. Let $B_1 = (A_1 \setminus \{x\}) \cup \{y\}$, $B = A_0 \cup (B_1)^i$ and $B_0 = B \cap V_0$ $(= A_0)$. Notice that $|A| = |B|$, $|A_i| = |B_i|$ for $i \in \{0, 1\}$, and

$$\int_{B_0} h_B^\beta d\mu < \int_{A_0} h^\beta d\mu,$$

because: with Z_B (resp. X_B) for $\mathrm{supp}(h_{B_0}) \setminus B_1$ (resp. $\mathrm{supp}(h_{B_0}) \cap B_1$), the location of y changes either from Z to X_B or from $A_0 \setminus (A_1 \cup Z)$ to $(B_0 \cap B_1) \setminus X_B$. In either case its contribution to the r.h.s. of (4.11) shrinks. So if $A_1 \not\subseteq A_0$, then we can shift it to a "worse" set. □

Let $\sigma = \int_Z h_0^\beta d\mu'$, $\gamma = \int_X h_0^\beta d\mu'$, $\alpha = \sigma + \gamma$ $(=\int h_0^\beta d\mu')$ and $\mu'(Z) = z$. Since $A_1 \subseteq A_0$, the r.h.s. of (4.11) is

$$\begin{aligned} \int_Z (h_0+1)^\beta d\mu' + \gamma + \mu'(A_0 \setminus (A_1 \cup Z)) &\geq (\sigma^{1/\beta} + z^{1/\beta})^\beta + \gamma + (a_0 - a_1 - z) \\ &= ((\alpha-\gamma)^{1/\beta} + z^{1/\beta})^\beta + \gamma + (a_0 - a_1 - z) \\ &\geq (\alpha^{1/\beta} + z^{1/\beta})^\beta + (a_0 - a_1 - z), \end{aligned} \tag{4.12}$$

where the first inequality is given by Lemma 4.9 and the second holds because $((\alpha - \gamma)^{1/\beta} + z^{1/\beta})^\beta + \gamma$ is increasing in γ.

So we are done if we show that the expression in (4.12) is at least

$$2(a_0 + a_1^2) - (a_0 + a_1)^2, \tag{4.13}$$

where we are entitled to assume

$$\alpha = \int h_0^\beta d\mu' \geq 2a_0(1 - a_0). \tag{4.14}$$

(see (4.7)) and

$$z \le \min\{\alpha, a_0 - a_1\} \tag{4.15}$$

(where the second bound holds since $Z \subseteq A_0 \setminus A_1$). We consider two cases depending on which of $a_0 - a_1$ and the r.h.s. of (4.14) is smaller.

Case 1. $2a_0(1 - a_0) \le a_0 - a_1$

Equivalently,

$$a_1 \le a_0(2a_0 - 1). \tag{4.16}$$

Also, since $0 \le a_0(2a_0 - 1)$, we have

$$a_0 \ge 1/2. \tag{4.17}$$

Note that (4.12) is decreasing in z and $z \le \alpha$ by (4.15), so recalling that $2^\beta = 3/2$ and using (4.14), we find that (4.12) is at least

$$\alpha/2 + a_0 - a_1 \ge a_0(1 - a_0) + a_0 - a_1. \tag{4.18}$$

Subtracting (4.13) from (4.18) gives

$$-a_1^2 + (2a_0 - 1)a_1,$$

which is nonnegative since

$$f(x, y) := -y^2 + (2x - 1)y \ge 0 \quad \text{for } x \in [\tfrac{1}{2}, 1] \text{ and } y \in [0, x(2x - 1)].$$

(Because: for any $y \ge 0$, $f(x, y)$ is nondecreasing in x, so it is enough to show the inequality holds when $y = x(2x - 1)$, in which case $f(x, y) = x(1 - x)(2x - 1)^2 \ge 0$.)

Case 2. $2a_0(1 - a_0) \ge a_0 - a_1$

Equivalently,

$$a_0(2a_0 - 1) \le a_1 \ (\le a_0). \tag{4.19}$$

Again using the fact that (4.12) is decreasing in z, now with $z \le a_0 - a_1$ by (4.15), we find that (4.12) is at least

$$(\alpha^{1/\beta} + (a_0 - a_1)^{1/\beta})^\beta, \tag{4.20}$$

which, in view of (4.14) (and the fact that (4.20) is increasing in α), is at least

$$((2a_0(1-a_0))^{1/\beta}+(a_0-a_1)^{1/\beta})^{\beta}. \tag{4.21}$$

Thus the proof that (4.12) is at least (4.13) in the present case is completed by the following proposition (applied with $x=a_0$ and $y=a_1$).

Proposition 4.11. *Let*

$$g(x,y)=((2x(1-x))^{1/\beta}+(x-y)^{1/\beta})^{\beta}-2(x+y^2)+(x+y)^2.$$

Then $g(x,y)\geq 0$ for $x,y\in[0,1]$ with $y\in[x(2x-1),x]$.

Proof. Observe that for $x\in[0,1]$,

$$g(x,x(2x-1))=x(1-x)(2x-1)^2\geq 0, \tag{4.22}$$

and

$$g(x,x)=0. \tag{4.23}$$

Also, the partial derivative of $g(x,y)$ with respect to y is

$$g_y(x,y)=-(x-y)^{\frac{1}{\beta}-1}((2x(1-x))^{\frac{1}{\beta}}+(x-y)^{\frac{1}{\beta}})^{\beta-1}+2(x-y).$$

Now, we claim that

$$\text{for given } x\in[0,1],\ g_y(x,y) \text{ is equal to zero for at most one } y\in[x(2x-1),x). \tag{4.24}$$

Indeed, let $A=x-y$ (>0) and $B=2x(1-x)$. Then

$$g_y(x,y)=0\Leftrightarrow A^{\frac{1}{\beta}}+B^{\frac{1}{\beta}}=2^{\frac{1}{\beta-1}}A^{\frac{2\beta-1}{\beta(\beta-1)}}. \tag{4.25}$$

Notice that $A^{\frac{1}{\beta}}+B^{\frac{1}{\beta}}$ is increasing in A while $2^{\frac{1}{\beta-1}}A^{\frac{2\beta-1}{\beta(\beta-1)}}$ is decreasing in A (since $\frac{2\beta-1}{\beta(\beta-1)}<0$). So we conclude that for any B, (4.25) holds at most once, which is (4.24).

Finally, we claim that

$$\text{for each } x\in(0,1),\text{ there is } c=c(x)>0 \text{ such that } g(x,y)>0 \text{ for all } y\in(x-c,x). \tag{4.26}$$

Note that Proposition 4.11 follows from the combination of (4.22), (4.23), (4.24), and (4.26).

Proof of (4.26). Given $x \in (0,1)$, for $c = c(x)$ TBA,

$$g(x, x-c) = ((2x(1-x))^{\frac{1}{\beta}} + c^{\frac{1}{\beta}})^{\beta} + 2x^2 - 2x - c^2,$$

so

$$g(x, x-c) > 0 \Leftrightarrow ((2x(1-x))^{\frac{1}{\beta}} + c^{\frac{1}{\beta}})^{\beta} > c^2 + 2x(1-x). \tag{4.27}$$

Now,

$$((2x(1-x))^{\frac{1}{\beta}} + c^{\frac{1}{\beta}})^{\beta} = 2x(1-x)\left(1 + \left(\frac{c}{2x(1-x)}\right)^{\frac{1}{\beta}}\right)^{\beta},$$

and if c is small enough,

$$\begin{aligned}\left(1 + \left(\frac{c}{2x(1-x)}\right)^{\frac{1}{\beta}}\right)^{\beta} &= \exp[\Theta(c^{1/\beta})\beta] \\ &= 1 + \Theta(c^{1/\beta}),\end{aligned}$$

which implies (4.27). □

4.3 Proof of Theorem 4.5

As noted at the end of Section 4.1.2, we prove Theorem 4.5 for $k = 1$ and then indicate what changes for $k = 2$. This seemed slightly clearer than proving them together, though the differences are minor. Extending to Theorem 4.8 is straightforward, though the counterpart of Proposition 4.14 is slightly more painful than the original.

As usual, $A \subseteq V$ is *increasing* if $x \in A$ and $y \geq x$ (with respect to the product order on V) imply $y \in A$ (and A is *decreasing* is defined similarly). For x, y with $x < y$, we write $x \lessdot y$ if $x \leq z \leq y$ implies $z \in \{x, y\}$. We will need Harris' Inequality [27]:

Theorem 4.12. *For any product measure ν on Q_n and increasing $A, B \subseteq V$,*

$$\nu(A \cap B) \geq \nu(A)\nu(B).$$

Recall that h_S was defined in (1.1) and, for disjoint $A, B \subseteq V$, set

$$h_{AB}(x) = \begin{cases} d_B(x) & \text{if } x \in A, \\ 0 & \text{if } x \notin A; \end{cases}$$

thus

$$\int_A h_{V\setminus B} d\mu = \int h_{AB}\ d\mu = 2^{-n}|\nabla(A,B)|.$$

We need the following easy consequence of Theorem 1.6.

Corollary 4.13. *If* (R,S,U) *is a partition of* V *with* $\mu(R\cup U)=\alpha$*, then*

$$(2^{-n}|\nabla(R,S)| = \int_R h_{R\cup U} d\mu \geq) \quad \int_R h^{\beta}_{R\cup U} d\mu \geq 2\alpha(1-\alpha) - n^{\beta}\mu(U).$$

Proof. Theorem 1.6 gives

$$2\alpha(1-\alpha) \leq \int h^{\beta}_{R\cup U} d\mu = \int_R h^{\beta}_{R\cup U} d\mu + \int_U h^{\beta}_{R\cup U} d\mu \leq \int_R h^{\beta}_{R\cup U} d\mu + n^{\beta}\mu(U),$$

and the corollary follows. □

In particular, taking $(R,S,U)=(B,A,W)$ gives Corollary 4.2. □

We now assume the situation of Theorem 4.5. Note that each of $\mu(A)$, $\mu(B)$ is $1/2 \pm O(\varepsilon)$. In what follows we (abusively) use "a.e." to mean "all but an $O(\varepsilon)$-fraction," so for example write "a.e. $x \in A$ satisfies Q" for "Q holds for all but an $O(\varepsilon)$-fraction of the members of A."

Proposition 4.14. *For a.e.* $x \in A$, $h_{AB}(x) = 1$.

Proof. Applying Corollary 4.13 with $(R,S,U)=(A,B,W)$ (and using (4.2)) gives

$$(1+\varepsilon)/2 \geq \int h_{AB} d\mu = \int_A h_{A\cup W} d\mu \geq \int_A h^{\beta}_{A\cup W} d\mu = 1/2 - O(\varepsilon). \tag{4.28}$$

In particular, $\int (h_{AB} - h^{\beta}_{AB}) d\mu = O(\varepsilon)$, which, since

$$\int (h_{AB} - h^{\beta}_{AB}) d\mu = \Omega\left(\mu(\{x \in A : h_{AB}(x) \notin \{0,1\}\})\right),$$

implies $h_{AB}(x) \in \{0,1\}$ for a.e. $x \in A$. □

The next observation will allow us to assume that A is increasing and B is decreasing.

Proposition 4.15. *For any partition* (A,B,W) *of* V *there is another partition* (A',B',W') *satisfying:*

1. $\mu(X) = \mu(X')$ *for* $X \in \{A,B,W\}$;

2. A' *is increasing and* B' *is decreasing;*

3. $|\nabla_i(A,B)| \geq |\nabla_i(A',B')|$ *for all* $i \in [n]$.

Proof. This is a typical "shifting" argument and we will be brief. For $i \in [n]$, the *i-shift* of a partition (A, B, W) is defined thus: let

$$V_0 = \{x \in V : x_i = 0\}, \quad V_1 = \{x \in V : x_i = 1\},$$

and for each $x \in V_0$ with $(x, x^i) \in (A,B), (A,W),$ or (W,B), switch the affiliations of x and x^i. This trivially does not change $|\nabla_i(A,B)|$, and it's easy to see that it does not increase $|\nabla_j(A,B)|$ for $j \in [n] \setminus \{i\}$. (Consider the contribution to $\nabla_j(A,B)$ of any quadruple $\{x, x^i, x^j, (x^i)^j\}$.)

It is also clear that no sequence of nontrivial shifts can cycle (e.g. since any such shift strictly increases $\sum_{x\in A}|x| - \sum_{x \in B}|x|$); so there is a sequence that arrives at an (A', B', W') stable under i-shifts (for all i), and this meets the requirements of the proposition. $\square$

Proof of Theorem 4.5. We first show there is an i as in (4.3). By Proposition 4.15, we may assume A is increasing and B is decreasing. For each $i \in [n]$, let $A_i = \{x \in A : x^i \in B\}$, and notice that

$$A_i \text{ is a decreasing subset of } A. \tag{4.29}$$

Indeed, given $x \in A_i$, consider any $y \in A$ satisfying $y \lessdot x$. Then $y^i \in B$ since $x^i \in B$ and B is decreasing, so $y \in A_i$.

By proposition 4.14,

$$\text{a.e. } x \in A \text{ is in exactly one } A_i; \tag{4.30}$$

in particular, if we let $A_0 = \{x \in A : d_B(x) = 0\}$, then $\mu(A_0) = O(\varepsilon)$.

Setting $\max \mu(A_i) = \mu(A) - \delta$, we just need to show that $\delta = O(\varepsilon)$.

Assume (w.l.o.g.) that $\max \mu(A_i) = \mu(A_1)$, and let $\tilde{A} = \cup_{i \neq 1} A_i$, $C_1 = A \setminus A_1$, and $\tilde{C} = A \setminus \tilde{A}$. By (4.30),

$$\mu(\tilde{C}) \geq \mu(A_1) - O(\varepsilon), \tag{4.31}$$

while $C_1 \cap \tilde{C} = A_0$ implies

$$\mu(C_1 \cap \tilde{C}) = O(\varepsilon).$$

Moreover, (4.29) and the fact that A is increasing imply that C_1 and $\tilde{C}$ are increasing (in V); so Theorem 4.12 gives

$$O(\varepsilon) = \mu(C_1 \cap \tilde{C}) \geq \mu(C_1)\mu(\tilde{C}) \geq \delta(\mu(A) - \delta - O(\varepsilon)), \tag{4.32}$$

whence

$$\delta = O(\varepsilon) \ \text{ or } \ \mu(A) - \delta - O(\varepsilon) = O(\varepsilon).$$

But $\delta = O(\varepsilon)$ is what we want, so we may assume for a contradiction that $\mu(A) - \delta - O(\varepsilon) = O(\varepsilon)$; equivalently, $\mu(A_1) = O(\varepsilon)$. In this case, $\mu(A_i) = O(\varepsilon)$ for all i, so there is a partition $[n] = I \cup J$ such that each of A_I $(:= \cup_{i \in I} A_i)$ and A_J has measure $\mu(A)/2 + O(\varepsilon)$. But then, setting $C_I = A \setminus A_I$ and $C_J = A \setminus A_J$, and again using Theorem 4.12, we have

$$O(\varepsilon) = \mu(C_I \cap C_J) \geq \mu(C_I)\mu(C_J) \geq \mu^2(A)/4 - O(\varepsilon),$$

which is impossible. □

For (4.4), let i be as above and for $\pi \in \{0,1\}$, let $C(i,\pi) = \{v : v_i = \pi\}$. If D is one of these subcubes then with $|A \cap D| = \delta 2^{n-1}$, Corollary 4.13 (applied in D with $R = A \cap D$ and $U = W \cap D$) gives at least $[2\delta(1-\delta) - O(\varepsilon)]2^{n-1}$ edges in $\nabla(A,B) \setminus \nabla_i A$, which with (4.2) and (4.3) forces δ to be either $O(\varepsilon)$ or $1 - O(\varepsilon)$. So exactly one, say C, has $\delta = 1 - O(\varepsilon)$, and this C satisfies (4.4). □

Changes for $k = 2$ (briefly). The only changes are to Proposition 4.14 and the final argument(s). For the former, the statement is now:

$$\text{for a.e. } x \in A, \ \ h_{AB}(x) = 2.$$

Set $f(x) = (4/3)x^{\beta}$. Theorem 1.6 gives $\int f(h_{A \cup W})d\mu \geq 1/2 - O(\varepsilon)$, leading to

$$\int f(h_{AB})d\mu \geq 1/2 - O(\varepsilon).$$

Now let $X(x) = h_{AB}(x)$ for $x \in A$ and write $\mathbb{E}$ for expectation w.r.t. uniform measure on A. Our assumptions on $\mu(A)$ and $|\nabla(A,B)|$ give

$$\mathbb{E}X = \frac{1}{\mu(A)}\int h_{AB}d\mu = \frac{|\nabla(A,B)|}{\mu(A)2^n} \leq 2 + O(\varepsilon),$$

so, using the concavity of f, we have

$$\int f(h_{AB})d\mu = \mu(A)\mathbb{E}f(X) \leq \mu(A)f(\mathbb{E}X) \leq 1/2 + O(\varepsilon).$$

It's then easy to see (if somewhat annoying to write) that concavity of f, with $\mathbb{E}f(X) - f(\mathbb{E}X) = O(\varepsilon)$ and $f(\mathbb{E}X) = 2 \pm O(\varepsilon)$ (and $X \in \mathbb{Z}$) implies, first, that there is a c such that $f(x) = c$ for a.e. $x \in A$, and, second, that $c = 2$.

For the step leading to (4.3) we may as well think of a general k. Thus we assume A and B are increasing and decreasing (resp.), with $n^\beta\mu(W) \leq \varepsilon$, $\mu(A) = (1 \pm \varepsilon)2^{-k}$, $|\nabla(A,B)| < (1+\varepsilon)k2^{n-k}$, and $h_{AB}(x) = k$ for a.e. $x \in A$, and want to show

$$\text{there is } I \subseteq [n] \text{ of size } k \text{ such that } |\nabla_i A| \geq (1 - O(\varepsilon))2^{n-k} \ \forall i \in I.$$

Here for each k-subset I of $[n]$ we set

$$A_I = \{x \in A : x^i \in B \ \forall i \in I\}.$$

Each A_I is decreasing in A and a.e. $x \in A$ is in exactly one A_I. We then assume $\max_I \mu(A_I) = \mu(A_{[k]}) = \mu(A) - \delta$ and continue essentially as before.

The step yielding (4.4) again takes no extra effort for general k: here we have 2^k subcubes corresponding to the members of $\{0,1\}^k$, and Corollary 4.13 (with (4.2) and (4.3)) shows that all but one of these meet A in sets of size $O(\varepsilon)2^{n-k}$ (and the one that doesn't is the promised C).

4.4 Proof of Proposition 4.4

Let $|A| = a$. For $z \in \{0,1\}^I$ let $V_z = \{x : x_i = z_i \ \forall i \in I\}$, $A_z = A \cap V_z$, $a_z = |A_z|$ and $\alpha_z = a_z/a$. Assume (w.l.o.g.) that a_z is maximum when $z = \underline{0}$. We have

$$\begin{aligned}\varepsilon a \geq |\nabla A \setminus \nabla_I A| &= \sum_z |\nabla(A_z, V_z \setminus A_z)| \geq \sum_z a_z \log_2(2^{n-k}/a_z) \\ &= a\left[H(\alpha_z : z \in \{0,1\}^I) + \log_2(2^{n-k}/a)\right] = aH(\alpha_z : z \in \{0,1\}^I) + O(\varepsilon)a,\end{aligned}$$

where H is binary entropy and the inequality is given by (4.1). It follows that each α_z is either $O(\varepsilon/\log(1/\varepsilon))$ or $1 - O(\varepsilon)$; so in fact $\alpha_{\underline{0}} = 1 - O(\varepsilon/\log(1/\varepsilon))$ and $V_{\underline{0}}$ is the promised subcube.

Chapter 5

The number of maximal independent sets in the Hamming cube

5.1 Lower bound and proof plan

We first briefly recall why the r.h.s. of (1.3) is an (asymptotic) lower bound. As usual an *induced matching* (IM) is an induced subgraph that is a matching. It is easy to see that the largest IM's of Q_n are of size $N/4$ and that there are exactly $2n$ of these, here called *canonical matchings* and denoted M^* (see below for a precise description). Each M^* gives rise to exactly $2^{N/4}$ MIS's, gotten by choosing one vertex from each edge of M^* and extending the resulting independent set to the (unique) MIS containing it. It is also easy to see (an argument is sketched at the end of this section) that the overlaps between the sets of MIS's gotten from different M^*'s are negligible, and the lower bound follows. In analogy with the problems mentioned above (beginning with Dedekind's) one may think of $2n$ "phases," one for each M^*. (E.g. for the simplest of the earlier instances—independent sets, or, in physics, the *hard-core model*—the vast majority of those sets consist almost entirely of vertices of a single parity, and the phases are "even" and "odd.")

So what Theorem 4.5 is really saying is that the number of MIS's *not* corresponding to canonical matchings is negligible. The proof of this goes roughly as follows. We first ("Step 1"; Lemma 5.1) show that almost every MIS is "associated" with some "large" IM. Step 2 (Lemma 5.2) then says that each "large" IM is close to some M^*. Finally, in Step 3 (Lemma 5.3), we show that the number of MIS's that are associated with an IM close to some M^* but are not obtained from M^* as above (that is, miss at least one edge of M^*) is small.

We use I and M for MIS's and IM's (respectively), $\mathcal{I}(G)$ for the set of MIS's in

G, and, in particular, $\mathcal{I}$ for $\mathcal{I}(Q_n)$. Write $I \sim M$ if each edge of M meets I. For bookkeeping purposes we fix a linear order "$\prec$" on the set of IM's of G and define $M_G(I)$ to be the first (in $\prec$) of the largest induced matchings M satisfying

$$I \sim M \tag{5.1}$$

and

$$\nabla(V(M), I \setminus V(M)) = \emptyset \tag{5.2}$$

(where $\nabla(X, Y)$ is the set of edges between X and Y and $V(M)$ is the set of vertices contained in edges of M). We also set $m_G(I) = |M_G(I)|$ and abbreviate $M_{Q_n}(I) = M(I)$ and $m_{Q_n}(I) = m(I)$.

A *canonical matching* of Q_n is the set of edges vv^i of parity ε, for some $i \in [n]$ and $\varepsilon \in \{0, 1\}$. Canonical matchings are denoted M^*. It is easy to see that (as mentioned earlier) the maximum size of an IM is $N/4$, and an IM is of this size iff it is canonical. We set $\mathcal{I}^* = \{I \in \mathcal{I} : I \sim M^* \text{ for some } M^*\}$.

We can now formalize our plan. Let

$$\mathcal{J} = \{I \in \mathcal{I} : m(I) > (1 - \log^3 n/n)N/4\}.$$

(The $\log^3 n/n$ is not optimal, but it is convenient and we have some room.)

Lemma 5.1.

$$|\mathcal{I} \setminus \mathcal{J}| = o(2^{N/4}).$$

(The actual bound will be $\log |\mathcal{I} \setminus \mathcal{J}| < (1 - \Omega(\log^3 n/n))N/4$.)

Lemma 5.2. *With $\beta = \log(3/2)$ (as in Chapter 4), if*

$$|M| = (1 - o(n^{-\beta}))N/4, \tag{5.3}$$

then there is an M^* with

$$|M \Delta M^*| = o(N)$$

(equivalently, $|M \cap M^*| = (1 - o(1))N/4$).

We use Lemma 5.2 to say that each I not covered by Lemma 5.1 (i.e. $I \in \mathcal{J}$) is closely tied to some M^*; precisely, for a suitable $\zeta = \zeta(n) = o(1)$, each $I \in \mathcal{J}$ satisfies

$$\text{there is an } M^* \text{ with } |M(I)\Delta M^*| < \zeta N.$$

Thus, the following lemma completes the proof of Theorem 4.5.

Lemma 5.3. *For any* M^**,*

$$|\{I \notin \mathcal{I}^* : |M(I)\Delta M^*| < \zeta N\}| = 2^{N/4-\omega(n/\log n)} \tag{5.4}$$

The following proposition is the promised lower bound discussion.

Proposition 5.4. *Let* M_1^*, M_2^* *be distinct canonical matchings and* $\mathcal{I}_j^* = \{I \in \mathcal{I} : I \sim M_j^*\}$ *for* $j = 1, 2$*. Then*

$$|\mathcal{I}_1^* \cap \mathcal{I}_2^*| \le 3^{N/8}.$$

Proof. This is easy and we just give an informal sketch. We may assume M_1^* and M_2^* use different directions, since otherwise $\mathcal{I}_1^* \cap \mathcal{I}_2^* = \{\mathcal{E}, \mathcal{O}\}$. We may further assume the two directions are $n-1$ and n, and consider the natural projection $\pi : \{0,1\}^{[n]} \to \{0,1\}^{[n-3]}$; thus the $\pi^{-1}(v)$'s are copies of Q_3 partitioning Q_n. It is then easy to see that for an $I \in \mathcal{I}_1^* \cap \mathcal{I}_2^*$ there are at most three possibilities for each $I \cap \pi^{-1}(v) \cap V(M_1^* \cup M_2^*)$ (and that these choices determine I), yielding the bound in the lemma. □

In what follows we will mainly be concerned with $I \in \mathcal{I}$ (recall this is {MIS's of Q_n}) having $m(I) \approx N/4$, for which the next little point will be helpful.

Observation 5.5. *If* $|M(I)| > (1-\varepsilon)N/4$*, then* $|I \setminus V(M(I))| < \varepsilon N$*.*

Proof. With $M = M(I)$, $W = V(M)$ and $Z = N(W) \setminus W$, we have $I \cap Z = \emptyset$ (by definition of $M(I)$) and

$$(n-1)|W| \;=\; |\nabla(W,Z)| \;\le\; (n-1)|Z|,$$

implying $|Z| \ge |W|$ and

$$|I \setminus V(M)| \;\le\; |V \setminus (W \cup Z)| \;<\; \varepsilon N.$$

□

5.1.1 Preview

In this section we preview the role of the [**Algorithm**] in Section 2.1.1. Recall that we use V for $V(Q_n)$.

Proposition 5.6. *For ξ running over binary strings, with $|\xi|$ denoting the length of ξ, and positive integers l and $r \le l/2$,*

$$\log|\{\xi : |\xi| \le l, |\text{supp}(\xi)| \le r\}| \ \le \ r\log(l/r) + O(r) + \log(l+1).$$

Proof. This follows from $\log \sum_{t \le r} \binom{l}{t} \le lH(r/l)$ (where H is binary entropy). □

Proposition 5.7. *If $Z \subseteq W \subseteq V$, $d_Z(x) \le d\ \forall x \in Z$ and $|\nabla W| \le L$, then*

$$|Z| \le (2n-d)^{-1}(n|W| + L).$$

Proof. This follows from

$$n|W \setminus Z| \ \ge \ |\nabla(Z, W \setminus Z)| \ \ge \ |Z|(n-d) - L.$$

□

In our uses of [**Algorithm**] one reason for stopping will usually be that degrees in X_i fall below some specified d; we then have a *tradeoff:*

(i) *Larger* d tends to mean smaller $\text{supp}(\xi)$: each $x_i \in I$ removes at least d vertices from consideration, so $|\text{supp}(\xi)| < |W|/d$. (And by Proposition 5.6, smaller $\text{supp}(\xi)$ means fewer possibilities for ξ.)

(ii) *Smaller* d tends to mean smaller X (by Proposition 5.7, applied with $Z = X$). Note the effect of varying d is not insignificant here since we are usually interested in $|X| - |W|/2$.

A simple but seemingly new idea that is one of the main drivers of the present work is that we can do better in (i) if we lower bound $d_{X_{i-1}}(x_i)$, not by the final cutoff d, but by whatever we get by plugging X_{i-1} in for Z in Proposition 5.7. We give two implementations of this idea; the first, in Section 5.2, is more elegant and precise, while the cruder version in Section 5.4 more simply illustrates the basic principle. (See also Remark 5.19.)

5.2 Proof of Lemma 5.1

In this section, I is always in $\mathcal{I}\setminus\mathcal{J}$. The eventual key here is Theorem 3.4, but we need to first reduce to a place where the theorem is helpful—so to a vertex set of size not much more than $N/2$ since we are interested in induced matchings of size around $N/4$. The algorithm of Section 2.1.1 provides a "cheap" way to do this.

For any subgraph H of Q_n, let

$$\mathrm{MIS}^*(H) = \{I \in \mathcal{I}(H) : m_H(I) \ \leq \ (1-\log^3 n/n)N/4\},$$

and $\mathrm{mis}^*(H) = |\mathrm{MIS}^*(H)|$. (Note the cutoff for $m_H(I)$ here is the one in the definition of $\mathcal{J}$.)

For the proof of Lemma 5.1 we run **[Algorithm]** with input our unknown I, stopping as soon as either

1. $|\mathrm{supp}(\xi)| \geq \frac{\log n}{2n}N$, or
2. $X_i = \emptyset$,

and let $X = X(I)$ and $H = H(I)$ $(= H(\xi))$ be as in Section 2.1.1. Notice that $I \in \mathcal{I}\setminus\mathcal{J}$ implies

$$I \cap X \in \mathrm{MIS}^*(H(I)),$$

so

$$|\mathcal{I}\setminus\mathcal{J}| \leq \sum_{\xi} \mathrm{mis}^*(H(\xi))$$

(where the sum runs over possible ξ's). Proposition 5.6 bounds the number of possible ξ's by

$$\exp_2\left[O\left(\log^2 n/n\right)N\right],$$

so that Lemma 5.1 will follow from

$$\log \mathrm{mis}^*(H(I)) \leq \left(1-\Omega\left(\frac{\log^3 n}{n}\right)\right)N/4 \ \text{ for all } \ I. \tag{5.5}$$

Proof of (5.5). Fix I and let $X = X(I)$ $(= V(H(I)))$. We first show that $|X|$ cannot be much larger than $N/2$. Let $d_i = \max\{d_{X_i}(v) : v \in X_i\}$ and $\overline{X}_i = V \setminus X_i$.

Observation 5.8. *For each i, $|X_i| \le (1 + d_i/n)N/2$.*

Proof. This follows from Proposition 5.7 with $Z = X_i$ and $W = V$ (and $L = 0$). □

Define α_i by

$$|X_i| = (1 + \alpha_i)N/2 \ ;$$

so $\alpha_0 = 1$ and Observation 5.8 says

$$d_i \ge \alpha_i n. \tag{5.6}$$

Observation 5.9. *If $\xi_i = 1$, then $\alpha_i < (1 - 2n/N)\alpha_{i-1}$.*

Proof. Using (5.6), we have

$$(1 + \alpha_i)N/2 = |X_i| = |X_{i-1}| - d_{i-1} - 1 < (1 + \alpha_{i-1})N/2 - \alpha_{i-1}n,$$

and the observation follows. □

Proposition 5.10.

$$|X| < (1 + 1/n)N/2. \tag{5.7}$$

Proof. Let α be the final α_i (so $|X| = (1+\alpha)N/2$). Assuming (as we may) that $X \neq \emptyset$, we have

$$|\mathrm{supp}(\xi)| \ge \frac{\log n}{2n} N,$$

so that Observation 5.9 (with $\alpha_0 = 1$ and the fact that α_i is decreasing in i) gives

$$\alpha \le (1 - 2n/N)^{\frac{\log n}{2n} N} < 1/n,$$

which is (5.7). □

Now, if

$$|X| < (1 - \Omega(\log^3 n/n))N/2$$

then (5.5) follows from Theorem 1.2; otherwise, applying Theorem 3.4 with $m = |X|$ and a suitable $\varepsilon = \Omega(\log^3 n/n)$ gives

$$\log \mathrm{mis}^*(H) < (1 - c\varepsilon)|X|/2 < (1 - \Omega(\log^3 n/n))N/4.$$

5.3 Proof of Lemma 5.2

Let M be as in Lemma 5.2. We may assume that

$$n-1 \text{ and } n \text{ are the two directions least used by } M. \tag{5.8}$$

Let $\pi : V \to V(Q_{n-2})$ be the natural projection, namely

$$\pi((\varepsilon_1, \ldots, \varepsilon_n)) = (\varepsilon_1, \ldots, \varepsilon_{n-2}),$$

and for $v \in V(Q_{n-2})$, let

$$U_v = \pi^{-1}(v) = \{(v, \varepsilon_{n-1}, \varepsilon_n) : \varepsilon_{n-1}, \varepsilon_n \in \{0,1\}\}.$$

For the rest of this section, "measure" refers to μ, the uniform measure on $V(Q_{n-2})$.

Say $v \in V(Q_{n-2})$ is *red* (or *in* R) if $U_v \cap V(M) = \{(v,0,0),(v,1,1)\}$ and *blue* (v in B) if $U_v \cap V(M) = \{(v,1,0),(v,0,1)\}$. (So $v \notin R \cup B$ iff U_v either contains an edge of M or meets $V(M)$ at most once.) Say $v \in R \cup B$ is *good* if there is a (necessarily unique) $v' \in N_v$ with the same color (R or B) as v; thus v is good iff U_v meets two edges of M and these have the same direction, and

$$\text{if } w \sim v \text{ are both good then they have the same color iff } w = v'. \tag{5.9}$$

Let X be the set of good vertices and $W = V(Q_{n-2}) \setminus X$ (the set of "bad" vertices).

Observation 5.11. $\mu(W) = o(n^{-\beta})$

Proof. As already noted, v is bad iff it satisfies one of: (i) U_v contains an edge of M; (ii) $|U_v \cap V(M)| \le 1$; (iii) v is red or blue and there is no vertex of the same color in N_v. It follows from (5.8) that the fraction of v's of the first type is $O(1/n)$, and from (5.3) that the fraction of the second type is $o(n^{-\beta})$.

For v as in (iii), let xy be one of the two M-edges meeting U_v, say with $x \in U_v$ and $y \in U_w$. Then $U_w \cap V(M) = \{y\}$, w is as in (ii), and v is the unique vertex of Q_{n-2} for which U_v and U_w are connected by an edge of M. Thus the number of vertices in (iii) is less than (actually at most half) the number in (ii), so these too make up an $o(n^{-\beta})$-fraction of the whole. $\square$

Recall that the parity of the edge vv^i is the parity of $\sum_{j\neq i} v_j$ and notice that

$$v \text{ and } vv^i \text{ have the same parity iff } v_i = 0. \tag{5.10}$$

It follows from (5.9) that $T := \{(v, v') : v \in X\}$ is a perfect matching of $Q_{n-2}[X]$.

Observation 5.12. *Each $e = vv' \in T$ corresponds to two edges of M ($((v, 0, 0), (v', 0, 0))$ and $((v, 1, 1), (v', 1, 1))$ if $v \in R$ and similarly if $v \in B$), and these edges have the same parity as e if $v \in R$ and the opposite parity if $v \in B$.*

Let $\Gamma = Q_{n-2}[X] - T$.

Observation 5.13. *For each $e \in T$, the ends of e are in different components of Γ. In particular no component of Γ has measure more than 1/2.*

Proof. Assume for a contradiction that $e = xy$ and $P = (x = x_0, x_1, x_2, \ldots, x_k = y)$ is a path in Γ. Notice that (5.9) implies x_i and x_{i+1} have different colors, while x and y have the same color. Thus $P \cup \{e\}$ is an odd cycle in Q_{n-2}, which is impossible. □

For the rest of this discussion we do not distinguish between components and their vertex sets.

Proposition 5.14. Γ *contains two components of measure* $1/2 - o(1)$.

(We really only need one such component, but for the same price can give the correct picture.)

Proof. This follows from Observation 5.13 and

$$\text{If } Z \text{ is a union of components of } \Gamma \text{ with } z := \mu(Z) \leq 1/2, \text{ then } z \text{ is either } o(1) \text{ or } 1/2 - o(1). \tag{5.11}$$

Proof of (5.11). Set $Y = X \setminus Z$. Since $\nabla(Z, Y) \subseteq T$ and T is a perfect matching of $Q_{n-2}[X]$, we have $h_{Z\cup W}(x) \in \{0, 1\}$ for $x \in Z$, which with Corollary 4.13 (applied in Q_{n-2} with $(R, S, U) = (Z, Y, W)$) and Observation 5.11 gives

$$z \geq \int_Z h_{Z\cup W} d\mu \geq 2z(1 - z) - o(1),$$

implying (5.11). □

□

Let Z be one of the two large components promised by Proposition 5.14 and $Y = X \setminus Z$. Again (as in the proof of (5.11)), we have $h_{Z\cup W}(x) \in \{0,1\}$ for $x \in Z$, which with Observation 5.11 and Theorem 4.5 implies that there are $i \in [n-2]$ with

$$|(\nabla_i Z) \cap T| = |\nabla_i(Z,Y)| \sim 2^{n-3} \tag{5.12}$$

and $\varepsilon \in \{0,1\}$ such that

all but $o(2^n)$ vertices of Z lie in the subcube $C(i,\varepsilon) = \{v : v_i = \varepsilon\} \quad (\subseteq V(Q_{n-2}))$.

Assume (w.l.o.g.) that $\varepsilon = 0$ and set

$$Z' = \{v \in Z \cap C(i,0) : vv^i \in T\}.$$

Connectivity of Z and (5.9) imply

any two vertices of Z either agree in both color and parity or disagree in both. (5.13)

Finally, for Lemma 5.2: For $v, w \in Z'$, Observation 5.12 and (5.10) imply that the edges of M corresponding to vv^i and ww^i have the same parity iff v and w either agree in both parity and color or disagree in both; but (5.13) says this is true for *any* $v, w \in Z'$. So *all* edges of M corresponding to edges of $\nabla_i(Z', Y)$ have the same parity and the lemma follows from (5.12).

5.4 Proof of Lemma 5.3

For the discussion in this section we fix a canonical matching M^* and show (proving Lemma 5.3)

$$|\{I \notin \mathcal{I}^* : |M(I)\Delta M^*| < \zeta N\}| = 2^{N/4-\omega(n/\log n)}. \tag{5.14}$$

Assume (w.l.o.g.) that M^* is the set of odd edges in direction n and let $\pi : V(Q_n) \to V(Q_{n-1})$ be the projection

$$\pi((\varepsilon_1, \dots, \varepsilon_n)) = (\varepsilon_1, \dots, \varepsilon_{n-1}).$$

Thus $\pi(V(M^*))$ is the set of odd vertices in Q_{n-1}, which we from now on denote by $\mathcal{O}$.

For $\varepsilon \in \{0,1\}$ let $V_\varepsilon = \{x \in V(Q_n) : x_n = \varepsilon\}$, and for $v \in V(Q_{n-1})$ let $\pi^{-1}(v) = \{v_0, v_1\}$ where $v_\varepsilon \in V_\varepsilon$. (We will not use the *coordinates* of v, so "v_ε" should cause no confusion.) For $I \in \mathcal{I}$, define the *labeling* $\sigma = \sigma(I)$ *of* $V(Q_{n-1})$ by:

$$\sigma_v = \begin{cases} 0 & \text{if} \quad v_0 \in I \\ 1 & \text{if} \quad v_1 \in I \\ \Lambda & \text{if} \quad I \cap \{v_0, v_1\} = \emptyset \end{cases}$$

Say v is *unoccupied* if $\sigma_v = \Lambda$, and *occupied* otherwise. Note that (since $I \in \mathcal{I}$)

$$\text{no two adjacent vertices have the same label from } \{0,1\} \tag{5.15}$$

and

$$\text{if } \sigma_v = \Lambda \text{ then both 0 and 1 appear on neighbors of } v. \tag{5.16}$$

Call a labeling $\sigma : V(Q_{n-1}) \to \{0, 1, \Lambda\}$ *legal* if it satisfies (5.15) and (5.16), and notice that $I \mapsto \sigma(I)$ is a bijection between $\mathcal{I}$ and the set of legal labelings. We will find both viewpoints useful in what follows and will assume, often without explicit mention, that when we are discussing I the labeling referred to is $\sigma(I)$.

For the rest of Section 5.4 we restrict to I as in (5.14), noting that then $\sigma = \sigma(I)$ satisfies

$$\text{all but a } o(1)\text{-fraction of odd vertices are occupied} \tag{5.17}$$

and, by Observation 5.5,

$$\text{only a } o(1)\text{-fraction of the even vertices are occupied.} \tag{5.18}$$

Notation below ($\mathcal{E}^*$, A_i and so on) is for a given I, which the notation suppresses. Write $\mathcal{E}^*$ for the set of occupied even vertices. Notice that $I \notin \mathcal{I}^*$ implies that there is at least one unoccupied $v \in \mathcal{O}$, which by (5.16) must have neighbors in both $\sigma^{-1}(0)$ and $\sigma^{-1}(1)$; in particular

$$\text{there is a non-singleton 2-component in } \mathcal{E}^*. \tag{5.19}$$

(Recall k-components were defined in Section 2.2.1.)

Notation.

- A_i's : non-singleton 2-components of $\mathcal{E}^*$
- $A = \cup A_i$
- $G_i = N(A_i)$, $G = N(A)$
- A_i (or simply i) is $\begin{cases} \textit{small} & \text{if } |G_i| < n^4 \text{ and} \\ \textit{large} & \text{otherwise} \end{cases}$
- $\hat{X} = \pi^{-1}(X)$ (for $X \subseteq V(Q_{n-1})$).

We usually (without comment) use lower case letters for the cardinalities of the sets denoted by the corresponding upper case letters, *except* that we use a for $|[A]|$ and a_i for $|[A_i]|$. (Recall the closure $[A]$ of A was defined in Section 2.2.3.) We also set $t_i = g_i - a_i$ and $t = g - a$, noting that $a \geq \sum a_i$ ($[A]$ *can* properly contain $\cup[A_i]$), so $t \leq \sum t_i$.

Before moving to lemmas we record two basic observations. The first says that in some sense all the action is in the $[A_i]$'s and G_i's (though this only approximately describes what will happen in the main argument; see (5.27)).

$$\text{All vertices of } \mathcal{O} \setminus G \text{ are occupied.} \tag{5.20}$$

Proof. All neighbors of the set in (5.20) are in $\mathcal{E} \setminus A$, and any occupied vertex from *this* set is a singleton 2-component of $\mathcal{E}^*$, so by (5.16) has all its neighbors occupied (with a common label). □

The second observation (this will be crucial; see (5.39)-(5.40) and (5.47), which leads *via* (5.48) to (5.57)) is

$$\text{for each } i\text{, each edge contained in } \hat{G}_i \text{ has a neighbor in } I \cap \hat{A}_i \tag{5.21}$$

(that is, one of its ends has such a neighbor; note these edges form an induced matching in Q_n).

5.4.1 Main lemma

We continue to restrict to I as in (5.14) and to suppress dependence on I in our notation. In what follows we use "cost of X" for the log of the number of possibilities for X.

Before turning to our main point, Lemma 5.16, we observe that there is not much to do when g is large:

Lemma 5.15. *The number of I's with $g = \Omega(N)$ is $2^{N/4-\Omega(N)}$.*

Proof. By (5.18), the cost of specifying A is at most $\log \binom{N/4}{\leq o(N)} = o(N)$, and that for labeling A is at most $|A| = o(N)$. But A and its labels determine G and *its* labels, while (5.20) says that the cost of labeling $\mathcal{O} \setminus G$ (given G) is at most $N/4 - g$ and that the labels for $\mathcal{O} \setminus G$ determine those for $\mathcal{E} \setminus N(G)$ (and all labels on $N(G) \setminus A$ are Λ). The lemma follows. □

We may thus assume from now on that (say)

$$g < N/4, \tag{5.22}$$

so that, by Proposition 2.6,

$$t = \Omega(g/\sqrt{n}) \quad \text{and} \quad t_i = \Omega(g_i/\sqrt{n}) \quad \text{for each } i. \tag{5.23}$$

This small but crucial point will be used repeatedly in what follows; indeed, one may say that the purpose of Lemmas 1.2 and 1.3 was to get us to (5.23). (Namely: Lemmas 1.2 and 1.3 lead to (5.18); (5.18) is the basis for Lemma 5.15; and Lemma 5.15 allows us to restrict to (5.22), where we have (5.23).)

Lemma 5.16. *For any $a \neq 0$ and $g < N/4$*

$$\log |\{I : |[A]| = a, |G| = g\}| = N/4 - \omega(t/\log n). \tag{5.24}$$

To see that this (with Lemma 5.15) gives Lemma 5.3, note that we always have $g \geq 2n-2$, and that if $g \leq n^2$ (say) then Proposition 2.7 gives $t \sim g$. Thus Lemma 5.16 and (5.23) bound the number of I's satisfying (5.22) by

$$2^{N/4}\left[n^4 2^{-\omega(n/\log n)} + \textstyle\sum_{g>n^2} g 2^{-\omega(g/(\sqrt{n}\log n))}\right] = 2^{N/4-\omega(n/\log n)}$$

(where the irrelevant n^4 and initial g in the sum are for choices of (g, a) and a respectively).

5.4.2 Proof of Lemma 5.16

Before beginning in earnest, we dispose of the minor cost of specifying the a_i's and g_i's (with $\sum a_i \leq a$, $\sum g_i = g$). The only thing to notice here is that, since $g_i \geq 2n - 2\ \forall i$, the number of i's is less than g/n. Thus Proposition 2.1 bounds the cost of the g_i's by $(g/n)\log(en)$ and that of the a_i's by

$$\begin{cases} (g/n)\log(en) & \text{if } (g >)\ a > 2g/n, \\ 2g/n & \text{if } a \leq 2g/n, \end{cases}$$

so also the overall "decomposition" cost by

$$O(g \log n/n) = O(t \log n/\sqrt{n}). \tag{5.25}$$

Preview and objective

It remains to specify A_i's (and thus G_i's and $[A_i]$'s) corresponding to the above parameters, and a labeling (σ) compatible with these specifications. For small i's it turns out to be easy to directly identify the A_i's and their labels (which also gives the associated G_i's and $[A_i]$'s and *their* labels).

For the large i's we think of "identification" and "labeling" phases, roughly corresponding to identifying the $[A_i]$'s (and G_i's), and then the restriction of σ to these sets—"roughly" because in the most interesting ("slack") case the first phase will not actually succeed in identifying the $[A_i]$'s. The identification phase takes place in the projection on Q_{n-1} and leans mainly on Lemma 2.11. For the labeling phase we return to Q_n and work with maximal independent sets rather than labelings (recall these are interchangeable), with arguments again based on the algorithm of Section 2.1.1. It is here that the crucial role of $\mathcal{J}$ will finally appear.

The large i's will be of two types, "tight" and "slack." The slack i's are treated last, when we already have full information on the small and tight i's. Here we produce a single pair $(S, F) \subseteq \mathcal{E} \times \mathcal{O}$ satisfying (*inter alia;* e.g. the role of F will appear later)

$$S \supseteq \cup\{[A_i] : i \text{ slack}\} \tag{5.26}$$

and

$$S \cup N(S) \text{ is disjoint from } \cup\{[A_i] \cup G_i : i \text{ small or tight}\},$$

and then specify labels for $S \cup N(S)$.

Since $N(S) \supseteq \cup\{G_i : i \text{ slack}\}$, (5.20) gives

$$\text{all vertices of } \mathcal{O} \setminus (\cup\{G_i : i \text{ small or tight}\} \cup N(S)) \text{ are occupied.} \tag{5.27}$$

Note also that

$$\text{a (legal) labeling is determined by its restriction to } \cup\{[A_i] : i \text{ small or tight}\} \cup S \cup \mathcal{O},$$

since each v not in this set (so $v \in \mathcal{E}$) has at least one occupied neighbor (for if all neighbors of v are unoccupied, then v is occupied and, by (5.16), $N^2(v)$ contains an occupied vertex, so v must be in some A_i).

Thus the cost of σ given its restriction to

$$\cup\{[A_i] \cup G_i : i \text{ small or tight}\} \cup S \cup N(S)$$

(so in particular the identity of this set) is at most

$$N/4 - \left[\sum\{g_i : i \text{ small or tight}\} + |N(S)|\right]. \tag{5.28}$$

This gives us a benchmark: for Lemma 5.16, the cost of the above information (through specification of labels for $S \cup N(S)$) should be less by $\omega(t/\log n)$ than the subtracted quantity in (5.28) (which in particular makes the decomposition cost (5.25) negligible). In the event, this will hold fairly locally: we will wind up paying $g_i - \Omega(t_i)$ for each small or tight i and $|N(S)| - \omega(t'/\log n)$ for (all) the slack i's, where $t' = \sum\{t_i : i \text{ slack}\}$. (We will repeat this last bit more precisely at the end of the section, following the proof of Lemma 5.20.)

Small i's. As suggested above, these are easy. Since $|A_i| \le a_i$ $(= |[A_i]|)$, the cost of identifying A_i, together with its labels, is at most

$$(n-2) + O(a_i \log n) + a_i \;=\; n + O(a_i \log n) \;<\; g_i - (1/2 - o(1))t_i. \tag{5.29}$$

Here the first two terms on the l.h.s., representing the cost of identifying A_i, are given by Proposition 2.4, and the final bound follows from $g_i \geq \max\{2n-2, t_i\}$ and $a_i = O(g_i/n)$, the latter holding for small i by Proposition 2.7.

But A_i and its labels determine G_i, $[A_i]$ and their labels (the labels since all vertices of $N(G_i) \setminus A_i$ are labeled Λ); so (5.29) actually bounds the total cost of identifying *and* labeling $[A_i] \cup G_i$.

Large i's. For a given large i, Lemma 2.11 gives $\mathcal{W} = \mathcal{W}(a_i, g_i)$, $\varphi = \varphi_i$, $S = S_i$ and $F = F_i$ (as in the lemma), at cost $O(t_i \log^2 n/\sqrt{n})$; so the cost of specifying these for all large i is

$$O(\textstyle\sum t_i \log^2 n/\sqrt{n}). \tag{5.30}$$

Let $\varepsilon = \varepsilon_n$ be a parameter satisfying

$$1 \gg \varepsilon \gg 1/\log n, \tag{5.31}$$

and say i is *tight* if (with ε as in (5.31))

$$g_i - f_i \leq \varepsilon t_i \tag{5.32}$$

and *slack* otherwise. (As usual we use $s_i = |S_i|$ and $f_i = |F_i|$. The role of ε is just to enable proper definitions of "tight" and "slack.")

For our purposes the most significant difference between these two possibilities is that specification of $([A_i], G_i)$ given (S_i, F_i) is cheap if i is tight, but becomes unaffordable as the difference in (5.32) grows; this leads to the following plan. We first treat tight i's, in each case paying for the full specification of $[A_i]$ (which determines G_i) and then the labels of $[A_i] \cup G_i$.

We then combine and slightly massage the remaining (slack) S_i's and F_i's, taking account of what we know so far, to produce a single pair (S, F) that in some sense approximates the slack parts of the configuration, and from (S, F) go directly to specification of labels (so we learn—implicitly—the identities of the slack $[A_i]$'s and G_i's only when we learn their labels.)

Tight i's. The next two lemmas bound the total cost of a tight i (so of $[A_i]$, G_i and their labels) by

$$g_i - \Omega(t_i). \tag{5.33}$$

Lemma 5.17. *For tight i, the cost of $([A_i], G_i)$ given (S_i, F_i) is $o(t_i)$.*

Lemma 5.18. *The cost of labeling a given $[A_i] \cup G_i$ is $g_i - \Omega(t_i)$.*

Remark. Lemma 5.18 does not require that i be tight.

Proof of Lemma 5.17. Given (S_i, F_i), fix some $A^* \in \varphi^{-1}(S_i, F_i)$. (Note A^* is closed. Note also that we are not considering possibilities for A^*, just naming a particular choice associated with (S_i, F_i)—e.g. the first member of $\varphi_i^{-1}(S_i, F_i)$ according to some order—so the specification costs nothing. This strangely helpful device is from [25].) The key (trivial) point here is that (given $H := A^*$)

$$(H \setminus G_i, G_i \setminus H) \text{ determines } (G_i, [A]).$$

So we should bound the costs of $H \setminus G_i$ and $G_i \setminus H$. Since $H \setminus G_i \subseteq H \setminus F_i$, the cost of $H \setminus G_i$ is at most $|H \setminus F_i| \leq \varepsilon t_i = o(t_i)$ (since i is tight).

On the other hand,

$$G_i \setminus H = N([A_i] \setminus A^*) \setminus H$$

(since each $x \in G_i \setminus H$ has a neighbor in $[A_i]$ and none in A^*); so we may specify $G_i \setminus H$ by specifying a $Y \subseteq [A_i] \setminus A^* \subseteq S_i \setminus A^*$ of size at most $|G_i \setminus H| \leq g_i - f_i = o(t_i)$ with $G_i \setminus H = N(Y) \setminus H$ (let Y contain one neighbor of x for each $x \in G_i \setminus H$). But, since $s_i < f_i + o(t_i) \leq g_i + o(t_i)$ (see (c) of Lemma 2.11), we have $|S_i \setminus A^*| = s_i - a_i \leq t_i + o(t_i)$; and the cost of specifying a subset of size $o(t_i)$ from a set of size $O(t_i)$ is $o(t_i)$. □

Proof of lemma 5.18. As promised earlier (see the discussion following (5.25)) we now return to Q_n and, with $W = \widehat{[A_i]} \cup \hat{G}_i$, bound the number of MIS's in $\Gamma := Q_n[W]$. (Note that since A_i is a 2-component of $\mathcal{E}^*$, $I \cap W$ is an MIS in Γ, possibilities for which correspond to possible (legal) labelings of $[A_i] \cup G_i$).

We run **[Algorithm]** (of Section 2.1.1) twice (or, really, once with a pause; here we index steps by j since i is already taken). For the first run (on all of Γ, with input the unknown I) we STOP as soon as

$$d_{X_j}(x) \leq n^{2/3} \text{ for all } x \in X_j.$$

This implies $|\text{supp}(\xi)| \leq 2(g_i+a_i)n^{-2/3}$ (note e.g. $|\hat{G}_i| = 2g_i$), so Proposition 5.6 bounds the cost of this run by

$$(2+o(1))(g_i+a_i)n^{-2/3}\log(n^{2/3}) = o(t_i), \tag{5.34}$$

where the "$o(t_i)$" uses (5.23). On the other hand, with Z_1 the final X_j from this run, Proposition 5.7 with $Z = Z_1$, $d = n^{2/3}$ and

$$L = |\nabla(W)| = 2(n-1)(g_i - a_i) \tag{5.35}$$

gives

$$\begin{aligned} |Z_1| &\leq (2n - n^{2/3})^{-1}(2n(g_i + a_i) + 2(n-1)(g_i - a_i)) \\ &< (2n - n^{2/3})^{-1}4ng_i \; < \; (1 + n^{-1/3})2g_i. \end{aligned} \tag{5.36}$$

We next run **[Algorithm]** on $Q_n[Z_1]$ and STOP as soon as either

(a) $d_{X_j}(x) \leq n^{1/3}$ for all $x \in X_j$ or

(b) $|X_j| \leq 2a_i$.

(Note we treat this as a fresh run rather than a continuation, and recycle X_j and ξ.)

Let Z_2 be the final X_j for this run. From (5.36) and (b) we have $z_1 - z_2 \leq 2t_i + 2n^{-1/3}g_i$, so in view of (a),

$$|\text{supp}(\xi)| \leq (z_1 - z_2)n^{-1/3} \leq 2t_i n^{-1/3} + 2g_i n^{-2/3} =: r.$$

Proposition 5.6 (with this r and $l = |W| \leq 4g_i$) then bounds the run cost by

$$O((t_i n^{-1/3} + g_i n^{-2/3})\log n + \log g_i) = o(t_i), \tag{5.37}$$

with the $o(t_i)$ given by (5.23).

Finally we consider the cost of specifying $I \cap Z_2$ (an MIS of $Q_n[Z_2]$). If the second run ends with $|Z_2| \leq 2a_i$ (as in (b)), then Theorem 1.2 bounds this cost by

$$a_i = g_i - t_i.$$

Suppose instead that the algorithm halts due to (a). In this case we again use Proposition 5.7, now with $Z = Z_2$, $d = n^{1/3}$ and L as in (5.35), to obtain (*cf.* (5.36))

$$|Z_2| < (1 + n^{-2/3})2g_i = 2g_i + o(t_i). \tag{5.38}$$

We now apply Theorem 3.4 in $\Gamma := Q_n[Z_2]$. The key here is (5.21), which implies

$$\text{no edge of } \hat{G}_i \text{ can belong to } M_\Gamma(I \cap Z_2) \tag{5.39}$$

(since the neighbor promised by (5.21) cannot come from $I \setminus Z_2$, which has no neighbors in Z_2). It follows that

$$m_\Gamma(I \cap Z_2) \le a_i \tag{5.40}$$

(each edge of $M_\Gamma(I \cap Z_2)$ meets (possibly meaning equals) one of the a_i edges of $\hat{A}_i$ and, since $M_\Gamma(I \cap Z_2)$ is an induced matching, the edges met are distinct). The combination of (5.38), (5.40) and Theorem 3.4 now again bounds the cost of $I \cap Z_2$ by $g_i - \Omega(t_i)$.

Summarizing, the cost of the two runs of **[Algorithm]** is $o(t_i)$ (see (5.34), (5.37)) and, regardless of how these end, the cost of $I \cap Z_2$ is $g_i - \Omega(t_i)$. The lemma follows. □

Remark 5.19. *Note—*cf. *the preview at the end of Section 2.1.1—the above argument does not work if we run **[Algorithm]** just once, stopping when degrees in X_j fall below $n^{1/3}$; for our bound on $|\text{supp}(\xi)|$ then becomes $2(g_i + a_i)n^{-1/3}$, so the cost bound in* (5.34) *increases to $\Theta(g_i n^{-1/3} \log n)$, which need not be small compared to t_i.*

Slack i's. At this point we have found and labeled

$$Y := \cup \{[A_i] \cup G_i : i \text{ small or tight}\},$$

so are left with the slack i's. As suggested above, these differ from tight i's in that the step that identifies the $([A_i], G_i)$'s is no longer affordable, and we instead go directly from the (S_i, F_i)'s to the labeling phase.

Set $Y_{\mathcal{E}} = Y \cap \mathcal{E}$ and $Y_{\mathcal{O}} = Y \cap \mathcal{O}$ (so $Y_{\mathcal{E}} = \cup \{[A_i] : i \text{ small or tight}\}$ and similarly for $Y_{\mathcal{O}}$). Writing $\cup^s$ and $\sum^s$ for union and sum over slack i's, set

$$S = (\cup^s S_i) \setminus N(Y_{\mathcal{O}}),\ F = \cup^s F_i,\ X = N(S) \setminus F$$

(note $N(Y_{\mathcal{O}}) \supseteq Y_{\mathcal{E}}$), $g' = \sum^s g_i$ and $t' = \sum^s t_i$. Notice that

$$g' - f > \varepsilon t' \tag{5.41}$$

and that with these definitions we still have the appropriate versions of (a)-(c) of Lemma 2.11, namely:

(a′) $S \supseteq \cup^s[A_i], F \subseteq \cup^s G_i$;

(b’) $d_F(u) \geq n - 1 - \sqrt{n}/\log n \quad \forall u \in S$;

(c′) $|S| \leq |F| + O(t'/(\sqrt{n}\log n))$.

Here (b’) is immediate from the corresponding statement for the (S_i, F_i)’s, as is (c′) once we observe that the F_i’s are disjoint (since the G_i’s are, and $F_i \subseteq G_i$). Similarly, (a′) holds because $S_i \supseteq [A_i]$ $(\forall i)$ and—the least uninteresting point here—$N(Y_{\mathcal{O}}) \cap (\cup^s[A_i]) = \emptyset$ (since there are no edges between $[A_i]$ and G_j if $i \neq j$).

The last ingredient in the proof of Lemma 5.16 is Lemma 5.20 below, before turning to which we need a few further observations.

First, we are about to return to Q_n (as in the proof of Lemma 5.18), where we will be running **[Algorithm]** on

$$W := \hat{S} \cup \hat{F}, \tag{5.42}$$

and for use in Proposition 5.7 will need a bound on $|\nabla W|$. Setting $\psi = \sqrt{n}/\log n$ (and for the moment still working in Q_{n-1}), we have (from (b’))

$$|\nabla S \setminus \nabla(S, F)| \leq s\psi \tag{5.43}$$

and

$$\begin{aligned} |\nabla F \setminus \nabla(S, F)| &\leq f(n-1) - s(n-1-\psi) \\ &= (f-s)(n-1) + s\psi, \end{aligned} \tag{5.44}$$

whence (now in Q_n)

$$L := |\nabla W| \leq 2(f-s)(n-1) + 4s\psi. \tag{5.45}$$

Set $U = \hat{S} \cup \widehat{N(S)}$. A second—crucial—observation is

$$I \cap U \textit{ is an MIS of } Q_n[U]. \tag{5.46}$$

Proof. Suppose instead that $x \in U \setminus (I \cup N(I \cap U))$. Then, since I is an MIS of Q_n, there are $y \sim x$ and $z \sim x^n$ with $y, z \in I$ and $y \notin U$. Note this implies $\pi(x) \in N(S)$ (as opposed to S), since otherwise $N(x) \subseteq U$. Now $\pi(y)$, $\pi(z)$ are distinct occupied neighbors of $\pi(x)$ (distinct since y and z, being in I, cannot be adjacent), meaning that $\pi(x) \in G_i$ for some slack i (slack because $N(S) \cap Y_{\mathcal{O}} = \emptyset$); but since A_i is a 2-component of $\mathcal{E}^*$, this implies $\pi(y) \in A_i$ and $y \in U$, a contradiction. □

Finally, we observe that

$$\text{the edges in } \widehat{N(S)} \text{ with neighbors in } I \cap U \text{ are precisely those in } \cup^s \hat{G}_i. \tag{5.47}$$

(We have already noted in (5.21) that edges in $\cup^s \hat{G}_i$ do have such neighbors (in $\cup^s \hat{A}_i$), so what (5.47) really says is that the remaining edges in $\widehat{N(S)}$ do not. This is because there are no occupied vertices in $S \setminus \cup^s A_i$: by (b') each v in S has a neighbor in F, so in some slack G_i, so if occupied must lie in A_i.) Of course at this point we don't know the G_i's, but what we *can* use from (5.47) is

$$\text{exactly } g' \text{ edges in } \widehat{N(S)} \text{ have neighbors in } I \cap U \text{ (so in } I \cap \hat{S}). \tag{5.48}$$

Lemma 5.20. *The cost of labeling $S \cup N(S)$ is at most*

$$f + x - \Omega(\varepsilon t') \quad (= |N(S)| - \Omega(\varepsilon t')) \tag{5.49}$$

(where x is the size of X, which was defined two lines before (5.41)).

Proof. This is similar to the proof of Lemma 5.18. We again run [**Algorithm**] in two stages, but this time only on W (defined in (5.42)). As before we STOP the first run when

$$d_{X_i}(x) \le n^{2/3} \quad \forall x \in X_i,$$

and let Z_1 be the (final) X_i produced by this stage. We then run the algorithm on $Q_n[Z_1]$, in this case stopping as soon as either

(a) $d_{X_i}(x) \le n^{1/3}$ for all $x \in X_i$ or

(b) $|X_i| \le 2(f - t')$

(of course (b) is possible only if $f \geq t'$), and letting Z_2 be the final X_i.

As before: the ξ produced by the first run has ($|\xi| \leq |W| = 2(s+f)$ and) $|\text{supp}(\xi)| \leq 2(s+f)n^{-2/3}$, so Proposition 5.6 bounds the cost of this run by

$$(2+o(1))(s+f)n^{-2/3}\log(n^{2/3}) = o(\varepsilon t') \tag{5.50}$$

(using $s+f \leq 2g'$, as follows from (c′) and (5.41), with (5.23) and (5.31)); Proposition 5.7 with $Z = Z_1$, $d = n^{2/3}$ and L as in (5.45) gives

$$\begin{aligned} |Z_1| &< (2n - n^{2/3})^{-1}[2n(s+f) + 2(f-s)n + 4s\psi] \\ &= (2n - n^{2/3})^{-1}[4nf + 4s\psi] \\ &\leq 2f(1+n^{-1/3}) + O(s\psi/n) \\ &= 2f(1+n^{-1/3}) + o(\varepsilon t') \end{aligned} \tag{5.51}$$

(using $s\psi/n = O(g'/(\sqrt{n}\log n)) = o(\varepsilon t')$, which follows from (5.23) and (5.31); this is the reason for the lower bound in (5.31)); (a), (b) and (5.51), now with the ξ from the second run, imply

$$|\text{supp}(\xi)| \leq (z_1 - z_2)n^{-1/3} \leq r := \begin{cases} (2fn^{-1/3} + O(t'))n^{-1/3} & \text{if } f \geq t', \\ (2+o(1))t'n^{-1/3} & \text{if } f < t'; \end{cases}$$

Proposition 5.6 with this r and $l = |W| = O(f)$ (note (b') implies $s < (1+o(1))f$) bounds the run cost by

$$O((fn^{-2/3} + t'n^{-1/3})\log n + \log f) = o(\varepsilon t'), \tag{5.52}$$

with the $o(\varepsilon t')$ given by (5.23) (and $f \leq g'$); and Proposition 5.7, with $Z = Z_2$, $d = n^{1/3}$ and, again, L as in (5.45), gives (*cf.* (5.51))

$$\begin{aligned} |Z_2| &< (2n - n^{1/3})^{-1}[2n(s+f) + 2(f-s)n + 4s\psi] \\ &\leq 2f(1+n^{-2/3}) + o(\varepsilon t') \;=\; 2f + o(\varepsilon t') \end{aligned} \tag{5.53}$$

(again—as in (5.51)—using $s\psi/n = o(\varepsilon t')$).

Let $P = I \cap (W \setminus Z_2)$ (the set of vertices that were "processed" in the two runs of the algorithm and turned out to be in I), $X' = \hat{X} \setminus N(P)$, $Z' = Z_2 \cup X'$ and $\Gamma = Q_n[Z']$. So we are down to identifying $I \cap Z'$ (Z' being the set of vertices of U whose membership in I is still in question). Noting that

$$I \cap Z' \text{ is an MIS of } \Gamma \tag{5.54}$$

(see (5.46)) and recalling that the run costs in (5.50) and (5.52) were $o(\varepsilon t')$, we find that Lemma 5.20 will follow from

$$\text{the cost of identifying } I \cap Z' \text{ is at most } f + x - \Omega(\varepsilon t'). \tag{5.55}$$

(Note we are still enforcing (5.48).)

If $|Z'| \leq 2(f + x) - \Omega(\varepsilon t')$ then (5.55) is given by Theorem 1.2 (and (5.54)). In particular this is true if the second run ends because of (b), since then $|Z'| \leq z_2 + 2x \leq 2(f + x - t')$.

So we are left with cases where the run is stopped by (a) and

$$|Z'| > 2(f + x) - o(\varepsilon t'),$$

which by (5.53) implies $x' = 2x - o(\varepsilon t')$, i.e.

$$|\hat{X} \setminus X'| = o(\varepsilon t'). \tag{5.56}$$

But (5.48) and the fact that each edge of $\hat{F}$ has a neighbor in $I \cap \hat{S}$ imply that exactly $g' - f > \varepsilon t'$ edges in $\hat{X}$ have neighbors in $I \cap \hat{S}$, which with (5.56) yields

$$(1 - o(1))\varepsilon t' \text{ edges in } X' \text{ have neighbors in } Z_2 \cap I \cap \hat{S}. \tag{5.57}$$

Now let $M = M_\Gamma(I \cap Z')$. According to the definition of M_Γ (see (5.2)) no edge as in (5.57) can be in M (*cf.* (5.39)), so M fails to cover at least one vertex from each of these edges (since, M being induced, $V(M)$ meets any edge not in M at most once). But then $z' \leq 2(f + x) + o(\varepsilon t')$ (which follows from (5.53) and $z' \leq z_2 + 2x$) implies

$$m_\Gamma(I \cap Z') = |M| < (2(f + x) - (1 - o(1))\varepsilon t')/2 = f + x - \Omega(\varepsilon t'),$$

and a final application of Theorem 3.4 (with the above bound on z') again gives (5.55), completing the proof of Lemma 5.20. □

In sum (making precise the discussion following (5.28)), we have paid:

1. $O(t \log n/\sqrt{n})$ for the decompositions of a and g (see (5.25));
2. $g_i - \Omega(t_i)$ for specification and labeling of $[A_i]$ and G_i for each small i (see (5.29));
3. $O(\sum t_i \log^2 n/\sqrt{n})$ for the (S_i, F_i)'s, i large (see (5.30));
4. for each tight i, $g_i - \Omega(t_i)$ for specification and labeling of $[A_i]$ and G_i, given (S_i, F_i) (see (5.33));
5. $|N(S)| - \Omega(\varepsilon t')$ for labeling $S \cup N(S)$, given (S, F) (which is determined by the (S_i, F_i)'s, together with the G_i's for small and tight i); see (5.49).

Finally, the sum of all these cost bounds is at most

$$|N(S)| + \sum\{g_i : i \text{ small or tight}\} + O(\sum t_i \log^2 n/\sqrt{n}) - \Omega\left(\sum\{t_i : i \text{ small or tight}\}\right) - \Omega(\varepsilon t'),$$

which (recalling $t' = \sum\{t_i : i \text{ slack}\}$, $t \le \sum t_i$ and $\varepsilon = \omega(1/\log n)$) is at most

$$|N(S)| + \sum\{g_i : i \text{ small or tight}\} - \omega(t/\log n);$$

and combining this with the additional cost in (5.28) (paid for the remaining labels in $\mathcal{O}$) gives Lemma 5.16.

Chapter 6

The number of 4-colorings of the Hamming cube

6.1 Lower bound and task

As discussed earlier, the asymptotic value in Theorem 1.9 is an obvious *lower* bound. We quickly show how this goes. Let $\{1,2,3,4\}$ be our set of colors, $\mathcal{C}$ the set of (six) ordered equipartitions of this set, and $\mathcal{F}$ the set of (proper) 4-colorings of Q_d. Say $f \in \mathcal{F}$ *agrees* with $(C,D) \in \mathcal{C}$ at v if

$$f_v \in C \Leftrightarrow v \in \mathcal{E}.$$

For given (C,D) and a fixed k, the number of colorings that *dis*agree with (C,D) at precisely k vertices is asymptotically $\binom{N}{k}2^{N-dk} \sim 2^N/k!$ (the k exceptional vertices—*flaws*—will typically have disjoint neighborhoods, whose colors are determined by those of the flaws), and summing over choices of (C,D) and k gives the value in Theorem 1.8. (See also the more general discussion in [10, Sec. 6.1]; in particular the case $q = 4$ of their Conjecture 6.2 is our Theorem 1.8. For $q \geq 5$, colorings will typically have *many* flaws, and the conjectured asymptotics are for $\log C_q(Q_d)$ rather than $C_q(Q_d)$ itself.)

In fact for almost every f there is some $(C,D) \in \mathcal{C}$ with which f agrees on all but a tiny fraction of the vertices; this special case of Theorem 1.1 of [10] is our point of departure:

Theorem 6.1. *There is a fixed $\alpha < 2$ such that for all but $|\mathcal{F}|2^{-\Omega(d)}$ f's in $\mathcal{F}$ there is some $(C,D) \in \mathcal{C}$ such that*

$$|\{v \in V : f \text{ disagrees with } (C,D) \text{ at } v\}| < \alpha^d.$$

For f and (C,D) as in Theorem 6.1, call (C,D) the *main phase* of f. (So not every f

has a main phase, but the number that do not is negligible.) We then write X_f for the set of vertices that disagree with (C, D) at f and call such vertices *bad* (for f).

Say a coloring (with a main phase) is *ideal* if any two of its bad vertices are at distance at least 3. The preceding lower bound discussion extends to say that the number of ideal colorings is less than $6\sum_k \binom{N}{k} 2^{N-dk} < 6e2^N$. So the asymptotic number of ideal colorings is $6e2^N$ and for Theorem 1.8 we should show that the number of *non-ideal* colorings is $o(2^N)$, which in view of Theorem 6.1 will follow if we show that the number of *non*-ideal colorings with a given main phase (C, D) is $o(2^N)$. (In fact it is $2^{N-\Omega(d)}$—see following (6.3)—which with Theorem 6.1 gives a similar value for what's lost in the "∼" of Theorem 1.8.)

We may specialize a little further: Let $\mathcal{F}^*$ be the set of non-ideal f's having main phase $(\{1, 2\}, \{3, 4\})$ and satisfying

$$|N(X_f \cap \mathcal{E})| \geq |N(X_f \cap \mathcal{O})| \tag{6.1}$$

(where N is neighborhood). Then Theorem 1.8 will follow from

$$|\mathcal{F}^*| = o(2^N), \tag{6.2}$$

and the rest of this chapter is concerned with proving this.

The tools that we will use are not unexpected, as both have been important in earlier work on questions of the present type, but the way they are combined here seems interesting. Specifically, what's perhaps most interesting is the use of entropy *following* application of [57, 22] (see Section 6.3.3). This is in contrast to, e.g., the use in [22] of an entropy-based result from [33] as a sort of preprocessing step (echoed in the role of the entropy-based argument in [22] here). Something similar in spirit—though not in implementation—to what we do appears in a recent breakthrough of Peled and Spinka [53] (on colorings of $\mathbb{Z}^d$ and related statistical physics models), which partly inspired our approach.

6.2 Main point

In what follows f will always be a (proper) coloring of Q_d. We use f_u for the value of f at u and f_U for the restriction of f to U.

We almost always use *lower case letters for the cardinalities of the sets denoted by the corresponding upper case letters* (thus $a = |A|$, $\hat{a} = \hat{A}$ and so on), usually without comment.

For $f \in \mathcal{F}^*$, we denote by A_f and $\hat{A}_f$ the unions of (resp.) the nonsingleton and singleton 2-components of $X_f \cap \mathcal{E}$, and set $G_f = N(A_f)$ and $\hat{G}_f = N(\hat{A}_f)$. Set

$$\mathcal{F}^*(g, \hat{g}) = \{f \in \mathcal{F}^* : |G_f| = g, |\hat{G}_f| = \hat{g}\}.$$

The next lemma is almost all of the story.

Lemma 6.2.

$$2^{-N}|\mathcal{F}^*(g,\hat{g})| = \begin{cases} 2^{-\Omega(d)} & \textit{if } g = 0 \textit{ and } \hat{g} \le d^2/\log d, \\ \exp[-\Omega(\frac{g}{\log d} + \frac{\hat{g}}{d}\log\frac{\hat{g}}{d})] & \textit{otherwise.} \end{cases}$$

We close this section with the derivation of (6.2) from Lemma 6.2. The lemma itself is proved in Section 6.3.

Proof of (6.2). We show

$$2^{-N}|\mathcal{F}^*| = 2^{-\Omega(d/\log d)}, \tag{6.3}$$

which a little more care with the bounds in Lemma 6.2 (see the remark following "A^l terms" in Section 6.3.4) would improve to $2^{-\Omega(d)}$. With $\sum^*$ running over $(g, \hat{g})$ satisfying $g \neq 0$ or $\hat{g} > d^2/\log d$, the lemma gives

$$2^{-N}|\mathcal{F}^*| = 2^{-N}\sum\nolimits^* |\mathcal{F}^*(g,\hat{g})| + (d^2/\log d)2^{-\Omega(d)};$$

so we are just interested in the sum, which we may bound by

$$\frac{d^2}{\log d}\sum_{g \ge d} 2^{-\zeta g/\log d} + \left(1 + \sum_{g\ge d} 2^{-\zeta g/\log d}\right)\left(\sum_{\hat{g} \ge d^2/\log d} 2^{-\zeta(\hat{g}/d)\log(\hat{g}/d)}\right), \tag{6.4}$$

where $\zeta > 0$ is the implied constant in the second line of Lemma 6.2. (Of course if g is not zero then it is at least d.)

With $x = 2^{-\zeta/\log d}$, the first sum in (6.4) is

$$x^d \sum_{i\geq 0} x^i = x^d/(1-x) = x^d \cdot O(\log d) = 2^{-\Omega(d/\log d)}.$$

Similarly, with $y = 2^{-\zeta \log(d/\log d)/d}$, the last sum in (6.4) is less than

$$\sum_{\hat{g}\geq d^2/\log d} y^{\hat{g}} = y^{d^2/\log d}/(1-y) = 2^{-\Omega(d)}.$$

So we have (6.3). □

6.3 Proof

6.3.1 Orientation

We first spend a little time trying to motivate what's happening below, hoping this makes the discussion easier to follow. For purposes of comparison we begin with a standardish entropy-based bound.

Given $\mathcal{G} \subseteq \mathcal{F}^*$, set

$$\begin{aligned} T(u) = T_{\mathcal{G}}(u) &= \tfrac{1}{d}H(f_{N_u}) + H(f_u|f(N_u)) \\ &= \tfrac{1}{d}[H(f(N_u)) + H(f_{N_u}|f(N_u))] + H(f_u|f(N_u)), \end{aligned} \tag{6.5}$$

where f is uniform from $\mathcal{G}$. (We will use this only with $u \in \mathcal{O}$.) Then

$$\begin{aligned} \log|\mathcal{G}| = H(f) &= H(f_{\mathcal{E}}) + H(f_{\mathcal{O}}|f_{\mathcal{E}}) \\ &\leq \frac{1}{d}\sum_{u\in\mathcal{O}} H(f_{N_u}) + \sum_{u\in\mathcal{O}} H(f_u|f(N_u)) = \sum_{u\in\mathcal{O}} T(u). \end{aligned}$$

Here the first two equalities are given by (a) and (b) of Lemma 2.12 and the inequality by Lemmas 2.13 and 2.12 (c,d), the former with

$$\alpha_S = \begin{cases} 1/d & \text{if } S = N_u \text{ for some } u \in \mathcal{O}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.6}$$

On the other hand, for each possible value c of $f(N_u)$,

$$\begin{aligned} H(f_u|f(N_u) = c) &\leq \log(4 - |c|), \\ H(f_{N_u}|f(N_u) = c) &\leq d\log|c|. \end{aligned}$$

Since $\log x + \log(4-x) \le 2$, this bounds $\frac{1}{d}H(f_{N_u}|f(N_u)) + H(f_u|f(N_u))$ (the main part of (6.5)) by

$$\sum_c \mathbb{P}(f(N_u)=c)[\tfrac{1}{d}H(f_{N_u}|f(N_u)=c) + H(f_u|f(N_u)=c)] \le 2,$$

yielding

$$T(u) \le 2 + O(1/d) \tag{6.7}$$

(since $H(f(N_u)) = O(1)$).

In particular, applying this with $\mathcal{G} = \mathcal{F}^*$ gives the easy bound

$$\log|\mathcal{F}^*| \le N + O(N/d), \tag{6.8}$$

whereas we want $\log|\mathcal{F}^*| < N - \omega(1)$; so what we do below may be thought of as fighting over this difference. (Note this argument makes no use of the fact that members of $\mathcal{F}^*$ are non-ideal, so can't give a bound less than N.)

We now very briefly sketch the actual argument. We think of $|\mathcal{F}^*|$ as the number of ways to specify $f \in \mathcal{F}^*$, which we do in two stages. The first of these identifies a "template," $\mathcal{T} = \mathcal{T}_f$, which provides some, usually incomplete, information on X_f (recall this is the set of vertices that are bad for f). In fact $\mathcal{T}$ will completely specify $X_f \cap \mathcal{E}$, but the information on $X_f \cap \mathcal{O}$ will typically be less precise.

The second ("coloring") stage then treats possibilities for f given $\mathcal{T}$. Thus we restrict to a set $\mathcal{G}$ of f's satisfying $\mathcal{T}_f = \mathcal{T}$, usually with some "cheap" part of f also specified, and return to the entropy approach leading to (6.8). The hope—and basic idea of the proof—is that what we save in the above argument by exploiting information from the template recovers (more exactly, *more than* recovers) what we've paid for said information.

In what follows we usually speak in terms of the *cost* of a choice, meaning the log of the number of possibilities for that choice, which we think of as the number of bits "paid" for the desired information.

6.3.2 Templates

A template will consist of two parts, the first specifying $X_f \cap \mathcal{E}$ and the second corresponding to, but not necessarily precisely identifying, the portion of $X_f \cap \mathcal{O}$ not adjacent to $X_f \cap \mathcal{E}$. (For perspective we note that the asymmetry between $\mathcal{E}$ and $\mathcal{O}$ corresponds to (6.1) in the definition of $\mathcal{F}^*$, an assumption we will use frequently below.)

Names for the sets involved will now be helpful; for a particular f we use the following notation, with dependence on f suppressed (so $X = X_f$, $A_i = A_i(f)$ and so on).

A_i's: non-singleton 2-components of $X \cap \mathcal{E}$;

$\hat{A}_i$'s: singleton 2-components of $X \cap \mathcal{E}$;

$G_i = N(A_i)$, $\hat{G}_i = N(\hat{A}_i)$;

$A = \cup A_i$ and similarly for $\hat{A}$, G and $\hat{G}$ (as in the passage preceding Lemma 6.2);

$\mathcal{R} = \mathcal{O} \setminus (G \cup \hat{G})$;

P_i's: 2-components of $X \cap \mathcal{R}$ meeting $N^2(G \cup \hat{G})$;

$\bar{P}_i$'s: non-singleton 2-components of $X \cap \mathcal{R}$ not meeting $N^2(G \cup \hat{G})$;

$\hat{P}_i$'s: singleton 2-components of $X \cap \mathcal{R}$ not in $N^2(G \cup \hat{G})$;

$Q_i = N(P_i)$, $\bar{Q}_i = N(\bar{P}_i)$ and $\hat{Q}_i = N(\hat{P}_i)$;

$P = \cup P_i$, $Q = \cup Q_i$ etc.

(See figure 6.1.) Note that the vertices of $Q \cup \bar{Q} \cup \hat{Q}$, not being in $A \cup \hat{A}$, are all good, while the template does not usually distinguish good and bad vertices of $G \cup \hat{G}$. (The one exception to this is in the treatment of the special case (6.9) in Section 6.3.5.) Note also that the G_i's and $\hat{G}_i$'s are pairwise disjoint and similarly for the Q_i's, $\bar{Q}_i$'s and $\hat{Q}_i$'s.

Treatment of the contributions (to our overall cost) of the above pieces will depend on their sizes, necessitating some further decomposition, as follows. (Recall $a_i = |A_i|$ and so on.) Say

$$A_i \text{ is } \begin{cases} \textit{small} & \text{if } g_i < \exp_2[\log^3 d] \text{ and} \\ \textit{large} & \text{otherwise,} \end{cases}$$

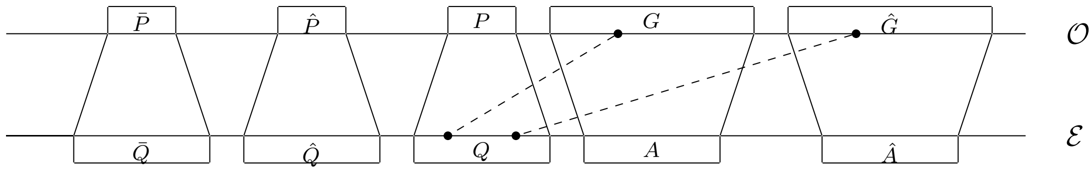

Figure 6.1:

and similarly for $\bar{P}_i$ and P_i. Let A^s, A^l be the unions of the small and large A_i's (resp.) and extend this notation in the natural ways; thus $G^s = N(A^s)$, P^l is the union of the large P_i's and so on. We also set $b = |B(A^l)|$ (see Section 2.2.3).

Remark. The choice $\exp_2[\log^3 d]$ is not delicate. The most serious constraint is in the discussion of (6.32), where we use $q^l = d^{\omega(\log d)}$. The other cutoffs could be smaller—we mainly need them to support application of Lemmas 2.8-2.9—but for simplicity we use one value for all.

Note that in proving Lemma 6.2 we are given g and $\hat{g}$. Analysis in Sections 6.3.4 and 6.3.5 will vary depending on these, but for now the discussion is general. It will be convenient to set $\mathbf{g} = g + \hat{g}$.

Before proceeding, we set aside the easy (but important) special case in which

$$a = \bar{p} = p = 0 \;\text{ and }\; \hat{g} \leq d^2/\log d. \tag{6.9}$$

This will be handled in Section 6.3.5, and *until then we restrict to f's that are not of this type.*

It will be helpful to have specified the sizes of some of the other sets above, namely

$$a^s \ (= |A^s|),\ g^s,\ p,\ q,\ \bar{p},\ \bar{q},\ \hat{p},\ i_{A^s},\ i_{\bar{P}^s},\ i_{P^s}, b,$$

which we may do at an (eventually negligible) cost of

$$O(\log \mathbf{g}). \tag{6.10}$$

(Most of these could be skipped, but it's easier to pay the above negligible cost up front than to waste time on this issue.)

We begin with costs associated with the non-large sets above. These choices are mostly treated as if made autonomously; that is, without trying to exploit proximity or non-proximity of different pieces. The one exception is in the cost of P^s, where we sometimes save substantially by choosing initial vertices for the 2-components from $N^2(G \cup \hat{G})$ rather than all of $\mathcal{O}$.

Claim 6.3. *The costs of identifying* $\hat{A}$,A^s,$\hat{P}$,$\bar{P}^s$ *and* P^s *are bounded by:*

[$\hat{A}$] $\log \binom{N/2}{\hat{a}} \leq \hat{g} - \Omega((\hat{g}/d)) \log(\hat{g}/d))$ *(using* $\hat{g} = \hat{a}d$*);*

[A^s] $i_{A^s}(d-1) + O(a^s \log d)$*;*

[$\hat{P}$] $\log \binom{N/2}{\hat{p}} \leq \hat{p}d = \hat{q}$*;*

[$\bar{P}^s$] $i_{\bar{P}^s}(d-1) + O(\bar{p}^s \log d)$*;*

[P^s] $i_{P^s} \log(e\mathbf{g}d^2/i_{P^s}) + O(p^s \log d)$.

Proof. The first and third of these are trivial and the others are instances of Proposition 2.5, with some relaxation of bounds. We use $Y = \mathcal{O}$ ($|Y| = N/2$) for all but [P^s], where, as mentioned above, we save significantly by taking $Y = N^2(G \cup \hat{G})$. □

For larger pieces we have the following bounds, which, in contrast to the elementary Claim 6.3, depend on the sophisticated results of Section 2.2.3.

Lemmas 2.9 and 2.8 bound the costs of A^l and $\bar{Q}^l$ by

$$g^l - b - \Omega(g^l/\log d) \tag{6.11}$$

and

$$\bar{q}^l - \Omega(\bar{q}^l/\log d). \tag{6.12}$$

(E.g. for (6.11): we first pay $O(g^l \log d/d)$ for the list of g_i's and (with the obvious meaning) b_i's corresponding to large A_i's (the cost bound given by Proposition 2.1, using $i_{A^l} < g^l/d$), and then apply Lemma 2.9 to the pieces, absorbing the initial $O(g^l \log d/d)$ and the 2^d from the lemma in the "Ω" term of (6.11).)

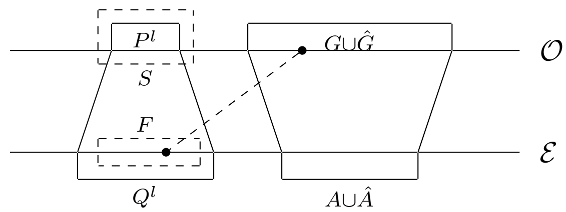

Figure 6.2: S and F

For P^l, perhaps the most interesting part of this story, the cost of full specification turns out to be more than we can afford, and we retreat to the approximations of Lemma 2.10. (As mentioned earlier, Lemma 2.10 was originally a *step* in the proof of Lemma 2.8; so its present appearance in a non-auxiliary role seems interesting.)

Here again we pay an initial

$$O(q^l \log d/d) \tag{6.13}$$

for $(q_i : i \in I)$, where I indexes the large P_i's. Then for $i \in I$ we slightly modify the output of Lemma 2.10 (applied here with the roles of $\mathcal{E}$ and $\mathcal{O}$ reversed), letting $(S_i', F_i) = \varphi_i(P_i) \in \mathcal{W}_i$, with $\varphi_i = \varphi_{q_i}$ and $\mathcal{W}_i = \mathcal{W}(q_i)$ as in the lemma, and setting $S_i = S_i' \setminus (G \cup G')$ (see Figure 2). Note (S_i, F_i) still enjoys the properties the lemma promised for (S_i', F_i); that is,

$$S_i \supseteq [P_i], \quad F_i \subseteq Q_i, \tag{6.14}$$

$$d_{F_i}(u) \geq d - d/\log d \quad \forall u \in S_i. \tag{6.15}$$

(The only thing to observe here—used for the first part of (6.14)—is that $[P_i] \cap (G \cup \hat{G}) = \emptyset$ follows from $N(P_i) \cap (A \cup \hat{A}) = \emptyset$. Incidentally, $S \supseteq P_i$ in (6.14) would be enough for our purposes.) Note also that the S_i's are pairwise disjoint (by (6.15) since the second part of (6.14) implies the F_i's are pairwise disjoint) and that, with $S = \cup S_i$, $F = \cup F_i$,

$$S \cap (G \cup \hat{G}) = \emptyset = F \cap (\bar{Q} \cup \hat{Q} \cup Q^s). \tag{6.16}$$

Lemma 2.10 bounds the cost of specifying the (S_i, F_i)'s by

$$O(q^l \log^2 d/d), \tag{6.17}$$

which absorbs the decomposition cost (6.13).

This completes the template stage (apart from the treatment of (6.9)). Formally—but we won't actually use this—we could say that $\mathcal{T} = \mathcal{T}_f$ is $(A, \hat{A}, \bar{P}, \hat{P}, P^s, S, F)$. (Note A determines A^s, A^l and similarly for $\bar{P}$.)

6.3.3 Colors

First notice that each $\bar{P}_i$, $\bar{Q}_i$, $\hat{P}_i$, $\hat{Q}_i$, P_i^s, Q_i^s (these being, of course, the 2-components of P^s and Q^s) and F_i is monochromatic. (The general observation is: if Z is a 2-linked subset of $\mathcal{O}$ or $\mathcal{E}$, all vertices of Z are bad and all vertices of $N(Z)$ are good, then each of Z, $N(Z)$ is monochromatic (and the color for Z determines the color for $N(Z)$).)

So we begin by paying

$$i_P + i_{\bar{P}} + \hat{p} \tag{6.18}$$

to specify the colors of these sets. (These are the "cheap" color choices mentioned earlier.) *We then restrict our discussion to the set $\mathcal{G}$ of f's agreeing with these specifications* (and the specified $\mathcal{T}$).

For appraising the cost of identifying a member of $\mathcal{G}$, we refine the discussion leading to (6.8). To begin, we will in each instance consider $T(u)$ (defined in (6.5)) only for the u's in some subset, say $\mathcal{U}$, of $\mathcal{O}$, with the rest of α (as in Lemma 2.13; *cf.* (6.6)) supported on singletons. (We use u and v for vertices of $\mathcal{O}$ and $\mathcal{E}$ respectively.) Thus we use

$$H(f) \;\leq\; \sum_{u\in\mathcal{U}} T(u) + \sum_{u\in\mathcal{O}\setminus\mathcal{U}} H(f_u|f(N_u)) + \sum_{v\in\mathcal{E}}(1 - d_{\mathcal{U}}(v)/d)H(f_v). \tag{6.19}$$

As noted earlier, $\mathcal{T}$ includes specification of $X_f \cap \mathcal{E}$, so we know which vertices of $\mathcal{E}$ are bad for f. A key ingredient in evaluating the first term in (6.19) is then the following variant of (6.7), in which—just to point out that this doesn't require uniform distribution—$T_\mu(u)$ is the natural generalization of $T_{\mathcal{G}}(u)$ to the probability distribution μ.

Proposition 6.4. *If $X \cup Y$ is a partition of N_u with $X, Y \neq \emptyset$, and f is chosen from some probability distribution μ on the set of colorings for which X is entirely good and Y entirely bad, then*

$$T_\mu(u) \leq 1 + O(1/d). \tag{6.20}$$

Proof. This is similar to the derivation of (6.7). Notice that $|f(N_u)|$ must be either 2 or 3 (it is at least 2 by our assumption on X, Y and at most 3 since $f(N_u) \not\ni f_u$), and that

$$f(N_u) \text{ determines } \begin{cases} f_{N_u} & \text{if } |f(N_u)| = 2, \\ f_u & \text{if } |f(N_u)| = 3, \end{cases}$$

so that $H(f_{N_u}|f(N_u) = c) = 0$ if $|c| = 2$ and $H(f_u|f(N_u) = c) = 0$ if $|c| = 3$. Moreover,

$$H(f_u|f(N_u) = c) \leq 1 \quad \text{if } |c| = 2,$$
$$H(f_{N_u}|f(N_u) = c) \leq d \quad \text{if } |c| = 3$$

(the d could be replaced by $\max\{|X|, |Y|\}$).

Thus $\frac{1}{d}H(f_{N_u}|f(N_u)) + H(f_u|f(N_u))$ (the main part of (6.5)) is

$$\textstyle\sum_c \mathbb{P}(f(N_u) = c)[\frac{1}{d}H(f_{N_u}|f(N_u) = c) + H(f_u|f(N_u) = c)] \leq 1,$$

and Proposition 6.4 follows since $H(f(N_u)) = O(1)$. □

Of course knowing $X_f \cap \mathcal{E}$ also bounds the last sum in (6.19) by

$$\sum_{v \in \mathcal{E}} (1 - d_{\mathcal{U}}(v)/d) = N/2 - |\mathcal{U}|,$$

which in cases where $\mathcal{G}$ specifies some of the f_v's, say those in $\mathcal{V} \subseteq \mathcal{E}$, improves to

$$N/2 - |\mathcal{U}| - \sum_{v \in \mathcal{V}} (1 - d_{\mathcal{U}}(v)/d) = N/2 - |\mathcal{U}| - |\nabla(\mathcal{V}, \mathcal{O} \setminus \mathcal{U})|/d. \tag{6.21}$$

So we will be evaluating (6.19) using (6.20) and (6.21) (with a small assist from (6.7)). From this point we take

$$\mathcal{U} = G \cup \hat{G} \ \text{ and } \ \mathcal{V} = \bar{Q} \cup \hat{Q} \cup Q^s \cup F$$

(so also $\mathcal{O} \setminus \mathcal{U} = \mathcal{R}$; recall colors for $\mathcal{V}$ were specified at (6.18)). We then have the following bounds for the three sums in (6.19).

The combination of (6.7) and (6.20) bounds the first by

$$g + b + \hat{g} + O(\mathbf{g}/d) \tag{6.22}$$

(using (6.7) for $u \in B$ and (6.20) for the rest).

We next claim that the second is at most

$$s \log 3 + N/2 - (g + \hat{g} + s) = N/2 - (g + \hat{g}) + s(\log 3 - 1) \tag{6.23}$$

(where $s = |S|$). Here we use $S \cap (G \cup \hat{G}) = \emptyset$ (see (6.16)) and

$$H(f_u|f(N_u)) \le H(f_u) \le \begin{cases} 1 & \text{if } u \in \mathcal{O} \setminus (G \cup \hat{G} \cup S), \\ \log 3 & \text{if } u \in S. \end{cases}$$

The second bound is trivial. For the first notice that we actually *know* f_u if $u \in \bar{P} \cup \hat{P} \cup P^s$ and in other cases know u is good (using $P^l \subseteq S$).

Finally, the last term in (6.19) is at most $N/2 - |\mathcal{U}| - |\nabla(\mathcal{V}, \mathcal{R})|/d$ (see (6.21)), which we rewrite as

$$N/2 - (g + \hat{g}) - [\bar{q} + \hat{q} + |\nabla(Q^s, \mathcal{R})|/d + |\nabla(F, \mathcal{R})|/d] \tag{6.24}$$

(using $N(\bar{Q} \cup \hat{Q}) \subseteq \mathcal{R}$ and $F \cap (\bar{Q} \cup \hat{Q} \cup Q^s) = \emptyset$ (see (6.16))).

6.3.4 In sum

It remains to check that the above cost bounds give Lemma 6.2 (in cases not covered by (6.9)). We are now playing the game mentioned near the end of Section 6.3.1, in which we try to balance costs from the template stage against what we have gained (relative to (6.8)) in the coloring stage (and need to come out slightly ahead).

The bounds are: from the template stage, (6.10) and the more serious bounds in Claim 6.3, (6.11), (6.12) and (6.17); and from the coloring stage, the minor (6.18) and the non-minor (6.22)-(6.24). We will recall the template bounds as we come to them. The total cost from the coloring stage is bounded by

$$N + b - (g + \hat{g}) + s(\log 3 - 1) - (\bar{q} + \hat{q}) - |\nabla(Q^s \cup F, \mathcal{R})|/d + O(\mathbf{g}/d), \tag{6.25}$$

gotten by summing (6.22)-(6.24) and absorbing (6.18) in the $O(\mathbf{g}/d)$.

Note that both the $O(\mathbf{g}/d)$ in (6.25) and the $O(\log \mathbf{g})$ in (6.10) are negligible relative to the bounds in Lemma 6.2. (The comparison is least drastic when $g = 0$ and $\hat{g}$ is not much more than $d^2/\log d$.) So we may safely ignore these terms and in particular, rearranging and slightly expanding, replace (6.25) by

$$N - \hat{g} - \hat{q} - g^s - g^l + b - \bar{q}^s - \bar{q}^l + s(\log 3 - 1) - |\nabla(Q^s \cup F, \mathcal{R})|/d. \tag{6.26}$$

The initial N will of course cancel the 2^{-N} in Lemma 6.2, and we want to show that the combination of the remaining terms in (6.26) and the template costs produces the savings the lemma promises. We consider terms in groups of two or three corresponding to the different constituents of the template, following the order in (6.26), with the expressions in curly brackets below representing template costs and those immediately following them taken from (6.26) (and the right hand sides the bounds we will use). We first collect all these bounds and then take stock.

$\underline{\hat{A}\text{ terms}}$: $\{\hat{g} - \Omega((\hat{g}/d))\log(\hat{g}/d))\} - \hat{g} = -\Omega((\hat{g}/d))\log(\hat{g}/d))$

$\underline{\hat{P}\text{ terms}}$: $\{\hat{q}\} - \hat{q} = 0$

$\underline{A^s\text{ terms}}$: $\{i_{A^s}(d-1) + O(a^s \log d)\} - g^s \le -(1/2 - o(1))g^s$

(since $g^s \ge \max\{2i_{A^s}(d-1), \Omega(a^s d/\log^2 d)\}$, the second bound by Lemma 2.7)

$\underline{A^l\text{ terms}}$: $\{g^l - b - \Omega(g^l/\log d)\} - g^l + b = -\Omega(g^l/\log d)$

Remark. Using the last two bounds, we could replace the second bound in Lemma 6.2 by $\exp[-\Omega(g^s + g^l/\log d + (\hat{g}/d)\log(\hat{g}/d))]$ and the bound in (6.3) by $2^{-\Omega(d)}$.

$\underline{\bar{P}^s\text{ terms}}$: $\{i_{\bar{P}^s}(d-1) + O(\bar{p}^s \log d)\} - \bar{q}^s \le -(1/2 - o(1))\bar{q}^s$

(as for the A^s terms).

$\underline{\bar{P}^l\text{ terms}}$: $\{\bar{q}^l - \Omega(\bar{q}^l/\log d)\} - \bar{q}^l \le 0$

The P terms require a little more care. Here we will sometimes incur a small loss—that is, a positive contribution—but can live with this provided these losses are negligible relative to

$$\mathbf{g}\min\{d^{-1}\log(\mathbf{g}/d),(\log d)^{-1}\} \tag{6.27}$$

since our current gain from $\hat{A}$, A^s and A^l is at least of this order. Recall from (6.24) that the last term in (6.26) is the same as $[|\nabla(Q^s,\mathcal{R})|+|\nabla(F,\mathcal{R})|]/d$.

$\underline{P^s \text{ terms}}$:

$$\{i_{P^s}\log(e\mathbf{g}d^2/i_{P^s})+O(p^s\log d)\}-|\nabla(Q^s,\mathcal{R})|/d. \tag{6.28}$$

Lemma 2.7 and the definition of "small" give

$$q^s>(1-o(1))p^s d/\log^2 d. \tag{6.29}$$

Set $k=\log_d q^s$ and suppose first that $k=d^{o(1)}$. Then Lemma 2.7 gives $|N(Q^s)|=\Omega(q^s d/k)$, implying that either

$$q^s=O(\mathbf{g}k/d) \tag{6.30}$$

or

$$|\nabla(Q^s,\mathcal{R})|\geq|N(Q^s)|-|\mathcal{U}|=\Omega(q^s d/k). \tag{6.31}$$

But if (6.30) holds then $i_{P^s}\leq q^s/d$ and (6.29) imply that the positive terms in (6.28) are negligible relative to (6.27). (Note this uses the fact that $x\log(A/x)$ is increasing on $(0,A/e]$.) If, on the other hand, (6.30) does not hold then by (6.31) those positive terms are dominated by the negative term.

If k is larger, then $\mathbf{g}\geq q^s$ implies that the first and second terms in (6.28) are (respectively) $O((\mathbf{g}/d)\log d)$ and (again using (6.29)) $O((\mathbf{g}/d)\log^3 d)$, both of which are dwarfed by the expression in (6.27).

$\underline{P^l \text{ terms}}$:

$$\{O(q^l\log^2 d/d)\}+s(\log 3-1)-|\nabla(F,\mathcal{R})|/d \tag{6.32}$$

Assuming $q^l\neq 0$, we have $\mathbf{g}\geq q^l\geq\exp_2[\log^3 d]$, so the first term in (6.32) is negligible relative to (6.27). On the other hand, (6.15) and $S\cap\mathcal{U}=\emptyset$ (see (6.16)) give

$$|\nabla(F,\mathcal{R})|/d\geq|\nabla(F,S)|/d\geq(1-1/\log d)s,$$

so the sum of the last two terms in (6.32) is at most $-(2-\log 3-1/\log d)s$.

Summary. In the second case of Lemma 6.2, the above gains from A and $\hat{A}$ give the promised bound (or the stronger $-\Omega(g^s+g^l/\log d+(\hat{g}/d)\log(\hat{g}/d))$ mentioned earlier).

If we are in the first case, the desired gain comes from $\bar{P}^s$ and/or P^s (at least one of which must be nonempty since we assume (6.9) does not hold; note $\mathbf{g} \le d^2/\log d$ implies $\bar{P}^l = P^l = \emptyset$). If $\bar{P}^s \neq \emptyset$ then the gain is at least $(1/2-o(1))\bar{q}^s = \Omega(d)$. If $P^s \neq \emptyset$, then we note that (6.30) is impossible, since $\mathbf{g} \le d^2/\log d$ and $q^s > d$; so (6.31) holds and we gain $\Omega(q^s d/k) = \Omega(d)$.

6.3.5 Finally

We return to the exceptional case (6.9), which we recall:

$$a = \bar{p} = p = 0 \ \text{ and } \ \hat{g} \le d^2/\log d. \tag{6.33}$$

Notice that if the first part of this holds then we must have

$$X \cap \hat{G} \neq \emptyset$$

(where $X = X_f$), since otherwise f is not ideal (so is not in $\mathcal{F}^*$). So it is enough to show that for each $x \in [1,\hat{g}]$, the number of possibilities for $f \in \mathcal{F}^*$ satisfying (6.33) and $|X_f \cap \hat{G}| = x$ is (suitably) small.

To begin (given x) we pay

$$\log\binom{N/2}{\hat{a}} + \log(\hat{g}/d) + \log\binom{N/2}{\hat{p}} + \log\binom{\hat{g}}{x} < \hat{g} + \hat{q} + O(x\log d) \tag{6.34}$$

for $\hat{A}$, $\hat{p}$, $\hat{P}$ and $X \cap \hat{G}$. We then assign colors to $\hat{A} \cup (X \cap \hat{G}) \cup \hat{P}$, noting that these determine the restriction of f to $\hat{G} \cup N(X \cap \hat{G}) \cup \hat{Q}$ (since $u \in \hat{G} \setminus X$ is colored by whichever of $3, 4$ is not assigned to its neighbor in $\hat{A}$, and similarly for $v \in (N(X \cap \hat{G}) \setminus \hat{A}) \cup \hat{Q}$). Thus, since vertices whose colors are not determined by these choices are good, the total coloring cost is at most

$$\hat{a} + x + \hat{p} + N - [\hat{g} + |N(X \cap \hat{G})| + \hat{q}] = N - [\hat{g} + \hat{q} + \Omega(xd)].$$

(For the r.h.s. note that $N(X \cap \hat{G}) \cap \hat{Q} = \emptyset$ (by the definition of $\hat{P}$) and that the bound on $\hat{g}$ in (6.33) implies $|N(X \cap \hat{G})| = \Omega(xd)$ (by Lemma 2.7) and $(\hat{p} \le)\ \hat{a} \le d/\log d$.)

Finally, combining with (6.34) and summing bounds the number of f's satisfying (6.33) by

$$\sum_{x\geq 1} 2^{N-\Omega(xd)} = 2^{N-\Omega(d)}.$$

Chapter 7

Thresholds versus fractional expectation-thresholds

7.1 More introduction

7.1.1 Thresholds

For a given finite set X and $p \in [0,1]$, μ_p is the product measure on 2^X (the power set of X) given by $\mu_p(S) = p^{|S|}(1-p)^{|X\setminus S|}$. An $\mathcal{F} \subseteq 2^X$ is *increasing* if $B \supseteq A \in \mathcal{F} \Rightarrow B \in \mathcal{F}$. If this is true (and $\mathcal{F} \neq 2^X, \emptyset$), then $\mu_p(\mathcal{F})(:= \sum\{\mu_p(S) : S \in \mathcal{F}\})$ is strictly increasing in p, and the *threshold*, $p_c(\mathcal{F})$, is the unique p for which $\mu_p(\mathcal{F}) = 1/2$. This is finer than the original Erdős–Rényi notion, according to which $p^* = p^*(n)$ is ***a*** threshold for $\mathcal{F} = \mathcal{F}_n$ if $\mu_p(\mathcal{F}) \to 0$ when $p \ll p^*$ and $\mu_p(\mathcal{F}) \to 1$ when $p \gg p^*$. (That $p_c(\mathcal{F})$ *is* always an Erdős–Rényi threshold follows from [4].)

Following [64, 65, 67], we say $\mathcal{F}$ is *p-small* if there is a $\mathcal{G} \subseteq 2^X$ such that $\mathcal{F} \subseteq \langle\mathcal{G}\rangle := \{T : \exists S \in \mathcal{G}, S \subseteq T\}$ and

$$\textstyle\sum_{S\in\mathcal{G}} p^{|S|} \leq 1/2. \tag{7.1}$$

Then $q(\mathcal{F}) := \max\{p : \mathcal{F} \text{ is } p\text{-small}\}$, which we call the *expectation-threshold* of $\mathcal{F}$ (note the term is used slightly differently in [35]), is a trivial lower bound on $p_c(\mathcal{F})$, since for $\mathcal{G}$ as above and T drawn from μ_p,

$$\mu_p(\mathcal{F}) \leq \mu_p(\langle\mathcal{G}\rangle) \leq \textstyle\sum_{S\in\mathcal{G}} \mu_p(T \supseteq S) = \sum_{S\in\mathcal{G}} p^{|S|} \quad (= \mathbb{E}[|\{S \in \mathcal{G} : S \subseteq T\}|]). \tag{7.2}$$

The following statement, the main conjecture (Conjecture 1) of [35], says that for *any* $\mathcal{F}$, this trivial lower bound on $p_c(\mathcal{F})$ is close to the truth.

Conjecture 7.1. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq Kq(\mathcal{F}) \log |X|.$$

We should emphasize how strong this is (from [35]: "It would probably be more sensible to conjecture that it is *not* true"). For example, it easily implies—and was largely motivated by—Erdős–Rényi thresholds for (a) perfect matchings in random r-uniform hypergraphs, and (b) appearance of a given bounded degree spanning tree in a random graph. These have since been resolved: the first—*Shamir's Problem*, circa 1980—in [32], and the second—a mid-90's suggestion of Kahn—in [51]. Both arguments are difficult and specific to the problems they address (e.g. they are utterly unrelated either to each other or to what we do here). See Section 7.7 for more on these and other consequences.

Talagrand [64, 67] suggests relaxing "p-small" by replacing the set system $\mathcal{G}$ above by what we may think of as a *fractional* set system, g: say $\mathcal{F}$ is *weakly p-small* if there is a $g : 2^X \to \mathbb{R}^+$ such that

$$\textstyle\sum_{S\subseteq T} g(S) \geq 1 \quad \forall T \in \mathcal{F} \quad \text{and} \quad \sum_{S\subseteq X} g(S)p^{|S|} \leq 1/2.$$

Then $q_f(\mathcal{F}) := \max\{p : \mathcal{F} \text{ is weakly } p\text{-small}\}$, the *fractional expectation-threshold* of $\mathcal{F}$, satisfies

$$q(\mathcal{F}) \leq q_f(\mathcal{F}) \leq p_c(\mathcal{F}) \tag{7.3}$$

(the first inequality is trivial and the second is similar to (7.2)), and Talagrand [67, Conjectures 8.3 and 8.5] proposes a sort of LP relaxation of Conjecture 7.1, and then a strengthening thereof. The first of these, the following, replaces q by q_f in Conjecture 7.1; the second, which adds replacement of $|X|$ by the smaller $\ell(\mathcal{F})$, is our Theorem 1.10.

Conjecture 7.2. *There is a universal K such that for every finite X and increasing $\mathcal{F} \subseteq 2^X$,*

$$p_c(\mathcal{F}) \leq K q_f(\mathcal{F}) \log |X|.$$

Talagrand further suggests the following "very nice problem of combinatorics," which implies *equivalence of* Conjectures 7.1 and 7.2, as well as of Theorem 1.10 and the corresponding strengthening of Conjecture 7.1.

Conjecture 7.3. *There is a universal K such that, for any increasing $\mathcal{F}$ on a finite set X, $q(\mathcal{F}) \geq q_f(\mathcal{F})/K$.*

(That is, weakly p-small implies (p/K)-small.)

Note the interest here is in Conjecture 7.3 for its own sake and as the most likely route to Conjecture 7.1; all applications of the latter that we are aware of follow just as easily from Theorem 1.10.

7.1.2 Spread hypergraphs and spread measures.

In what follows a *hypergraph* on the (*vertex*) set X is a collection $\mathcal{H}$ of subsets of X (*edges* of $\mathcal{H}$), with *repeats allowed.* For $S \subseteq X$, we use $\langle S\rangle$ for $\{T \subseteq X : T \supseteq S\}$, and for a hypergraph $\mathcal{H}$ on X, we write $\langle \mathcal{H}\rangle$ for $\cup_{S\in\mathcal{H}}\langle S\rangle$. We say $\mathcal{H}$ is *ℓ-bounded* (resp. *ℓ-uniform* or an *ℓ-graph*) if each of its members has size at most (resp. exactly) ℓ, and *κ-spread* if

$$|\mathcal{H} \cap \langle S\rangle| \leq \kappa^{-|S|}|\mathcal{H}| \quad \forall S \subseteq X. \tag{7.4}$$

(Note that edges are counted with multiplicities on both sides of (7.4).)

A major advantage of the fractional versions (Conjecture 7.2 and Theorem 1.10) over Conjecture 7.1—and the source of the present relevance of [2]—is that they admit, via linear programming duality, reformulations in which the specification of $q_f(\mathcal{F})$ gives a usable starting point. Following [67], we say a probability measure ν on 2^X is *q-spread* if

$$\nu(\langle S\rangle) \leq q^{|S|} \quad \forall S \subseteq X.$$

Thus a hypergraph $\mathcal{H}$ is κ-spread iff uniform measure on $\mathcal{H}$ is q-spread with $q = \kappa^{-1}$.

As observed by Talagrand [67], the following is an easy consequence of duality.

Proposition 7.4. *For an increasing family $\mathcal{F}$ on X, if $q_f(\mathcal{F}) \leq q$, then there is a $(2q)$-spread probability measure on 2^X supported on $\mathcal{F}$.* □

This allows us to reduce Theorem 1.10 to the following alternate (actually, equivalent) statement. In this chapter *with high probability* (w.h.p.) means with probability tending to 1 as $\ell \to \infty$.

Theorem 7.5. *There is a universal K such that for any ℓ-bounded, κ-spread hypergraph $\mathcal{H}$ on X, a uniformly random $((K\kappa^{-1}\log \ell)|X|)$-element subset of X belongs to $\langle \mathcal{H} \rangle$ w.h.p.*

The easy reduction is given in Section 7.2.

7.1.3 Assignments

The second main result of this chapter provides upper bounds on the minima of a large class of hypergraph-based stochastic processes, somewhat in the spirit of [66] (see also [65, 68]), saying that in "smoother" settings, the logarithmic corrections of Conjecture 7.2 and Theorem 1.10 are not needed.

For a hypergraph $\mathcal{H}$ on X, let ξ_x $(x \in X)$ be independent random variables, each uniform from $[0, 1]$, and set

$$\xi_{\mathcal{H}} = \min_{S \in \mathcal{H}} \sum_{x \in S} \xi_x \tag{7.5}$$

and $Z_{\mathcal{H}} = \mathbb{E}[\xi_{\mathcal{H}}]$.

Theorem 7.6. *There is a universal K such that for any ℓ-bounded, κ-spread hypergraph $\mathcal{H}$, we have $Z_{\mathcal{H}} \le K\ell/\kappa$, and $\xi_{\mathcal{H}} \le K\ell/\kappa$ w.h.p.*

These bounds are often tight (again up to the value of K). The distribution of the ξ_x's is not very important; e.g. it's easy to see that the same statement holds if they are Exp(1) random variables, as in the next example.

Theorem 7.6 was motivated by work of Frieze and Sorkin [20] on the "axial" version of the *random d-dimensional assignment problem.* This asks (for fixed d and large n) for estimation of

$$Z_d^A(n) = \mathbb{E}\left[\min \sum_{x \in S} \xi_x\right], \tag{7.6}$$

where the ξ_x's $(x \in X := [n]^d)$ are independent Exp(1) weights and S ranges over "axial assignments," meaning $S \subseteq X$ meets each *axis-parallel hyperplane* $(\{x \in X : x_i = a\}$ for some $i \in [d]$ and $a \in [n])$ exactly once. For $d = 2$ this is classical; see [20] for its rather glorious history. For $d = 3$ the deterministic version was one of Karp's [41] original

NP-complete problems. Progress on the random version has been limited; see [20] for a guide to the literature.

Frieze and Sorkin show (regarding bounds; they are also interested in algorithms) that for suitable $c_1 > 0$ and c_2,

$$c_1 n^{-(d-2)} < Z_d^A(n) < c_2 n^{-(d-2)} \log n. \tag{7.7}$$

(The lower bound is easy and the upper bound follows from the Shamir bound of [32].)

In present language, $Z_d^A(n)$ is essentially (that is, apart from the difference in the distributions of the ξ_x's) $Z_{\mathcal{H}}$, with $\mathcal{H}$ the set of perfect matchings of the complete, balanced d-uniform d-partite hypergraph on dn vertices (that is, the collection of d-sets meeting each of the pairwise disjoint n-sets $V_1, \ldots, V_d$). This is easily seen to be κ-spread with $\kappa = (n/e)^{d-1}$ (apart from the nearly irrelevant d-particity, it is the $\mathcal{H}$ of Shamir's Problem), so the correct bound is an instance of Theorem 7.6:

Corollary 7.7. $Z_d^A(n) = \Theta(n^{-(d-2)})$.

Frieze and Sorkin also considered the "planar" version of the problem, in which S in (7.6) meets each *line* ($\{x \in X : x_j = y_j \ \forall j \neq i\}$ for some $i \in [d]$ and $y \in X$) exactly once; and one may of course generalise from hyperplanes/lines to k-dimensional "subspaces" for a given $k \in [d-1]$. It's easy to see what to expect here, and one may hope Theorem 7.6 will eventually apply, but we at present lack the technology to say the relevant $\mathcal{H}$'s are suitably spread.

7.2 Little things

Usage. As is usual, we use $\binom{X}{r}$ for the family of r-element subsets of X, and $[S,T]$ for $\{R : S \subseteq R \subseteq T\}$. Our default universe is X, with $|X| = n$.

In what follows we assume ℓ and n are somewhat large (when there is an ℓ it will be at most n), as we may do since smaller values can by handled by adjusting the K's in Theorems 7.5 and 7.6. Asymptotic notation referring to some parameter λ (usually ℓ) is used in the natural way: implied constants in $O(\cdot)$ and $\Omega(\cdot)$ are independent of λ, and $f = o(g)$ (also written $f \ll g$) means f/g is smaller than any given $\varepsilon > 0$ for large enough values of λ. Following a standard abuse, we usually pretend large numbers are integers.

For $p \in [0,1]$ and $m \in [n]$, X_p and X_m are (respectively) a p-random subset of X (drawn from μ_p) and a uniformly random m-element subset of X. The latter is not entirely kosher, since we will also see *sequences* X_i; but we will never see both interpretations in close proximity, and the overlap should cause no confusion.

In a couple places it will be helpful to assume uniformity, which we will justify using the next little point.

Observation 7.8. *If $\mathcal{H}$ is ℓ-bounded and κ-spread, and we replace each $S \in \mathcal{H}$ by M new edges, each consisting of S plus $\ell - |S|$ new vertices (each used just once), then for large enough M the resulting ℓ-graph $\mathcal{G}$ is again κ-spread.*

Derivation of Theorem 1.10 from Theorem 7.5. Let $\mathcal{F}$ be as in Theorem 1.10 with $\mathcal{G}$ its set of minimal elements, let ℓ with $\ell(\mathcal{F}) \leq \ell = O(\ell(\mathcal{F}))$ be large enough that the exceptional probability in Theorem 7.5 is less than $1/4$ and let ν be the $(2q)$-spread probability measure promised by Proposition 7.4, where $q = q_f(\mathcal{F})$. We may assume ν is supported on $\mathcal{G}$ (since transferring weight from S to $T \subseteq S$ doesn't destroy the spread condition) and that ν takes values in $\mathbb{Q}$. We may then replace $\mathcal{G}$ by $\mathcal{H}$ whose edges are copies of edges of $\mathcal{G}$, and ν by uniform measure on $\mathcal{H}$.

Setting $m = ((2Kq\log\ell)n)$ and $p = 2m/n$ (with $n = |X|$ and K as in Theorem 7.5), we then have (using Theorem 7.5 with $\kappa = 1/(2q)$)

$$\mu_p(\mathcal{F}) \geq \mathbb{P}(X_p \in \langle\mathcal{H}\rangle) \geq \mathbb{P}(|X_p| \geq m)\mathbb{P}(X_m \in \langle\mathcal{H}\rangle) \geq 3\mathbb{P}(|X_p| \geq m)/4 > 1/2,$$

implying $p_c(\mathcal{F}) < p = 4Kq\log\ell$. (Note $\mathcal{H}$ q-spread with $\emptyset \notin \mathcal{H}$ implies $q \geq 1/n$, so that m is somewhat large and $\mathbb{P}(|X_p| \geq m) \approx 1$.) □

Remark 7.9. *This was done fussily to cover smaller ℓ in Theorem 1.10; if $\ell \to \infty$, then it gives $\mathbb{P}(X_p \in \langle\mathcal{H}\rangle) \to 1$.*

7.3 Main Lemma

Let γ be a slightly small constant (e.g. $\gamma = 0.1$ suffices), and let C_0 be a constant large enough to support the estimates that follow. Let $\mathcal{H}$ be an r-bounded, κ-spread hypergraph on a set X of size n, with $r, \kappa \geq C_0^2$. Set $p = C/\kappa$ with $C_0 \leq C \leq \kappa/C_0$ (so $p \leq 1/C_0$), $r' = (1-\gamma)r$ and $N = \binom{n}{np}$. Finally, fix $\psi : \langle\mathcal{H}\rangle \to \mathcal{H}$ satisfying $\psi(Z) \subseteq Z$ for all $Z \in \langle\mathcal{H}\rangle$; set, for $W \subseteq X$ and $S \in \mathcal{H}$,

$$\chi(S, W) = \psi(S \cup W) \setminus W;$$

and say the pair (S, W) is *bad* if $|\chi(S, W)| > r'$ and *good* otherwise.

The heart of our argument is the following lemma (an improvement of [2, Lemma 5.7]), regarding which a little orientation may be helpful. We will (in Theorems 7.5 and 7.6) be choosing a random subset of X in small increments and would like to say we are likely to be making good progress toward containing some $S \in \mathcal{H}$. Of course such progress is not to be expected for a *typical* S, but this is not the goal: having chosen a portion W of our eventual set, we just need the remainder to contain *some* $S \setminus W$, and may focus on those that are more likely (basically meaning small). The key idea (introduced in [2] and refined here) is that a general $S \setminus W$, while not itself small, will, in consequence of the spread assumption, typically *contain* some small $S' \setminus W$. (In fact $\chi(S, W)$ will usually be one of these: an $S' \setminus W$ contained in $S \setminus W$ will *typically* be small, so we don't need to steer this choice.) We then replace each "good" $S \setminus W$ by $\chi(S, W)$ and iterate, a second nice feature of the spread condition being that it is not much affected by this substitution.

Lemma 7.10. *For $\mathcal{H}$ as above, and W chosen uniformly from $\binom{X}{np}$,*

$$\mathbb{E}[|\{S \in \mathcal{H} : (S, W) \text{ is bad}\}|] \leq |\mathcal{H}|C^{-r/3}.$$

Proof. It is enough to show, for $s \in (r', r]$,

$$\mathbb{E}\left[|\{S \in \mathcal{H} : (S, W) \text{ is bad and } |S| = s\}|\right] \leq (\gamma r)^{-1}|\mathcal{H}|C^{-r/3}, \tag{7.8}$$

or, equivalently, that

$$|\{(S, W) : (S, W) \text{ is bad and } |S| = s\}| \leq (\gamma r)^{-1}N|\mathcal{H}|C^{-r/3}. \tag{7.9}$$

(Note $\gamma r = r - r'$ bounds the number of s for which the set in question can be nonempty, whence the negligible factors $(\gamma r)^{-1}$.)

We now use $\mathcal{H}_s = \{S \in \mathcal{H} : |S| = s\}$. Let $B = \sqrt{C}$ and for $Z \supseteq S \in \mathcal{H}_s$ say (S, Z) is *pathological* if there is $T \subseteq S$ with $t := |T| > r'$ and

$$|\{S' \in \mathcal{H}_s : S' \in [T, Z]\}| > B^r |\mathcal{H}| \kappa^{-t} p^{s-t}. \tag{7.10}$$

From now on we will always take $Z = W \cup S$ (with W as in Lemma 7.10); thus $|Z|$ is typically roughly np and, since $\mathcal{H}$ is κ-spread, $|\mathcal{H}|\kappa^{-t}p^{s-t}$ is a natural upper bound on what one might expect for the l.h.s. of (7.10).

Note that in proving (7.9) we may assume $s \leq n/2$: we may of course assume $|\mathcal{H}_s|$ is at least the r.h.s. of (7.8); but then for an $S \in \mathcal{H}_s$ of largest multiplicity, say m, we have

$$m \leq \kappa^{-s}|\mathcal{H}| \leq \kappa^{-s}\gamma r C^{r/3}|\mathcal{H}_s| \leq \kappa^{-s}\gamma r C^{r/3} m 2^n,$$

which is less than m if $s > n/2$ (since $\kappa > C$).

We bound the nonpathological and pathological parts of (7.9) separately; this (with the introduction of "pathological") is the source of our improvement over [2].

Nonpathological contributions. We first bound the number of (S, W) in (7.9) with (S, Z) nonpathological. This basically follows [2], but "nonpathological" allows us to bound the number of possibilities in Step 3 below by the r.h.s. of (7.10), where [2] settles for something like $|\mathcal{H}|\kappa^{-t}$.

Step 1. There are at most

$$\sum_{i=0}^{s} \binom{n}{np+i} \leq \binom{n+s}{np+s} \leq N p^{-s} \tag{7.11}$$

choices for $Z = W \cup S$.

Step 2. Given Z, let $S' = \psi(Z)$. Choose $T := S \cap S'$, for which there are at most $2^{|S'|} \leq 2^r$ possibilities, and set $t = |T| > r'$. (If $t \leq r'$ then (S, W) cannot be bad, as $\chi(S, W) = S' \setminus W \subseteq T$.)

Step 3. Since we are only interested in nonpathological choices, the number of possibilities for S is now at most

$$B^r |\mathcal{H}| \kappa^{-t} p^{s-t}.$$

Step 4. Complete the specification of (S, W) by choosing $W \cap S$, the number of possibilities for which is at most 2^s.

In sum, since $s \le r$ and $t > r' = (1-\gamma)r$, the number of nonpathological possibilities is at most

$$2^{r+s} N|\mathcal{H}|B^r (p\kappa)^{-t} \le N|\mathcal{H}|(4B)^r C^{-t} < N|\mathcal{H}|[4BC^{-(1-\gamma)}]^r. \tag{7.12}$$

Pathological contributions. We next bound the number of (S, W) as in (7.9) with (S, Z) pathological. The main point here is Step 4.

Step 1. There are at most $|\mathcal{H}|$ possibilities for S.

Step 2. Choose $T \subseteq S$ witnessing the pathology of (S, Z) (i.e. for which (7.10) holds); there are at most 2^s possibilities for T.

Step 3. Choose $U \in [T, S]$ for which

$$|\mathcal{H}_s \cap [U, (Z \setminus S) \cup U]| > 2^{-(s-t)} B^r |\mathcal{H}| \kappa^{-t} p^{s-t}. \tag{7.13}$$

(Here the left hand side counts members of $\mathcal{H}_s$ in Z whose intersection with S is precisely U. Of course, existence of U as in (7.13) follows from (7.10).) The number of possibilities for this choice is at most 2^{s-t}.

Step 4. Choose $Z \setminus S$, the number of choices for which is less than $N(2/B)^r$. To see this, write Φ for the r.h.s. of (7.13). Noting that $Z \setminus S$ must belong to $\binom{X\setminus S}{np} \cup \binom{X\setminus S}{np-1} \cup \cdots \cup \binom{X\setminus S}{np-s}$, we consider, for Y drawn uniformly from this set,

$$\mathbb{P}(|\mathcal{H}_s \cap [U, Y \cup U]| > \Phi). \tag{7.14}$$

Set $|U| = u$. We have

$$|\mathcal{H}_s \cap \langle U \rangle| \le |\mathcal{H} \cap \langle U \rangle| \le |\mathcal{H}|\kappa^{-u},$$

while, for any $S' \in \mathcal{H}_s \cap \langle U \rangle$,

$$\mathbb{P}(Y \supseteq S' \setminus U) \le \left(\frac{np}{n-s}\right)^{s-u}$$

(of course if $S' \cap S \ne U$ the probability is zero); so

$$\vartheta := \mathbb{E}\left[|\mathcal{H}_s \cap [U, Y \cup U]|\right] \le |\mathcal{H}|\kappa^{-u} \left(\frac{np}{n-s}\right)^{s-u} \le |\mathcal{H}|\kappa^{-u} (2p)^{s-u}$$

(since $n-s \geq n/2$). Markov's Inequality then bounds the probability in (7.14) by ϑ/Φ, and this bounds the number of possibilities for $Z \setminus S$ by $N(\vartheta/\Phi)$ (*cf.* (7.11)), which is easily seen to be less than $N(2/B)^r$.

Step 5. Complete the specification of (S, W) by choosing $S \cap W$, which can be done in at most 2^s ways.

Combining (and slightly simplifying), we find that the number of pathological possibilities is at most

$$|\mathcal{H}|N(16/B)^r. \tag{7.15}$$

Finally, the sum of the bounds in (7.12) and (7.15) is less than the $(\gamma r)^{-1}N|\mathcal{H}|C^{-r/3}$ of (7.9). □

7.4 Small uniformities

As in [2, Lemma 5.9], very small set sizes are handled by a simple Janson bound:

Lemma 7.11. *For an r-bounded, κ-spread $\mathcal{G}$ on Y, and $\alpha \in (0,1)$,*

$$\mathbb{P}(Y_\alpha \notin \langle \mathcal{G} \rangle) \leq \exp\left[-\left(\sum_{t=1}^{r}\binom{r}{t}(\alpha\kappa)^{-t}\right)^{-1}\right]. \tag{7.16}$$

Proof. We may assume $\mathcal{G}$ is r-uniform, since modifying it according to Observation 7.8 doesn't decrease the probability in (7.16). Denote members of $\mathcal{G}$ by S_i and set $\zeta_i = \mathbf{1}_{\{Y_\alpha \supseteq S_i\}}$. Then

$$\mu := \sum \mathbb{E}[\zeta_i] = |\mathcal{G}|\alpha^r$$

and

$$\Lambda := \sum\sum\{\mathbb{E}[\zeta_i\zeta_j] : S_i \cap S_j \neq \emptyset\} \leq |\mathcal{G}|\sum_{t=1}^{r}\binom{r}{t}\kappa^{-t}|\mathcal{G}|\alpha^{2r-t} = \mu^2\sum_{t=1}^{r}\binom{r}{t}(\alpha\kappa)^{-t}$$

(where the inequality holds because $\mathcal{G}$ is κ-spread), and Janson's Inequality (e.g. [31, Thm. 2.18(ii)]) bounds the probability in (7.16) by $\exp[-\mu^2/\Lambda]$. □

Corollary 7.12. *Let $\mathcal{G}$ be as in Lemma 7.11, let $t = \alpha|Y|$ be an integer with $\alpha\kappa \geq 2r$, and let $W = Y_t$. Then*

$$\mathbb{P}(W \notin \langle \mathcal{G} \rangle) \leq 2\exp[-\alpha\kappa/(2r)].$$

Proof. Lemma 7.11 gives

$$\exp[-\alpha\kappa/(2r)] \geq \mathbb{P}(Y_\alpha \notin \langle\mathcal{G}\rangle) \geq \mathbb{P}(|Y_\alpha| \leq t)\mathbb{P}(W \notin \langle\mathcal{G}\rangle) \geq \mathbb{P}(W \notin \langle\mathcal{G}\rangle)/2,$$

where we use the fact that any binomial ξ with $\mathbb{E}[\xi] \in \mathbb{Z}$ satisfies $\mathbb{P}(\xi \leq \mathbb{E}[\xi]) \geq 1/2$; see e.g. [48]. □

7.5 Proof of Theorem 7.5

It will be (very slightly) convenient to prove the theorem assuming $\mathcal{H}$ is (2κ)-spread. Let γ and C_0 be as in Section 7.3 and $\mathcal{H}$ as in the statement of Theorem 7.5, and recall that asymptotics refer to ℓ. We may of course assume that $\kappa \geq 2\gamma^{-1}C_0 \log \ell$ (or the result is trivial with a suitably adjusted K).

Fix an ordering "$\prec$" of $\mathcal{H}$. In what follows we will have a sequence $\mathcal{H}_i$, with $\mathcal{H}_0 = \mathcal{H}$ and

$$\mathcal{H}_i \subseteq \{\chi_i(S, W_i) : S \in \mathcal{H}_{i-1}\},$$

where W_i and χ_i will be defined below (with χ_i a version of the χ of Section 7.3). We then order $\mathcal{H}_i$ by setting

$$\chi_i(S, W_i) \prec_i \chi_i(S', W_i) \Leftrightarrow S \prec_{i-1} S'.$$

(So each member of $\mathcal{H}_i$ ultimately inherits its position in $\prec_i$ from some member of $\mathcal{H}$. This is not very important: we will be applying Lemma 7.10 repeatedly, and the present convention just provides a concrete ψ for each stage of the iteration.)

Set $C = C_0$ and $p = C/\kappa$, define m by $(1-\gamma)^m = \sqrt{\log \ell}/\ell$, and set $q = \log \ell/\kappa$. Then $\gamma^{-1}\log \ell \sim m \leq \gamma^{-1}\log \ell$ and Theorem 7.5 will follow from the next assertion.

Claim 7.13. *If W is a uniform $((mp+q)n)$-subset of X, then $W \in \langle\mathcal{H}\rangle$ w.h.p.*

Proof. Set $\delta = 1/(2m)$. Let $r_0 = \ell$ and $r_i = (1-\gamma)r_{i-1} = (1-\gamma)^i r_0$ for $i \in [m]$. Let $X_0 = X$ and, for $i = 1, \ldots, m$, let W_i be uniform from $\binom{X_{i-1}}{np}$ and set $X_i = X_{i-1} \setminus W_i$. (Note the assumption $\kappa \geq 2\gamma^{-1}C_0 \log \ell$ ensures $|X_m| \geq n/2$.)

For $S \in \mathcal{H}_{i-1}$ let $\chi_i(S, W_i) = S' \setminus W_i$, where S' is the first member of $\mathcal{H}_{i-1}$ contained in $W_i \cup S$ (with $\mathcal{H}_{i-1}$ ordered by $\prec_{i-1}$). Say S is *good* if $|\chi_i(S, W_i)| \leq r_i$ (and *bad* otherwise), and set

$$\mathcal{H}_i = \{\chi_i(S, W_i) : S \in \mathcal{H}_{i-1} \text{ is good}\}.$$

Thus $\mathcal{H}_i$ is an r_i-bounded collection of subsets of X_i and inherits the ordering $\prec_i$ as described above.

Finally, choose W_{m+1} uniformly from $\binom{X_m}{nq}$. Then $W := W_1 \cup \cdots \cup W_{m+1}$ is as in Claim 7.13. Note also that $W \in \langle \mathcal{H} \rangle$ whenever $W_{m+1} \in \langle \mathcal{H}_m \rangle$. (More generally, $W_1 \cup \cdots \cup W_i \cup Y \in \langle \mathcal{H} \rangle$ whenever $Y \subseteq X_i$ lies in $\langle \mathcal{H}_i \rangle$.)

So to prove the claim, we just need to show

$$\mathbb{P}(W_{m+1} \in \langle \mathcal{H}_m \rangle) = 1 - o(1) \tag{7.17}$$

(where the $\mathbb{P}$ refers to the entire sequence $W_1, \ldots, W_{m+1}$).

For $i \in [m]$ call W_i *successful* if $|\mathcal{H}_i| \geq (1 - \delta)|\mathcal{H}_{i-1}|$, call W_{m+1} successful if it lies in $\langle \mathcal{H}_m \rangle$, and say a sequence of W_i's is successful if each of its entries is. We show a little more than (7.17):

$$\mathbb{P}(W_1, \ldots, W_{m+1} \text{ is successful}) = 1 - \exp\left[-\Omega(\sqrt{\log \ell})\right]. \tag{7.18}$$

For $i \in [m]$, according to Lemma 7.10 (and Markov's Inequality),

$$\mathbb{P}(W_i \text{ is } \textit{not} \text{ successful} \,|\, W_1, \ldots, W_{i-1} \text{ is successful}) < \delta^{-1} C^{-r_{i-1}/3},$$

since $W_1, \ldots, W_{i-1}$ successful implies $|\mathcal{H}_{i-1}| > (1-\delta)^m |\mathcal{H}| > |\mathcal{H}|/2$, which, since $|\mathcal{H}_{i-1} \cap \langle I \rangle| \leq |\mathcal{H} \cap \langle I \rangle|$ and we assume $\mathcal{H}$ is (2κ)-spread), gives the spread condition (7.4) for $\mathcal{H}_{i-1}$. Thus

$$\mathbb{P}(W_1, \ldots, W_m \text{ is successful}) > 1 - \delta^{-1} \sum_{i=1}^{m} C^{-r_{i-1}/3} > 1 - \exp\left[-\sqrt{\log \ell}\right] \tag{7.19}$$

(using $r_m = \sqrt{\log \ell}$).

Finally, if $W_1, \ldots, W_m$ is successful, then Corollary 7.12 (applied with $\mathcal{G} = \mathcal{H}_m$, $Y = X_m$, $\alpha = nq/|Y| \geq q$, $r = r_m$, and $W = W_{m+1}$) gives

$$\mathbb{P}(W_{m+1} \notin \langle \mathcal{H}_m \rangle) \leq 2 \exp\left[-\sqrt{\log \ell}/2\right], \tag{7.20}$$

and we have (7.18) and the claim. □

7.6 Proof of Theorem 7.6

We assume the setup of Theorem 7.6 with γ and C_0 as in Section 7.3 and $\kappa \geq C_0^2$ (or there is nothing to prove). We may assume $\mathcal{H}$ is ℓ-uniform, since the construction of Observation 7.8 produces an ℓ-uniform, κ-spread $\mathcal{G}$ with $\xi_{\mathcal{G}} \geq \xi_{\mathcal{H}}$. In particular this gives

$$|\mathcal{H}|\ell = \sum_{x \in X} |\mathcal{H} \cap \langle x \rangle| \leq n\kappa^{-1}|\mathcal{H}|. \tag{7.21}$$

We first assume κ is *slightly* large, precisely

$$\kappa \geq \log^3 \ell; \tag{7.22}$$

the similar but easier argument for smaller values will be given at the end. (The bound in (7.22) is convenient but there is nothing delicate about this choice.)

Claim 7.14. *For κ as in (7.22) and $C_0 \leq C \leq \gamma\kappa/(4\log \ell)$,*

$$\mathbb{P}(\xi_{\mathcal{H}} > (3C/\gamma)\ell/\kappa) < \exp[-(\log \ell \log C)/4].$$

Proof of Theorem 7.6 in regime (7.22) given Claim 7.14. The "w.h.p." statement is immediate (take $C = C_0$). For the expectation, $Z_{\mathcal{H}}$, set $t = (3C_0/\gamma)\ell/\kappa$ and $T = 3\ell/(4\log \ell)$. By Claim 7.14 we have, for all $x \in [t, T]$,

$$\mathbb{P}(\xi_{\mathcal{H}} > x) \leq f(x) := \exp\left[-\log \ell \log(\gamma\kappa x/3\ell)/4\right] = (bx)^a = b^a x^a,$$

where $a = -(\log \ell)/4$ and $b = \gamma\kappa/3\ell$. Noting that $\xi_{\mathcal{H}} \leq \ell$, we then have

$$Z_{\mathcal{H}} \leq t + \int_t^T \mathbb{P}(\xi_{\mathcal{H}} > x)dx + \ell\mathbb{P}(\xi_{\mathcal{H}} > T) \leq t + \int_t^T f(x)dx + \ell f(T) = O(\ell/\kappa).$$

Here $t = O(\ell/\kappa)$ and the other terms are much smaller: the integral is less than $-1/(a+1)b^a t^{a+1} = O(1/\log \ell)C_0^a t$, while (7.22) easily implies that $f(T) = (\gamma\kappa/(4\log \ell))^a$ is $o(1/\kappa)$. □

Proof of Claim 7.14. Terms not defined here (beginning with $p = C/\kappa$ and W_i; note C is now as in Claim 7.14, rather than set to C_0) are as in Section 7.5, but we (re)define m by $(1-\gamma)^m = \log \ell/\ell$ and set $q = \log C \log^2 \ell/\kappa$, noting that (7.21) gives $p \geq C\ell/n$.

It's now convenient to generate the W_i's using the ξ_x's in the natural way: let

$$a_i = \begin{cases} (ip)n & \text{if } i \in \{0\} \cup [m], \\ (mp+q)n & \text{if } i = m+1, \end{cases}$$

and let W_i consist of the x's in positions $a_{i-1}+1, \ldots, a_i$ when X is ordered according to the ξ_x's.

Proposition 7.15. *With probability* $1 - e^{-\Omega(C\ell)}$,

$$\xi_x \le \varepsilon_i := \left\{ \begin{array}{ll} 2ip & \text{if } i \in \{0\} \cup [m] \\ 2(mp+q) & \text{if } i = m+1 \end{array} \right\} \text{ for all } i \text{ and } x \in W_i. \tag{7.23}$$

Proof. Failure at $i \ge 1$ implies

$$|\xi^{-1}[0, \varepsilon_i]| < a_i. \tag{7.24}$$

But $|\xi^{-1}[0, \varepsilon_i]|$ is binomial with mean $\varepsilon_i n = 2a_i \ge 2C\ell$, so the probability that (7.24) occurs for some i is less than $\exp[-\Omega(C\ell)]$ (see e.g. [31, Theorem 2.1]). □

We now write $\overline{W}_i$ for $W_1 \cup \cdots \cup W_i$.

Proposition 7.16. *If* $W_{m+1} \in \langle \mathcal{H}_m \rangle$, *then* W *contains some* $S \in \mathcal{H}$ *with*

$$|S \setminus \overline{W}_i| \le r_i \;\forall i \in [m].$$

Proof. Suppose $W \supseteq S_m \in \mathcal{H}_m$. By construction (of the $\mathcal{H}_i$'s) there are $S_{m-1}, \ldots, S_1, S_0 =: S$ with $S_i \in \mathcal{H}_i$ and $S_i = S_{i-1} \setminus W_i$, whence $S_i = S \setminus \overline{W}_i$ for $i \in [m]$; and $S_i \in \mathcal{H}_i$ then gives the proposition. □

We now define "success" for $(\xi_x : x \in X)$ to mean that $W_1, \ldots, W_{m+1}$ is successful in our earlier sense *and* (7.23) holds. Notice that with our current values of m and q (and $r_m = \ell(1-\gamma)^m = \log \ell$), we can replace the error terms in (7.19) and (7.20) by essentially $\delta^{-1} C^{-\log \ell/3}$ and $e^{-\log C \log \ell/2}$, which with Proposition 7.15 bounds the probability that $(\xi_x : x \in X)$ is *not* successful by (say) $\exp[-(\log \ell \log C)/4]$.

We finish with the following observation.

Proposition 7.17. *If* $(\xi_x : x \in X)$ *is successful then* $\xi_{\mathcal{H}} \le (3C/\gamma)\ell/\kappa$.

Proof. For S as in Proposition 7.16, we have (with $W_0 = \emptyset$ and $\varepsilon_0 = 0$)

$$\begin{aligned}\xi_{\mathcal{H}} \le \sum_{i=1}^{m+1} \varepsilon_i |S \cap W_i| \;&=\; \sum_{i=1}^{m+1} (\varepsilon_i - \varepsilon_{i-1})|S \setminus \overline{W}_{i-1}| \\ &\le 2\left[\sum_{i=1}^{m} (1-\gamma)^{i-1} p + (1-\gamma)^m q\right]\ell \\ &\le 2[C/(\gamma\kappa) + (\log \ell/\ell)(\log C \log^2 \ell/\kappa)]\ell < (3C/\gamma)\ell/\kappa. \qquad \square\end{aligned}$$

This completes the proof of Claim 7.14 (and of Theorem 7.6 when κ satisfies (7.22)).

$\square$

Finally, for κ below the bound in (7.22) (actually, for κ up to about $\ell/\log \ell$), a subset of the preceding argument suffices. We proceed as before, but now only with $C = C_0$ (so $p = C_0/\kappa$), stopping at m defined by $(1-\gamma)^m = 1/\kappa$ (so $m \approx \gamma^{-1} \log \kappa$). The main difference here is that there is no "Janson" phase: $W_1, \ldots, W_m$ is successful with probability $1 - \exp[-\Omega(\ell/\kappa)]$, and when it *is* successful we have (as in the proof of Proposition 7.17, now just taking $W_{m+1} = X \setminus \overline{W}_m$)

$$\xi_{\mathcal{H}} \le \sum_{i=1}^{m} (\varepsilon_i - \varepsilon_{i-1})|S \setminus \overline{W}_{i-1}| + |S \cap W_{m+1}| < 2(C_0/(\gamma\kappa))\ell + \ell/\kappa$$

(so also $Z_{\mathcal{H}} \le O(\ell/\kappa) + \exp[-\Omega(\ell/\kappa)]\ell = O(\ell/\kappa)$).

7.7 Applications

Much of the significance of Theorem 1.10—and of the skepticism with which Conjecture 7.1 was viewed in [35]—derives from the strength of their consequences, a few of which we discuss (*briefly*) here.

For this discussion, $\mathcal{K}_n^r = \binom{V}{r}$ is the complete r-graph on $V = [n]$, and $\mathcal{H}_{n,p}^r$ is the r-uniform counterpart of the usual binomial random graph $G_{n,p}$. Given r, n and an r-graph H, we use $\mathcal{G}_H$ for the collection of (unlabeled) copies of H in $\mathcal{K}_n^r$ and $\mathcal{F}_H$ for $\langle \mathcal{G}_H \rangle$. As usual, Δ is maximum degree.

As noted earlier, Conjecture 7.1 was motivated especially by Shamir's Problem (since resolved in [32]), and the conjecture that became Montgomery's theorem [51]. Very briefly: for n running over multiples of a given (fixed) r, Shamir's Problem asks for estimation of $p_c(\mathcal{F}_H)$ when H is a perfect matching (n/r disjoint edges), and [32] proves the natural conjecture that this threshold is $\Theta(n^{-(r-1)}\log n)$; and [51] shows that for fixed d, the threshold for $G_{n,p}$ to contain a given n-vertex tree with maximum degree d is $\Theta(n^{-1}\log n)$, where the implied constant in the upper bound depends on d (though it probably shouldn't). See [32, 51] for some account of the history of these problems. In both cases—and in most of the other examples mentioned following Theorem 7.18 (all but the one from [46])—the lower bounds derive from the coupon-collectorish requirement that the (hyper)edges cover the vertices, and it is the upper bounds that are of interest.

In fact, Theorem 1.10 gives not just Montgomery's theorem, but its natural extension to r-graphs and more. (Strictly speaking, Montgomery proves more than the original conjecture and we are not so far recovering this stronger result.) Say an r-graph F is a *forest* if it contains no *cycle*, meaning distinct vertices $v_1,\ldots,v_k$ and distinct edges $e_1,\ldots,e_k$ such that $v_{i-1},v_i\in e_i\ \forall i$ (with subscripts mod k). A *spanning tree* is then a forest of size $(n-1)/(r-1)$. For a (general) r-graph F, let $\rho(F)$ be the maximum size of a forest in F and set

$$\varphi(F)=\max\{1-\rho(F')/|F'|:\emptyset\neq F'\subseteq F\}.$$

Theorem 7.18. *For each r and c there is a K such that if H is an r-graph on $[n]$ with $\Delta(H)\le d$ and $\varphi(H)\le c/\log n$, then*

$$p_c(\mathcal{F}_H)<Kdn^{-(r-1)}\log|H|.$$

This gives $p_c(\mathcal{F}_H)=\Theta(n^{-(r-1)}\log n)$ if H is a perfect matching (as in Shamir's Problem), or a "loose Hamiltonian cycle" (a result of [6], to which we refer for definitions and history of the problem), and $p_c(\mathcal{F}_H)<Kdn^{-(r-1)}\log n$ if H is a spanning tree with $\Delta(H)\le d$. For fixed d the latter is the aforementioned r-graph generalization of [51] (or a slight improvement thereof in that the dependence on d—which, again, is probably unnecessary—is explicit), and for $d=n^{\Omega(1)}$ it is a result of Krivelevich [46, Theorem 1], which is again tight up to the value of K (see [46, Theorem 2]).

The last application we discuss here was suggested by Simon Griffiths and Rob Morris. Set $c_d=(d!)^{2/(d(d+1))}$ and $p^*(d,n)=c_dn^{-2/(d+1)}(\log n)^{2/(d(d+1))}$.

Theorem 7.19. *For fixed d and H any graph on $[n]$ with $\Delta(H) \le d$,*

$$p_c(\mathcal{F}_H) < (1 + o(1))p^*(d, n). \tag{7.25}$$

When $(d+1) \mid n$ and H is a K_{d+1}-factor (that is, $n/(d+1)$ disjoint K_{d+1}'s), $p^*(d, n)$ is the asymptotic value of $p_c(\mathcal{F}_H)$. Here (7.25) with $O(1)$ in place of $1+o(1)$ was proved in [32], while the asymptotics are given by the combination of [34] and [56, 28]; we state this in a form convenient for use below:

Theorem 7.20. *For fixed d and $\varepsilon > 0$, and n ranging over multiples of $d + 1$, if $p > (1 + \varepsilon)p^*(d, n)$, then $G_{n,p}$ contains a K_{d+1}-factor w.h.p.* □

Interest in $p_c(\mathcal{F}_H)$ for H as in Theorem 7.19 dates to at least 1992, when Alon and Füredi [1] showed the upper bound $O(n^{-1/d}(\log n)^{1/d})$, and has intensified since [32], motivated by the idea that K_{d+1}-factors should be the worst case. See [13, 14] for history and the most recent results; with $O(1)$ in place of $1 + o(1)$, Theorem 7.19 is conjectured in [13] and in the stronger "universal" form in [14].

Theorem 7.20 probably extends to r-graphs and d of the form $\binom{s-1}{r-1}$. This just needs extension of Theorem 1 of [56] to r-graphs (suggested at the end of [56]), which (with [34]) would give asymptotics of the threshold for $\mathcal{H}^r_{n,p}$ to contain a K^r_s-factor (where K^r_s is the complete r graph on s vertices).

Each of Theorems 7.18 and 7.19 begins with the following easy observations. (The first, an approximate converse of Proposition 7.4, is the trivial direction of LP duality.)

Observation 7.21. *If an increasing $\mathcal{F}$ supports a q-spread measure, then $q_f(\mathcal{F}) < q$.*

(More precisely, $q_f(\mathcal{F})$ is the least q such that $\mathcal{F}$ supports a probability measure ν with $\nu(\langle S\rangle) \le 2q^{|S|}$ $\forall S$.)

Observation 7.22. *Uniform measure on $\mathcal{G}_H$ is q-spread if and only if: for $S \subseteq \mathcal{K}^r_n$ isomorphic to a subhypergraph of H, σ a uniformly random permutation of V, and $H_0 \subseteq \mathcal{K}^r_n$ a given copy of H,*

$$\mathbb{P}(\sigma(S) \subseteq H_0) \le q^{|S|}. \tag{7.26}$$

Proving Theorem 7.18 is now just a matter of verifying (7.26) with $q = O(dn^{-(r-1)})$, which we leave to the reader. (It is similar to the proof of (7.28).)

Proof of Theorem 7.19. The next assertion is the main thing we need to check here.

Lemma 7.23. *There is $\varepsilon = \varepsilon_d > 0$ such that if H is as in Theorem 7.19 and has no component isomorphic to K_{d+1}, then*

$$q_f(\mathcal{F}_H) \le n^{-(2/(d+1)+\varepsilon)} =: q. \tag{7.27}$$

Proof. We just need to show (7.26) for q as in (7.27) and S, H_0 as in Observation 7.22, say with $W = V(S)$, $s = |S|$, and f the size of a spanning forest of S. We may of course assume S has no isolated vertices, so $w := |W| \le 2f$. We show

$$\mathbb{P}(\sigma(S) \subseteq H_0) < (e^2 d/n)^f \tag{7.28}$$

and

$$\frac{f}{s} \ge \frac{2(d+1)}{(d+2)d} = \frac{2}{d+1} + \varepsilon_0, \tag{7.29}$$

where $\varepsilon_0 = 1/((d+2)(d+1)d)$, implying that for any $\varepsilon < \varepsilon_0$, (7.26) holds for large enough n.

Proof of (7.28). Let $\alpha, \beta : W \to V$ be, respectively, a uniform injection and a uniform map. Then

$$\begin{aligned}(d/n)^f \ge \mathbb{P}(\beta(S) \subseteq H_0) \;&\ge\; \mathbb{P}(\beta \text{ is injective})\mathbb{P}(\beta(S) \subseteq H_0 | \beta \text{ is injective}) \\ = (n)_w n^{-w}\mathbb{P}(\alpha(S) \subseteq H_0) \;&>\; e^{-2f}\mathbb{P}(\sigma(S) \subseteq H_0).\end{aligned}$$ □

Proof of (7.29). We may of course assume S is connected, in which case we have $f = w - 1$ and upper bounds on s: $\binom{w}{2}$ if $w \le d$; $\binom{d+1}{2} - 1$ if $w = d+1$; and $wd/2$ if $w \ge d+2$. The corresponding lower bounds on f/s are $2/d$, $2d/((d+2)(d+1)-2)$ and $2(d+1)/((d+2)d)$, the smallest of which is the last. □

This completes the proof of Lemma 7.23. □

We are now ready for Theorem 7.19. Let $\varsigma = \varsigma_n$ be some slow $o(1)$ (e.g. $1/\log n$). By Theorem 7.20 there is $p_1 \sim p^*(d,n)$ such that if $(d+1)\,|\,m > (1-\varsigma)n$ then G_{m,p_1} contains a K_{d+1}-factor w.h.p., while by Lemma 7.23 and Theorem 1.10 (or, more precisely, Remark 7.9), there is p_2 with $p^*(d,n) \gg p_2 \gg n^{-(2/(d+1)+\varepsilon)}$ such that if $m \geq \varsigma n$ then for any given m-vertex H' with $\Delta(H') \leq d$, G_{m,p_2} contains (a copy of) H' w.h.p.

Let H_1 be the union of the copies of K_{d+1} in H (each of which must be a component of H), $H_2 = H - H_1$, and $n_i = |V(H_i)|$ (so $n_1 + n_2 = n$). Let $G_1 \sim G_{n,p_1}$ and $G_2 \sim G_{n,p_2}$ be independent on the common vertex set $V = [n]$ and $G = G_1 \cup G_2$. Then $G \sim G_{n,p}$ with $p = 1 - (1 - p_1)(1 - p_2) \sim p^*(d,n)$, and we just need to show $G \supseteq H$ w.h.p. In fact we find each H_i in the corresponding G_i, in order depending on n_2: if $n_2 \geq \varsigma n$, then w.h.p. G_1 contains H_1, say on vertex set V_1, and w.h.p. $G_2[V \setminus V_1]$ contains H_2; and if $n_2 < \varsigma n$, then w.h.p. G_2 contains H_2 on some V_2, and w.h.p. $G_1[V \setminus V_2]$ contains H_1. □

References

[1] N. Alon and Z. Füredi, *Spanning subgraphs of random graphs,* Graphs Combin. **8** (1992), 91–94. 96

[2] R. Alweiss, S. Lovett, K. Wu, and J. Zhang, *Improved bounds for the sunflower lemma*, Preprint, arXiv:1908.08483v1. 82, 86, 87, 89

[3] B. Bollobás, *Random graphs*, second ed., Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001. 5

[4] B. Bollobás and A. Thomason, *Thresholds functions*, Combinatorica **7** (1987), 35-38. 5, 80

[5] F.R.K. Chung, P. Frankl, R. Graham and J.B. Shearer, *Some intersection theorems for ordered sets and graphs,* J. Comb. Theory Ser. A. **48**, (1986), 23-37. 13

[6] A. Dudek, A. Frieze, P. Loh, and S. Speiss, *Optimal divisibility conditions for loose Hamilton cycles in random hypergraphs,* Electon. J. Combin. **19** (2012), Paper 44, 17. 95

[7] D. Duffus, P. Frankl and V. Rödl, *Maximal independent sets in the covering graph of the cube,* Discrete Appl. Math Vol. **161** Issue 9, (2013), 1203-1208. 4

[8] D. Ellis, *Almost Isoperimetric Subsets of the Discrete Cube,* Combin. Probab. and Comput. **20** (2011), 363-380. 28

[9] D. Ellis, N. Keller and N. Lifshitz, *On the structure of subsets of the discrete cube with small edge boundary,* Discrete Analysis 2018:9. 28

[10] J. Engbers and D. Galvin, *H-coloring tori,* J. Comb. Theory Ser. B **102**, (2012), 1110–1133. 4, 64

[11] P. Erdős, D.J. Kleitman and B.L. Rothschild, *Asymptotic enumeration of K_n-free graphs,* Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973) Tomo II, Atti dei Convegni Lincei, No. 17, Accad. Naz. Lincei, Rome, (1976), 19–27. 4

[12] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **5** (1960), 17–61. 5

[13] A. Ferber, G. Kronenberg, and K. Luh, *Optimal threshold for a random graph to be 2-universal,* Trans. Amer. Math. Soc. **372** (2019), 4239–4262. 96

[14] A. Ferber, K. Luh, and O. Nguyen, *Embedding large graphs into a random graph,* Bull. Lond. Math. Soc. **49** (2017), 784–797. 96

[15] J. Flum and M. Grohe, *Parameterized Complexity Theory,* Springer, New York, 2006. 14

[16] K. Frankston, J. Kahn, B. Narayanan, and J. Park, *Thresholds versus fractional expectation-thresholds,* arXiv:1910.13433 [math.CO]. iv

[17] E. Friedgut, *Sharp thresholds of graph properties, and the k-sat problem*, J. Amer. Math. Soc. **12** (1999), 1017–1054, With an appendix by J. Bourgain. 5

[18] E. Friedgut, *Hunting for sharp thresholds*, Random Structures Algorithms **26** (2005), 37–51. 5

[19] E. Friedgut, G. Kalai and A. Naor, *Boolean functions whose Fourier transform is concentrated on the first two levels,* Adv. Appl. Math. **29** (2002), 427–437. 28

[20] A. Frieze and G.B.Sorkin, *Efficient algorithms for three-dimensional axial and planar random assignment problems* Random Structures Algorithms **46** (2015), 160–196. 6, 83, 84

[21] Z. Füredi, *The number of maximal independent sets in connected graphs,* J. Graph Theory **11**, (1987), 463–470. 7

[22] D. Galvin, *On homomorphisms from the Hamming cube to Z,* Israel J. Math **138**, (2003), 189–213. 5, 10, 11, 12, 65

[23] D. Galvin, *Independent sets in the discrete hypercube,* arXiv:1901.01991 [math.CO] 10, 12

[24] D. Galvin and J. Kahn, *On phase transition in the hard-core model on $\mathbb{Z}^d$,* Combin. Probab. Comput. **13**, (2004), 137–164. 5

[25] A. Hamm and J. Kahn, *On Erdős-Ko-Rado for random hypergraphs II,* Combin. Probab. Comput., to appear. 56

[26] L. H. Harper, *Optimal numberings and isoperimetric problems on graphs, J. Combinatorial Theory* **1** (1966), 385–393. 27

[27] T. E. Harris, *A lower bound for the critical probability in a certain percolation process,* Proc. Cambridge. Philos. Soc. **56** (1960), 13–20. 35

[28] A. Heckel, *Random triangles in random graphs,* Preprint, arXiv:1802.08472. 96

[29] M. Hujter and Z. Tuza, *The number of maximal independent sets in triangle-free graphs,* SIAM J. Discrete Math **6**, (1993), 284–288. 1

[30] L. Ilinca and J. Kahn, *Counting maximal antichains and independent sets,* Order **30**, (2013), 427–435. 3, 4

[31] S. Janson, T.Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000. 5, 89, 93

[32] A. Johansson, J. Kahn, and V. Vu, *Factors in random graphs*, Random Structures Algorithms **33** (2008), 1–28. 81, 84, 94, 95, 96

[33] J. Kahn, *Range of cube-indexed random walk,* Isr. J. Math. **124** (2001), 189–201. 65

[34] J. Kahn, *Asymptotics for Shamir's problem,* Preprint, arXiv:1909.06834. 96

[35] J. Kahn and G. Kalai, *Thresholds and expectation thresholds,* Combin. Probab. Comput. **16** (2007), 495–502. 5, 80, 81, 94

[36] J. Kahn, G. Kalai and N. Linial, *The influence of variables on Boolean functions: Extended abstract,* Proc. 29th IEEE Symposium on Foundations of Computer Science, 1988. 28

[37] J. Kahn and J. Park, *The number of 4-colorings of the Hamming cube,* Israel J. Math, to appear. iv

[38] J. Kahn and J. Park, *Stability for maximal independent sets,* Electron. J. Combin. **27(1)** (2020), P1.59. iv

[39] J. Kahn and J. Park, *An isoperimetric inequality for the Hamming cube and some consequences,* arXiv:1909.04274 [math.CO]. iv

[40] J. Kahn and J. Park, *The number of maximal independent sets in the Hamming cube,* arXiv:1909.04283 [math.CO]. iv

[41] R.M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, New York.), 1972, pp. 85–103. 83

[42] D. Knuth, *The art of computer programming Vol. I,* Addison Wesley, London, 1969. 10

[43] A.D. Korshunov, *The number of monotone Boolean functions,* Problemy Kibernet **38**, (1980), 5–108. (Russian) 5

[44] A.D. Korshunov and A.A. Sapozhenko, *The number of binary codes with distance 2,* Problemy Kibernet **40**, (1983), 111–130. (Russian) 5, 10

[45] J. Körner and V. Wei, *Odd and even Hamming spheres also have minimum boundary,* Discrete Math. **51** (1984), 147–165. 11

[46] M. Krivelevich, *Embedding spanning trees in random graphs,* SIAM J. Discrete Math. **24** (2010), 1495–1500. 95

[47] J. H. Lindsey II, *Assignment of numbers to vertices,* Amer. Math. Monthly **71** (1964), 508–516. 27

[48] N. Lord, *Binomial averages when the mean is an integer,* Math. Gaz. **94** (2010), 331–332. 90

[49] G.A. Margulis, *Probabilistic characteristics of graphs with large connectivity,* Probl. Peredachi Inf. **10** (1974), 101–108. English version: *Problems Info. Transmission* **10** (1977), 174-179. 27

[50] R.J. McEliece, *The Theory of Information and Coding,* Addison-Wesley, London, 1977. 13

[51] R. Montgomery, *Spanning trees in random graphs*, Adv. Math. **356** (2019), 106793, 92. 81, 94, 95

[52] J.W. Moon and L. Moser, *On cliques in graphs,* Israel J. Math. **3**, (1965), 23–28. 1

[53] R. Peled and Y. Spinka, *Rigidity of proper colorings of* $\mathbb{Z}^d$*,* arXiv: 1808.03597 [math.PR] 5, 65

[54] Y. Rabinovich, personal communication to Jeff Kahn, 2013. 1

[55] Y. Rabinovich, J.A. Telle and M. Vatshelle, *Upper Bounds on Boolean-Width with Applications to Exact Algorithms,* IPEC 2013: Parameterized and Exact Computation, Springer, (2013), 308–320. 2

[56] O. Riordan, *Random cliques in random graphs,* Preprint, arXiv:1802.01948. 96

[57] A.A. Sapozhenko, *On the number of connected subsets with given cardinality of the boundary in bipartite graphs,* Metody Diskret. Analiz. **45**, (1987), 42–70. (Russian) 12, 65

[58] A.A. Sapozhenko, *The number of antichains in multilayered ranked sets,* Diskret. Mat. **1**, (1989), 110–128. (Russian; translation in Discrete Math. Appl. 1 (1991), no. 2, 149–169) 5

[59] A.A. Sapozhenko, *The number of antichains in ranked partially ordered sets,* Diskret. Mat. **1**, (1989), 74–93. (Russian; translation in Discrete Math. Appl. 1 (1991), no. 1, 35–58) 5

[60] A.A. Sapozhenko, *On the Number of Independent Sets in Extenders,* Diskret. Matem. **13** (1), (2001), 56. (in Russian) 2

[61] A.A. Sapozhenko, *The Number of Independent Sets in Graphs,* Moscow Univ. Math. Bull. **62**, (2007), 116–118. 2, 8

[62] M. Talagrand, *Isoperimetry, Logarithmic Sobolev Inequalities on the Discrete Cube, and Margulis' Graph Connectivity Theorem,* Geometric and Functional Analysis **3**, (1993), 295–314. 3, 4, 27

[63] M. Talagrand, *On Russo's approximate 0-1 law,* Ann. Probab. **22** (1994), 1576–1587. 28

[64] M. Talagrand, *Are all sets of positive measure essentially convex?*, Geometric aspects of functional analysis (Israel, 1992–1994), Oper. Theory Adv. Appl., vol. 77, Birkhäuser, Basel, 1995, pp. 295–310. 80, 81

[65] M. Talagrand, *The generic chaining,* Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005. 80, 83

[66] M. Talagrand, *Selector processes on classes of sets*, Probab. Theory Related Fields **135** (2006), 471–486. 83

[67] M. Talagrand, *Are many small sets explicitly small?,* Proceedings of the 2010 ACM International Symposium on Theory of Computing, (2010), 13–35. 5, 80, 81, 82

[68] M. Talagrand, *Upper and lower bounds for stochastic processes*, A Series of Modern Surveys in Mathematics, vol. 60, Springer, Heidelberg, 2014. 83