Anticoncentration for subgraph statistics

Matthew Kwan [*]   Benny Sudakov[†]   Tuan Tran[‡]

Abstract

Consider integers k, ℓ such that $0 \le \ell \le \binom{k}{2}$. Given a large graph G, what is the fraction of k-vertex subsets of G which span exactly ℓ edges? When G is empty or complete, and ℓ is zero or $\binom{k}{2}$, this fraction can be exactly 1. On the other hand, if ℓ is far from these extreme values, one might expect that this fraction is substantially smaller than 1. This was recently proved by Alon, Hefetz, Krivelevich and Tyomkyn who intiated the systematic study of this question and proposed several natural conjectures.

Let $\ell^* = \min\{\ell, \binom{k}{2} - \ell\}$. Our main result is that for any k and ℓ, the fraction of k-vertex subsets that span ℓ edges is at most $\log^{O(1)}(\ell^*/k)\sqrt{k/\ell^*}$, which is best-possible up to the logarithmic factor. This improves on multiple results of Alon, Hefetz, Krivelevich and Tyomkyn, and resolves one of their conjectures. In addition, we also make some first steps towards some analogous questions for hypergraphs.

Our proofs involve some Ramsey-type arguments, and a number of different probabilistic tools, such as polynomial anticoncentration inequalities, hypercontractivity, and a coupling trick for random variables defined on a "slice" of the Boolean hypercube.

1   Introduction

For an n-vertex graph G and some $0 \le k \le n$, consider a uniformly random set of k vertices $A \subseteq V(G)$ and define the random variable $X_{G,k} = e(G[A])$ to be the number of edges induced by the random k-set A. The point probability $\Pr(X_{G,k} = \ell)$ is then the fraction of k-vertex subsets of G which induce exactly ℓ edges. If G is an empty graph and $\ell = 0$, or if G is a complete graph and $\ell = \binom{k}{2}$, this probability is exactly one. However, if ℓ is far from these extreme values, and G is sufficiently large, one might expect $\Pr(X_{G,k} = \ell)$ to be small. For example, Ramsey's theorem tells us that all sufficiently large graphs must have induced k-vertex subgraphs that are empty or complete, so if $\ell \ne \{0, \binom{k}{2}\}$ and G is sufficiently large then certainly $\Pr(X_{G,k} = \ell) < 1$. In general, what upper bounds can we give on $\Pr(X_{G,k} = \ell)$ for large G?

Recently, Alon, Hefetz, Krivelevich and Tyomkyn [4] initiated the systematic study of this question, motivated by its connections to graph inducibility[1]. They proved some upper bounds on $\Pr(X_{G,k} = \ell)$ for various values of k and ℓ, and made some appealing conjectures. To state these, we recall some of their notation. Let $I(n, k, \ell) = \max\{\Pr(X_{G,k} = \ell) : |V(G)| = n\}$ be the maximum

[1]Roughly speaking, the *inducibility* of a graph H measures the maximum number of induced copies of H a large graph can have. This notion was introduced in 1975 by Pippenger and Golumbic [38], and has enjoyed a recent surge of interest; see for example [7, 21, 42, 27].

value of $\Pr(X_{G,k} = \ell)$ over all n-vertex graphs, and let $\mathrm{ind}(k,\ell) = \lim_{n\to\infty} I(n,k,\ell)$ (one can use a standard averaging argument to show that $I(n,k,\ell)$ is a monotone nonincreasing function of n). The following three conjectures appear as [4, Conjecture 1.1, Conjecture 6.1 and Conjecture 6.2].

Conjecture 1. *For all $0 < \ell < \binom{k}{2}$ we have $\mathrm{ind}(k,\ell) \leq 1/e + o_k(1)$.*

Conjecture 2. *For all k, ℓ satisfying $\min\left\{\ell, \binom{k}{2} - \ell\right\} = \omega_k(k)$, we have $\mathrm{ind}(k,\ell) = o_k(1)$.*

Conjecture 3. *For all k, ℓ satisfying $\min\left\{\ell, \binom{k}{2} - \ell\right\} = \Omega_k(k^2)$, we have $\mathrm{ind}(k,\ell) = O(k^{-1/2})$.*

The authors of [4] proved some partial results for all of these conjectures. Specifically, under the assumptions of Conjecture 1 they proved that $\mathrm{ind}(k,\ell) = 1 - \Omega_k(1)$, under the assumptions of Conjecture 2 they proved that $\mathrm{ind}(k,\ell) \leq 1/2 + o_k(1)$, and under the assumptions of Conjecture 3 they proved that $\mathrm{ind}(k,\ell) = O(k^{-0.1})$.

Our main result is the following theorem, simultaneously implying Conjecture 2 and an asymptotic version of Conjecture 3. This improves two of the aforementioned results in [4].

Theorem 1.1. *For all k and all $0 \leq \ell \leq \binom{k}{2}$, let $\ell^* = \min\left\{\ell, \binom{k}{2} - \ell\right\}$. We have*

$$\mathrm{ind}(k,\ell) \leq \log^{O(1)}(\ell^*/k)\sqrt{\frac{k}{\ell^*}}.$$

We remark that we allow the "$O(1)$" term to equal zero, so the above statement still makes sense (and is in fact trivial) if $\ell^* \leq k$. Note that up to the logarithmic factor Theorem 1.1 is essentially best-possible. Indeed, for any $s \leq k$ and any $f = \omega(1)$, let $n = fk$ and consider the n-vertex complete bipartite graph $G = K_{fs,fk-fs}$. Then for $\ell = s(k-s)$ we have

$$\Pr(X_{G,k} = \ell) \geq \frac{\binom{fs}{s}\binom{fk-fs}{k-s}}{\binom{fk}{k}} = \Theta\left(\sqrt{\frac{k}{s(k-s)}}\right) = \Theta\left(\sqrt{\frac{k}{\ell}}\right).$$

We prove Theorem 1.1 in Section 3. Our proof depends on a polynomial anticoncentration inequality due to Meka, Nguyen and Vu [33], which itself depends on a weak version of the so-called Gotsman–Linial conjecture in the theory of Boolean functions, proved by Kane [24]. Any improvements to this anticoncentration inequality, potentially via progress towards the Gotsman–Linial conjecture, would result in corresponding improvements to Theorem 1.1. We discuss this further in Section 6.

It is also interesting to study related questions for hypergraphs; indeed, in [4] the authors specifically suggested that a natural analogue of Conjecture 1 might also hold for r-uniform hypergraphs. We make a first step in this direction, generalising a result in [4]. For $0 < \ell < \binom{k}{r}$ and an r-uniform hypergraph G with at least k vertices, we may define $X_{G,k}$, $I_r(n,k,\ell)$ and $\mathrm{ind}_r(k,\ell)$ in the obvious way: $X_{G,k}$ is the number of edges induced by a uniformly random k-vertex subset of G, $I_r(n,k,\ell)$ is the maximum value of $\Pr(X_{G,k} = \ell)$ over n-vertex r-uniform hypergraphs G, and $\mathrm{ind}_r(k,\ell) = \lim_{n\to\infty} I_r(n,k,\ell)$.

Theorem 1.2. *For any r there exists $\varepsilon = \varepsilon(r) > 0$ such that for any $0 < \ell < \binom{k}{r}$ we have $\mathrm{ind}_r(k,\ell) \leq 1 - \varepsilon$.*

A proof of this theorem for graphs appears as [4, Theorem 1.3], and proceeds via a long and complicated fourth-moment calculation. We give a short proof of Theorem 1.2 in Section 4 using a hypercontractive inequality. For concreteness, we remark that Theorem 1.2 holds with $\varepsilon = 2^{-4/3}3^{-16r}$.

Finally, the natural hypergraph generalisation of Conjecture 3 is that for any fixed r and any k, ℓ satisfying $\min\{\ell, \binom{k}{2} - \ell\} = \Omega_k(k^r)$, we have $\mathrm{ind}_r(k,\ell) = O(k^{-1/2})$. This problem appears to be quite difficult; we make a first step in the case $r = 3$.

2

Theorem 1.3. *For all k, ℓ satisfying $\ell^* = \min\left\{\ell, \binom{k}{3} - \ell\right\} = \Omega_k\left(k^3\right)$, we have*

$$\mathrm{ind}_3(k, \ell) \leq \frac{\log^{O(1)} k}{\sqrt{k}}.$$

We prove Theorem 1.3 in Section 5.

1.1 Discussion and main ideas of the proofs

Let $A = (a_{xy})_{x,y}$ be the adjacency matrix of a graph G. We can express $X_{G,k}$ as a homogeneous quadratic polynomial

$$\frac{1}{2}\boldsymbol{\xi} A \boldsymbol{\xi}^T = \sum_{1 \leq x < y \leq n} a_{xy} \xi_x \xi_y,$$

where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ is a uniformly random length-n zero-one vector with exactly k ones. To prove Theorem 1.1 we need to upper-bound $\mathrm{ind}(k, \ell)$, which essentially comes down to upper-bounding probabilities of the form $\Pr\left(\boldsymbol{\xi} A \boldsymbol{\xi}^T = 2\ell\right)$.

This point of view suggests the application of *quadratic anticoncentration inequalities*. Indeed, initially motivated by applications in random matrix theory [10], several authors [35, 40, 9, 33] have studied probabilities of the form $\Pr\left(\boldsymbol{\gamma} A \boldsymbol{\gamma}^T = x\right)$, for $\boldsymbol{\gamma}$ a sequence of *independent* random variables. The general theme is that if there are many nonzero entries arranged appropriately in A, then this probability is small.

Of course, due to the condition that $\boldsymbol{\xi}$ has exactly k ones, it is not a sequence of independent random variables, but one might hope that the dependencies are not too severe. For example, $\boldsymbol{\xi}$ is in some sense quite similar to the random vector $\boldsymbol{\xi}_{\mathrm{Ber}} = (\gamma_1, \ldots, \gamma_n)$ where each γ_i is independently (k/n)-Bernoulli-distributed[2]. If A has few nonzero entries, then one can prove using a concentration inequality that $\boldsymbol{\xi}_{\mathrm{Ber}} A \boldsymbol{\xi}_{\mathrm{Ber}}^T$ is likely to be small (and therefore not equal to ℓ, unless ℓ is itself small). It is therefore very straightforward to apply a quadratic anticoncentration inequality to prove a variant of Theorem 1.1 with $\boldsymbol{\xi}_{\mathrm{Ber}}$ in place of $\boldsymbol{\xi}$ (meaning that $X_{G,k}$ is the number of edges in a (k/n)-Bernoulli random set, instead of a uniformly random set of exactly k vertices). Actually, in general, for any r-uniform hypergraph G, the random variable $X_{G,k}$ can be expressed as a homogeneous degree-r polynomial of $\boldsymbol{\xi}$. So, using a cubic anticoncentration inequality we can similarly give an easy proof of the "Bernoulli version" of Theorem 1.3, and using the Bonami–Beckner hypercontractive inequality we can give an easy proof of the "Bernoulli version" of Theorem 1.2.

However, in the setting of this paper, approximating $\boldsymbol{\xi}$ with $\boldsymbol{\xi}_{\mathrm{Ber}}$ is quite unsatisfactory, because in addition to the "genuine" anticoncentration coming from the combinatorial structure of G, there is also spurious anticoncentration arising from fluctuation in the number of ones in $\boldsymbol{\xi}_{\mathrm{Ber}}$. For example, if G is a graph clique then $\boldsymbol{\xi} A \boldsymbol{\xi}^T$ is constant, while $\boldsymbol{\xi}_{\mathrm{Ber}} A \boldsymbol{\xi}_{\mathrm{Ber}}^T$ is anticoncentrated purely because the number of vertices in a (k/n)-Bernoulli random set is itself anticoncentrated.

In the setting of Theorem 1.2 it is straightforward to overcome this issue: we merely apply a different hypercontractive inequality in place of the Bonami–Beckner inequality. Despite the widespread utility of the Bonami–Beckner inequality, the wider theory of hypercontractive inequalities does not seem to be well-known in the combinatorics community. In our case the necessary inequality is essentially due to Lee and Yau [30].

For Theorems 1.1 and 1.3, we use a coupling argument: it turns out that there is a natural way to realise the distribution of $\boldsymbol{\xi}$ as a function of a random permutation σ and a certain sequence $\boldsymbol{\gamma}$ of i.i.d. random variables. If we condition on any outcome of σ, then $X_{G,k}$ can be viewed as a (non-homogeneous) degree-r polynomial $f_\sigma(\boldsymbol{\gamma})$ of $\boldsymbol{\gamma}$, to which we can apply standard anticoncentration

---

[2]We say that γ has the p-Bernoulli distribution if $\Pr(\gamma = 1) = p$ and $\Pr(\gamma = 0) = 1 - p$.

inequalities. For a non-homogeneous polynomial, anticoncentration inequalities tend to give bounds depending on the nonzero coefficients of maximum degree, so the remaining difficulty lies in studying the nonzero maximium-degree coefficients in f_σ (which depend on σ).

It turns out that these coefficients have a combinatorial interpretation: for example, if G is a graph (as in Theorem 1.1), then the nonzero degree-2 coefficients in f_σ in some sense arise from 4-tuples of vertices (x, x', y, y') such that

$$a_{xy} - a_{xy'} - a_{x'y} + a_{x'y'} \neq 0.$$

In the special case where $\ell^* = \Omega(k^2)$, we can use a simple Ramsey-type argument to show that G has $\Omega(k^4)$ such tuples (this turns out to follow from the fact that 2-edge-coloured complete graphs with many blue and red edges have many alternating paths of length 3). This allows us to show that f_σ is likely to have many nonzero coefficients, allowing us to deduce Theorem 1.1 via a quadratic anticoncentration inequality. For the general case of Theorem 1.1 we need to use a more refined anticoncentration inequality due to Meka, Nguyen and Vu [33] for which it suffices to find a large matching in an auxiliary graph defined in terms of the nonzero degree-2 coefficients. In the proof of Theorem 1.1 this auxiliary graph will be a random graph depending on σ. We will carefully define a greedy procedure that finds the required matching with high probability.

The situation for hypergraphs is much less straightforward than for graphs, which is why Theorem 1.3 is so much weaker than Theorem 1.1. In contrast to the graph case, even in the setting of Theorem 1.3 where G is a 3-uniform hypergraph with $\ell^* = \Omega(k^3)$, it may happen that $f_\sigma(\gamma)$ has very few degree-3 coefficients, which prevents us from directly applying an anticoncentration inequality. To overcome this, we prove a variant of the Meka–Nguyen–Vu anticoncentration inequality which (under certain specific circumstances) allows us to take coefficients of non-maximum degree into account. We then prove an approximate classification of 3-uniform hypergraphs G such that $f(\gamma)$ has few nonzero degree-3 coefficients (using a theorem of Fox and Sudakov on "unavoidable patterns" and the induced hypergraph removal lemma), and we study the lower-degree coefficients of f for all such G. This unfortunately involves some slightly complicated case analysis.

1.2 Notation

We use standard asymptotic notation throughout. For functions $f = f(n)$ and $g = g(n)$ we write $f = O(g)$ to mean there is a constant C such that $|f| \leq C|g|$, we write $f = \Omega(g)$ to mean there is a constant $c > 0$ such that $f \geq c|g|$ for sufficiently large n, we write $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $g = \omega(f)$ to mean that $f/g \to 0$ as $n \to \infty$. All asymptotics are as $n \to \infty$ unless specified otherwise (specifically, notation of the form $o_k(1)$ indicates that asymptotics are as $k \to \infty$).

For a positive integer n, we write $[n]$ to mean the set $\{1, \ldots, n\}$. For a set S we write $\binom{S}{k}$ for the collection of all subsets of S of size exactly k, and we write $\binom{S}{\leq k}$ for the collection of all subsets of size at most k. Less standardly, for a zero-one sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$, we write $|\boldsymbol{x}|$ for the number of ones in $\boldsymbol{x}$. For any sequence $\boldsymbol{x} = (x_1, \ldots, x_n)$, and any $I \subseteq [n]$, we write $\boldsymbol{x}^I$ to mean the monomial $\prod_{i \in I} x_i$.

We also use standard (hyper)graph theoretic notation. In particular, for a hypergraph G on the vertex set V and a set of vertices $S \subseteq V$, let $\deg(S)$ be the number of edges $e \in E(G)$ such that $S \subseteq e$. Also, we write "r-graph" as shorthand for "r-uniform hypergraph".

2   Probabilistic Tools

For any $0 \le k \le n$ let $\mathrm{BL}(n,k)$ be the uniform distribution on sequences $\boldsymbol{x} \in \{0,1\}^n$ with $|\boldsymbol{x}| = k$. This is precisely the distribution of $\boldsymbol{\xi}$ as outlined in Section 1.1. It is sometimes informally known as the uniform distribution "on the slice", and is also known as the limiting distribution of the Bernoulli–Laplace model of diffusion. For an r-graph G on the vertex set $[n]$, note that the random variable $X_{G,k}$ can be interpreted as a homogeneous degree-r polynomial of $\boldsymbol{\xi} \in \mathrm{BL}(n,k)$: if $a_S = \mathbf{1}_{S \in E(G)}$ indicates the presence of an edge $S \subseteq [n]$ in G, then we can write

$$X_{G,k} = \sum_S a_S \boldsymbol{\xi}^S.$$

In this section we collect a number of general results about $\mathrm{BL}(n,k)$ which will be useful for our proofs. Some of these are known, and some are new.

2.1   Concentration

There are a number of well-known concentration results which can be applied to functions of $\mathrm{BL}(n,k)$-distributed random zero-one sequences (see for example [20, Corollary 2.2]). In the proof of Theorem 1.1 we will use the following "non-uniform" concentration inequality, which we were not able to find elsewhere in the literature.

Lemma 2.1. *Consider $f : \{0,1\}^n \to \mathbb{R}$ such that*

$$|f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n)| \le c_i$$

for all $\boldsymbol{x} \in \{0,1\}^n$ and all $i \in [n]$. Let $\boldsymbol{\xi} \in \mathrm{BL}(n,k)$. Then

$$\Pr(f(\boldsymbol{\xi}) - \mathbb{E}f(\boldsymbol{\xi}) \ge t) \le \exp\left(-\frac{t^2}{8\sum_{i=1}^n c_i^2}\right).$$

Proof. We may assume without loss of generality that $c_1 \ge \cdots \ge c_n$. Consider the Doob martingale $Z_i = \mathbb{E}[f(\boldsymbol{\xi})|\xi_1, \ldots \xi_i]$, so $Z_0 = \mathbb{E}f(\boldsymbol{\xi})$ and $Z_n = Z_{n-1} = f(\boldsymbol{\xi})$. Let $\mathcal{L}(x_1, \ldots, x_i)$ be the conditional distribution of $\boldsymbol{\xi}$ given $\xi_1 = x_1, \ldots, \xi_i = x_i$.

We want to show that

$$|\mathbb{E}[f(\mathcal{L}(x_1, \ldots, x_{i-1}, 0))] - \mathbb{E}[f(\mathcal{L}(x_1, \ldots, x_{i-1}, 1))]| \le 2c_i$$

for all feasible $x_1, \ldots, x_{i-1} \in \{0,1\}$; this will imply that $|Z_i - Z_{i-1}|$ is uniformly bounded by $2c_i$, so the desired result will follow from the Azuma–Hoeffding inequality (see for example [23, Theorem 2.25]).

If $\boldsymbol{\xi}$ is distributed as $\mathcal{L}(x_1, \ldots, x_{i-1}, 0)$, we can change ξ_i to 1 and then randomly choose one of the ones among $\xi_{i+1}, \ldots, \xi_n$ and change it to 0; we thereby obtain the distribution $\mathcal{L}(x_1, \ldots, x_{i-1}, 1)$. This provides a coupling between $\mathcal{L}(x_1, \ldots, x_{i-1}, 0)$ and $\mathcal{L}(x_1, \ldots, x_{i-1}, 1)$ that differs in only two coordinates i and $j > i$, and since $c_j \le c_i$ this implies the required bound. $\square$

2.2   "Weak" anticoncentration via hypercontractivity

The key ingredient in our proof of Theorem 1.2 is the following "weak" anticoncentration inequality.

Lemma 2.2. *Let f be an n-variable polynomial of degree d, and let $\boldsymbol{\xi} \in \mathrm{BL}(n,n/2)$. Suppose that the random variable $f(\boldsymbol{\xi})$ is not constant. Then, for any $\ell \in \mathbb{R}$, we have*

$$\Pr(f(\boldsymbol{\xi}) = \ell) \le 1 - e^{-O(d)}.$$

To prove Lemma 2.2 it will suffice to control the fourth moment of low-degree polynomials of $\mathrm{BL}(n,k)$-distributed random vectors. This is due to the fact (also observed in [4]) that if a random variable has fourth moment comparable to its variance squared then it is reasonably likely to have fluctuations comparable to its standard deviation. The following lemma is a corollary of [3, Lemma 3.2 (i)].

Lemma 2.3. *Let Z be a non-constant real random variable satisfying $\mathbb{E}Z = 0$ and $\mathbb{E}Z^4 \leq b(\mathbb{E}Z^2)^2$. Then for any $\ell \in \mathbb{R}$, $\Pr(Z \neq \ell) \geq 1/(2^{4/3}b)$.*

Now, to bound the fourth moment of a low-degree polynomial of a $\mathrm{BL}(n,k)$-distributed random vector, we will want a hypercontractive inequality that can be applied to $\mathrm{BL}(n,k)$ analogously to standard applications of the Bonami–Beckner hypercontractive inequality (see for example [37]) in discrete Fourier analysis. We will use a hypercontractive inequality for (the Markov semigroup of) Bernoulli-Laplace diffusion, which can be deduced (see for example [11]) from a log-Sobolev inequality proved by Lee and Yau [30]. We will present this hypercontractive inequality in a convenient self-contained form due to Filmus [15].

First, we need the notion of a *harmonic* polynomial, originally introduced by Dunkl [12, 13]: a polynomial g in the variables $x_1, \ldots, x_n$ is said to be harmonic if

$$\sum_{i=1}^{n} \frac{\partial g}{\partial x_i} = 0.$$

It turns out that every random variable of the form $f(\boldsymbol{\xi})$, for $\boldsymbol{\xi} \in \mathrm{BL}(n,k)$, can be represented in the form $g(\boldsymbol{\xi})$, for g a harmonic multilinear polynomial. We will moreover need the fact that if f is a polynomial of degree d, then g also has degree at most d. The following lemma effectively appears as [17, Lemma 3.17].

Lemma 2.4. *Let f be an n-variable polynomial of degree d, and let $\boldsymbol{\xi} \in \mathrm{BL}(n,k)$. Then there is a harmonic multilinear polynomial g such that $f(\boldsymbol{\xi}) = g(\boldsymbol{\xi})$; this polynomial has degree at most $\min\{d, k, n-k\}$.*

Now, for a harmonic multilinear polynomial f, let $f^{=d}$ be the d-th homogeneous part of f consisting of terms with degree exactly d. Each of these parts is "orthogonal" in the sense that for $\boldsymbol{\xi} \in \mathrm{BL}(n,k)$ and $d \neq d'$ we have $\mathbb{E}\left[f^{=d}(\boldsymbol{\xi})f^{=d'}(\boldsymbol{\xi})\right] = 0$ (see [15, Theorem 3.1]). For each t, let H_t be the operator on n-variable harmonic multilinear polynomials defined as follows. For a harmonic multilinear polynomial f, let

$$H_t f = \sum_{i=0}^{n} \exp\left(-t\frac{2i(n+1-i)}{n(n-1)}\right) f^{=i}.$$

We are now in a position to state the promised hypercontractive inequality, which essentially appears as [15, Proposition 6.2].

Proposition 2.5. *Let $\boldsymbol{\xi} \in \mathrm{BL}(n, pn)$ and let*

$$\rho = -\frac{2}{n \log 2 \log(p(1-p))}.$$

Then for any $t \geq 0$ and $q \geq 2$ satisfying $q - 1 \leq e^{2\rho t}$, and any n-variable harmonic multilinear polynomial g, we have

$$\mathbb{E}[|H_t g(\boldsymbol{\xi})|^q]^{2/q} \leq \mathbb{E}g(\boldsymbol{\xi})^2.$$

The only reason we need Proposition 2.5 is for the following corollary.

Corollary 2.6. *Let $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$ and let f be an n-variable polynomial of degree d. Then*

$$\mathbb{E}f(\boldsymbol{\xi})^4 = e^{O(d)}\left(\mathbb{E}f(\boldsymbol{\xi})^2\right)^2.$$

Proof. By Lemma 2.4, there is a harmonic multilinear polynomial g of degree at most d such that $f(\boldsymbol{\xi}) = g(\boldsymbol{\xi})$. By Proposition 2.5 with $p = 1/2$, $q = 4$ and $t = \Theta(n)$, and orthogonality of the different homogeneous parts, we have

$$\left(\mathbb{E}f(\boldsymbol{\xi})^4\right)^{1/2} \leq \mathbb{E}(H_{-t}f(\boldsymbol{\xi}))^2$$

$$= \sum_{i=0}^{d} e^{O(i)}\mathbb{E}\left(f^{=i}(\boldsymbol{\xi})\right)^2$$

$$= e^{O(d)}\mathbb{E}(f(\boldsymbol{\xi}))^2. \qquad \square$$

Lemma 2.2 is now a direct consequence of Corollary 2.6 and Lemma 2.3, applied to $f(\boldsymbol{\xi}) - \mathbb{E}f(\boldsymbol{\xi})$.

2.3 Coupling

Many standard probabilistic tools are designed to work for product distributions, where independence can be exploited. Although $\mathrm{BL}(n, k)$ is not a product distribution, there is a well-known way to approximate $\mathrm{BL}(n, k)$ with a product of Bernoulli-distributed random variables, and for many purposes these distributions can be considered essentially equivalent (see for example [23, Corollary 1.16] and the invariance principles in [16, 17]). However, for the purposes of proving Theorems 1.1 and 1.3 this kind of approximation is too coarse. Instead we will use a non-standard coupling between $\mathrm{BL}(n, n/2)$ and $\mathrm{Rad}^{n/2}$, where Rad is the Rademacher distribution (that is, the uniform distribution on $\{-1, 1\}$). The following observation essentially appears in the proof of [31, Proposition 4.10] (a similar coupling also appears in [4, p. 15]).

Fact 2.7. *The distribution $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$ can be obtained as follows. Let σ be a uniformly random permutation of $[n]$ and let $\boldsymbol{\gamma} \in \mathrm{Rad}^{n/2}$ be a sequence of $n/2$ i.i.d. Rademacher random variables. Then set $\xi_{\sigma(i)} = 1$ for all i such that $\gamma_i = 1$ and set $\xi_{\sigma(i+n/2)} = 1$ for all i such that $\gamma_i = -1$. For all other indices j set $\xi_j = 0$.*

In order to use Fact 2.7, we need to translate polynomials of $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$ into polynomials of $\boldsymbol{\gamma} \in \mathrm{Rad}^{n/2}$.

Lemma 2.8. *Consider a random variable X of the form*

$$X = \sum_{S \in \binom{[n]}{d}} a_S \boldsymbol{\xi}^S,$$

where $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$. Under the coupling in Fact 2.7, X is a function of $\boldsymbol{\gamma}, \sigma$. If we condition on any outcome of σ, then X is a multilinear polynomial in the γ_i with degree at most d. Consider a subset $I \subseteq [n/2]$ of size at least $d - 1$, and write $I = \{i_1, \ldots, i_q\}$; then the coefficient g_I of $\boldsymbol{\gamma}^I$ is

$$\frac{1}{2^d} \sum_{\boldsymbol{b} \in \{0,1\}^q} (-1)^{|\boldsymbol{b}|} a\left(\left\{\sigma\left(i_j + b_j\frac{n}{2}\right) : 1 \leq j \leq q\right\}\right),$$

where for $R \subseteq [n]$, $a(R)$ is the sum of all a_S with $S \supseteq R$.

7

Proof. Given a permutation σ of $[n]$, define the functions $\alpha : [n] \to \{0, 1\}$ and $\tilde{\sigma} : [n] \to [n/2]$ by

$$(\alpha(x), \tilde{\sigma}(x)) = \begin{cases} (0, \sigma^{-1}(x)) & \sigma^{-1}(x) \leq n/2 \\ (1, \sigma^{-1}(x) - n/2) & \text{otherwise.} \end{cases}$$

Also, for $S \subseteq [n]$, let $|\alpha(S)|$ be the number of $x \in S$ for which $\alpha(x) = 1$.

Now, observe that $\xi_x = \left(1 + (-1)^{\alpha(x)} \gamma_{\tilde{\sigma}(x)}\right)/2$. We may write

$$X = \frac{1}{2^d} \sum_{S \in \binom{[n]}{d}} a_S \prod_{x \in S} \left(1 + (-1)^{\alpha(x)} \gamma_{\tilde{\sigma}(x)}\right)$$

$$= \frac{1}{2^d} \sum_{S \in \binom{[n]}{d}} a_S \sum_{R \subseteq S} (-1)^{|\alpha(R)|} \prod_{x \in R} \gamma_{\tilde{\sigma}(x)}.$$

Consider any $I = \{i_1, \ldots, i_q\} \subseteq [n/2]$ with $q \geq d-1$. We have $\boldsymbol{\gamma}^I = \prod_{x \in R} \gamma_{\tilde{\sigma}(x)}$ if and only if R is of the form $\{\sigma(i_j + b_j n/2) : 1 \leq j \leq q\}$ for some $\boldsymbol{b} \in \{0, 1\}^q$, in which case $|\alpha(R)| = |\boldsymbol{b}|$. (If $q < d - 1$ then there are other possibilities for R due to the fact that $\gamma_i^2 = 1$). The desired result follows. $\square$

As an illustration of Lemma 2.8, we consider the special case where X is of the form $\sum_{i<j} a_{ij} \xi_i \xi_j$, for $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$. If G is an n-vertex graph with adjacency matrix (a_{ij}), then this random variable has precisely the distribution of $X_{G,n/2}$. Under the coupling in Fact 2.7, if we condition on any outcome of σ, then X is a quadratic polynomial in the γ_i, and the coefficient of $\gamma_i \gamma_j$ is

$$\frac{1}{4} \left(a_{\sigma(i)\sigma(j)} - a_{\sigma(i)\sigma(j+n/2)} - a_{\sigma(i+n/2)\sigma(j)} + a_{\sigma(i+n/2)\sigma(j+n/2)}\right).$$

2.4 Polynomial Anticoncentration

In the proof of Theorems 1.1 and 1.3 we will use an anticoncentration inequality for polynomials of Rademacher random variables proved by Meka, Nguyen and Vu [33]. For $\boldsymbol{x} = (x_1, \ldots, x_n)$, consider a degree-$d$ polynomial

$$f(\boldsymbol{x}) = \sum_{S \in \binom{[n]}{\leq d}} f_S \boldsymbol{x}^S$$

in $\boldsymbol{x}$. The *rank* of f is the size of the largest matching in the d-uniform hypergraph on the vertex set $[n]$ obtained by putting an edge $S \subseteq [n]$ whenever $|S| = d$ and $f_S \neq 0$. The following theorem is a direct corollary of [33, Theorem 1.6].

Theorem 2.9. *Fix $d \in \mathbb{N}$ and let $\boldsymbol{\gamma} \in \mathrm{Rad}^n$. Let f be a degree-d polynomial with rank r. Then for any $\ell \in \mathbb{R}$,*

$$\Pr(f(\boldsymbol{\gamma}) = \ell) \leq \frac{\log^{O(1)}(r)}{\sqrt{r}}.$$

A drawback of Theorem 2.9 is that it ignores coefficients of f that are not of maximum degree. We deduce the following result, which allows us to take these coefficients into account under certain circumstances. We will need this for the proof of Theorem 1.3.

Corollary 2.10. *Fix $d \in \mathbb{N}$, let $\boldsymbol{\gamma} \in \mathrm{Rad}^n$, and consider a degree-$d$ polynomial*

$$f(\boldsymbol{x}) = \sum_{S \in \binom{[n]}{\leq d}} f_S \boldsymbol{x}^S.$$

8

Let $m_d = \max\{|f_S| : |S| = d\}$ be the maximum coefficient of degree d, and let H' be the $(d-1)$-uniform hypergraph with edge set $\{S : |S| = d-1, |f_S| \geq rm_d\}$. Suppose that H' has a matching of size r. Then for any $\ell \in \mathbb{R}$,

$$\Pr(f(\boldsymbol{\gamma}) = \ell) \leq \frac{\log^{O(1)}(r)}{\sqrt{r}}.$$

Proof. Let H be the d-uniform hypergraph used to define the rank of f, with an edge for every nonzero degree-d coefficient. If H has a matching of size $r/(2d)$ then we are done by Theorem 2.9. Otherwise, H has an independent set I of size larger than $n - r/2$. Condition on any outcome of the values γ_i for $i \notin I$. Now, $f(\boldsymbol{\gamma})$ can be expressed as a polynomial $g\big((\gamma_i)_{i \in I}\big)$ of the remaining γ_i, depending on the values of γ_i for $i \notin I$. This polynomial has degree at most $d-1$. Specifically, for $S \subseteq I$ with size $d-1$, the coefficient of S in g is

$$f_S + \sum_{i \notin I} f_{S \cup \{i\}} \gamma_i.$$

If $f_S \geq rm_d$, then this coefficient is nonzero (in fact, $f_S \geq (r/2)m_d$ suffices), so each edge of $H'[I]$ corresponds to a nonzero degree-$(d-1)$ coefficient in g. Moreover, by assumption $H'[I]$ has a matching of size at least $r - (n - |I|) \geq r/2$. Therefore, g has rank at least $r/2$, so the desired result again follows from Theorem 2.9. $\qquad \square$

3 Anticoncentration in graphs

In this section we present the proof of Theorem 1.1. It suffices to prove that $\Pr(X_{G,k} = \ell) = O\big(\log^{O(1)}(\ell^*/k)/\sqrt{\ell^*/k}\big)$ for every graph G with $2k$ vertices, because $I(n, k, \ell)$ is a monotone non-increasing function of n. So, let G be a graph on the vertex set $[n]$, for $n = 2k$. Let $X = X_{G,k}$. We may assume that $e(G) \leq \binom{n}{2}/2$, by taking the complement of G if necessary. We express X in the form $X(\boldsymbol{\xi}) = \sum_{1 \leq x < y \leq n} a_{xy} \xi_x \xi_y$, where the a_{xy} are the entries of the adjacency matrix of G and $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$.

Note that $\mathbb{E}X \approx e(G)/4$. We first want to use Lemma 2.1 to show that if $e(G)$ is not at least of the same order as ℓ then $\Pr(X_{G,k} = \ell)$ is very small.

Claim 3.1. *For any constant $\varepsilon > 0$, if $\ell \geq (1 + \varepsilon)\mathbb{E}X$ or $\ell \leq (1 - \varepsilon)\mathbb{E}X$ then*

$$\Pr(X = \ell) \leq \exp\left(-\Omega\left(\frac{\varepsilon^2 \ell}{k}\right)\right).$$

Proof. Note that X is of the form required to apply Lemma 2.1, with $c_x = \deg(x)$. Then

$$\begin{aligned}
\Pr(X = \ell) &\leq \exp\left(-\Omega\left(\frac{\varepsilon^2 e(G)^2}{\sum_{x=1}^n \deg(x)^2}\right)\right) \\
&\leq \exp\left(-\Omega\left(\frac{\varepsilon^2 e(G)^2}{n^2(e(G)/n)}\right)\right) \\
&\leq \exp\left(-\Omega\left(\frac{\varepsilon^2 \ell}{k}\right)\right),
\end{aligned}$$

where in the second inequality we have used the upper bound on $\sum_{x=1}^n \deg(x)^2$ obtained by taking $\deg(x) = n$ for as many x as possible, and in the third inequality we have used the fact that $\ell \leq (1 - \varepsilon)\mathbb{E}X = O(e(G))$. $\qquad \square$

From now on we will assume that $e(G) = \Omega(\ell)$, which also implies that $\ell = \Theta(\ell^*)$. Now, let σ be a uniformly random permutation of $[n]$, and let H be the (random) graph on the vertex set $[k]$ with an edge between i and j if

$$a_{\sigma(i)\sigma(j)} - a_{\sigma(i)\sigma(j+k)} - a_{\sigma(i+k)\sigma(j)} + a_{\sigma(i+k)\sigma(j+k)} \neq 0.$$

The heart of the proof of Theorem 1.1 is the following claim.

Claim 3.2. *The graph H has a matching of size $\Omega(\ell/k)$, with probability $1 - O(k/\ell)$.*

Before proving Claim 3.2, we will show how it implies Theorem 1.1.

Proof of Theorem 1.1. Let $\mathcal{E}$ be the event that H has a matching of size $\Omega(\ell/k)$. We learn from Lemma 2.8 (see the discussion at the end of Section 2.3) that X is a quadratic polynomial in $\gamma \in \text{Rad}^{n/2}$, and the coefficient of $\gamma_i \gamma_j$ is $\frac{1}{4}\left(a_{\sigma(i)\sigma(j)} - a_{\sigma(i)\sigma(j+k)} - a_{\sigma(i+k)\sigma(j)} + a_{\sigma(i+k)\sigma(j+k)}\right)$. Hence the rank of X (as a polynomial in the γ_i) is equal to the size of a maximum matching in H. Thus $\Pr(X = \ell | \mathcal{E}) \leq \log^{O(1)}(\ell/k)\sqrt{\frac{k}{\ell}}$, by Theorem 2.9. Combined with Claim 3.2, we obtain

$$\Pr(X = \ell) \leq \Pr(\overline{\mathcal{E}}) + \Pr(X = \ell | \mathcal{E}) \leq \log^{O(1)}(\ell/k)\sqrt{\frac{k}{\ell}}. \qquad \square$$

In order to prove Theorem 1.1, it remains to show Claim 3.2. As a warm-up, we first sketch the proof of Claim 3.2 in the regime where $\ell = \Omega(k^2)$. A key observation is that if $a_{vw} = a_{v'w'} \neq a_{v'w}$ (that is, the path $vwv'w'$ alternates between edges and non-edges) then $a_{vw} - a_{vw'} - a_{v'w} + a_{v'w'} \neq 0$. That is to say, edges in H arise from alternating paths of length 3 in G. When $\ell = \Omega(k^2)$, we can show that G has $\Omega(k^4)$ alternating 3-paths. Roughly speaking, the reason is that G can be divided into two parts V_1 and V_2 such that all vertices in V_1 have reasonably high degree and all vertices in V_2 have reasonably high non-degree. If there were many non-edges in V_1 or many edges in V_2 this would immediately give us many alternating 3-paths. Otherwise V_1 is almost a clique and V_2 is almost an independent set, in which case almost every pair of edges between V_1 and V_2 gives rise to an alternating 3-path through V_2, and almost every pair of non-edges between V_1 and V_2 gives rise to an alternating 3-path through V_1, and there must be many alternating 3-paths of at least one of these two types. Now, if G has $\Omega(k^4)$ alternating 3-paths, it follows from a concentration inequality that H is very likely to have $\Omega(k^2)$ edges, and hence a matching of size k.

In the general case this simplistic approach does not suffice, and the way we find our matching in H will differ slightly depending on the structure of G. Let $U \subseteq [n]$ be the set of "high-degree" vertices with degree at least $0.9n$. We divide the proof of Claim 3.2 into two cases: the case when many edges are incident to U will be handled in Section 3.1, and the case where many edges avoid U will be treated in Section 3.2.

3.1 Case 1: many edges are incident to the high-degree vertices

First, consider the case where $e(G)/2$ edges are incident to U. In this case, $2k|U| \geq e(G)/2$, so $|U| = \Omega(\ell/k)$. We can in fact assume that $|U| \geq 3$, because if $\ell = O(k)$ the statement of Theorem 1.1 is trivial. Now consider the following procedure to iteratively build a matching M in H. At step t, let V_t be the set of vertices $v \in [n]$ such that the value of $\sigma^{-1}(v)$ has not yet been revealed, choose any distinct $u, w \in U \cap V_t$, and reveal the values of $i = \sigma^{-1}(u)$ and $j = \sigma^{-1}(w)$. If $i, j \leq k$ and the values of $\sigma(i + k)$ and $\sigma(j + k)$, have not already been revealed, reveal them, and if we find that $\{i, j\}$ is an edge in H then we add $\{i, j\}$ to M.

The above procedure can continue while $4t + 2 \leq |U|$ (that is, $t \leq (|U| - 2)/4$). Let $T = \min\{(|U| - 2)/4, 0.01n\} = \Omega(\ell/k)$. We claim that every step $t \leq T$ has probability $\Omega(1)$ of successfully adding an edge to M.

10

Claim 3.3. *For any $t \leq T = \Omega(\ell/k)$, condition on any outcome of the information revealed before step t. Then the probability that an edge is added to M in step t is at least 0.02.*

Proof. Let Q_t be the set of indices $q \in [k]$ such that $\sigma(q)$ or $\sigma(q + k)$ have already been revealed in previous steps (that is, $\sigma(q) \notin V_t$ or $\sigma(q + k) \notin V_t$). Observe that $|Q_t| \leq 8T$. Now, reveal the values of $i = \sigma^{-1}(u)$ and $j = \sigma^{-1}(w)$. The probability that $i, j \leq k$ and $\sigma(i + k), \sigma(j + k) \in V_t$ is at least

$$\frac{k - |Q_t|}{n} \cdot \frac{k - |Q_t| - 1}{n} \geq \left(\frac{k - 8T - 1}{n}\right)^2 \geq 0.2. \tag{1}$$

Condition on such an outcome of i, j. Note that $|V_t| \geq n - 4t \geq 0.9n$, and recall that u and w have degree at least $0.9n$ (as vertices in U). It follows that $|N_{V_t}(u)|, |N_{V_t}(w)| \geq 0.8n$. Let $P \subseteq N_{V_t}(w) \times N_{V_t}(u)$ be the set of distinct ordered pairs of vertices (u', w') with $u' \in N_{V_t}(w) \setminus \{u\}$ and $w' \in N_{V_t}(u) \setminus \{w\}$, so that $|P| \geq (0.8n - 1)(0.8n - 2) \geq 0.6n^2$. Recall that we are assuming $e(G) \leq \binom{n}{2}/2$, so at most $n^2/2$ of the pairs in P are edges of G. Let P' be the set of pairs of P which are not edges, so that $|P'| \geq 0.1n^2$. Observe that if $(\sigma(i + k), \sigma(j + k)) \in P'$ then the vertices $\sigma(i), \sigma(j + k), \sigma(i + k), \sigma(j)$ form an alternating path, so

$$a_{\sigma(i)\sigma(j)} - a_{\sigma(i)\sigma(j+k)} - a_{\sigma(i+k)\sigma(j)} + a_{\sigma(i+k)\sigma(j+k)} = a_{\sigma(i)\sigma(j)} - 2 \neq 0,$$

meaning that $\{i, j\}$ is an edge of H and can be added to M. The probability of this is at least $|P'|/n^2 \geq 0.1$. Recall that this is a conditional probability, so we multiply by (1) for the desired result. $\qquad\square$

Claim 3.3 implies that the eventual size of M stochastically dominates the binomial distribution $\mathrm{Bin}(T, 0.02)$, so by the Chernoff bound we have $|M| \geq 0.01T$ with probability $1 - e^{-\Omega(\ell/k)}$.

3.2 Case 2: many edges avoid the high-degree vertices

Let $\overline{U} = [n] \setminus U$; it remains to consider the case where $\overline{U}$ induces at least $e(G)/2 = \Omega(\ell)$ edges. Observe that in $G[\overline{U}]$ we can greedily find a matching of size at least $s := e(G[\overline{U}])/k = \Omega(\ell/k)$; let $S \subseteq [n]^2$ be such a matching. Now consider the following procedure to iteratively build a matching M in H. At step t, let V_t be the set of vertices $v \in [n]$ such that the value of $\sigma^{-1}(v)$ has not yet been revealed, choose any $(u, v) \in S \cap V_t^2$, and reveal the values of $i = \sigma^{-1}(u)$ and $j = \sigma^{-1}(w)$. If $i, j \leq k$ and the values of $\sigma(i + k)$ and $\sigma(j + k)$, have not already been revealed, reveal them, and if we find that $\{i, j\}$ is an edge in H then we add $\{i, j\}$ to M.

The above procedure can continue while $4t + 1 \leq |S|$ (that is, $t \leq (|S| - 1)/4$). Let $T = \min\{(|S| - 1)/4, 0.01n\} = \Omega(\ell/k)$. As in Section 3.1, we claim that every step $t \leq T$ has probability $\Omega(1)$ of successfully adding an edge to M.

Claim 3.4. *For any $t \leq T = \Omega(\ell/k)$, condition on any outcome of the information revealed before step t. Then the probability that an edge is added to M in step t is at least 0.0004.*

Proof. Reveal the values of $i = \sigma^{-1}(u)$ and $j = \sigma^{-1}(w)$. As in Case 1, the probability that $i, j \leq k$ and $\sigma(i + k), \sigma(j + k) \in V_t$ is at least $((k - 8T - 1)/n)^2 \geq 0.2$. Condition on such an outcome of i, j. Note that $|V_t| \geq n - 4t \geq 0.95n$, and recall that u and w have degree at most $0.9n$, so $|V_t \setminus N(u)|, |V_t \setminus N(w)| \geq 0.05n$. Let $P \subseteq V_t^2$ be the set of distinct ordered pairs of vertices (u', w') with $u' \in (V_t \setminus N(u)) \setminus \{u, w\}$ and $w' \in (V_t \setminus N(w)) \setminus \{u, w\}$, so that $|P| \geq (0.05n - 2)(0.05n - 3) \geq 0.002n^2$. Observe that if $(\sigma(j + k), \sigma(i + k)) \in P$ then the vertices $\sigma(i + k), \sigma(j), \sigma(i), \sigma(j + k)$ form an alternating path and

$$a_{\sigma(i)\sigma(j)} - a_{\sigma(i)\sigma(j+k)} - a_{\sigma(i+k)\sigma(j)} + a_{\sigma(i+k)\sigma(j+k)} = 1 + a_{\sigma(i+k)\sigma(j+k)} \neq 0,$$

11

meaning that $\{i, j\}$ is an edge of H and can be added to M. The probability of this is at least $|P|/n^2 \geq 0.002$. $\qquad\square$

As in Section 3.1, it follows that the eventual size of M stochastically dominates the binomial distribution $\mathrm{Bin}(T, 0.0004)$, so by the Chernoff bound we have $|M| \geq 0.0002T$ with probability $1 - e^{-\Omega(\ell/k)}$.

4 "Weak" anticoncentration in hypergraphs

Theorem 1.2 will be an almost immediate consequence of Lemma 2.2. The only combinatorial fact we need is as follows.

Lemma 4.1. *The following holds for any r. Let G be an r-graph on $2k$ vertices. Then either G induces a clique on k vertices, or an independent set on k vertices, or else it induces two k-vertex subgraphs with different numbers of edges.*

Proof. We can assume $k \geq r$, because otherwise G trivially induces an independent set on k vertices. We will assume that every k-vertex subset of G induces the same number of edges, and prove that G is a clique or independent set.

We claim that for all $s \leq k$, $\deg(S)$ takes a constant value among vertex subsets S of size s. This will imply the desired result, because if S has r vertices, then $\deg(S)$ is either zero or one, depending on whether S is an edge in G. We prove our desired claim by induction on s, so assume it holds for all sizes less than some s. For $i < s$, let d_i be the common value of $\deg(S)$ among S of size i.

Now, let $\mathrm{St}(v)$ be the set of edges of G which contain a vertex v. For a set S of s vertices, by the inclusion-exclusion principle, the number of edges of G which intersect S is

$$\left| \bigcup_{v \in S} \mathrm{St}(v) \right| = \sum_{i=1}^{s} (-1)^{i+1} \left(\sum_{S' \in \binom{S}{i}} \left| \bigcap_{v \in S'} \mathrm{St}(v) \right| \right) = \sum_{i=1}^{s-1} (-1)^{i+1} \binom{s}{i} d_i + (-1)^{s+1} \deg(S).$$

So, if we had $S_1, S_2 \subseteq V$ with $|S_1| = |S_2| = s$ and $(-1)^{s+1} \deg(S_1) < (-1)^{s+1} \deg(S_2)$, it would imply that $e(V \setminus S_1) > e(V \setminus S_2)$. Since we are assuming $k \geq s$ and $|V| = 2k$, we would have $|V \setminus S_1| = |V \setminus S_2| \geq k$, so by averaging there would be k-vertex subsets $U_1 \subseteq V \setminus S_1$ and $U_2 \subseteq V \setminus S_2$ such that $e(U_1) > e(U_2)$. This would be a contradiction. $\qquad\square$

Now we can prove Theorem 1.2. Similarly to the proof of Theorem 1.1, by monotonicity it suffices to show that $\Pr(X_{G,k} = \ell) \leq 1 - \varepsilon$, for some ε depending only on r, whenever G is an r-graph on $n = 2k$ vertices. If $X_{G,k}$ is identically equal to 0 or $\binom{n}{k}$ then we are done, and otherwise, by Lemma 4.1, $X_{G,k}$ must be supported on at least two values (meaning that it is not a constant).

Then, note that we can express $X_{G,k}$ in the form $f(\boldsymbol{\xi})$, for f a polynomial of degree at most r and $\boldsymbol{\xi} \in \mathrm{BL}(n, k)$. The desired result therefore follows from Lemma 2.2.

5 Anticoncentration in dense 3-graphs

In this section we prove Theorem 1.3. Let G be a 3-graph on the vertex set $[n]$, for $n = 2k$; as in Section 3, it suffices to prove that $\Pr(X_{G,k} = \ell) \leq \log^{O(1)} k / \sqrt{k}$. Also, with the same arguments as in Claim 3.1, we can assume that $\min\{e(G), \binom{n}{3} - e(G)\} = \Omega(n^3)$.

Now, we express $X = X_{G,k}$ as

$$\sum_{1 \le x < y < z \le n} a_{xyz} \xi_x \xi_y \xi_z,$$

where $a_{xyz} = \mathbf{1}_{\{x,y,z\} \in E(G)}$ and $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$. One might hope that Conjecture 5 would follow from a 3-graph generalisation of the arguments in Theorem 1.1. Indeed, it would suffice to show that G has $\Omega(n^6)$ "good" 6-tuples of vertices (x, x', y, y', z, z') such that

$$a_{xyz} - a_{xyz'} - a_{xy'z} - a_{x'yz} + a_{xy'z'} + a_{x'yz'} + a_{x'y'z} - a_{x'y'z'} \tag{2}$$

is nonzero. Unfortunately, in contrast with the 2-uniform case in Theorem 1.1, there exist 3-graphs G satisfying $\min\{e(G), \binom{n}{3} - e(G)\} = \Omega(n^3)$ which have no good 6-tuple at all. It will be useful to classify all such 3-graphs, which we will do after making some definitions. Fix a set $\{x, x', y, y', z, z'\}$ of size 6, and let $\mathcal{F}$ be the set of all 3-graphs on this vertex set such that the expression in (2) is nonzero. We say that a 3-graph is $\mathcal{F}$-*free* if it induces no 3-graph from $\mathcal{F}$. Also, we define a family of $\mathcal{F}$-free 3-graphs as follows. For two disjoint sets A and B and a set of disjoint pairs $M \subseteq A \times B$, let $G_{A,B,M}$ be the 3-graph on the vertex set $A \cup B$, whose edges are the triples which intersect both A and B, except those triples which include a pair from M.

Lemma 5.1. *Consider an $\mathcal{F}$-free n-vertex 3-graph G with $\min\{e(G), \binom{n}{3} - e(G)\} = \Omega(n^3)$. Then, provided n is sufficiently large, G or its complement is of the form $G_{A,B,M}$, for some partition $A \cup B$ of the vertex set of G and some set of disjoint pairs $M \subseteq A \times B$.*

The proof of Lemma 5.1 involves some somewhat complicated casework, so we defer it to Section 5.1.

Now, under the coupling in Fact 2.7, X is a function of a random permutation $\sigma : [n] \to [n]$ and a random vector $\boldsymbol{\gamma} \in \mathrm{Rad}^{n/2}$. By Lemma 2.8, for any outcome of σ, X is a polynomial in the γ_i of degree at most 3. The coefficient g_{ijq} of $\gamma_i \gamma_j \gamma_q$ is

$$\sum_{\boldsymbol{b} \in \{0,1\}^3} (-1)^{|\boldsymbol{b}|} a_{\sigma(i+kb_1)\,\sigma(j+kb_2)\,\sigma(q+kb_3)}$$

(note that $|g_{ijq}| \le 4$) and the coefficient g_{ij} of $\gamma_i \gamma_j$ is

$$\deg(\sigma(i), \sigma(j)) - \deg(\sigma(i+k), \sigma(j)) - \deg(\sigma(i), \sigma(j+k)) + \deg(\sigma(i+k), \sigma(j+k)).$$

Let H be the random 3-graph on the vertex set $[n/2]$ with an edge $\{i, j, q\}$ whenever $g_{ijq} \ne 0$. First suppose that G contains $\Omega(n^6)$ induced subgraphs from $\mathcal{F}$, and let $N = \Omega(n^6)$ be the number of ordered 6-tuples (x, x', y, y', z, z') such that the expression in (2) is nonzero. Then

$$\mathbb{E}e(H) = \frac{n}{2}\left(\frac{n}{2} - 2\right)\left(\frac{n}{2} - 4\right)\frac{N}{n(n-1)\dots(n-5)} = \Theta(n^3)$$

Also, note that modifying σ by a transposition changes $e(H)$ by at most $2\binom{n}{2}$. By a McDiarmid-type concentration inequality for random permutations (see for example [32, Section 3.2]), we therefore have

$$\Pr(e(H) \le \mathbb{E}e(H)/2) = \exp\left(-\Omega\left(\frac{(n^3)^2}{n \cdot \binom{n}{2}^2}\right)\right) = e^{-\Omega(n)}.$$

But if $e(H) \ge \mathbb{E}e(H)/2 = \Omega(n^3)$ then we can greedily find a matching of size $\Omega(n)$ in H, and Theorem 2.9 finishes the proof.

It remains to consider the case where G contains $o\bigl(n^6\bigr)$ induced subgraphs from $\mathcal{F}$. In this case, by the induced hypergraph removal lemma (see [41, Theorem 6]), we can add and remove $o\bigl(n^3\bigr)$ edges from G to obtain a 3-graph with no induced subgraphs from $\mathcal{F}$. By Lemma 5.1, we can assume this 3-graph is of the form $G_{A,B,M}$. Note that only $O(|M|n) = o\bigl(n^3\bigr)$ edges of $G_{A,B,M}$ can involve a pair in M, so we can actually obtain $G' := G_{A,B,\emptyset}$ by adding and removing $o\bigl(n^3\bigr)$ edges from G. Recall that $\min\bigl\{e(G), \binom{n}{3} - e(G)\bigr\} = \Omega\bigl(n^3\bigr)$, so we must have $|A|, |B| = \Omega(n)$. Also observe that if $x, y' \in A$ and $x', y \in B$, then

$$\bigl|\deg_{G'}(x,y) - \deg_{G'}(x',y) - \deg_{G'}(x,y') + \deg_{G'}(x',y')\bigr| = \bigl|2(n-2) - |A| - |B|\bigr| = n - 4. \quad (3)$$

There are $\Omega\bigl(n^4\bigr)$ such choices of (x, x', y, y'). We claim that in fact there are $N' = \Omega\bigl(n^4\bigr)$ choices of (x, x', y, y') such that

$$\deg_G(x,y) - \deg_G(x',y) - \deg_G(x,y') + \deg_G(x',y') \geq n/2. \quad (4)$$

Indeed, recall that G' is obtained from G by adding and removing $o\bigl(n^3\bigr)$ edges, and adding or removing an edge from G can affect the value of the above expression by at most 1, for at most $O\bigl(n^2\bigr)$ 4-tuples (x, x', y, y'). Therefore there can be only $o\bigl(n^4\bigr)$ 4-tuples which satisfy (3) but not (4).

Now, let H' be the random graph on the vertex set $[n/2]$ with an edge $\{i, j\}$ if $g_{ij} \geq n/2$. We have

$$\mathbb{E}e\bigl(H'\bigr) = \frac{n}{2}\Bigl(\frac{n}{2} - 2\Bigr)\frac{N'}{n(n-1)(n-2)(n-3)} = \Theta\bigl(n^2\bigr),$$

and, as in the previous case, by a concentration inequality $e(H') = \Omega\bigl(n^2\bigr)$ with probability $1 - e^{-\Omega(n)}$, in which case H' has a matching of size $m = \Omega(n)$. The desired result then follows from Corollary 2.10, with $d = 3$ and $r = \min\{m, n/8\}$.

5.1 Characterising $\mathcal{F}$-free 3-graphs

In this subsection we prove Lemma 5.1, as an inductive consequence of the following two lemmas. Let $K_{a,b}^{(3)}$ be the complete bipartite 3-graph with parts of sizes a and b (meaning that the vertex set can be partitioned into two parts of sizes a and b, and the edges are those triples which intersect both parts).

Lemma 5.2. *Under the conditions of Lemma 5.1, G contains a copy of $K_{5,5}^{(3)}$ or its complement.*

Lemma 5.3. *Consider an $\mathcal{F}$-free 3-graph G and let v be one of its vertices. Suppose that $G - v = G_{A,B,M}$ for some partition $A \cup B$ of the vertex set of $G - v$ and some set of disjoint pairs $M \subseteq A \times B$. Suppose also that $|A|, |B| \geq 5$. Then there is a partition $A' \cup B'$ of the vertex set of G, and a set of disjoint pairs $M' \subseteq A' \times B'$, satisfying $A' \supseteq A$, $B' \supseteq B$, $M' \supseteq M$, such that $G = G_{A',B',M'}$.*

To prove Lemma 5.2, we use the following theorem due to Fox and Sudakov [18, Theorem 4.2], which states that certain natural "patterns" are unavoidable in hypergraphs with density bounded away from zero and one. This theorem was actually stated in [18] without proof (Fox and Sudakov were mainly concerned about an analogous theorem for graphs), so for completeness we include a proof in Appendix A.

Theorem 5.4. *Consider a red-blue colouring of the edges of the complete n-vertex 3-graph, with $\Omega\bigl(n^3\bigr)$ red edges and $\Omega\bigl(n^3\bigr)$ blue edges, and consider any $t \in \mathbb{N}$.*

If n is sufficiently large then G contains disjoint vertex subsets V_1, V_2, V_3 each of size t, such that for every function $f : \{1, 2, 3\} \to \{1, 2, 3\}$, all the edges $\{v_1, v_2, v_3\}$ with $v_i \in V_{f(i)}$ for $i \in \{1, 2, 3\}$ have the same colour, but the edge colouring of $V_1 \cup V_2 \cup V_3$ is not monochromatic.

Now we prove Lemma 5.2 and Lemma 5.3. Note that n-vertex 3-graphs are equivalent to red-blue colourings of the complete n-vertex 3-graph; we switch between these points of view interchangeably.

Proof of Lemma 5.2. We may assume that G contains disjoint vertex subsets V_1, V_2, V_3 satisfying the conclusion of Theorem 5.4, with $t = 5$. Suppose first that for some i_1, i_2, the set $V_{i_1} \cup V_{i_2}$ is not monochromatic and the induced colouring is not isomorphic to $K_{5,5}^{(3)}$ or its complement. For all remaining possibilities of the induced colouring, one can check that if $x, y, z \in V_{i_1}$ and $x', y', z' \in V_{i_2}$ then the expression in (2) is nonzero, meaning that G is not $\mathcal{F}$-free, which is a contradiction. Alternatively, if each $V_{i_1} \cup V_{i_2}$ is monochromatic, then there is only one possibility (up to swapping colours) for the colouring of $V_1 \cup V_2 \cup V_3$, and one can check that if $x, y, z \in V_1$, $x', y' \in V_2$ and $z' \in V_3$, then the expression in (2) is nonzero, which is again a contradiction. □

Proof of Lemma 5.3. First, consider $x_a, x'_a \in A$ and $x_b, x'_b \in B$ such that $(x_a, x'_b), (x'_a, x_b) \notin M$. By the assumption that $|B| \geq 5$, there is $x_b^* \in B \setminus \{x_b, x'_b\}$ such that $(x'_a, x_b^*), (x_a, x_b^*) \notin M$. If $(x, x', y, y', z, z') = (v, x_b^*, x_a, x_b, x'_a, x'_b)$ then the expression in (2) is equal to

$$1 + a_{vx_a x'_a} + a_{vx_b x'_b} - a_{vx_a x'_b} - a_{vx'_a x_b} = 0. \tag{5}$$

This implies that at most one of $\{v, x_a, x'_a\}$ and $\{v, x_b, x'_b\}$ can be an edge. But for any $x_a, x'_a \in A$ and any two vertices in B, we can assign those two vertices the labels x_b and $x_{b'}$ in such a way that $(x_a, x'_b), (x'_a, x_b) \notin M$. So either there is no edge containing v and two vertices from A, or there is no edge containing v and two vertices from B. Without loss of generality, suppose the former is the case.

Now, we wish to study the edges of the form $\{v, x_a, x_b\}$, for $x_a \in A$ and $x_b \in B$. Let Γ be the auxiliary bipartite graph on the vertex set $A \cup B$ with an edge $(x_a, x_b) \in A \times B$ if $\{v, x_a, x_b\}$ is *not* an edge in G.

Claim. *Either $\Gamma = M$ or Γ is obtained from M by adding every edge (x_a, x^*) incident to a single vertex $x^* \in B$ which does not appear in any pair of M.*

Proof of claim. For any $x_a, x'_a \in A$, $x_b, x'_b \in B$ with $(x_a, x'_b), (x'_a, x_b) \notin M$, by (5) at least one of $\{v, x_a, x'_b\}$ and $\{v, x'_a, x_b\}$ is an edge. So $\Gamma \setminus M$ does not have a matching of size 2, and by Kőnig's theorem the edges in $\Gamma \setminus M$ are all incident to a single vertex x^*, which may be in A or B. (If $\Gamma \setminus M$ consists of just one edge, then we can take $x^* \in B$, and if $\Gamma \setminus M$ has no edges, we set $x^* = \emptyset$).

Consider any $x_a, x'_a \in A$ and $x_b \in B$. Choose $x_b^* \in B \setminus \{x_b\}$ and $x_a^* \in A \setminus \{x_a, x'_a\}$ such that $(x_a^*, x_b), (x_a^*, x_b^*), (x_a, x_b^*), (x'_a, x_b^*) \notin M$. If $(x, x', y, y', z, z') = (v, x_b^*, x_b, x_a^*, x_a, x'_a)$ then the expression in (2) is equal to

$$\mathbf{1}_{(x'_a, x_b) \notin M} - \mathbf{1}_{(x_a, x_b) \notin M} + a_{vx_a x_b} - a_{vx'_a x_b} = 0. \tag{6}$$

If $(x_a, x_b) \in M$, then $(x'_a, x_b) \notin M$, so (6) immediately implies that $\{v, x'_a, x_b\}$ is an edge of G, and $\{v, x_a, x_b\}$ is not. This implies that $M \subseteq \Gamma$, and it also allows us to rule out the possibility that x^* is a vertex in B that appears in some pair of M (in this case we would have proved that $\Gamma \setminus M$ has no edges incident to x^*, which would contradict the choice of x^*).

Also, if $(x_a, x_b), (x'_a, x_b) \notin M$ then (6) implies that (x_a, x_b) is an edge of Γ if and only if (x'_a, x_b) is an edge of Γ. This rules out the possibility that $x^* \in A$, and proves that if $x^* \neq \emptyset$ then actually $(x_a, x^*) \in \Gamma$ for all x_a, finishing the proof of the claim □

Given the above claim, it now remains to consider edges of the form $\{v, x_b, x'_b\}$, for $x_b, x'_b \in B$. For any $x_b, x'_b \in B \setminus \{x^*\}$ we may choose $x_a, x'_a \in A$ such that $(x_a, x'_b), (x'_a, x_b) \notin M$, so by (5), $\{v, x_b, x'_b\}$ is an edge of G. If $\Gamma = M$ we may now conclude that $G = G_{A', B, M}$ for $A' = A \cup \{v\}$.

Alternatively, if $\Gamma \neq M$, we may similarly deduce from (5) that for any $x_b \in B \setminus \{x^*\}$, $\{v, x_b, x^*\}$ is not an edge of G. We may then conclude that $G = G_{A',B,M'}$ for $A' = A \cup \{v\}$ and $M' = M \cup \{(v, x^*)\}$. $\qquad\square$

6 Further directions of research

6.1 Possible generalisations and improvements

An obvious conjecture is that the logarithmic term in Theorem 1.1 is unnecessary, as follows.

Conjecture 4. *For all k and all $0 \leq \ell \leq \binom{k}{2}$, let $\ell^* = \min\left\{\ell, \binom{k}{2} - \ell\right\}$. We have*

$$\mathrm{ind}(k, \ell) = O\left(\sqrt{k/\ell^*}\right).$$

We remark that the logarithmic term in Theorem 1.1 arises purely because of the corresponding logarithmic term in Theorem 2.9. In turn, this logarithmic term arises from an estimate due to Kane for the so-called Gotsman–Linial conjecture, as follows. For an n-variable Boolean function $f : \{-1, 1\}^n \to \mathbb{R}$, the *average sensitivity* of f is defined by

$$\mathrm{AS}(f) = \sum_{i=1}^{n} \Pr(f(\gamma_1, \ldots, \gamma_{i-1}, \gamma_i, \gamma_{i+1}, \ldots, \gamma_n) \neq f(\gamma_1, \ldots, \gamma_{i-1}, -\gamma_i, \gamma_{i+1}, \ldots, \gamma_n)),$$

where $\boldsymbol{\gamma} \in \mathrm{Rad}^n$. The Gotsman–Linial conjecture essentially[3] states that if a Boolean function f has the form $f(\boldsymbol{x}) = \mathbf{1}_{p(\boldsymbol{x}) > 0}$ for a degree-d polynomial p (that is, f is a degree-d *polynomial threshold function*), then $\mathrm{AS}(f) = O(d\sqrt{n})$. Kane [24] proved that if f is a degree-$O(1)$ threshold function then $\mathrm{AS}(f) = \sqrt{n}\log^{O(1)} n$. To prove Conjecture 4 via the methods in [33] and the methods in Section 3, it would suffice to prove that $\mathrm{AS}(f) = O(\sqrt{n})$ under the same assumptions. In particular, a bound of the form $\mathrm{AS}(f) \leq g(n)$ would imply that under the conditions of Conjecture 4 we have $\mathrm{ind}(k, \ell) = O(g(\ell^*/k)k/\ell^*)$.

Next, the appropriate generalisation of Conjecture 4 to hypergraphs seems to be as follows.

Conjecture 5. *For any r, k and any $0 \leq \ell \leq \binom{k}{r}$, let $\ell^* = \min\left\{\ell, \binom{k}{r} - \ell\right\}$. We have*

$$\mathrm{ind}_r(k, \ell) = O\left(\sqrt{k^{r-1}/\ell^*}\right).$$

It is likely that the arguments in Section 5 can be pushed to hypergraphs with uniformity higher than three, but our proof of Theorem 1.3 involves some rather complicated checking of cases, which would likely be even more complicated for higher uniformities. To be specific, the part of the proof that really depends on the uniformity-3 assumption is Lemma 5.1, which (essentially) classifies the 3-graphs such that

$$a_{xyz} - a_{xyz'} - a_{xy'z} - a_{x'yz} + a_{xy'z'} + a_{x'yz'} + a_{x'y'z} - a_{x'y'z'} = 0$$

for every choice of distinct vertices (x, x', y, y', z, z'). In general, one might try to classify the r-graphs such that

---

[3]The original conjecture due to Gotsman and Linial [19] was slightly sharper than what is stated here, but it was recently disproved [8, 26].

$$\sum_{\boldsymbol{b} \in \{0,1\}^r} (-1)^{|\boldsymbol{b}|} a_{z_1^{b_1} \dots z_r^{b_r}} = 0$$

for every choice of distinct vertices $(z_1^0, z_1^1, z_2^0, z_2^1, \dots, z_r^0, z_r^1)$. As a natural family of r-graphs with this property, consider the complete $(r-1)$-partite r-graphs $K_{n_1,\dots,n_{r-1}}^{(r)}$, whose vertex set is partitioned into $r-1$ parts of sizes $n_1, \dots, n_{r-1}$, and whose edge set is defined to consist of all sets of r vertices which intersect every part. Is it true that every r-graph with the aforementioned property resembles some $K_{n_1,\dots,n_{r-1}}^{(r)}$? If so, using the same approach as for Theorem 1.3 one could prove that if $\ell^* = \Omega_k(k^r)$ then

$$\mathrm{ind}_r(k, \ell) \leq \frac{\log^{O(1)} k}{\sqrt{k}}.$$

However, since our proof uses the induced hypergraph removal lemma, new ideas would be necessary to address the "sparse" case where $\ell^* = o_k(k^r)$.

As observed in [4], it also makes sense to generalise Conjecture 1 to hypergraphs:

Conjecture 6. *For all $0 < \ell < \binom{k}{r}$ we have $\mathrm{ind}_r(k, \ell) \leq 1/e + o_k(1)$.*

It would be interesting even to prove a bound of the form $\mathrm{ind}_r(k, \ell) \leq 1 - \varepsilon + o_k(1)$ for any $\varepsilon > 0$ that does not depend on r; perhaps this can be accomplished with a careful bare-hands fourth-moment computation as in [4]. We remark however that the bound $\mathrm{ind}_r(k, \ell) \leq 1 - 2^{-4/3} 3^{-16r}$ in Theorem 1.2 is clearly not best-possible; instead of using hypercontractivity for $\mathrm{BL}(n, n/2)$ via Lemma 2.2, it would have been possible to give a less direct proof using the much more developed theory of hypercontractivity for Rad^n, then appealing to the invariance principle in [17]. This would have given the bound $\mathrm{ind}_r(k, \ell) \leq 1 - 9^{-r} + o_k(1)$. Of course, this still depends exponentially on r.

6.2 Ramsey graphs

It is also interesting to study the probabilities $\Pr(X_{G,k} = \ell)$ for restricted classes of (hyper)graphs G. If these probabilities are small it would seem to give some evidence that the graphs in question are very "diverse" or "disordered". In particular, say that a graph is *C-Ramsey* if it has no clique or independent set of size $C \log_2 n$. There has been a lot of work on diversity of Ramsey graphs from various points of view; in particular, Kwan and Sudakov [28] recently resolved a conjecture of Erdős, Faudree and Sós which effectively says that if G is an $O(1)$-Ramsey graph then for many values of k, the random variables $X_{G,k}$ have large support (see also [2, 6, 5, 1, 34, 29] for related work). It would be very interesting to study the probabilities $\Pr(X_{G,k} = \ell)$ for Ramsey graphs. For example, if G is an n-vertex $O(1)$-Ramsey graph, is it true that

$$\Pr\big(X_{G,n/2} = \ell\big) = O(1/n)$$

for all ℓ?

As mentioned in [29], it would also be interesting to ask a more tractable version of this question for $X = e(G[A])$, where A is a uniformly random vertex subset of an $O(1)$-Ramsey graph G. In this case X can be interpreted as a quadratic polynomial of a random vector $\boldsymbol{\gamma} \in \mathrm{Rad}^n$, so this question is closely related to a conjecture of Costello [9, Conjecture 3] attempting to characterise the quadratic polynomials f in n variables with point probabilities $\Pr(f(\boldsymbol{\gamma}) = \ell)$ much larger than $1/n$ (see also a related inverse theorem of Nguyen [35]).

6.3 Anticoncentration "on the slice"

Historically, almost all the work on anticoncentration has focused on sums or low-degree polynomials of independent random variables (for example, see the survey of Nguyen and Vu [36] concerning the Littlewood–Offord problem and its variants). A recent exception is a Littlewood–Offord-type theorem "on the slice" due to Litvak, Lytova, Tikhomirov, Tomczak-Jaegermann and Youssef [31, Proposition 4.10]. The aforementioned authors proved an upper bound on the point probabilities $\Pr(f(\boldsymbol{\xi}) = \ell)$ for an n-variable degree-1 polynomial f of a random vector $\boldsymbol{\xi} \in \mathrm{BL}(n, n/2)$, using the coupling in Fact 2.7, and used this result to study the singularity probability of a random zero-one matrix with fixed row and column sums. Specifically, they were able to show that $f(\boldsymbol{\xi})$ has good anticoncentration if f has many pairs of different coefficients, which means that f is far from a multiple of the polynomial $x_1 + \cdots + x_n$.

In this paper we generalised the above methods to higher-degree polynomials, and effectively showed that for a degree-d polynomial f, the random variable $f(\boldsymbol{\xi})$ has good anticoncentration if there are many $2d$-tuples of coefficients which satisfy a certain inequality. This is in some sense a combinatorial criterion, and it would be interesting if an algebraic criterion could be proved to also suffice. For example, does $f(\boldsymbol{\xi})$ have good anticoncentration whenever f is in some sense far from a polynomial with $(x_1 + \cdots + x_n)$ as a factor?

There is also the possibility that more natural anticoncentration theorems could be stated in terms of harmonic polynomials (recall the definition from Section 2.2), which are in some sense the correct representation of functions on the slice (see for example [17]). For harmonic polynomials of $\mathrm{BL}(n, pn)$ random vectors, the invariance principle in [17] can be used to apply standard anticoncentration theorems, but the error terms in this invariance principle prevent one from obtaining optimal bounds in this way.

Finally, one could also study anticoncentration phenomena for more general combinatorial random variables; for example, functions of random permutations, as in Hoeffding's combinatorial central limit theorem [22].

References

[1] N. Alon, J. Balogh, A. Kostochka, and W. Samotij, *Sizes of induced subgraphs of Ramsey graphs*, Combin. Probab. Comput. **18** (2009), no. 4, 459–476.

[2] N. Alon and B. Bollobás, *Graphs with a small number of distinct induced subgraphs*, Discrete Math. **75** (1989), no. 1-3, 23–30, Graph theory and combinatorics (Cambridge, 1988).

[3] N. Alon, G. Gutin, M. Krivelevich, *Algorithms with large domination ratio*, J. Algorithms **50** (2004), no. 1, 118–131.

[4] N. Alon, D. Hefetz, M. Krivelevich, and M. Tyomkyn, *Edge-statistics on large graphs*, arXiv preprint arXiv:1805.06848 (2018).

[5] N. Alon and A. V. Kostochka, *Induced subgraphs with distinct sizes*, Random Structures Algorithms **34** (2009), no. 1, 45–53.

[6] M. Axenovich and J. Balogh, *Graphs having small number of sizes on induced k-subgraphs*, SIAM J. Discrete Math. **21** (2007), no. 1, 264–272.

[7] J. Balogh, P. Hu, B. Lidický, and F. Pfender, *Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle*, European J. Combin. **52** (2016), no. part A, 47–58.

[8] B. Chapman, *The Gotsman-Linial conjecture is false*, Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA, 2018, pp. 692–699.

[9] K. P. Costello, *Bilinear and quadratic variants on the Littlewood-Offord problem*, Israel J. Math. **194** (2013), no. 1, 359–394.

[10] K. P. Costello, T. Tao, and V. Vu, *Random symmetric matrices are almost surely nonsingular*, Duke Math. J. **135** (2006), no. 2, 395–413.

[11] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab. **6** (1996), no. 3, 695–750.

[12] C. F. Dunkl, *A Krawtchouk polynomial addition theorem and wreath products of symmetric groups*, Indiana Univ. Math. J. **25** (1976), no. 4, 335–358.

[13] C. F. Dunkl, *Orthogonal functions on some permutation groups*, Relations between combinatorics and other parts of mathematics (Proc. Sympos. Pure Math., Ohio State Univ., Columbus, Ohio, 1978), Proc. Sympos. Pure Math., XXXIV, Amer. Math. Soc., Providence, R.I., 1979, pp. 129–147.

[14] P. Erdős, *On extremal problems of graphs and generalized graphs*, Israel J. Math. **2** (1964), no. 3, 183–190.

[15] Y. Filmus, *An orthogonal basis for functions over a slice of the Boolean hypercube*, Electron. J. Combin. **23** (2016), no. 1, Paper 1.23, 27.

[16] Y. Filmus, G. Kindler, E. Mossel, and K. Wimmer, *Invariance principle on the slice*, 31st Conference on Computational Complexity, LIPIcs. Leibniz Int. Proc. Inform., vol. 50, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016, pp. Art. No. 15, 10.

[17] Y. Filmus and E. Mossel, *Harmonicity and invariance on slices of the Boolean cube*, 31st Conference on Computational Complexity, LIPIcs. Leibniz Int. Proc. Inform., vol. 50, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016, pp. Art. No. 16, 13.

[18] J. Fox and B. Sudakov, *Unavoidable patterns*, J. Combin. Theory Ser. A **115** (2008), no. 8, 1561–1569.

[19] C. Gotsman and N. Linial, *Spectral properties of threshold functions*, Combinatorica **14** (1994), no. 1, 35–50.

[20] C. Greenhill, M. Isaev, M. Kwan, and B. D. McKay, *The average number of spanning trees in sparse graphs with given degrees*, European J. Combin. **63** (2017), 6–25.

[21] D. Hefetz and M. Tyomkyn, *On the inducibility of cycles*, J. Combin. Theory Ser. B **133** (2018), 243–258.

[22] W. Hoeffding, *A combinatorial central limit theorem*, Ann. Math. Statistics **22** (1951), 558–566.

[23] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Cambridge University Press, 2000.

[24] D. M. Kane, *The correct exponent for the Gotsman-Linial conjecture*, Comput. Complexity **23** (2014), no. 2, 151–175.

[25] T. Kővári, V. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloq. Math. **1** (1954), no. 3, 50–57.

[26] H. W. Kim, C. Maldonado, and J. Wellens, *On graphs and the Gotsman-Linial conjecture for $d = 2$*, arXiv preprint arXiv:1709.06650 (2017).

[27] D. Král', S. Norin, and J. Volec, *A bound on the inducibility of cycles*, J. Combin. Theory Ser. A **161**, arXiv preprint arXiv:1801.01556 (2018).

[28] M. Kwan and B. Sudakov, *Proof of a conjecture on induced subgraphs of Ramsey graphs*, Trans. Amer. Math. Soc., to appear, arXiv preprint arXiv:1712.05656 (2017).

[29] M. Kwan and B. Sudakov, *Ramsey graphs induce subgraphs of quadratically many sizes*, Int. Math. Res. Not. (IMRN), to appear, arXiv preprint arXiv:1711.02937 (2017).

[30] T.-Y. Lee and H.-T. Yau, *Logarithmic Sobolev inequality for some models of random walks*, Ann. Probab. **26** (1998), no. 4, 1855–1873.

[31] A. E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef, *Adjacency matrices of random digraphs: singularity and anti-concentration*, J. Math. Anal. Appl. **445** (2017), no. 2, 1447–1491.

[32] C. McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms Combin., vol. 16, Springer, Berlin, 1998, pp. 195–248.

[33] R. Meka, O. Nguyen, and V. Vu, *Anti-concentration for polynomials of independent random variables*, Theory Comput. **12** (2016), Paper No. 11, 16.

[34] B. Narayanan, J. Sahasrabudhe, and I. Tomon, *Ramsey graphs induce subgraphs of many different sizes*, Combinatorica, to appear, arXiv preprint arXiv:1609.01705 (2016).

[35] H. H. Nguyen, *Inverse Littlewood-Offord problems and the singularity of random symmetric matrices*, Duke Math. J. **161** (2012), no. 4, 545–586.

[36] H. H. Nguyen and V. H. Vu, *Small ball probability, inverse theorems, and applications*, Erdös centennial, Bolyai Soc. Math. Stud., vol. 25, János Bolyai Math. Soc., Budapest, 2013, pp. 409–463.

[37] R. O'Donnell, *Analysis of Boolean functions*, Cambridge University Press, New York, 2014.

[38] N. Pippenger and M. C. Golumbic, *The inducibility of graphs*, J. Combin. Theory Ser. B **19** (1975), no. 3, 189–203.

[39] H. J. Prömel, *Ramsey Theory for Discrete Structures*, Springer, New York, 2013.

[40] A. Razborov and E. Viola, *Real advantage*, ACM Trans. Comput. Theory **5** (2013), no. 4, Art. 17, 8.

[41] V. Rödl and M. Schacht, *Generalizations of the removal lemma*, Combinatorica **29** (2009), no. 4, 467–501.

[42] R. Yuster, *On the exact maximum induced density of almost all graphs and their inducibility*, J. Combin. Theory Ser. B, to appear, arXiv preprint arXiv:1801.01047 (2018).

A Unavoidable patterns in Hypergraphs

For the sake of completeness and the convenience of the reader, in this section we provide a proof of (a generalisation of) Theorem 5.4. This theorem appears as [18, Theorem 4.2], but was stated without proof.

Theorem A.1. *For each $\varepsilon > 0$ and positive integers r and t, there is a positive integer $N = N(t, r, \varepsilon)$ such that the following holds. Consider any red-blue colouring of the complete r-graph with $n \geq N$ vertices which has at least εn^r edges in each colour. Then there are disjoint vertex subsets $V_1, \ldots, V_r$, each of size t, such that for every function $f : \{1, \ldots, r\} \to \{1, \ldots, r\}$, all the edges $\{v_1, \ldots, v_r\}$ with $v_i \in V_{f(i)}$ for $i \in \{1, \ldots, r\}$ have the same colour, but the edge colouring of $V_1 \cup \cdots \cup V_r$ is not monochromatic.*

We remark that our proof will actually give the slightly stronger statement that the edges contained in V_1 have a different colour than the edges with a vertex in each V_i.

At the heart of our proof of Theorem A.1 is the following lemma.

Lemma A.2. *For each $\varepsilon > 0$ and positive integers r and q, there is a positive integer $M = M(q, r, \varepsilon)$ such that the following holds. Consider any red-blue colouring of the complete r-graph with $n \geq M$ vertices which has at least εn^r edges in each colour. Then there exist disjoint vertex subsets $V_1, \ldots, V_{r-1}, R, B$, each of size q, such that*
  (i) *$\{v_1, \ldots, v_{r-1}, v\}$ is coloured red for $v_1 \in V_1, \ldots, v_{r-1} \in V_{r-1}, v \in R$, and*
  (ii) *$\{v_1, \ldots, v_{r-1}, v\}$ is coloured blue whenever $v_1 \in V_1, \ldots, v_{r-1} \in V_{r-1}, v \in B$.*

We will deduce Lemma A.2 from the following lemma together with some classical extremal results.

Lemma A.3. *For any $r \in \mathbb{N}$ and $\varepsilon > 0$, let $\alpha_r(\varepsilon) = (\varepsilon/3)^{4^r}$. Consider a red-blue colouring of $\binom{[n]}{r}$ with at least εn^r edges in each colour. Then there are $\alpha_r(\varepsilon)n^{r-1}$ $(r-1)$-sets of vertices that are simultaneously contained in $\alpha_r(\varepsilon)n$ red edges and $\alpha_r(\varepsilon)n$ blue edges.*

Proof. We will prove by induction the stronger version of this statement for *partial* red-blue colourings where we allow $\alpha_{r+1}(\varepsilon)n^r$ edges to be uncoloured. The base case $r = 1$ is trivial, so assume that this claim holds for all uniformities less than some $r \geq 2$.

Suppose to the contrary that this statement is false. Let $\mathcal{R}$ contain the $(r-1)$-sets of vertices which are in fewer than $\alpha_r(\varepsilon)n$ blue edges, and let $\mathcal{B}$ contain the $(r-1)$-sets which are in fewer than $\alpha_r(\varepsilon)n$ red edges. Then let $\mathcal{S}$ contain the remaining $(r-1)$-sets, which are each simultaneously contained in $\alpha_r(\varepsilon)n$ red edges and $\alpha_r(\varepsilon)n$ blue edges. We are assuming that $|\mathcal{S}| < \alpha_r(\varepsilon)n^{r-1}$. The number of red edges is then at most $|\mathcal{R}|n + |\mathcal{S}|n + \alpha_r(\varepsilon)n|\mathcal{B}| \leq |\mathcal{R}|n + \alpha_r(\varepsilon)n^r + \alpha_r(\varepsilon)n^r$, so $|\mathcal{R}| \geq (\varepsilon - \alpha_r(\varepsilon) - \alpha_r(\varepsilon))n^{r-1} \geq (\varepsilon/3)n^{r-1}$. We may similarly deduce that $|\mathcal{B}| \geq (\varepsilon/3)n^{r-1}$.

Now, consider the red-blue colouring of the edges of the complete n-vertex $(r-1)$-graph, where we colour an edge red if it is in $\mathcal{R}$ and blue if it is in $\mathcal{B}$. By induction, since at most $|\mathcal{S}| < \alpha_r(\varepsilon)n^{r-1}$ edges are uncoloured, there are $\alpha_{r-1}(\varepsilon/3)n^{r-2}$ $(r-2)$-sets that are simultaneously contained in $\alpha_{r-1}(\varepsilon/3)n$ sets from $\mathcal{R}$ and $\alpha_{r-1}(\varepsilon/3)n$ sets from $\mathcal{B}$. Since fewer than $\alpha_{r+1}(\varepsilon)n^r$ edges are uncoloured in total, we can find such an $(r-2)$-set Z with the extra property that Z is contained in at most

$$\frac{\binom{r}{2}\alpha_{r+1}(\varepsilon)n^r}{\alpha_{r-1}(\varepsilon/3)n^{r-2}} < \alpha_r(\varepsilon)n^2$$

uncoloured edges. Let $\mathcal{Q}$ be the collection of all r-sets of the form $X \cup Y$, for $Z \subseteq X \in \mathcal{R}$ and $Z \subseteq Y \in \mathcal{B}$. Note that X, Y are uniquely determined by their union because we can write

$(X \cup Y) \setminus Z = \{x, y\}$, where $\{x\} \cup Z = X$ is red and $\{y\} \cup Z = Y$ is blue. So $|\mathcal{Q}| \geq (\alpha_{r-1}(\varepsilon/3)n)^2$. On the other hand, let $N_Z \leq n$ be the number of $X \in \mathcal{R}$ which include Z. By the choice of $\mathcal{R}$, fewer than $N_Z \alpha_r(\varepsilon)n \leq \alpha_r(\varepsilon)n^2$ of elements of $\mathcal{Q}$ are blue, and by the choice of $\mathcal{B}$ fewer than $\alpha_r(\varepsilon)n^2$ are red. But we have seen above that fewer than $\alpha_r(\varepsilon)n^2$ are uncoloured, so $|\mathcal{Q}| \leq 3\alpha_r(\varepsilon)n^2$. One can check that

$$\alpha_{r-1}(\varepsilon/3)^2 > 3\alpha_r(\varepsilon),$$

yielding our desired contradiction. $\qquad\square$

Proof of Lemma A.2. By Lemma A.3, there is a collection $\mathcal{S}$ of $\Omega(n^{r-1})$ $(r-1)$-sets of vertices that are simultaneously contained in $\Omega(n)$ red edges and $\Omega(n)$ blue edges. Let G_{red} be the bipartite graph with vertex set $V(G_{\mathrm{red}}) = \mathcal{S} \cup [n]$ and edge set $E(G_{\mathrm{red}}) = \{(S, v) : S \in \mathcal{S}, v \in [n], S \cup \{v\} \text{ is red}\}$, and define G_{blue} in exactly the same way, using blue edges instead of red edges. As $e(G_{\mathrm{red}}) = \Omega(|\mathcal{S}|n) = \Omega(n^r)$, it follows from the Kővári–Sós–Turán theorem [25] that G_{red} must contain a complete bipartite graph with parts $\mathcal{S}' \subseteq \mathcal{S}$ and $R \subseteq [n]$ satisfying $|\mathcal{S}'| = |\mathcal{S}|^{1-o(1)} = n^{r-1-o(1)}$ and $|R| = q$. Similarly, applying the Kővári–Sós–Turán theorem to the induced subgraph $G_{\mathrm{blue}}[\mathcal{S}' \cup [n]]$, we can find a complete bipartite subgraph of G_{blue} with parts $\mathcal{S}'' \subset \mathcal{S}'$ and $B \subset [n]$ such that $|\mathcal{S}''| = |\mathcal{S}'|^{1-o(1)} = n^{r-1-o(1)}$ and $|B| = q$. Since $\mathcal{S}''$ is an $(r-1)$-graph on $[n]$ with $n^{r-1-o(1)}$ edges, a result due to Erdős [14, Theorem 1] (essentially generalising the Kővári–Sós–Turán theorem to hypergraphs) tells us that $\mathcal{S}''$ contains a complete $(r-1)$-partite $(r-1)$-graph whose parts $V_1, \ldots, V_{r-1}$ have the same size q. Clearly, the vertex subsets $V_1, \ldots, V_{r-1}, R, B$ have the desired properties. $\qquad\square$

To prove Theorem A.1 we also need the following Ramsey-type result.

Lemma A.4. *For all $r, t \in \mathbb{N}$, there is $Q_r(t) \in \mathbb{N}$ such that the following holds. Consider a red-blue colouring of the edges of the complete r-graph, and consider vertex sets $V_1', \ldots, V_r'$ each of size at least $Q_r(t)$, such that all the edges with a vertex in each V_i' have the same colour, and all the edges within V_1' have the other colour. Then there are subsets $V_i \subseteq V_i'$ of size t satisfying the conclusion of Theorem A.1.*

Proof. For every function $f : \{1, \ldots, r\} \to \{1, \ldots, r\}$, we say that $V_1, \ldots, V_r$ is *f-good* if all the edges $\{v_1, \ldots, v_r\}$ with $v_i \in V_{f(i)}$ for $i \in \{1, \ldots, r\}$ have the same colour. Provided $Q_r(t)$ is large enough, we can iteratively apply the Product Ramsey Theorem (see for example [39, Theorem 9.2]) to shrink the V_i' until they are f-good for every f. $\qquad\square$

We are finally ready to prove Theorem A.1.

Proof of Theorem A.1. By Ramsey's theorem, there is a function $R_r : \mathbb{N} \to \mathbb{N}$ such that every red-blue colouring of the edges of the complete r-graph on $R_r(k)$ vertices has a monochromatic k-clique.

We apply Lemma A.2 to obtain vertex subsets $V_1'', V_2'' \ldots, V_{r-1}'', R, B$ each of size $R_r(Q_r(t))$. Let V_1' be a monochromatic $Q_r(t)$-clique in V_1'', and assume without loss of generality that it is red. Choose $V_2' \subset V_2'', \ldots, V_{r-1}' \subset V_{r-1}''$ and $V_r' \subset B$ such that $|V_i'| = Q_r(t)$ for every $2 \leq i \leq r$. Then apply Lemma A.4. $\qquad\square$