

NEARLY TIGHT BOUNDS FOR MAXCUT IN HYPERGRAPHS

OLIVER JANZER* AND JULIEN PORTIER†

ABSTRACT. An r -cut of a k -uniform hypergraph is a partition of its vertex set into r parts, and the size of the cut is the number of edges which have at least one vertex in each part. The study of the possible size of the largest r -cut in a k -uniform hypergraph was initiated by Erdős and Kleitman in 1968. For graphs, a celebrated result of Edwards states that every m -edge graph has a 2-cut of size $m/2 + \Omega(m^{1/2})$, which is sharp. In other words, there exists a cut which exceeds the expected size of a random cut by the order of $m^{1/2}$. Conlon, Fox, Kwan and Sudakov proved that any k -uniform hypergraph with m edges has an r -cut whose size is $\Omega(m^{5/9})$ larger than the expected size of a random r -cut, provided that $k \geq 4$ or $r \geq 3$. They further conjectured that this can be improved to $\Omega(m^{2/3})$, which would be sharp. Recently, R  ty and Tomon improved the bound $m^{5/9}$ to $m^{3/5-o(1)}$ when $r \in \{k-1, k\}$. Using a novel approach, we prove the following approximate version of the Conlon–Fox–Kwan–Sudakov conjecture: for each $\varepsilon > 0$, there is some $k_0 = k_0(\varepsilon)$ such that for all $k > k_0$ and $2 \leq r \leq k$, in every k -uniform hypergraph with m edges there exists an r -cut exceeding the random one by $\Omega(m^{2/3-\varepsilon})$. Moreover, we show that (if $k \geq 4$ or $r \geq 3$) every k -uniform linear hypergraph has an r -cut exceeding the random one by $\Omega(m^{3/4})$, which is tight and proves a conjecture of R  ty and Tomon.

1. INTRODUCTION

Given a graph G , a *cut* refers to a partition of the vertex set into two subsets, with the *size* of the cut defined as the number of edges that have one vertex in each of these subsets. The MaxCut of a graph G is the maximum size of a cut over all possible bipartitions of the vertex set of G , and is denoted by $\text{mc}(G)$. The MaxCut problem has been central in both combinatorics and theoretical computer science, and appears in Karp’s famous list of 21 NP-complete problems [27]. As a simple probabilistic argument shows that $\text{mc}(G) \geq m/2$ for any graph G with m edges, the focus has been on proving bounds on the *surplus* $\text{sp}(G)$ of a graph, defined as $\text{sp}(G) = \text{mc}(G) - m/2$. A celebrated result by Edwards [13, 14] shows that $\text{sp}(G) \geq (\sqrt{8m+1}-1)/8$, which is tight for a complete graph with an odd number of vertices. Very recently, an inverse theorem for MaxCut has been established by Balla, Hambardzumyan and Tomon [6], stating that if a graph G with m edges has $\text{sp}(G) = O(\sqrt{m})$, then G contains a clique of order $\Omega(\sqrt{m})$. Much work has been devoted to establishing stronger lower bounds for $\text{sp}(G)$ under the assumption that G is H -free, for some fixed graph H . Those efforts were initiated by Erdős and Lov  sz in the 70s (see [15]), and a substantial amount of research has since focused in this direction, see for instance [1, 3, 4, 10, 19, 7, 26]. Despite remarkable recent breakthroughs on this topic with far-reaching applications beyond Graph Theory (see, e.g. [26]), tight bounds seem to remain out of reach in most cases.

A central problem in the area concerns the extension of Edwards’ result to hypergraphs. An r -cut of a k -uniform hypergraph G is a partition of the vertex set of G into r parts, and the *size* of the cut is defined as the number of hyperedges having at least one vertex in each part of the cut. We are then interested in the maximum size of an r -cut over all possible partitions of the vertex set of G , which is denoted by $\text{mc}_r(G)$. As shown by Erdős and Kleitman [16], by assigning each vertex to one of the r parts independently and uniformly at random, it follows that $\text{mc}_r(G) \geq \frac{S(k,r)r!}{r^k}m$ for every graph G with m edges, where $S(k,r)$ is the Stirling number

*Institute of Mathematics, EPFL, Lausanne, Switzerland. oliver.janzer@epfl.ch

†Institute of Mathematics, EPFL, Lausanne, Switzerland. julien.portier@epfl.ch

of second kind, denoting the number of unlabelled partitions of $\{1, \dots, k\}$ into r nonempty sets. The r -surplus (or r -excess) of G is then defined as $\text{sp}_r(G) = \text{mc}_r(G) - \frac{S(k,r)r!}{r^k}m$. Similarly, we say that the surplus of an r -cut in a k -graph with m edges is the size of the cut minus $\frac{S(k,r)r!}{r^k}m$. Conlon, Fox, Kwan and Sudakov [12] showed that for hypergraphs, Edwards' bound can be significantly improved.

Theorem 1.1 (Conlon–Fox–Kwan–Sudakov [12]). *For every $2 \leq r \leq k$ with $(r, k) \neq (2, 2)$ and $(r, k) \neq (2, 3)$, every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{5/9})$.*

They noted that for $(r, k) = (2, 3)$, Steiner triple systems show that there are 3-uniform hypergraphs on m edges with 2-surplus $\Theta(\sqrt{m})$. They furthermore observed that with high probability the binomial random k -graph $G_k(n, n^{3-k})$ has r -surplus $O(m^{2/3})$, where m is the number of its edges. This disproved a conjecture of Scott [35], which had predicted that the complete k -graph has the smallest 2-surplus.

Conlon et al. [12] conjectured that this random hypergraph is asymptotically optimal whenever $(r, k) \neq (2, 2), (2, 3)$.

Conjecture 1.2 (Conlon–Fox–Kwan–Sudakov [12]). *For any $2 \leq r \leq k$ with either $k \geq 4$ or $r \geq 3$, every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{2/3})$.*

In the particular case where $r \in \{k-1, k\}$, Raty and Tomon [34] recently improved Conlon, Fox, Kwan and Sudakov's bound to $\Omega(m^{3/5-o(1)})$ by using spectral techniques.

Theorem 1.3 (Raty–Tomon [34]). *If $r \in \{k-1, k\}$ and $r \geq 3$, then every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{3/5-o(1)})$.*

Raty and Tomon pointed out that their methods do not extend to the range $2 \leq r \leq k-2$. Obtaining strong results in the case $r = 2$ is of particular interest for several reasons. Firstly, it is this case that has been the most extensively studied in Theoretical Computer Science (see, e.g. [22, 21]). This problem is known as *Max E_k -Set Splitting*, or, in the case the hypergraph is not uniform, as *Max Set Splitting*. Secondly, this case has close connections to other well-studied problems such as negative (or positive) discrepancy [9], bisection width [33] and hypergraph colourings [29]. In the concluding remarks we will mention our upcoming work which extends the techniques in this paper to the setting of positive discrepancy and bisection width.

Using a novel approach, we prove the following approximate version of the Conlon–Fox–Kwan–Sudakov conjecture for all values of r .

Theorem 1.4. *For any $\varepsilon > 0$, there exists some $k_0 = k_0(\varepsilon)$ such that for any $k > k_0$ and $2 \leq r \leq k$, every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$.*

In the case of linear hypergraphs, still for $r \in \{k-1, k\}$, Raty and Tomon [34] proved a stronger bound $\Omega(m^{3/4-o(1)})$. They noted that this is asymptotically optimal, by constructing k -graphs with r -surplus $O(m^{3/4})$ in which each pair of vertices is in at most $O(\log m)$ edges (i.e., the hypergraphs are nearly linear). They conjectured that the $o(1)$ term can be removed and that the result can be extended to all $2 \leq r \leq k$.

We completely resolve this conjecture.

Theorem 1.5. *For any $2 \leq r \leq k$ with either $k \geq 4$ or $r \geq 3$, every k -uniform linear hypergraph with m edges has an r -cut of surplus $\Omega(m^{3/4})$.*

We also show that this is tight up to a constant factor by constructing linear (not just nearly linear) k -graphs with r -surplus $O(m^{3/4})$.

Proposition 1.6. *Fix some integers $2 \leq r \leq k$. There is a constant $\alpha = \alpha(k)$ such that, for every $m \in \mathbb{N}$, there exists an m -edge k -uniform linear hypergraph with r -surplus at most $\alpha m^{3/4}$.*

Finally, while Conlon, Fox, Kwan and Sudakov's construction of a 4-uniform hypergraph with 2-surplus $O(m^{2/3})$ is probabilistic, we show that highly structured constructions can achieve the same bound, which perhaps demonstrates the difficulty of proving tight lower bounds. Recall that an $(n, s, 2)$ -Steiner system is an n -element set V together with a set of s -element subsets of V (called *blocks*) with the property that each pair of elements in V is contained in exactly one block. Note that finite projective planes are examples of such Steiner systems with $s = \Theta(\sqrt{n})$.

Proposition 1.7. *Let $s \geq 4$ and let $S_1, \dots, S_t$ be the blocks of an $(n, s, 2)$ -Steiner system on ground set V . Let G be the 4-uniform hypergraph on vertex set V obtained by placing a complete 4-graph in each S_i . Then $e(G) = \Theta(n^2 s^2)$ and $\text{sp}(G) = O(n^2 + ns^2)$. In particular, when $s = \Theta(\sqrt{n})$ (which holds for projective planes), we have $\text{sp}(G) = O(e(G)^{2/3})$.*

Note that in any $(n, s, 2)$ -Steiner system, we have either $s = n$ or $s = O(\sqrt{n})$, so in almost all cases the n^2 term dominates the ns^2 term.

2. PROOF IDEAS

While our proofs use certain ideas from both [12] and [34], they are mostly different from these previous arguments. One similarity to [34] is that we also rely on the semidefinite programming (SDP) approach. This influential technique was developed by Goemans and Williamson [20] to find approximation algorithms for MaxCut in graphs. Carlson et al. [10] adapted this method to prove lower bounds for the MaxCut of H -free graphs, and this approach was further developed in a series of papers (see, e.g. [19, 7, 32]).

Let us first recall the basic SDP approach, after which we explain how we will use it and what will be our main innovations. Note that finding a cut with large surplus in a graph G amounts to finding an assignment $f : V(G) \rightarrow \{-1, +1\}$ such that $\sum_{uv \in E(G)} f(u)f(v)$ is quite negative. The SDP approach aims to instead find a unit vector $\mathbf{x}_u \in \mathbb{R}^N$ for each $u \in V(G)$ (where N is arbitrary but is the same for all $u \in V(G)$) such that $\sum_{uv \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle$ is quite negative. Clearly, the minimum of this “relaxed” optimization problem is no larger than the minimum for the original problem. Moreover, a powerful generalization of the classical Grothendieck inequality (proved in [30, 11, 5]) states that the solution of these two optimization problems are the same within a $\log(|V(G)|)$ factor, so, to prove lower bounds for the surplus of G , it suffices to construct unit vectors $\mathbf{x}_u$ for each $u \in V(G)$ such that $\sum_{uv \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle$ is quite negative.

At this point we remark that there have been two rather different approaches for constructing the vectors $\mathbf{x}_u$. The first approach is combinatorial, using the edges of the graph as an input. This was the approach taken by [10] and it was significantly further developed in [19]. The second approach uses algebraic and spectral information to define the vectors: the algebraic viewpoint was first used in [7], and a new, spectral approach was introduced in the influential paper [32] by Raty, Sudakov and Tomon. It is the spectral approach which was used also in the paper of Raty and Tomon [34] on MaxCut in hypergraphs.

2.1. The basics of our approach. In this paper, we use a combinatorial/probabilistic SDP approach, combined with ingredients from modular arithmetics. Moreover, we develop new tools that make the SDP technique widely applicable in the setting of hypergraphs. Let us first see how finding large 2-cuts in graphs (or hypergraphs) can help finding large r -cuts in k -graphs. We start with the case $r = k$, as this one turns out to be the easiest to deal with. We also assume that $k \geq 4$; the $k = 3$ case requires a slightly different model. By simple reductions, finding a large k -cut in a k -graph H can be reduced to finding a large 2-cut in a certain auxiliary random multigraph G associated to H . The reason is as follows.

Partition the vertex set of H as $V \cup C$, where each vertex in H is independently placed in V with probability $2/k$ and in C with probability $(k-2)/k$. We will commit to select the parts $1, \dots, k-2$ in our k -cut from C and the parts $k-1$ and k from V . Note that at this stage we may restrict our attention to those hyperedges e in H which have precisely 2 vertices in V (and $k-2$ vertices in C), otherwise e will certainly not be cut. The next step is to expose the

partition on C . After this, we are only interested in edges e whose $k-2$ vertices in C were “cut perfectly”, i.e. which had one vertex assigned to each part in $[k-2]$. It is only these edges which may still be cut, and our goal is to maximize (by choosing the best 2-partition in V) the number of such edges which end up being cut.

Using this procedure, our task can be reformulated as follows. Let G_0 be a multigraph on vertex set V in which each edge has a set of $k-2$ colours from a set C . Two edges in G_0 with the same vertex set can have different sets of colours. To each colour $c \in C$, assign a number from $[k-2]$ uniformly at random. Now let G be the (random) submultigraph of G_0 obtained by keeping only those edges whose $k-2$ colours receive all the $k-2$ numbers in $[k-2]$. To prove good lower bounds on the k -surplus of k -graphs, it suffices to prove good lower bounds for the expected 2-surplus of the random multigraph G .

We will now explain our approach for defining the vectors $\mathbf{x}_u$, for each $u \in V$. We start with a naive version, which is not sufficient to prove our main results on its own (but could be used to give an alternative proof of Theorem 1.3 by R  ty and Tomon). We will later explain our two main ideas, which allow us to prove the stronger results we have in this paper. More precisely, the first main idea, presented below in Section 2.2, allows us to prove an $m^{2/3-o_{k \rightarrow \infty}(1)}$ lower bound for the k -surplus in k -graphs, while the second main idea, presented in Section 2.3, allows us to extend all our results to general r -cuts in k -graphs (and has further applications, to positive discrepancy and bisection width, that are closely related to 2-cuts and will be discussed in the concluding remarks).

For an edge $e \in E(G_0)$ with vertices u, v , we let $B(e, u, v)$ be the indicator of the event that the colours of e receive all $k-2$ numbers in $[k-2]$ (otherwise, if e does not contain u and v , let $B(e, u, v) = 0$). We let $B(u, v) = \sum_{f \in E(G_0)} B(f, u, v)$, let $L(e, u, v) = B(e, u, v) - \mathbb{E}[B(e, u, v)]$ and let $L(u, v) = B(u, v) - \mathbb{E}[B(u, v)]$. For each $u \in V$, we define $\mathbf{x}_u \in \mathbb{R}^V$ by taking

$$\mathbf{x}_u(v) = \begin{cases} -1, & \text{if } v = u \\ \frac{L(u, v)}{\beta}, & \text{otherwise,} \end{cases}$$

where β is a suitable normalization factor not dependent on u (but dependent on G_0), chosen so that each $\mathbf{x}_u$ has length $O(1)$ (the requirement that $\mathbf{x}_u$ is a unit vector can be relaxed to this). Recall that we want to prove that G has large surplus in expectation, which amounts to proving that the expression

$$\mathbb{E} \left[\sum_{e=uv \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle \right] = \sum_{e=uv \in E(G_0)} \mathbb{E} \left[\mathbf{1}_{e \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle \right]$$

is very negative.

Now note that for some edge $e = uv \in E(G_0)$,

$$\mathbf{1}_{e \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle = B(e, u, v) \langle \mathbf{x}_u, \mathbf{x}_v \rangle = B(e, u, v) \left(-2 \frac{L(u, v)}{\beta} + \sum_{w \in V \setminus \{u, v\}} \frac{L(u, w)L(v, w)}{\beta^2} \right).$$

We would like to prove that the first term, namely

$$-2\beta^{-1}B(e, u, v)L(u, v), \tag{1}$$

is negative in expectation (for each e), while the total sum corresponding to the second term, namely

$$\beta^{-2} \sum_{e=uv \in E(G_0)} \sum_{w \in V \setminus \{u, v\}} \mathbb{E}[B(e, u, v)L(u, w)L(v, w)], \tag{2}$$

is not too positive. For the first term, notice that

$$\begin{aligned}\mathbb{E}[B(e, u, v)L(u, v)] &= \mathbb{E}\left[B(e, u, v) \sum_{f \in E(G_0)} (B(f, u, v) - \mathbb{E}[B(f, u, v)])\right] \\ &= \sum_{f \in E(G_0)} \mathbb{E}[B(e, u, v)(B(f, u, v) - \mathbb{E}[B(f, u, v)])].\end{aligned}$$

Note that in the sum we only consider edges f with vertex set $\{u, v\}$, otherwise $B(f, u, v) = 0$ with probability 1. An important observation is that for any e and f with vertex set $\{u, v\}$, we have $\mathbb{E}[B(e, u, v)(B(f, u, v) - \mathbb{E}[B(f, u, v)])] \geq 0$. Indeed, conditioning on the event that the $k - 2$ colours of e receive all numbers in $[k - 2]$ cannot make it less likely that the $k - 2$ colours of f receive all numbers in $[k - 2]$. Moreover, when $f = e$, then $\mathbb{E}[B(e, u, v)(B(f, u, v) - \mathbb{E}[B(f, u, v)])] = \text{Var}(B(e, u, v))$ is at least some positive constant depending on k . Thus, for each edge e , $\mathbb{E}[B(e, u, v)L(u, v)]$ is at least a positive constant. This implies that the terms (1) contribute very negatively to $\mathbb{E}[\sum_{e=uv \in E(G)} \langle \mathbf{x}_u, \mathbf{x}_v \rangle]$.

We now turn to upper bounding (2), which we view as an error term. Note that

$$\begin{aligned}&\sum_{e=uv \in E(G_0)} \sum_{w \in V \setminus \{u, v\}} \mathbb{E}[B(e, u, v)L(u, w)L(v, w)] \\ &= \sum_{e=uv \in E(G_0)} \sum_{w \in V \setminus \{u, v\}} \sum_{f, f' \in E(G_0)} \mathbb{E}[B(e, u, v)L(f, u, w)L(f', v, w)].\end{aligned}$$

Observe that the term $\mathbb{E}[B(e, u, v)L(f, u, w)L(f', v, w)]$ vanishes provided that the colour sets of e , f and f' are pairwise disjoint, as in that case the random variables $B(e, u, v)$, $B(f, u, w)$, $B(f', v, w)$ are mutually independent. Unfortunately, if the colour sets of e, f, f' overlap, this term may not vanish¹. Note that if (say) f and f' share at least one colour, they correspond to hyperedges in H that intersect in at least two vertices: one (namely w) in V , and another one in C . Hence, if we knew that our original hypergraph H has no pairs of vertices with large codegree, we could get good bounds on the error term (2). For example, with this approach we could recover the nearly tight lower bound $m^{3/4-o(1)}$ for linear hypergraphs obtained by Raty and Tomon [34] (and in fact we could remove the $o(1)$ term), as well as their lower bound $m^{3/5-o(1)}$ in the general (i.e. not necessarily linear) case. Using the observation in the footnote, with this approach and some more ideas, we can in fact improve the $3/5$ exponent for k -cut in k -graphs for all $k \geq 4$.

2.2. First main idea: a modular definition of the SDP vectors. Unfortunately, large codegrees of triples always cause significant issues with the above approach, and the bottleneck of the method is an exponent bounded away from the desired $2/3$. Our first main idea is to modify the definition of the vectors $\mathbf{x}_u$ as follows. For an edge e with vertex set $\{u, v\}$, we let $B^*(e, u, v)$ be the indicator of the event that the sum of the numbers that the colours of e receive is $\sum_{i=1}^{k-2} i$ modulo $k - 2$. Note that $B^*(e, u, v)$ positively correlates with $B(e, u, v)$: indeed, $B(e, u, v) = 1$ implies $B^*(e, u, v) = 1$. On the other hand, the major advantage is that the distribution of $B^*(e, u, v)$, conditioned on some partial assignment of the colours of e , is the same as the unconditioned distribution, unless we have exposed *all* colours of e . (This is not the case for $B(e, u, v)$, where assigning numbers to two colours already changes the distribution).

We now define $\mathbf{x}_u$ as before, except we use $B^*(f, u, v)$ instead of $B(f, u, v)$. More formally, let $B^*(u, v) = \sum_{f \in E(G_0)} B^*(f, u, v)$, let $L^*(e, u, v) = B^*(e, u, v) - \mathbb{E}[B^*(e, u, v)]$ and let $L^*(u, v) = B^*(u, v) - \mathbb{E}[B^*(u, v)]$. For each $u \in V$, we define $\mathbf{x}_u \in \mathbb{R}^V$ by taking

$$\mathbf{x}_u(v) = \begin{cases} -1, & \text{if } v = u \\ \frac{L^*(u, v)}{\beta}, & \text{otherwise.} \end{cases}$$

¹In our setting, where $k \geq 4$, single overlaps in fact do not cause problems, but if say f and f' have at least two colours in common, that already makes the term not vanish.

Thanks to the positive correlation between $B(e, u, v)$ and $B^*(e, u, v)$, terms (1) provide a similarly negative contribution as before. But the key difference is that now we have a much better bound for terms (2), as $\mathbb{E}[B(e, u, v)L^*(f, u, w)L^*(f', v, w)]$ vanishes unless the colour set of f' is contained in the union of the colour set of e and the colour set of f . In particular, it is not hard to prove that the main error term will now come from triples (e, f, f') where f and f' have *identical* colour sets – a much stricter condition than intersecting in at least two elements when k is large. With tools we develop in this paper, we are able to upper bound rather efficiently the contribution of such terms.

We conclude this subsection by remarking that in order to execute the above strategy, it is necessary to assume some mild control on the maximum degree and the maximum codegree of the hypergraph whose surplus we are bounding. Indeed, this bound on the maximum degree is needed to choose the normalization factor β efficiently when defining the vectors $\mathbf{x}_u(v)$, and the codegree bound is essential to control the various error terms arising. It had already been demonstrated in the work of Conlon et al. [12] that the general case can be essentially reduced to the case where we have good bounds on the maximum degree and maximum codegree of the hypergraph. We will collect the necessary reduction lemmas (see, e.g., Lemma 3.7) in Section 3.

2.3. Second main idea: an SDP bound for 2-cuts in hypergraphs. So far, we have only dealt with the case $r = k$ (and with $r = k - 1$, as it has a simple reduction to the $r = k$ case). The case $r < k - 1$ is fundamentally different.² While it is still true that for every r , lower bounding the r -surplus in a k -graph H can be reduced to lower bounding the expected 2-surplus in a certain auxiliary random coloured multihypergraph G , in general this G is not a multigraph (and is not even uniform). Prior to our work, there was no widely applicable result in the literature that allowed one to obtain large 2-cuts in *hypergraphs* from vectors assigned to each vertex. A result along these lines was proved by R  ty and Tomon in their paper on bisection width and positive discrepancy (see [33], Lemma 3.1), but we need something significantly stronger. Our result, which extends the usual graph SDP bound to hypergraphs, is as follows. (Here and below, a mixed k -graph is a hypergraph whose edges are of size at most k . The surplus of a 2-cut in such a hypergraph is the difference between the size of the cut and the expected size of a random 2-cut.)

Theorem 2.1. [*SDP bound for 2-cuts in hypergraphs*] Fix $k \in \mathbb{N}$. Let $G = (V, E)$ be a mixed k -multigraph. Let $N \in \mathbb{N}$, and for every $u \in V$, let x_u be a vector in $\mathbb{R}^N$ such that $\|x_u\| \in [1, 2]$. Then G has a 2-cut of surplus at least

$$- \sum_{e \in E} C_{|e|} \sum_{u \neq v \in e} \langle x_u, x_v \rangle - O_k \left(\sum_{e \in E} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2 \right),$$

where $C_{|e|}$ are positive constants, depending only on the size of e .

Note that in this result we do not have a $\log n$ loss like in the general SDP rounding result of [30, 11, 5] (and, of course, that result does not even apply in the hypergraph setting). This helps us to get a tight bound for linear hypergraphs.

2.4. Organization of this paper. In the next section we collect some preliminary results that we will use in our proofs. In Section 4, we prove our SDP bound for hypergraphs (Theorem 2.1). In the following two sections we prove our results about linear hypergraphs: in Section 5, we prove the lower bound, Theorem 1.5, while in Section 6, we prove the matching upper bound, Proposition 1.6. Our main result about general hypergraphs, Theorem 1.4, is proved in Section 7. In Section 8, we prove Proposition 1.7. Finally, in Section 9, we present several concluding remarks and mention further applications and upcoming work.

²Recall that the method of R  ty and Tomon [34] does not work in this case.

3. PRELIMINARIES

Notation. For two functions f and g , we write $f = O(g)$ to say that there is a constant C such that $|f| \leq Cg$, and we write $f = \Omega(g)$ to say that there is a constant $c > 0$ such that $f \geq cg$. We write $O_t(g)$ and $\Omega_t(g)$ if the implied constants may depend on some parameter t . We sometimes omit the subscript if the dependence is on the uniformity k , which is always a constant throughout the paper. A k -graph is a hypergraph whose hyperedges have size k . A *mixed k -graph* is a hypergraph whose hyperedges have size at most k . A *mixed k -multigraph* is a multihypergraph whose hyperedges have size at most k . For a mixed k -multigraph H , the *surplus* of an r -cut is the difference between the size of the cut and the expected size of a random r -cut. The r -surplus of H , denoted by $\text{sp}_r(H)$, is the maximum surplus of an r -cut of H . We also write $\text{sp}(H)$ for the 2-surplus of H . For a hypergraph H and distinct vertices u and v , we write $\deg_H(u, v)$ to denote the number of edges containing both u and v .

We first recall the Azuma-Hoeffding inequality (see, for instance, Theorem 2.25 in [24]).

Lemma 3.1 (Azuma-Hoeffding inequality). *Let $(X_i)_{i \geq 0}$ be a martingale and let $c_i > 0$ for each $i \geq 1$. If $|X_i - X_{i-1}| \leq c_i$ for each $i \geq 1$, then, for each $n \geq 1$,*

$$\Pr(|X_n - X_0| \geq t) \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right).$$

We will need some basic properties of cuts. We start with the two following simple lemmas, which can be found as Lemma 2.5 and Lemma 2.6 in [34] for instance. For a multihypergraph H , the *underlying ℓ -multigraph* of H is the multihypergraph on the same vertex set as H obtained by connecting every set of ℓ vertices by an edge as many times as it appears in an edge of H .

Lemma 3.2. *Let G be a multigraph, and let $V_1, \dots, V_k \subseteq V(G)$ be disjoint sets. Then*

$$\text{sp}(G) \geq \sum_{i=1}^k \text{sp}(G[V_i]).$$

Lemma 3.3. *Let H be a k -multigraph, and let H' be the underlying $(k-1)$ -multigraph. Then*

$$2k \cdot \text{sp}_k(H) \geq \text{sp}_{k-1}(H) \geq \frac{1}{2} \cdot \text{sp}_{k-1}(H').$$

We will also make use of the following lemma, which can be found in [12] as Lemma 5.1.

Lemma 3.4. *For any fixed $2 \leq r \leq k$, consider an n -vertex, m -edge k -multigraph H . For any q, D, Δ , at least one of the following holds:*

- *there is an r -cut with surplus $\Omega(Dq)$, or*
- *there is a vertex subset U with $|U| \geq n - 2q - km/\Delta$ such that for every distinct $u, v \in U$, $\deg_H(v) \leq \Delta$ and $\deg_H(u, v) \leq D$.*

Moreover, we need the following slight strengthening of Lemma 5.6 from [12] which can be obtained via the same proof.

Lemma 3.5. *Fix $2 \leq r \leq k$ and consider an n -vertex, m -edge k -multigraph H . Suppose U is a vertex subset with $|U| = n - b$. Then, for every $\varepsilon > 0$, at least one of the following holds:*

- *there is an r -cut with surplus $\Omega(\varepsilon m/b)$, or*
- *there are at least $(1 - \varepsilon)m$ edges which have at least $k - 1$ of their vertices in U .*

We will also need the following result.

Lemma 3.6. *Let H be an n -vertex, m -edge k -multigraph, and let $2 \leq r \leq k$. There exists some $\varepsilon = \varepsilon(k) > 0$ such that the following holds. Let U be a vertex subset, and suppose that at least $(1 - \varepsilon)m$ edges have at least $k - 1$ of their vertices in U . Then at least one of the following holds:*

- *H has r -surplus $\Omega(m)$,*

- $H[U]$ has $\Omega(m)$ edges.

Proof. Suppose that $H[U]$ has less than εm edges. Thus there are at least $(1 - 2\varepsilon)m$ edges which have exactly $k - 1$ of their vertices in U . Let ϕ be a random r -cut obtained by assigning each vertex in U to each part $\{1, \dots, r - 1\}$ uniformly at random and independently, and assigning each vertex in U^c to part r . For integers $2 \leq y \leq x$, we let $p(x, y)$ be the probability that an edge of size x is cut by a uniformly random y -cut. We remark that $p(x - 1, y - 1) > p(x, y)$; indeed once the part (in the cut) of the first vertex of an edge of size x has been revealed, the edge is cut if and only if the parts of the other $x - 1$ vertices include the $y - 1$ parts left to be hit, an event which happens with probability strictly less than $p(x - 1, y - 1)$. Then the expected size of ϕ is at least $p(k - 1, r - 1)(1 - 2\varepsilon)m \geq p(k, r)m + \Omega(m)$ for small enough ε , which concludes the proof. $\square$

Combining the previous lemmas we obtain the following result.

Lemma 3.7. *Let G be a n -vertex, m -edge k -multigraph, and let $2 \leq r \leq k$. Let $\varepsilon > 0$. Then either G has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$, or there exists $U \subseteq V(G)$ such that*

- $G[U]$ has $\Omega(m)$ edges,
- for every distinct $u, v \in U$, $\deg_G(v) \leq m^{2/3-\varepsilon}$ and $\deg_G(u, v) \leq m^{1/3-2\varepsilon}$.

Proof. By applying Lemma 3.4 for $\Delta = m^{2/3-\varepsilon}$, $q = m^{1/3+\varepsilon}$ and $D = m^{1/3-2\varepsilon}$, if G has no r -cut of surplus $\Omega(m^{2/3-\varepsilon})$, then there exists some vertex subset U with $|U| \geq n - O(m^{1/3+\varepsilon})$ such that for all distinct $u, v \in U$, $\deg_G(v) \leq \Delta$ and $\deg_G(u, v) \leq D$. Fix some $\delta > 0$ depending only on k . By Lemma 3.5, we may assume that there are at least $(1 - \delta)m$ edges which have at least $k - 1$ of their vertices in U (otherwise we have an r -cut of desired size). Finally by choosing δ sufficiently small and applying Lemma 3.6, we obtain either an r -cut of desired size, or that $G[U]$ has $\Omega(m)$ edges, as wanted. $\square$

The next lemma shows that, to obtain an r -cut with large surplus, it is sufficient to select $r - q$ parts of the cut uniformly at random, and then find a q -cut with large surplus in a naturally defined auxiliary hypergraph. The proof is straightforward, and therefore deferred to Section A.

Lemma 3.8. *Let $q, r \geq 2$ be integers, such that $q \leq r - 1$, and let H be a hypergraph. Let σ assign independently to each vertex $v \in V(H)$ either one of the numbers in $\{q + 1, \dots, r\}$, each with probability $1/r$, or $*$, with probability q/r . Let H_σ be the multihypergraph on vertex set $\sigma^{-1}(*)$ and edge set $\{e \cap \sigma^{-1}(*) : e \in E(H), \forall j \in [q + 1, r], |e \cap \sigma^{-1}(j)| \geq 1\}$. Then H has an r -cut of surplus at least $\mathbb{E}_\sigma[\text{sp}_q(H_\sigma)]$.*

We will also need the following slightly more involved reduction lemma, whose proof is essentially the same as Lemma 7.1 from [12], and therefore deferred to Section B.

Lemma 3.9. *Let $H = (V, E)$ be a multihypergraph. Let $V = U \cup U^c$. Let $\omega : V \rightarrow [r]$ be chosen uniformly at random and let $P\omega$ be the function from V to $\{*, 1, \dots, r\}$ such that*

$$P\omega(v) = \begin{cases} \omega(v) & \text{if } v \in U^c \text{ or } \omega(v) \in \{3, \dots, r\}, \\ * & \text{otherwise.} \end{cases}$$

Let $H_{P\omega}$ be the multihypergraph on vertex set $(P\omega)^{-1}()$ and edge set $\{e \cap (P\omega)^{-1}(*) : e \in E, \{3, \dots, r\} \subseteq P\omega(e), \{1, \dots, r\} \not\subseteq P\omega(e)\}$, where two copies of $e \cap (P\omega)^{-1}(*)$ are added to $H_{P\omega}$ if $P\omega(e) \cap \{1, \dots, r\} = \{3, \dots, r\}$. Then H has r -surplus at least $\mathbb{E}_\omega[\text{sp}(H_{P\omega})]/2$.*

We rewrite here the previous lemma in the special case $r = 2$ as we will use this version in a couple of places. More precisely it shows that, when looking for a 2-cut with large surplus, one may assign to some vertices one part of the cut uniformly at random, and then find a 2-cut with large surplus in a naturally defined auxiliary hypergraph.

Corollary 3.10. *Let $H = (V, E)$ be a multihypergraph, and let $W \subseteq V$. Let $\sigma : W^c \rightarrow \{1, 2\}$ be taken uniformly at random. Let H_σ be the multihypergraph on vertex set W such that for every edge $e \in E(H)$, if all vertices of e are in W then we add two copies of e , otherwise, if $|\sigma(e \setminus W)| = 1$ and $|e \cap W| > 0$, we add one copy of $e \cap W$. Then*

$$\text{sp}(H) \geq \mathbb{E}_\sigma[\text{sp}(H_\sigma)]/2.$$

The next lemma reduces the problem of finding large r -cuts in k -graphs to finding large 2-cuts in hypergraphs. The proof follows closely the argument in the proof of Theorem 1.4 in [12]. Therefore, we defer this proof to Section C.

Lemma 3.11. *Let $\varepsilon > 0$ be fixed. Let $q \geq 2$, $k \geq 4$ be some fixed integers. Suppose that every mixed k -multigraph with maximum degree $O(m^{2/3-\varepsilon})$, maximum codegree $O(m^{1/3-2\varepsilon})$, $O(m)$ hyperedges in total, and $\Omega(m)$ hyperedges of size $q+2$ has a 2-cut of surplus $\Omega(m^{2/3-\varepsilon})$.*

Then for every $2 \leq r \leq k-q$, every m -edge k -multigraph has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$.

A similar (but simpler) reduction holds for linear hypergraphs. This can be proven the same way as the previous lemma, and therefore its proof is left to the reader.

Lemma 3.12. *Fix $k \geq 4$ and $\alpha \in (0, 1)$. Suppose that every linear mixed k -graph with $O(m)$ hyperedges in total and $\Omega(m)$ hyperedges of size at least 4 has a 2-cut of surplus $\Omega(m^\alpha)$.*

Then for every $2 \leq r \leq k-2$, every linear k -graph with $O(m)$ hyperedges in total and $\Omega(m)$ hyperedges of size at least $r+2$ has an r -cut of surplus $\Omega(m^\alpha)$.

We recall the following standard integral estimates.

Lemma 3.13. *For any $a > 0$, we have*

$$\int_{-\infty}^0 e^{-az^2} dz = \frac{1}{2} \sqrt{\frac{\pi}{a}}, \text{ and } \int_{-\infty}^0 z e^{-az^2} dz = -\frac{1}{2a}.$$

From the previous estimates, the following lemma is immediate. In what follows, we write Z for the vector $(z_1, \dots, z_k)$.

Lemma 3.14. *Fix a positive integer k . Let $M \in [0, 1/2)$ and let A be a k by k diagonal matrix, such that all diagonal entries have values between $1-M$ and $1+M$. Then, for any $1 \leq s < \ell \leq k$,*

$$\frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp\left(-\frac{1}{2} Z^T A Z\right) dZ = -C + O_k(M),$$

where $C = \frac{1}{\pi 2^k}$.

Proof. Note that

$$\begin{aligned} & \frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp\left(-\frac{1}{2} Z^T A Z\right) dZ \\ &= -\frac{1}{2} \frac{1}{(2\pi)^{k/2}} \prod_{i=1}^k \int_{z_i=-\infty}^0 f_i(z_i) \exp\left(-\frac{1}{2} A_{i,i} z_i^2\right) dz_i, \end{aligned}$$

where $f_i(x) = x$ if $i = s$ or $i = \ell$, and $f_i(x) = 1$ otherwise. Applying Lemma 3.13, we obtain

$$\frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp\left(-\frac{1}{2} Z^T A Z\right) dZ = -\frac{1}{2} \frac{1}{\sqrt{2\pi} A_{s,s}} \cdot \frac{1}{\sqrt{2\pi} A_{\ell,\ell}} \cdot \prod_{i \neq s, \ell} \frac{1}{2} \sqrt{\frac{1}{A_{i,i}}}.$$

Since $1-M \leq A_{i,i} \leq 1+M$ for every i , we obtain

$$\frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp\left(-\frac{1}{2} Z^T A Z\right) dZ = -C + O_k(M),$$

for $C = \frac{1}{\pi 2^k}$, as wanted. $\square$

4. AN SDP BOUND FOR 2-CUT IN HYPERGRAPHS

The goal of this section is to establish Theorem 2.1, which, we recall, is a hypergraph generalization of the standard SDP bound for MaxCut in graphs and will be used repeatedly throughout the proofs. For the reader's convenience, we restate it below.

Theorem 2.1. *[SDP bound for 2-cuts in hypergraphs] Fix $k \in \mathbb{N}$. Let $G = (V, E)$ be a mixed k -multigraph. Let $N \in \mathbb{N}$, and for every $u \in V$, let x_u be a vector in $\mathbb{R}^N$ such that $\|x_u\| \in [1, 2]$. Then G has a 2-cut of surplus at least*

$$-\sum_{e \in E} C_{|e|} \sum_{u \neq v \in e} \langle x_u, x_v \rangle - O_k\left(\sum_{e \in E} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2\right),$$

where $C_{|e|}$ are positive constants, depending only on the size of e .

We will establish this result via the following geometric lemma.

Lemma 4.1. *Fix some positive integer k . Let $N \in \mathbb{N}$, and let $a_1, \dots, a_k$ be points on $\mathbb{S}^{N-1}$. Suppose that for some real $M \in [0, 1]$ we have that $|\langle a_i, a_j \rangle| \leq M$ for every distinct $i, j \in [k]$. Take a random point on $\mathbb{S}^{N-1}$, and let $\vec{r}$ be the associated random vector. Then the probability that all inner products $\langle a_i, \vec{r} \rangle$ are negative is $(1/2)^k + C \sum_{i \neq j} \langle a_i, a_j \rangle + O_k(M^2)$, where $C = \frac{1}{\pi 2^k}$.*

We remark that a similar (but significantly weaker) result was obtained in [33] (see their Lemma 3.1). Our result implies theirs but is both more general and more accurate: we do not have an assumption saying that all $\langle a_i, a_j \rangle$ have the same sign (which would not be satisfied in our setting), and our error term is smaller.

We now turn to the proof of Lemma 4.1. We first rewrite the desired probability as an integral.

Lemma 4.2. *Let k, N be positive integers, let $a_1, \dots, a_k$ be points on $\mathbb{S}^{N-1}$ and let $0 \leq M \leq 1$. Suppose that $|\langle a_i, a_j \rangle| \leq M$ for every distinct $i, j \in [k]$. Take a random point on $\mathbb{S}^{N-1}$, and let $\vec{r}$ be the associated random vector. Then the probability that all inner products $\langle a_i, \vec{r} \rangle$ are negative can be written as*

$$\int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 \frac{1}{(2\pi)^{k/2}} \exp\left(-\frac{1}{2} Z^T A Z\right) dZ + O_k(M^2),$$

for some symmetric matrix A such that

- for every i , we have $A_{i,i} = 1 + O_k(M^2)$,
- for every $i \neq j$, we have $A_{i,j} = -\langle a_i, a_j \rangle + O_k(M^2)$.

Proof. First, observe that we may assume that $M \leq c(k)$ for a sufficiently small positive constant $c(k)$, otherwise the statement is trivial by choosing $A = I$ and sufficiently large implicit constants in the $O_k(\cdot)$ terms. It is well known (see for instance [31]) that by sampling uniformly random independent standard Gaussians $W_1, \dots, W_N$, the random vector $\vec{r}$ has the same distribution as

$$\left(\frac{W_1}{\sqrt{\sum_{i=1}^N W_i^2}}, \dots, \frac{W_N}{\sqrt{\sum_{i=1}^N W_i^2}} \right).$$

Let $\vec{s} = (W_1, \dots, W_N)$. It is clear that $(\langle a_1, \vec{s} \rangle, \dots, \langle a_k, \vec{s} \rangle)$ is a Gaussian vector of covariance matrix G and mean vector 0, where G is the Gram matrix of $a_1, \dots, a_k$. Therefore,

$$\begin{aligned} \mathbb{P}(\langle a_1, \vec{r} \rangle < 0, \dots, \langle a_k, \vec{r} \rangle < 0) &= \mathbb{P}(\langle a_1, \vec{s} \rangle < 0, \dots, \langle a_k, \vec{s} \rangle < 0) \\ &= \int_{z_1 < 0} \cdots \int_{z_k < 0} \frac{1}{(2\pi)^{k/2} \det(G)^{1/2}} \exp\left(-\frac{1}{2} Z^T G^{-1} Z\right) dZ. \end{aligned}$$

We let $G = I + E$, where I is the identity matrix. Note that $E_{i,i} = 0$ and $E_{i,j} = \langle a_i, a_j \rangle$ for every $i \neq j$, so in particular all entries of E are bounded by M in absolute value. By the standard definition of the determinant, it is easy to see that $\det(G) = 1 + O_k(M^2)$ since each product

in the expression for $\det(G)$, except the one which is the product of all diagonal entries, has at least two non-diagonal terms.

Moreover, we have by Neumann series expansion (which converges since $0 \leq M \leq c(k)$ for a small constant $c(k)$ by our assumption at the beginning) that $G^{-1} = I - E + K$, where $K = \sum_{\ell \geq 2} (-E)^\ell$, and therefore each entry of K is at most $O_k(M^2)$ in absolute value. Thus $A = G^{-1}$ satisfies the required conditions. $\square$

The next lemma bounds the integral obtained in Lemma 4.2.

Lemma 4.3. *Fix some positive integer k . Let $0 \leq M \leq \frac{1}{4k}$, let A be a symmetric k by k matrix, with non-diagonal entries at most M in absolute value, and diagonal entries between $1 - M^2$ and $1 + M^2$.*

Let $\Phi(A) = \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 \frac{1}{(2\pi)^{k/2}} \exp(-\frac{1}{2} Z^T A Z) dZ$. Then, for $C = \frac{1}{\pi 2^k}$, we have

$$\Phi(A) = (1/2)^k - C \left(\sum_{i \neq j} A_{i,j} \right) + O_k(M^2).$$

Proof. We first make the following claim.

Claim 4.4. *For any two coordinates $s \neq \ell$, we have*

$$\frac{\partial \Phi(A)}{\partial A_{s,\ell}} = -C + O_k(M),$$

where $C = \frac{1}{\pi 2^k}$.

Proof of Claim 4.4. First, we note that by applying AM-GM inequality, we have for any $Z \in \mathbb{R}^n$ that

$$\begin{aligned} \exp(-\frac{1}{2} Z^T A Z) &= \exp\left(-\sum_{i,j} \frac{1}{2} A_{i,j} z_i z_j\right) \leq \exp\left(-\sum_i \frac{1}{2} A_{i,i} z_i^2 + \sum_{i \neq j} \frac{1}{2} |A_{i,j}| \left(\frac{z_i^2 + z_j^2}{2}\right)\right) \\ &= \exp\left(-\sum_i \left(\frac{1}{2} A_{i,i} - \sum_{j \neq i} \frac{1}{2} |A_{i,j}|\right) z_i^2\right) \\ &= \exp(-\frac{1}{2} Z^T A_{diag}^+ Z), \end{aligned}$$

where the matrix A_{diag}^+ is diagonal, with diagonal entries between $1 - kM - M^2$ and $1 + kM + M^2$. Similarly, there exists a diagonal matrix A_{diag}^- , with diagonal entries between $1 - kM - M^2$ and $1 + kM + M^2$, such that for any $Z \in \mathbb{R}^n$, we have

$$\exp(-\frac{1}{2} Z^T A Z) \geq \exp(-\frac{1}{2} Z^T A_{diag}^- Z).$$

Next we have by definition of Φ that

$$\frac{\partial \Phi(A)}{\partial A_{s,\ell}} = \frac{1}{(2\pi)^{k/2}} \cdot \frac{\partial \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 \exp(-\frac{1}{2} Z^T A Z) dZ}{\partial A_{s,\ell}}.$$

Note that

$$\left| \frac{\partial \exp(-\frac{1}{2} Z^T A Z)}{\partial A_{s,\ell}} \right| = \left| \frac{1}{2} z_s z_\ell \exp(-\frac{1}{2} Z^T A Z) \right| \leq \left| z_s z_\ell \exp(-\frac{1}{2} Z^T A_{diag}^+ Z) \right|.$$

As the function $\left| z_s z_\ell \exp(-\frac{1}{2} Z^T A_{diag}^+ Z) \right|$ is integrable, we may invert the integration and derivation signs (see for instance Theorem 2.27 in [18]) and obtain that

$$\begin{aligned} \frac{\partial \Phi(A)}{\partial A_{s,\ell}} &= \frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 \frac{\partial \exp(-\frac{1}{2} Z^T A Z)}{\partial A_{s,\ell}} dZ \\ &= \frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp(-\frac{1}{2} Z^T A Z) dZ. \end{aligned}$$

As $\exp(-\frac{1}{2} Z^T A_{diag}^- Z) \leq \exp(-\frac{1}{2} Z^T A Z) \leq \exp(-\frac{1}{2} Z^T A_{diag}^+ Z)$, it follows that

$$\begin{aligned} \frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp(-\frac{1}{2} Z^T A_{diag}^- Z) dZ &\leq \frac{\partial \Phi(A)}{\partial A_{s,\ell}} \\ &\leq \frac{1}{(2\pi)^{k/2}} \int_{z_1=-\infty}^0 \cdots \int_{z_k=-\infty}^0 -\frac{1}{2} z_s z_\ell \exp(-\frac{1}{2} Z^T A_{diag}^+ Z) dZ. \end{aligned}$$

Applying Lemma 3.14, we obtain

$$\frac{\partial \Phi(A)}{\partial A_{s,\ell}} = -C + O_k(M),$$

where $C = \frac{1}{\pi^{2k}}$, as wanted. $\square$

From Claim 4.4, the result then follows by a classical multivariate Taylor expansion over the variables $A_{s,\ell}$ for all $s \neq \ell$, as it is easy to see that for any $s \neq \ell$ and $s' \neq \ell'$, $\frac{\partial^2 \Phi(A)}{\partial A_{s,\ell} \partial A_{s',\ell'}}$ is uniformly bounded (by the same computation as the first partial derivative in Claim 4.4), and $\Phi(A_{diag}) = (1/2)^k + O_k(M^2)$ by Lemma 3.13, where A_{diag} is the diagonal matrix obtained from A by removing the non-diagonal entries. $\square$

Lemma 4.1 then follows as a simple combination of Lemma 4.2 and Lemma 4.3 (noting, once again, that the statement is trivial unless $M \leq c(k)$ for a small positive constant $c(k)$). With Lemma 4.1 in hand, it is easy to derive Theorem 2.1.

Proof of Theorem 2.1. First, without loss of generality, we can assume, by adding dummy coordinates without changing the cross-inner products, that all vectors have norm 2. Further, we can assume, by rescaling the vectors, that all x_u have same norm 1. We take a random point r on the sphere $\mathbb{S}^{N-1}$, and we let $\vec{r}$ be the associated vector. Let H^+ be the set of points in $\mathbb{R}^N$ whose inner product with $\vec{r}$ is non-negative. Let ϕ be the 2-cut associated to $\vec{r}$ by putting every vertex u in class 1 if x_u is in H^+ , and in class 2 otherwise. Let S_ϕ be the size of ϕ . We now compute $\mathbb{E}[S_\phi]$. Clearly, we have that $S_\phi = \sum_{e \in E} 1_e$ is cut, and therefore $\mathbb{E}[S_\phi] = \sum_{e \in E} \mathbb{P}(e \text{ is cut})$.

Let $e \in E$. By Lemma 4.1, we have

$$\mathbb{P}(e \text{ is cut}) \geq 1 - \frac{1}{2^{|e|-1}} - C_{|e|} \sum_{u \neq v \in e} \langle x_u, x_v \rangle - O_k \left(\sum_{u \neq v \in e} \langle x_u, x_v \rangle^2 \right).$$

Summing over all hyperedges $e \in E$ gives the desired result. $\square$

5. LINEAR HYPERGRAPHS: TIGHT LOWER BOUND

In this section we prove Theorem 1.5. Before presenting the result from which we will deduce Theorem 1.5, we introduce an important definition.

Definition 5.1 (Random (m, Δ, k) -multihypergraph of linear type). Fix some positive integer k and a vertex set V . We say that a random multihypergraph $H = (V, R)$ is a *random (m, Δ, k) -multihypergraph of linear type* if it satisfies the following properties:

- (i) There exists a multiset X consisting of subsets of V of size at most k such that $|X| = O(m)$, and R is a random subset of X defined as follows: there exist a set $\mathcal{S}$ and some independent random variables $(Z_i)_{i \in \mathcal{S}}$ such that for every $e \in X$, there exists a set $S_e \subseteq \mathcal{S}$ of size at most k such that the event $e \in R$ is a function of $\{Z_i : i \in S_e\}$.
(We will call the elements of X *potential hyperedges* in the hypergraph H , as whether they belong to the edge set R of H depends on the random variables $(Z_i)_{i \in \mathcal{S}}$.)
- (ii) For any potential hyperedges $e, f \in X$, we have $\mathbb{P}(e \in R, f \in R) \geq \mathbb{P}(e \in R)\mathbb{P}(f \in R)$.
- (iii) There is a set $Y \subseteq X$ of size $|Y| = \Omega(m)$ such that for every $e \in Y$, we have $|e| \geq 2$ and $\Omega(1) = \mathbb{P}(e \in R) = 1 - \Omega(1)$.
- (iv) For every vertex $u \in V$, there are at most Δ potential hyperedges $e \in X$ such that $u \in e$.
- (v) There exists $D = O(1)$ such that for any distinct vertices $u, v \in V$, there are at most D potential hyperedges $e \in X$ such that $u, v \in e$, and for every $u \in V$ and $i \in \mathcal{S}$, there are at most D potential hyperedges $e \in X$ such that $u \in e$ and $i \in S_e$.

Remark 5.2. In Definition 5.1, since X is a multiset, it is possible that two potential hyperedges e and f have the same vertex set but $S_e \neq S_f$. To avoid clutter in the proof, we will often refer to the set S_e as the *set of colours* of the potential hyperedge e . In what follows, we will call the elements of X hyperedges without writing *potential*.

Remark 5.3. Roughly speaking, a random (m, Δ, k) -multihypergraph of linear type arises from exposing a random partial cut on a (deterministic) linear k -graph H_0 with roughly m edges and maximum degree at most Δ . The hyperedges of H_0 are precisely $e \cup S_e$ for each $e \in X$, where the random cut is being exposed on S_e . Note that (iv) encodes the maximum degree condition on H_0 , whereas (v) encodes the (approximate) linearity condition for H_0 .

The key result to prove Theorem 1.5 is the following.

Theorem 5.4. *Fix $k \in \mathbb{N}$. Let $m, \Delta \geq 1$, and let H be a random (m, Δ, k) -multihypergraph of linear type. Then we have*

$$\mathbb{E}[\text{sp}(H)] = \Omega\left(\frac{m}{\sqrt{\Delta}}\right).$$

Proof. For a hyperedge $e \in X$ containing vertices u, v , we let $B(e, u, v)$ be the indicator of the event $e \in R$ (otherwise, if e does not contain u and v , let $B(e, u, v) = 0$ whether or not $e \in R$). We let $B(u, v) = \sum_{f \in X} B(f, u, v)$, let $L(e, u, v) = B(e, u, v) - \mathbb{E}[B(e, u, v)]$ and let $L(u, v) = B(u, v) - \mathbb{E}[B(u, v)]$. Let c be a sufficiently small positive constant. For each $u \in V$, we define $x_u \in \mathbb{R}^V$ by

$$x_u(v) = \begin{cases} -1, & \text{if } v = u \\ c \cdot \frac{L(u, v)}{\sqrt{\Delta}}, & \text{otherwise.} \end{cases}$$

Observe that by (iv) and (v), we have $\|x_u\|^2 \leq 1 + O(D^2 c^2) \leq 2$ for c sufficiently small.

We apply Theorem 2.1 to H and get

$$\text{sp}(H) \geq - \sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle - O\left(\sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2\right).$$

Taking expectation, this gives

$$\begin{aligned} \mathbb{E}[\text{sp}(H)] &\geq \mathbb{E}\left[- \sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle\right] - O\left(\mathbb{E} \sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2\right) \\ &\geq \sum_{e \in X} C_e \sum_{u \neq v \in e} \mathbb{E}[-\mathbf{1}_{e \in R} \langle x_u, x_v \rangle] - O\left(\mathbb{E} \sum_{e \in X} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2\right). \end{aligned} \quad (3)$$

Our goal is now to show that the RHS of the previous expression is $\Omega\left(\frac{m}{\sqrt{\Delta}}\right)$. We start by giving a lower bound of the first term.

For a given edge $e \in X$, and distinct vertices $u, v \in e$, we have

$$\begin{aligned} \mathbb{E}[-\mathbf{1}_{e \in R} \langle x_u, x_v \rangle] &= \mathbb{E}[-B(e, u, v) \left(x_u(u)x_v(u) + x_u(v)x_v(v) + \sum_{w \neq u, v} x_u(w)x_v(w) \right)] \\ &= \mathbb{E}[B(e, u, v) 2c \cdot \frac{L(u, v)}{\sqrt{\Delta}} - \sum_{w \neq u, v} c^2 B(e, u, v) \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}}]. \end{aligned} \quad (4)$$

We first focus on the first subterm. Note that

$$\mathbb{E}[2c \cdot B(e, u, v) \frac{L(u, v)}{\sqrt{\Delta}}] = \mathbb{E}[2c \cdot B(e, u, v) \frac{L(e, u, v)}{\sqrt{\Delta}} + 2c \cdot B(e, u, v) \frac{\sum_{f \in X \setminus \{e\}} L(f, u, v)}{\sqrt{\Delta}}]. \quad (5)$$

By definition of Y in (iii), we have

$$\mathbb{E}[2c \cdot B(e, u, v) \frac{L(e, u, v)}{\sqrt{\Delta}}] = \mathbb{E}[2c \cdot \frac{\text{Var}(B(e, u, v))}{\sqrt{\Delta}}] = \Omega\left(\frac{c \mathbf{1}_{e \in Y}}{\sqrt{\Delta}}\right).$$

Furthermore, by (ii), we have

$$\mathbb{E}[2c \cdot B(e, u, v) \frac{\sum_{f \in X \setminus \{e\}} L(f, u, v)}{\sqrt{\Delta}}] = \mathbb{E}\left[2c \cdot \frac{\sum_{f \in X \setminus \{e\}} \text{Cov}(B(e, u, v), B(f, u, v))}{\sqrt{\Delta}}\right] \geq 0.$$

Plugging the two above bounds in (5), we get

$$\mathbb{E}[2c \cdot B(e, u, v) \frac{L(u, v)}{\sqrt{\Delta}}] = \Omega\left(\frac{c \mathbf{1}_{e \in Y}}{\sqrt{\Delta}}\right). \quad (6)$$

We now consider the second subterm. Note that

$$\mathbb{E}\left[\sum_{w \neq u, v} c^2 B(e, u, v) \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}}\right] = \sum_{w \neq u, v} \sum_{f, f' \in X} c^2 \mathbb{E}[B(e, u, v) \frac{L(f, u, w)}{\sqrt{\Delta}} \cdot \frac{L(f', v, w)}{\sqrt{\Delta}}].$$

We remark that when the random variables $B(e, u, v)$, $B(f, u, w)$ and $B(f', v, w)$ are mutually independent, the corresponding expectation vanishes, which is always the case unless at least two of S_e , S_f and $S_{f'}$ share at least one colour. Suppose that S_e and S_f share a colour i . Then we have $O(m)$ choices for (e, u, v) . As $|S_e| = O(1)$, there are $O(1)$ choices for i , and therefore by (v), we have $O(D)$ choices for (f, w) . Finally, by (v) again, we have $O(D)$ choices for f' , thereby $O(mD^2)$ choices in this case. The two other cases can be dealt with similarly, and we find that there are at most $O(mD^2)$ tuples (e, f, f', u, v, w) for which $\mathbb{E}[B(e, u, v) \frac{L(f, u, w)}{\sqrt{\Delta}} \cdot \frac{L(f', v, w)}{\sqrt{\Delta}}]$ does not vanish, and thus (since $B(e, u, v) \frac{L(f, u, w)}{\sqrt{\Delta}} \cdot \frac{L(f', v, w)}{\sqrt{\Delta}} = O(\frac{1}{\Delta})$ with probability 1)

$$\sum_{e \in X} \sum_{u \neq v \in e} \mathbb{E}\left[\sum_{w \neq u, v} c^2 B(e, u, v) \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}}\right] = O\left(\frac{c^2 m D^2}{\Delta}\right). \quad (7)$$

Combining (4), (6) and (7), together with $|Y| = \Omega(m)$ from (iii), we obtain

$$\sum_{e \in X} C_e \sum_{u \neq v \in e} \mathbb{E}[-\mathbf{1}_{e \in R} \langle x_u, x_v \rangle] = \Omega\left(\frac{cm}{\sqrt{\Delta}}\right) - O\left(\frac{c^2 m D^2}{\Delta}\right). \quad (8)$$

We now turn to the second term of (3). We fix some hyperedge $e \in X$ and distinct vertices $u, v \in e$. Using the inequality $(a + b)^2 \leq 2a^2 + 2b^2$, we have

$$\begin{aligned} \mathbb{E}[\langle x_u, x_v \rangle^2] &= \mathbb{E} \left[\left(2c \cdot \frac{L(u, v)}{\sqrt{\Delta}} - \sum_{w \in V \setminus \{u, v\}} c^2 \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}} \right)^2 \right] \\ &\leq \mathbb{E} \left[2 \left(2c \frac{L(u, v)}{\sqrt{\Delta}} \right)^2 + 2 \left(\sum_{w \in V \setminus \{u, v\}} c^2 \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}} \right)^2 \right] \\ &= \mathbb{E} \left[8c^2 \frac{L(u, v)^2}{\Delta} \right] + \mathbb{E} \left[2c^4 \sum_{w, w' \in V \setminus \{u, v\}} \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}} \cdot \frac{L(u, w')}{\sqrt{\Delta}} \cdot \frac{L(v, w')}{\sqrt{\Delta}} \right]. \end{aligned}$$

The first term is trivially $O(\frac{D^2 c^2}{\Delta})$ by (v). We therefore focus on the second term, and we want to bound

$$\sum_{e \in X} \sum_{u \neq v \in e} \sum_{f_1, f_2, f_3, f_4} \sum_{w, w' \in V \setminus \{u, v\}} \mathbb{E}[L(f_1, u, w) L(f_2, v, w) L(f_3, u, w') L(f_4, v, w')] \quad (9)$$

For a given term not to vanish in expectation, each f_i must have at least one colour that also belongs to some f_j with $j \neq i$. For the tuple (f_1, u, w) , we have $O(m)$ choices. Then f_1 must share a colour with f_2 , f_3 or f_4 . If it is f_2 , then we have $O(D)$ choices for (f_2, v) by (v). We then have $O(D)$ choices for e , $O(\Delta)$ choices for (f_3, w') by (iv), and $O(D)$ choices for f_4 , giving in total $O(mD^3\Delta)$ choices. If f_1 shares a colour with f_3 , then we have $O(\Delta)$ choices for (f_2, v) by (iv), $O(D)$ choices for e , $O(D)$ choices for (f_3, w') , and $O(D)$ choices for f_4 , thereby giving a total of $O(mD^3\Delta)$ possibilities. The case where f_1 shares a colour with f_4 is nearly identical. We conclude that

$$\sum_{e \in X} \sum_{u \neq v \in e} \mathbb{E} \left[2c^4 \sum_{w, w' \in V \setminus \{u, v\}} \frac{L(u, w)}{\sqrt{\Delta}} \cdot \frac{L(v, w)}{\sqrt{\Delta}} \cdot \frac{L(u, w')}{\sqrt{\Delta}} \cdot \frac{L(v, w')}{\sqrt{\Delta}} \right] = O \left(\frac{c^4 m D^3}{\Delta} \right).$$

It follows that

$$\mathbb{E} \left[\sum_{e \in X} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2 \right] = O \left(\frac{D^2 c^2 m}{\Delta} \right) + O \left(\frac{c^4 m D^3}{\Delta} \right). \quad (10)$$

By plugging in (8) and (10) into (3), we obtain by taking c sufficiently small that

$$\mathbb{E}[\text{sp}(H)] = \Omega \left(\frac{m}{\sqrt{\Delta}} \right),$$

as wanted. $\square$

With Theorem 5.4 in hand, it is relatively easy to establish Theorem 1.5 using the tools developed by Conlon, Fox, Kwan, and Sudakov [12] and Raty and Tomon [34]. For instance, the following lemma can be obtained from Theorem 5.4 in a similar manner as done by Raty and Tomon [34]. We therefore defer its proof to the Section D.

Lemma 5.5. *Let H be a 3-multigraph with $O(m)$ edges and maximum codegree $O(1)$. Assume that there is a set $S \subset V(H)$ such that $H[S]$ has m hyperedges, and every vertex in S has degree at most Δ in H . Then*

$$\text{sp}_3(H) = \Omega \left(\frac{m}{\sqrt{\Delta}} \right).$$

Similarly, from Theorem 5.4, it is easy to obtain the following lemma by using the work of Conlon, Fox, Kwan, and Sudakov [12]. We therefore also defer its proof to the Section E.

Lemma 5.6. *Fix $k \geq 4$, and let H be a linear mixed k -graph with $O(m)$ edges. Assume that there is a set $S \subset V(H)$ such that there are at least m hyperedges $e \in E(H)$ with $|e| \geq 4$ and $|e \cap S| \geq 2$, and every vertex in S has degree at most Δ in H . Then*

$$\text{sp}_2(H) = \Omega\left(\frac{m}{\sqrt{\Delta}}\right).$$

We are now ready to prove Theorem 1.5.

Proof of Theorem 1.5. We start with the case where $r \in \{k-1, k\}$. Let H be a k -uniform linear hypergraph with m edges. Let $H_k = H$, and for $i = k-1, \dots, 3$, we let H_i be the underlying i -graph of H_{i+1} . Then by Lemma 3.3, we have $\text{sp}_{i+1}(H_{i+1}) \geq \frac{1}{4(i+1)} \text{sp}_i(H_i)$ for all $3 \leq i \leq k-1$. Furthermore, by the same lemma, we have $\text{sp}_{k-1}(H_k) \geq \frac{1}{2} \text{sp}_{k-1}(H_{k-1})$. It follows that in either case $r \in \{k-1, k\}$, we have $\text{sp}_r(H) = \Omega_r(\text{sp}_3(H_3))$. Note that H_3 is a multihypergraph with maximum codegree $D = O(1)$ and $\tilde{m} = \Theta(m)$ hyperedges.

Let G be the underlying multigraph of H_3 . Thus $e(G) = 3\tilde{m}$. We let $\Delta = 100D^{1/2}\tilde{m}^{1/2}$, let T be the set of vertices of G which have degree more than Δ and let $S = V(G) \setminus T$. Since $6\tilde{m} = 2e(G) \geq |T|\Delta$, we have $|T| \leq 0.1\tilde{m}^{1/2}D^{-1/2}$. Since the multiplicity of every edge of G is at most D , we thus have $e(G[T]) \leq D|T|^2/2 \leq \tilde{m}/200$.

If $e_G(S, T) \geq 1.6\tilde{m}$, we have $\text{sp}(G) \geq 0.1\tilde{m}$. Since by Lemma 3.3, we have $\text{sp}_3(H_3) \geq \text{sp}(G)/12$, we are done in this case.

If $e_G(S, T) < 1.6\tilde{m}$, then $e(G[S]) = 3\tilde{m} - e(G[T]) - e_G(S, T) \geq 1.3\tilde{m}$. However,

$$e(G[S]) \leq 3e(H_3[S]) + e_{H_3}(S, T) \leq 3e(H_3[S]) + \frac{1}{2}e_G(S, T)$$

(where $e_{H_3}(S, T)$ is the number of edges in H_3 intersecting both S and T), and therefore $e(H_3[S]) \geq 0.1\tilde{m}$. Since each vertex in S has degree at most $\Delta = 100D^{1/2}\tilde{m}^{1/2}$ in H_3 , Lemma 5.5 implies that $\text{sp}_3(H_3) = \Omega(m^{3/4})$, from which we conclude that $\text{sp}_r(H) = \Omega(m^{3/4})$.

We now turn to the remaining cases, i.e. $2 \leq r \leq k-2$. By Lemma 3.12, it is enough to prove that if H' is a linear mixed k -graph with $O(m)$ hyperedges in total and at least m hyperedges of size at least 4, then $\text{sp}(H') = \Omega(m^{3/4})$. We start similarly as in the proof of the cases $r = k = 3$ by letting $\Delta = km^{1/2}$, T be the set of vertices of H' which have degree more than Δ and $S = V(H') \setminus T$. Since $km \geq |T|\Delta$, we have $|T| \leq m^{1/2}$. Therefore the number of edges in H' with at least 2 vertices in T is at most $|T|^2/2 \leq m/2$. Thus, the number of edges in H' with at least 4 vertices in total and at least 2 vertices in S is $\Omega(m)$, and every vertex in S has degree at most Δ in H' . Hence, by Lemma 5.6, we have $\text{sp}(H') = \Omega(m^{3/4})$. $\square$

6. UPPER BOUND FOR THE SURPLUS OF LINEAR HYPERGRAPHS

In this section we prove Proposition 1.6.

Proof of Proposition 1.6. Note that we may assume that m is sufficiently large compared to k . Given m , let $n = Cm^{1/2}$ for some large constant C that may depend on k . Let m' be another parameter to be specified. Let $e_1, e_2, \dots, e_{m'}$ be independent and uniformly random k -sets in $[n]$. Let $E = \{e_i : i \in [m'] \text{ such that } \forall j \in [m'] \setminus \{i\}, |e_i \cap e_j| \leq 1\}$, i.e. the set of e_i which do not intersect any other e_j in at least two vertices. Let H be the k -uniform hypergraph with vertex set $[n]$ and edge set E . By definition, H is linear. Let $f(m')$ be the expected number of edges in H . It is clear that $f(m') = m'(1-p)^{m'-1}$, where $p = \frac{\sum_{i=2}^k \binom{k}{i} \binom{n-k}{k-i}}{\binom{n}{k}} = \frac{\beta_k}{n^2} + O(\frac{1}{n^3})$ for some constant β_k . Therefore, for C large enough, there exists $m' = \Theta_k(m)$ such that $m + m^{2/3} < f(m') < m + 2m^{2/3}$. Note that $|E|$ is $\left(\binom{k}{2} + 1\right)$ -Lipschitz with respect to revealing each e_i , and therefore by Azuma's inequality (see Lemma 3.1), for any constant $\gamma > 0$ we have with high probability that

$$||E| - f(m')| \leq \gamma(m')^{2/3}.$$

As $m' = O(m)$, it follows by choosing γ sufficiently small that with high probability we have $m \leq |E| \leq m + 3m^{2/3}$. We delete edges if necessary in an arbitrary way to obtain an m -edge subgraph G . We now show that with high probability G has r -surplus $O(m^{3/4})$. Clearly, it suffices to prove that with high probability H has r -surplus $O(m^{3/4})$. We argue that this holds by taking a union bound over all r -cuts. We fix some r -cut $\psi : [n] \rightarrow [r]$. Let ϕ be the random variable associated to the size of ψ as an r -cut of H . Then ϕ is clearly $\binom{k}{2} + 1$ -Lipschitz with respect to revealing each e_i , and therefore by Azuma's inequality,

$$\mathbb{P}(|\phi - \mathbb{E}[\phi]| \geq C'm^{3/4}) \leq \exp(-2C(\log r)m^{1/2}) \quad (11)$$

for some constant $C' = C'(k)$. Let q_ψ be the probability that ψ cuts a random k -set in $[n]$. For $i \in [r]$, we let $n_i = |\psi^{-1}(i)|$. We claim that q_ψ is maximal when ψ is balanced, i.e., when $|n_i - n_j| \leq 1$ for each i, j . To see this, assume that $n_i \geq n_j + 2$ for some i, j . Consider an r -cut ψ' which agrees with ψ on all but one vertex v , but has $\psi'(v) = j$ whereas $\psi(v) = i$. Now if e is a random k -set in $[n]$, reveal $e \cap \left(\bigcup_{\ell \in [r] \setminus \{i, j\}} \psi^{-1}(\ell)\right)$. Conditioning on any outcome, the probability that ψ' cuts e is at least as large as the probability that ψ cuts e . Hence, q_ψ is indeed maximal when ψ is balanced.

Let $q = \frac{S(k, r)r!}{r^k}$ be the probability that a uniformly random r -cut cuts a given hyperedge. It is easy to see that when ψ is balanced, then $q_\psi = (1 + O(1/n))q$, and therefore we have $q_\psi \leq (1 + O(1/n))q$ for any fixed r -cut ψ . Now observe that $\mathbb{E}[\phi] = q_\psi \mathbb{E}[e(H)]$. Indeed, $\phi = \sum_{i=1}^{m'} \mathbf{1}_{e_i \in H} \mathbf{1}_{e_i \text{ is cut by } \psi}$, so

$$\begin{aligned} \mathbb{E}[\phi] &= \sum_{i=1}^{m'} \mathbb{P}(e_i \in H \text{ and } e_i \text{ is cut by } \psi) = \sum_{i=1}^{m'} \mathbb{P}(e_i \in H) \mathbb{P}(e_i \text{ is cut by } \psi) = q_\psi \sum_{i=1}^{m'} \mathbb{P}(e_i \in H) \\ &= q_\psi \mathbb{E}[e(H)]. \end{aligned}$$

Therefore

$$\mathbb{E}[\phi] = q_\psi \mathbb{E}[e(H)] = q_\psi f(m') \leq \left(1 + O\left(\frac{1}{n}\right)\right) q(m + 2m^{2/3}) = \frac{S(k, r)r!}{r^k} m + O(m^{2/3}).$$

Plugging this in (11), we find that there exists a constant $C'' = C''(k)$ such that

$$\mathbb{P}\left(\phi - \frac{S(k, r)r!}{r^k} m \geq C''m^{3/4}\right) \leq \exp(-2C(\log r)m^{1/2}).$$

Since with high probability, H has at least m edges, it follows by a union bound over all $r^n = r^{Cm^{1/2}}$ possible r -cuts ψ of $[n]$ that with high probability H has r -surplus $O(m^{3/4})$. $\square$

7. GENERAL HYPERGRAPHS

In this section we prove Theorem 1.4. We start with the following analogue of Definition 5.1. Some explanation is given after the definition.

Definition 7.1 (Random $(m, \Delta, D, \Lambda, \lambda, k)$ -multihypergraph of general type). Fix some positive integer k and a vertex set V . We say that a random multihypergraph $H = (V, R)$ is a *random $(m, \Delta, D, \Lambda, \lambda, k)$ -multihypergraph of general type* if it satisfies the following properties:

- (i) There are multisets X and Y consisting of subsets of V of size at most k , and R is a random subset of $X \cup Y$, defined as follows: there exist a set $\mathcal{S}$ and some independent random variables $(Z_i)_{i \in \mathcal{S}}$ such that for every $e \in X \cup Y$, there exists a set $S_e \subseteq \mathcal{S}$ of size at most k such that the event $e \in R$ is a function of $\{Z_i : i \in S_e\}$.
(We will call the elements of $X \cup Y$ *potential hyperedges* in the hypergraph H , as whether they belong to the edge set R of H depends on the random variables $(Z_i)_{i \in \mathcal{S}}$.)
- (ii) $|X \cup Y| = O(m)$ and $|Y| = \Omega(m/\log m)$.
- (iii) There is a random subset $T \subseteq Y$ such that for every $e \in Y$, the event $e \in T$ is a function of $\{Z_i : i \in S_e\}$.

- (iv) For any $e_1, \dots, e_j, f \in Y$, the event $f \in T$ is independent³ of the events $e_1 \in T, \dots, e_j \in T$ unless $S_f \subseteq S_{e_1} \cup \dots \cup S_{e_j}$.
- (v) For any $e \in X \cup Y$ and $f, g \in Y$, the event $g \in T$ is independent of the events $e \in R, f \in T$ unless $S_g \subseteq S_e \cup S_f$.
- (vi) For any $e \in X \cup Y$ and $f \in Y$, we have $\mathbb{P}(e \in R, f \in T) \geq \mathbb{P}(e \in R)\mathbb{P}(f \in T)$.
- (vii) For any $e, f \in Y$ with $S_e = S_f$, we have $\mathbb{P}(e \in R, f \in T) - \mathbb{P}(e \in R)\mathbb{P}(f \in T) \geq \Omega(1)$.
- (viii) For every $e \in Y$, we have $|e| = 2$, and $|S_e|$ is the same for each $e \in Y$.
- (ix) For every vertex $u \in V$, there are at most Δ potential hyperedges $e \in X \cup Y$ such that $u \in e$.
- (x) For any distinct vertices $u, v \in V$, there are at most D potential edges $e \in X \cup Y$ such that $u, v \in e$, and for every $u \in V$ and $i \in S$, there are at most D potential edges $e \in Y$ such that $u \in e$ and $i \in S_e$.
- (xi) For every $e \in Y$, the number of potential edges $f \in Y$ with $S_e = S_f$ is at most $\Lambda\lambda$.
- (xii) For each $e \in Y$, there are at least $\lambda/2$ but at most λ potential edges $f \in Y$ with $V(e) = V(f)$ (i.e., such that e and f have the same vertex set) and $S_e = S_f$.
- (xiii) there exists some constant C such that with probability at least $1 - m^{-10}$, for every distinct $u, v \in V$, the following holds: letting $B(u, v)$ be the number of potential edges in T with vertex set $\{u, v\}$ and letting $e_g(u, v)$ be the number of potential edges in Y with vertex set $\{u, v\}$, we have $|B(u, v) - \mathbb{E}[B(u, v)]| \leq C(\log m)^{k/2} \lambda^{1/2} e_g(u, v)^{1/2}$.

Remark 7.2. In Definition 7.1, since X and Y are multisets, it is possible that two potential hyperedges e and f have the same vertex set but $S_e \neq S_f$. To avoid clutter in the proof, we will often refer to the set S_e as the *set of colours* of the potential hyperedge e . In what follows, we will often call the elements of $X \cup Y$ hyperedges without writing *potential*.

Remark 7.3. Roughly speaking, a random $(m, \Delta, D, \Lambda, \lambda, k)$ -multihypergraph of general type will arise from revealing a random partial cut on a (deterministic) k -multigraph H_0 , with the following degree conditions:

- the maximum degree of H_0 is at most Δ
- the maximum codegree of a pair of vertices in H_0 is at most D
- for a large proportion of hyperedges in H_0 , the multiplicity is roughly λ
- the maximum codegree of a set of $k - 2$ vertices in H_0 is at most $\Lambda\lambda$ (or at most Λ if multiplicities are ignored).

Properties (ix), (x), (xii) and (xi) (respectively) make this intuition formal. The hyperedges of H_0 are $e \cup S_e$ for each $e \in X \cup Y$.

Compared to Definition 5.1, the most significant difference in Definition 7.1 is that we now have two sets X and Y of potential hyperedges in H . Roughly speaking, Y is the set of potential hyperedges (in fact potential *edges*: see property (viii)) that we can “control well” – note that there are several conditions in Definition 7.1 that apply only to elements of Y and not to X . Property (ii) states that the number of such “controlled” potential edges is comparable to the total number of potential edges. The set T of potential edges is sampled from this good set Y , and these are the edges that we will use to define our SDP vectors x_u , for each $u \in V$. Although this is not a condition, one should think of T as a superset of $R \cap Y$ – in particular, whether a potential edge $e \in Y$ gets sampled to R will correlate with the event $e \in T$ (see property (vi) applied with $f = e$).

The key result to prove Theorem 1.4 is the following.

³For events $A, A_1, \dots, A_\ell$, we say that A is independent of $A_1, \dots, A_\ell$ if for any events $(B_i)_{i \in [\ell]}$ such that $B_i = A_i$ or $B_i = A_i^c$ for each $i \in [\ell]$, we have $\mathbb{P}(A, B_1, \dots, B_\ell) = \mathbb{P}(A)\mathbb{P}(B_1, \dots, B_\ell)$.

Theorem 7.4. Fix $k \in \mathbb{N}$ and let $m, \Delta, D, \Lambda, \lambda \geq 1$ be some reals. Let $H = (V, R)$ be a random $(m, \Delta, D, \Lambda, \lambda, k)$ -multihypergraph of general type. Then we have

$$\mathbb{E}[\text{sp}(H)] = \Omega\left(\frac{m\lambda}{\beta \log m}\right) - O\left(\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4}\right),$$

whenever $\beta \geq (\log m)^{(k+1)/2} \sqrt{\Delta\lambda}$.

Before we prove Theorem 7.4, we establish the following counting lemma, which we will use to get a good control on the error terms in the proof of Theorem 7.4.

Lemma 7.5. Fix $k \in \mathbb{N}$ and let $m, \Delta, D, \Lambda, \lambda$ be some positive integers. Let $H = (V, R)$ be a random $(m, \Delta, D, \Lambda, \lambda, k)$ -multihypergraph of general type. Let Q be the number of tuples $(u, v, w, w', f_1, f_2, f_3, f_4, e)$ such that $f_1, f_2, f_3, f_4 \in Y$ satisfy $S_{f_i} \subseteq \bigcup_{j \neq i} S_{f_j}$ for every $i \in [4]$, $e \in X \cup Y$ and $u, v, w, w' \in V$ satisfy that u, v, w are distinct, u, v, w' are distinct, $u, w \in f_1$, $v, w \in f_2$, $u, w' \in f_3$, $v, w' \in f_4$ and $u, v \in e$. Then

$$Q = O(m\Lambda\lambda^2\Delta D + mD^3\lambda).$$

We remark that the proof of Lemma 7.5 uses only items (i), (ii), (viii), (ix), (x), (xi), and (xii) in Definition 7.1 of a random multihypergraph of general type.

Proof. Firstly, we note that for such a tuple, once $(u, v, w, w', f_1, f_2, f_3, f_4)$ is fixed, we have $O(D)$ choices for e by (x), and therefore we simply focus on bounding the number of tuples $(u, v, w, w', f_1, f_2, f_3, f_4)$ such that $f_1, f_2, f_3, f_4 \in Y$ satisfy $S_{f_i} \subseteq \bigcup_{j \neq i} S_{f_j}$ for every $i \in [4]$, and $u, v, w, w' \in V$ satisfy that u, v, w are distinct, u, v, w' are distinct, $u, w \in f_1$, $v, w \in f_2$, $u, w' \in f_3$, and $v, w' \in f_4$.

For the tuple (f_1, u, w) , we have $O(m)$ choices by (i) and (ii). Since $S_{f_1} \subseteq \bigcup_{j \neq 1} S_{f_j}$, then f_1 must share at least one colour with either f_2, f_3 or f_4 .

Suppose first that f_1 shares at least one colour with f_2 , and we distinguish between two further cases, whether $S_{f_1} = S_{f_2}$ or not.

In the first subcase (i.e. if $S_{f_1} = S_{f_2}$), we have $O(\Lambda\lambda)$ choices for (f_2, v) by (xi). We then have $O(\Delta)$ choices for (f_3, w') by (ix), and since $S_{f_4} \subseteq \bigcup_{j \neq 4} S_{f_j}$, we have $O(1)$ choices for S_{f_4} , and therefore by (xii), we have $O(\lambda)$ choices for f_4 . This gives $O(m\Lambda\lambda^2\Delta)$ choices in total in this subcase.

In the second subcase (i.e. if $S_{f_1} \neq S_{f_2}$), we have $O(D)$ choices for (f_2, v) by (x). Since $S_{f_1} \subseteq \bigcup_{j \neq 1} S_{f_j}$, $S_{f_1} \neq S_{f_2}$, and all S_f for $f \in Y$ have the same size by (viii), we must have some colour i such that $i \in S_{f_1} \cap S_{f_\ell}$ for some $\ell = 3, 4$. Therefore, we have $O(D)$ choices for (f_ℓ, w') by (x). Let h be such that f_h has not been chosen yet. As $S_{f_h} \subseteq \bigcup_{j \neq h} S_{f_j}$, we have $O(1)$ choices for S_h , and thus $O(\lambda)$ choices for f_h by (xii). Thus we have $O(mD^2\lambda)$ choices in total in this subcase.

The case where f_1 shares at least one colour with f_3 is symmetric to the case where f_1 shares at least one colour with f_2 , and thus we have the same bounds here.

We now argue in the last remaining case, that is, where f_1 shares a colour with f_4 . As we have already dealt with the cases where f_1 shares a colour with f_2 or f_3 , we may assume that $S_{f_1} = S_{f_4}$. We have $O(\Delta)$ choices for (f_2, v) by (ix). Then we have $O(\Lambda\lambda)$ choices for (f_4, w') by (xi), and $O(\lambda)$ choices for f_3 by (xii) (as there are $O(1)$ choices for S_{f_3}). Thus we have $O(m\Lambda\lambda^2\Delta)$ choices in total in this subcase.

In total, taking into account the $O(D)$ choices for e , we find that $Q = O(m\Lambda\lambda^2\Delta D + mD^3\lambda)$, as desired. $\square$

We are now ready to prove Theorem 7.4.

Proof of Theorem 7.4. For a hyperedge $e \in Y$ such that $e = \{u, v\}$, we let $B(e, u, v)$ be the indicator of the event $e \in T$ (otherwise, if $e \neq \{u, v\}$, let $B(e, u, v) = 0$ whether or not $e \in T$). We let $B(u, v) = \sum_{f \in Y} B(f, u, v)$, let $L(e, u, v) = B(e, u, v) - \mathbb{E}[B(e, u, v)]$ and let $L(u, v) =$

$B(u, v) - \mathbb{E}[B(u, v)]$.

By (xiii), for some constant C , we have that, with probability at least $1 - m^{-10}$, for every u, v the following holds:

$$|L(u, v)| = |B(u, v) - \mathbb{E}[B(u, v)]| \leq C(\log m)^{k/2} \lambda^{1/2} e_g(u, v)^{1/2}, \quad (12)$$

where $e_g(u, v)$ is the number of edges $e \in Y$ with vertex set $\{u, v\}$. We let Ψ be the associated event.

For each $u \in V$, we define $x_u \in \mathbb{R}^V$ by taking

$$x_u(v) = \begin{cases} -1, & \text{if } v = u \\ \frac{L(u, v)}{\beta}, & \text{otherwise.} \end{cases} \quad (13)$$

If Ψ holds, then by (12) and using that $\beta \geq (\log m)^{(k+1)/2} \sqrt{\Delta \lambda}$, we have for every u that

$$1 \leq \|x_u\|^2 = 1 + \sum_{v \in V, v \neq u} \frac{L(u, v)^2}{\beta^2} \leq 1 + \sum_{v \in V, v \neq u} \frac{C^2 (\log m)^k \lambda e_g(u, v)}{\Delta \lambda (\log m)^{k+1}} \leq 2,$$

for large enough m , where the last inequality uses (ix).

Applying Theorem 2.1, we get

$$\text{sp}(H) \geq \mathbf{1}_\Psi \left(- \sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle - O\left(\sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2 \right) \right).$$

Taking expectation, we obtain

$$\mathbb{E}[\text{sp}(H)] \geq -\mathbb{E}[\mathbf{1}_\Psi \sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle] - O(\mathbb{E}[\mathbf{1}_\Psi \sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2]) \quad (14)$$

We will now argue that we may remove both occurrences of $\mathbf{1}_\Psi$ in the previous equation and still get a valid lower bound on $\mathbb{E}[\text{sp}(H)]$, up to a tiny error term. Indeed, for every $u, v \in V$, by Cauchy-Schwarz we have $|\langle x_u, x_v \rangle| \leq \|x_u\| \cdot \|x_v\|$. Since $|L(u, v)| = O(m)$ by (ii) and $\beta \geq 1$, we have $\|x_u\| = O(m^{3/2})$ for every $u \in V$ (with probability 1), and therefore $|\langle x_u, x_v \rangle| = O(m^3)$ and $\langle x_u, x_v \rangle^2 = O(m^6)$. Therefore

$$\begin{aligned} -\mathbb{E}[\mathbf{1}_\Psi \sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle] - O(\mathbb{E}[\mathbf{1}_\Psi \sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2]) &= O(\mathbb{E}[\mathbf{1}_\Psi m^4] + \mathbb{E}[\mathbf{1}_\Psi m^7]) \\ &= O(m^7/m^{10}) = O(m^{-3}). \end{aligned}$$

It then follows from (14) that

$$\begin{aligned} \mathbb{E}[\text{sp}(H)] &\geq -\mathbb{E}[\sum_{e \in R} C_e \sum_{u \neq v \in e} \langle x_u, x_v \rangle] - O(\mathbb{E}[\sum_{e \in R} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2]) - O(m^{-3}) \\ &\geq \sum_{e \in X \cup Y} C_e \sum_{u \neq v \in e} \mathbb{E}[-\mathbf{1}_{e \in R} \langle x_u, x_v \rangle] - O(\mathbb{E}[\sum_{e \in X \cup Y} \sum_{u \neq v \in e} \langle x_u, x_v \rangle^2]) - O(m^{-3}). \end{aligned} \quad (15)$$

Our goal is now to give a lower bound of the RHS of the previous expression. We start by giving a lower bound of the first term of (15). For a hyperedge $e \in X \cup Y$ and $u, v \in e$, we let $K(e, u, v)$ be the indicator of the event $e \in R$ (otherwise, if $e \not\supset \{u, v\}$, let $K(e, u, v) = 0$ whether or not $e \in R$). For a given hyperedge $e \in X \cup Y$, and distinct vertices $u, v \in e$, we have

$$\mathbb{E}[-\mathbf{1}_{e \in R} \langle x_u, x_v \rangle] = \mathbb{E}[2K(e, u, v) \frac{L(u, v)}{\beta}] - \mathbb{E}[\sum_{w \neq u, v} K(e, u, v) \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta}]. \quad (16)$$

We first focus on the first subterm, and write

$$\mathbb{E}[K(e, u, v) \frac{L(u, v)}{\beta}] = \mathbb{E}[K(e, u, v) \sum_{f \in M_e^1} \frac{L(f, u, v)}{\beta} + K(e, u, v) \sum_{f \in M_e^2} \frac{L(f, u, v)}{\beta}], \quad (17)$$

where, if $e \in Y$, then M_e^1 is the set of edges $f \in Y$ containing u and v such that $S_e = S_f$, and M_e^2 is the set of edges $f \in Y$ containing u and v such that $S_e \neq S_f$, and, if $e \in X$, then $M_e^1 = \emptyset$ and $M_e^2 = Y$. By (vii), we have

$$\mathbb{E}[K(e, u, v) \sum_{f \in M_e^1} \frac{L(f, u, v)}{\beta}] = \Omega\left(\frac{|M_e^1|}{\beta}\right).$$

By (xii), we have that $|M_e^1| \geq \lambda/2$ if $e \in Y$, and therefore

$$\mathbb{E}[K(e, u, v) \sum_{f \in M_e^1} \frac{L(f, u, v)}{\beta}] = \Omega\left(\frac{\lambda \mathbf{1}_{e \in Y}}{\beta}\right).$$

Furthermore, by (vi), we have

$$\mathbb{E}[K(e, u, v) \sum_{f \in M_e^2} \frac{L(f, u, v)}{\beta}] \geq 0.$$

Plugging the two above bounds in (17), we get

$$\mathbb{E}[K(e, u, v) \frac{L(u, v)}{\beta}] = \Omega\left(\frac{\lambda \mathbf{1}_{e \in Y}}{\beta}\right). \quad (18)$$

We now consider the second subterm in (16). Note that

$$\begin{aligned} \sum_{e \in X \cup Y} C_e \sum_{u \neq v \in e} \mathbb{E} \left[\sum_{w \neq u, v} K(e, u, v) \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta} \right] = \\ \sum_{e \in X \cup Y} C_e \sum_{u \neq v \in e} \sum_{w \neq u, v} \sum_{f, f' \in Y} \mathbb{E} \left[K(e, u, v) \frac{L(f, u, w)}{\beta} \cdot \frac{L(f', v, w)}{\beta} \right]. \end{aligned} \quad (19)$$

We remark that, by (v), unless $S_f \subseteq S_e \cup S_{f'}$ and $S_{f'} \subseteq S_e \cup S_f$, we have

$$\mathbb{E}[K(e, u, v) L(f, u, w) L(f', v, w)] = 0.$$

We now count those terms $K(e, u, v) L(f, u, w) L(f', v, w)$ in (19) which do not vanish in expectation. We first distinguish between two cases:

- $S_f = S_{f'}$,
- there exists some colour i such that $i \in S_f$ and $i \in S_e$,

which cover all possibilities as, by (viii), we have $|S_f| = |S_{f'}|$ for all $f, f' \in Y$. In the first case, we have $O(m\Lambda\lambda)$ possibilities for (f, f', u, v, w) by (xi). Then, by (x), we have at most $O(D)$ choices for e , thereby we have $O(m\Lambda\lambda D)$ choices for (e, f, f', u, v, w) in total. In the second case, we have at most $O(m)$ possibilities for (e, u, v) , then $O(D)$ possibilities for (f, w) . As $S_{f'} \subseteq S_e \cup S_f$, we have $O(1)$ possibilities for $S_{f'}$, and therefore $O(\lambda)$ possibilities for f' by (xii). In total, this gives $O(mD\lambda)$ possibilities. Therefore, as $\Lambda \geq 1$, combining all this gives that

$$\sum_{e \in X \cup Y} C_e \sum_{u \neq v \in e} \mathbb{E} \left[\sum_{w \neq u, v} K(e, u, v) \frac{L(u, w)}{\beta} \frac{L(v, w)}{\beta} \right] = O\left(\frac{m\Lambda\lambda D}{\beta^2}\right). \quad (20)$$

Combining (16), (18) and (20), together with $|Y| = \Omega(m/\log m)$ from (ii), we obtain

$$\sum_{e \in X \cup Y} C_e \mathbb{E} \left[\sum_{u \neq v \in e} -\mathbf{1}_{e \in R} \langle x_u, x_v \rangle \right] = \Omega\left(\frac{m\lambda}{\beta \log m}\right) - O\left(\frac{m\Lambda\lambda D}{\beta^2}\right). \quad (21)$$

We now turn to the second term of (15). We fix some hyperedge $e \in X \cup Y$ and distinct vertices $u, v \in e$. Using the inequality $(a + b)^2 \leq 2a^2 + 2b^2$, we have

$$\begin{aligned} \mathbb{E}[\langle x_u, x_v \rangle^2] &= \mathbb{E} \left[\left(-2 \frac{L(u, v)}{\beta} + \sum_{w \in V \setminus \{u, v\}} \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta} \right)^2 \right] \\ &\leq \mathbb{E} \left[2 \left(2 \frac{L(u, v)}{\beta} \right)^2 + 2 \left(\sum_{w \in V \setminus \{u, v\}} \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta} \right)^2 \right] \\ &= \mathbb{E} \left[8 \frac{L(u, v)^2}{\beta^2} \right] + \mathbb{E} \left[2 \sum_{w, w' \in V \setminus \{u, v\}} \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta} \cdot \frac{L(u, w')}{\beta} \cdot \frac{L(v, w')}{\beta} \right]. \quad (22) \end{aligned}$$

For the first term, we have

$$\mathbb{E}[L(u, v)^2] = \sum_{f, f' \in Y} \mathbb{E}[L(f, u, v) L(f', u, v)].$$

By (iv), we have $\mathbb{E}[L(f, u, v) L(f', u, v)] = 0$ unless $S_f = S_{f'}$, and therefore by (xii), we have for every fixed f that $\sum_{f' \in Y} \mathbb{E}[L(f, u, v) L(f', u, v)] = O(\lambda)$. Summing over $O(D)$ choices for f by (x) then gives

$$\mathbb{E} \left[\frac{L(u, v)^2}{\beta^2} \right] = O \left(\frac{\lambda D}{\beta^2} \right).$$

Turning to the second term in (22), we want to bound

$$\sum_{e \in X \cup Y} \sum_{u \neq v \in e} \sum_{w, w' \in V \setminus \{u, v\}} \sum_{f_1, f_2, f_3, f_4 \in Y} \mathbb{E}[L(f_1, u, w) L(f_2, v, w) L(f_3, u, w') L(f_4, v, w')] \quad (23)$$

For a given expectation term not to vanish, by (iv), we must have for every $i \in [4]$ that $S_{f_i} \subseteq \bigcup_{j \neq i} S_{f_j}$. Let Q be the number of tuples $(u, v, w, w', f_1, f_2, f_3, f_4, e)$ such that $f_1, f_2, f_3, f_4 \in Y$ satisfy $S_{f_i} \subseteq \bigcup_{j \neq i} S_{f_j}$ for every $i \in [4]$, $e \in X \cup Y$ and $u, v, w, w' \in V$ satisfy that u, v, w are distinct, u, v, w' are distinct, $u, w \in f_1$, $v, w \in f_2$, $u, w' \in f_3$, $v, w' \in f_4$ and $u, v \in e$. By Lemma 7.5, we have $Q = O(m\Lambda\lambda^2\Delta D + mD^3\lambda)$, and therefore there are at most $O(m\Lambda\lambda^2\Delta D + mD^3\lambda)$ terms which do not vanish in (23). Since each of those terms is $O(1)$, we get

$$\sum_{e \in X \cup Y} \sum_{u \neq v \in e} \mathbb{E} \left[\sum_{w, w' \in V \setminus \{u, v\}} \frac{L(u, w)}{\beta} \cdot \frac{L(v, w)}{\beta} \cdot \frac{L(u, w')}{\beta} \cdot \frac{L(v, w')}{\beta} \right] = O \left(\frac{m\Lambda\lambda^2\Delta D + mD^3\lambda}{\beta^4} \right).$$

It follows that

$$\sum_{e \in X \cup Y} \sum_{u \neq v \in e} \mathbb{E}[\langle x_u, x_v \rangle^2] = O \left(\frac{m\lambda D}{\beta^2} + \frac{m\Lambda\lambda^2\Delta D + mD^3\lambda}{\beta^4} \right). \quad (24)$$

Plugging in (21) and (24) into (15), we obtain

$$\mathbb{E}[\text{sp}(H)] = \Omega \left(\frac{m\lambda}{\beta \log m} \right) - O \left(\frac{m\Lambda\lambda D}{\beta^2} + \frac{m\Lambda\lambda^2\Delta D + mD^3\lambda}{\beta^4} \right) - O(m^{-3}).$$

As $\beta \geq (\log m)^{(k+1)/2} \sqrt{\Delta\lambda}$, we have $\frac{m\Lambda\lambda^2\Delta D}{\beta^4} = O(\frac{m\Lambda\lambda D}{\beta^2})$, and therefore we can ignore the $O(\frac{m\Lambda\lambda^2\Delta D}{\beta^4})$ term. We can also ignore the $O(m^{-3})$ term using that $\mathbb{E}[\text{sp}(H)] = \Omega(1)$. Therefore,

$$\mathbb{E}[\text{sp}(H)] = \Omega \left(\frac{m\lambda}{\beta \log m} \right) - O \left(\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4} \right),$$

as wanted. $\square$

The following lemma will be used to verify that condition (xiii) is satisfied in our applications of Theorem 7.4.

Lemma 7.6. *Let b, k be positive integers, and let H be a k -uniform multihypergraph on a vertex set of size at most n . Suppose that we give independently to each vertex a colour uniformly at random in $\mathbb{Z}_b$. Fix $a \in \mathbb{Z}_b$, and let X_a be the random variable counting the number of hyperedges in which the sum of the colours of the vertices is equal to a modulo b . Then for every constant C_1 , there exists a constant $C_2 = C_2(k, b, C_1)$, such that for n sufficiently large, we have with probability at least $1 - n^{-C_1}$ that*

$$|X_a - \mathbb{E}[X_a]| \leq C_2(\log n)^{k/2} \cdot \left(\sum_{e \in E(H)} \lambda_e^2 \right)^{1/2},$$

where λ_e is the multiplicity of the hyperedge e .

Proof. We prove this lemma by induction on k . For $k = 1$, the result follows by a standard application of Azuma's inequality (see Lemma 3.1). Suppose now that the result is true for $k - 1$. We may assume that $|V(H)| = n$ by adding, if necessary, isolated vertices. We take an arbitrary ordering $v_1, \dots, v_n$ of the vertices of H , and reveal their colour one by one. We construct a martingale $(M_j)_{0 \leq j \leq n}$ as follows. First, we set $M_0 = 0$. Before we reveal the colour of vertex v_j , we let H_j be the hypergraph on vertex set $\{v_1, \dots, v_{j-1}\}$ and edge set consisting of $e \setminus v_j$ for every hyperedge $e \in E(H)$ with $j = \max\{i : v_i \in e\}$. We consider the following process: if there exists a colour $\ell \in \mathbb{Z}_b$ such that in H_j , the number X_ℓ^j of hyperedges on which the sum of the colours is congruent to ℓ modulo b satisfies $|X_\ell^j - \frac{e(H_j)}{b}| > C'_2(\log n)^{(k-1)/2} \cdot (\sum_{e \in E(H): v_j \in e} \lambda_e^2)^{1/2}$, where $C'_2 = C(k - 1, b, C_1 + 2)$, we say we terminate with failure, and set $M_n = \dots = M_j$. Otherwise, we set $M_{j+1} = M_j + X_z^j - \frac{e(H_j)}{b}$, where z is the congruence class of $a - c(v_j)$, and $c(v_j)$ is the colour assigned to v_j . We then move on to reveal the colour of v_{j+1} . We remark that, by induction hypothesis, we terminate with failure at step j with probability at most n^{-C_1-2} . Therefore, by union bound, the probability of terminating with failure is at most n^{-C_1-1} . We also note that $(M_j)_{0 \leq j \leq n}$ is a martingale, whose j -th increment is bounded from above by $C'_2(\log n)^{(k-1)/2} (\sum_{e \in E(H): v_j \in e} \lambda_e^2)^{1/2}$. By Azuma's inequality (see Lemma 3.1), we get

$$\begin{aligned} \mathbb{P}(|M_n| > C(\log n)^{k/2} (\sum_{e \in E(H)} \lambda_e^2)^{1/2}) &\leq 2 \exp \left(- \frac{(C(\log n)^{k/2} (\sum_{e \in E(H)} \lambda_e^2)^{1/2})^2}{2 \sum_{j=1}^n (C'_2(\log n)^{(k-1)/2} \cdot (\sum_{e \in E(H): v_j \in e} \lambda_e^2)^{1/2})^2} \right) \\ &= 2 \exp \left(- \frac{C^2 \log n \sum_{e \in E(H)} \lambda_e^2}{2(C'_2)^2 \sum_{j=1}^n \sum_{e \in E(H): v_j \in e} \lambda_e^2} \right). \end{aligned}$$

Since $\sum_{j=1}^n \sum_{e \in E(H): v_j \in e} \lambda_e^2 = k \sum_{e \in E(H)} \lambda_e^2$, it follows that for C sufficiently large

$$\mathbb{P}(|M_n| > C(\log n)^{k/2} (\sum_{e \in E(H)} \lambda_e^2)^{1/2}) \leq 2 \exp \left(- \frac{C^2 \log n}{2k(C'_2)^2} \right) \leq n^{-C_1-1}. \quad (25)$$

Since $M_n = X_a - \mathbb{E}[X_a]$ unless we terminate with failure, we have that $M_n = X_a - \mathbb{E}[X_a]$ holds with probability at least $1 - n^{-C_1-1}$. This together with (25) give the desired conclusion. $\square$

The next lemma and its corollary will be used to verify that conditions (xi) and (xii) are satisfied in our applications of Theorem 7.4.

Lemma 7.7. *Let $\Gamma = \binom{k}{2} m^{\frac{1}{\lceil k/2 \rceil}}$ and let G be a k -uniform hypergraph with m edges. Then there exist an ordering $e^1, \dots, e^m$ of the edges of G and some $u^i, v^i \in e^i$ (with $u^i \neq v^i$) such that for each $1 \leq i \leq m - 1$, the number of $j > i$ for which e^j contains $e^i \setminus \{u^i, v^i\}$ is at most Γ .*

Proof. It suffices to prove that there exist some $e \in E(G)$ and distinct $u, v \in e$ with the property that the number of $f \in E(G) \setminus \{e\}$ containing $e \setminus \{u, v\}$ is at most Γ . Indeed, if such e and u, v exist, then we can just set $e^i = e$, $u^i = u$, $v^i = v$ and continue with the hypergraph $G - e$.

Now assume that such e and u, v do not exist. Let $f^0 = \{x_1, x_2, \dots, x_k\}$ be an arbitrary hyperedge in G . By assumption, there are more than Γ edges $f^1 \in E(G)$ containing $f^0 \setminus \{x_1, x_2\}$. For each such f^1 , by assumption there are more than Γ edges $f^2 \in E(G)$ containing $f^1 \setminus \{x_3, x_4\}$. Continue in this manner, defining, for each choice of f^i , more than Γ different f^{i+1} s, all in $E(G)$. Observe that, for a given f^0 , the number of different choices for the sequence $(f^1, f^2, \dots, f^{\lfloor k/2 \rfloor})$ is greater than $\Gamma^{\lfloor k/2 \rfloor}$. Moreover, for any given f^0 and $f^{\lfloor k/2 \rfloor}$, the number of possibilities for $(f^1, \dots, f^{\lfloor k/2 \rfloor})$ is at most $\binom{k}{2}^{\lfloor k/2 \rfloor}$ since $f^{i+1} \setminus (f^i \setminus \{x_{2i+1}, x_{2i+2}\}) \subset f^{\lfloor k/2 \rfloor}$ for each i . Since $f^{\lfloor k/2 \rfloor} \in E(G)$, it follows that $\Gamma^{\lfloor k/2 \rfloor} < \binom{k}{2}^{\lfloor k/2 \rfloor} m$, which is a contradiction. $\square$

From Lemma 7.7, we deduce the following corollary.

Corollary 7.8. *Let H be a k -uniform multihypergraph with m edges. Then there exists a sub-multihypergraph H' of H with at least $\Omega\left(\frac{m}{\log m}\right)$ edges such that there are some positive integers λ, Λ with $\Lambda = O(m^{\frac{1}{\lfloor k/2 \rfloor}})$ and a subset $S(e) \subset e$ of size $k - 2$ for every $e \in E(H')$ such that*

- *the multiplicity of every edge $e \in E(H')$ is at least $\lambda/2$ and at most λ ,*
- *for every edge $e \in E(H')$, the number of edges $f \in E(H')$ with $S(e) = S(f)$ is at most $\lambda\Lambda$.*

Proof. As the multiplicity of every edge is between 1 and m , a simple dyadic pigeon holing argument shows that there exists a sub-multihypergraph H^* of H with at least $\Omega\left(\frac{m}{\log m}\right)$ edges such that the multiplicity of every edge $e \in E(H^*)$ is at least $\lambda/2$ and at most λ , for some integer λ . We now consider the hypergraph $\tilde{H}$ obtained by keeping only one copy of every edge of H^* . For $M = |E(\tilde{H})|$ we choose an ordering $e^1, e^2, \dots, e^M$ of $E(\tilde{H})$ and vertices $u^i, v^i \in e^i$ for every $i \in [M]$ according to Lemma 7.7, and let $S(e^i) = e^i \setminus \{u^i, v^i\}$. If for some $i < j$, we have $S(e^i) = S(e^j)$, then in particular $e^i \setminus \{u^i, v^i\} = S(e^i) \subset e^j$. By construction, the number of such e^j is at most Γ , for any given e^i , for some $\Gamma = O(M^{\frac{1}{\lfloor k/2 \rfloor}})$. Therefore we have at most $M\Gamma$ pairs $(e, f) \in E(\tilde{H})^2$ such that $S(e) = S(f)$. Let R be the set of edges $e \in E(\tilde{H})$ such that there are at least 4Γ edges $f \in E(\tilde{H})$ with $S(e) = S(f)$. Then $2\Gamma|R| \leq M\Gamma$, giving $|R| \leq M/2$. Letting $\Lambda = 4\Gamma = O(m^{\frac{1}{\lfloor k/2 \rfloor}})$ and H' be the sub-multihypergraph of H^* consisting of those edges (with their multiplicity in H^*) that are not in R , we have that $|E(H')| = \Omega\left(\frac{m}{\log m}\right)$ and for every edge $e \in E(H')$, there are at most $\Lambda\lambda$ edges $f \in E(H')$ such that $S(e) = S(f)$. $\square$

We now turn to the proof of Theorem 1.4. We distinguish between two cases: $r \leq k - q$ for some large constant $q = q(\varepsilon)$, and $r > k - q$.

7.1. The case of small r . We first deal with the case where r is not too large in Theorem 1.4.

Theorem 7.9. *For every $\varepsilon > 0$, there exist some positive integer q , such that for every $2 \leq r \leq k - q$, every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$.*

Proof. First of all, by Lemma 3.11, it suffices to show that for every $\varepsilon > 0$, there exists some positive integer q such that every mixed k -multigraph H with maximum degree at most $\Delta = \Theta(m^{2/3-\varepsilon})$, maximum codegree $D = O(m^{1/3-2\varepsilon})$, $O(m)$ hyperedges in total, and $\Omega(m)$ hyperedges of size $q + 2$ has a 2-cut of surplus $\Omega(m^{2/3-\varepsilon})$. More precisely, we will show that every sufficiently large even integer q will be suitable. Clearly, we may assume that m is sufficiently large.

We apply Corollary 7.8 to the subhypergraph of H consisting of the edges of size exactly $q + 2$ and obtain a set $B \subseteq E(H)$ of hyperedges such that $|B| = \Omega\left(\frac{m}{\log m}\right)$ and there exist positive integers λ, Λ with $\Lambda = O(m^{\frac{1}{\lfloor (q+2)/2 \rfloor}})$ and a subset $S(e) \subset e$ of size q for every $e \in B$ such that

- the multiplicity of every edge $e \in B$ is at least $\lambda/2$ and at most λ ,
- for every edge $e \in B$, the number of edges $f \in B$ with $S(e) = S(f)$ is at most $\lambda\Lambda$.

We now take a random partition $W \cup W^c$ of $V(H)$ by putting each vertex with probability $1/2$ in W , and in W^c otherwise, independently for each vertex. We say that a hyperedge $e \in E(H)$ is *good* if $e \in B$ and $e \cap W^c = S(e)$, and e is *bad* otherwise. Note that the expected number of good hyperedges is $\Omega(m/\log m)$, and therefore we may fix an outcome of W such that the number of good hyperedges is $\Omega(m/\log m)$.

Let $\sigma : W^c \rightarrow \{1, 2\}$ be chosen uniformly at random. We now construct the following auxiliary multihypergraph H_σ on vertex set W : for every edge $e \in E(H)$, if all vertices of e are in W then we add two copies of e , otherwise, if $|\sigma(e \setminus W)| = 1$ and $|e \cap W| > 0$, we add one copy of $e \cap W$. Let $R = E(H_\sigma)$. By Corollary 3.10, we have

$$\text{sp}(H) \geq \mathbb{E}_\sigma[\text{sp}(H_\sigma)]/2. \quad (26)$$

We now claim that $H_\sigma = (W, R)$ is a random $(m, 2\Delta, 2D, \Lambda, \lambda, k)$ -multihypergraph of general type. Indeed, it suffices to verify that all conditions in Definition 7.1 are satisfied.

- (i) We let $Y = \{e \cap W : e \in E(H) \text{ is good}\}$, and we let $X = \{e \cap W : e \in E(H) \text{ is bad and } e \cap W \neq \emptyset\}$, where, if $e \subset W$, the element $e \cap W = e$ appears with multiplicity 2. For every hyperedge $e \in E(H)$ with representative $e_W := e \cap W \in X \cup Y$, we let $S_{e_W} = e \cap W^c$, and for every vertex $a \in W^c$, we let $Z_a = \sigma(a)$. Then the event $\{e_W \in R\}$ is a function of $\{Z_i : i \in S_{e_W}\}$: more precisely, we have $\{e_W \in R\}$ if and only if all the Z_i for $i \in S_{e_W}$ have the same value.
- (ii) We have that $|X \cup Y| \leq 2e(H) = O(m)$, and by our choice of W , we have $|Y| = \Omega(m/\log m)$.
- (iii) We let T consist of all $e \in Y$ for which $\sum_{v \in S_e} \sigma(v)$ is congruent to 0 modulo 2.
- (iv) Immediate from the definition of T .
- (v) Immediate from the definition of T .
- (vi) If $S_f \not\subseteq S_e$, then we have $\mathbb{P}(e \in R, f \in T) = \mathbb{P}(e \in R)\mathbb{P}(f \in T)$. If $S_f \subseteq S_e$, then we let $A = S_e \setminus S_f$. For $i = 1, 2$, we let U_A^i be the event that $\sigma(v) = i$ for every $v \in A$, and $U_{S_f}^i$ be the event that $\sigma(v) = i$ for every $v \in S_f$. Then

$$\mathbb{P}(e \in R, f \in T) = \mathbb{P}(U_A^1 \cap U_{S_f}^1 \cap \{f \in T\}) + \mathbb{P}(U_A^2 \cap U_{S_f}^2 \cap \{f \in T\}).$$

As for $i = 1, 2$, $U_{S_f}^i \subseteq \{f \in T\}$ (as q is even) and U_A^i is independent from $U_{S_f}^i$, we have

$$\begin{aligned} \mathbb{P}(e \in R, f \in T) &= \mathbb{P}(U_A^1)\mathbb{P}(U_{S_f}^1) + \mathbb{P}(U_A^2)\mathbb{P}(U_{S_f}^2) = 2^{-|A|}2^{-|S_f|} + 2^{-|A|}2^{-|S_f|} \\ &= 2^{-|A|-|S_f|+1} \geq \mathbb{P}(e \in R)\mathbb{P}(f \in T). \end{aligned}$$

- (vii) As $\{e \in R\} \subseteq \{e \in T\} = \{f \in T\}$ (where the inclusion is due to q being even), we are left to prove that $\mathbb{P}(e \in R)(1 - \mathbb{P}(e \in T)) = \Omega(1)$. Since it is easy to check that $\mathbb{P}(e \in T) = 1 - \Omega(1)$ and $\mathbb{P}(e \in R) = \Omega(1)$, we get the desired property.
- (viii) It is immediate from the definition of Y that $|e| = 2$ and $|S_e| = q$ for each $e \in Y$.
- (ix) Immediate.
- (x) Immediate.
- (xi) Immediate from the definition of Y .
- (xii) Immediate from the definition of Y .
- (xiii) Let u, v be distinct vertices in W . We let $H_{u,v}$ be the q -uniform multihypergraph on vertex set W^c , where we add the hyperedge S_e for every $e \in Y$ with vertex set $\{u, v\}$. Note that $e_g(u, v) = |E(H_{u,v})|$ and that the multiplicity of every hyperedge in $H_{u,v}$ is at most λ by the previous item. Note further that σ induces a uniformly random colouring of the vertex set of $H_{u,v}$ with colours in $\mathbb{Z}_2$, and that $B(u, v)$ corresponds to the number of hyperedges in $H_{u,v}$ in which the sum of the colours of the vertices is equal to 0 in $\mathbb{Z}_2$. Ignoring isolated vertices, the number of vertices in $H_{u,v}$ is $O(m)$. Thus, by applying Lemma 7.6 for $H_{u,v}$ with $a = 0$, $b = 2$, $k = q$ and $C_1 = 13$, there exists a constant

$C = C(q, b, C_1)$ such that, with probability at least $1 - \Omega(m^{-13})$, we have

$$|B(u, v) - \mathbb{E}[B(u, v)]| \leq C(\log O(m))^{q/2} \lambda^{1/2} e_g(u, v)^{1/2} \leq C(\log m)^{k/2} \lambda^{1/2} e_g(u, v)^{1/2}.$$

Taking a union bound over all distinct $u, v \in W$ (for which there are at most $O(m^2)$ choices, ignoring isolated vertices in H), we obtain the desired result.

We now apply Theorem 7.4 and get

$$\mathbb{E}_\sigma[\text{sp}(H_\sigma)] = \Omega\left(\frac{m\lambda}{\beta \log m}\right) - O\left(\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4}\right),$$

where $\beta = (\log m)^{(k+1)/2} \sqrt{2\Delta\lambda}$. Recalling that $\Delta = \Theta(m^{2/3-\varepsilon})$, $D = O(m^{1/3-2\varepsilon})$ and $\Lambda = O(m^{\frac{1}{\lfloor (q+2)/2 \rfloor}})$, we get that

$$\begin{aligned} \frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4} &= \frac{m\lambda}{\beta \log m} \left(\frac{\Lambda D \log m}{\beta} + \frac{D^3 \log m}{\beta^3} \right) \\ &= \frac{m\lambda}{\beta \log m} \cdot O\left(\frac{\Lambda D}{\sqrt{\Delta}} + \frac{D^3}{\Delta^{3/2}}\right) \\ &= \frac{m\lambda}{\beta \log m} \cdot O\left(m^{\frac{1}{\lfloor (q+2)/2 \rfloor} - 3\varepsilon/2} + m^{-9\varepsilon/2}\right). \end{aligned}$$

For q sufficiently large, we have $m^{\frac{1}{\lfloor (q+2)/2 \rfloor} - 3\varepsilon/2} + m^{-9\varepsilon/2} = o(1)$, which shows that

$$\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4} = o\left(\frac{m\lambda}{\beta \log m}\right),$$

and therefore

$$\mathbb{E}_\sigma[\text{sp}(H_\sigma)] = \Omega\left(\frac{m\lambda}{\beta \log m}\right) = \Omega\left(m^{2/3}\right).$$

Together with (26), we obtain the desired conclusion. $\square$

7.2. The case of large r . We now deal with the case of large r in Theorem 1.4, by showing the following.

Theorem 7.10. *For every $\varepsilon > 0$, and every positive integer q , there exists an integer k_0 , such that for every $k \geq k_0$ and $k - q \leq r \leq k$, every k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$.*

In the proof, we will use the following lemma, whose proof will be given in Section 7.3.

Lemma 7.11. *For every nonnegative integer t , there exists $r_0 \in \mathbb{N}$ such that for every prime $r \geq r_0$, the following holds. We let $f : [r+t] \rightarrow [r]$ be a uniformly random surjective function. Then the probability that $\sum_{i=1}^r f(i)$ is congruent to $r(r+1)/2$ modulo r is greater than $1/r$.*

Note that Lemma 7.11 is trivial for $t = 0$. When $r = k$, we only need this special case of the lemma, so the reader who is not interested in the case $k - q \leq r \leq k - 1$ may skip Section 7.3.

We first prove the following lemma which ensures that some auxiliary hypergraph arising in the proof of Theorem 7.10 is random of general type.

Lemma 7.12. *Let q be a fixed positive integer. There exists p_0 such that the following holds. Let $p \geq p_0$ be prime and let κ be some integer such that $p + 2 \leq \kappa \leq p + q + 2$. Let $H = (V, E)$ be a mixed κ -multigraph such that $|E| \geq m$, where m is sufficiently large. Let $W \subseteq W' \subseteq V$, and let $W^c = V \setminus W$. Suppose that there exist a subset $\mathcal{B}$ of E such that all vertices of every edge of $\mathcal{B}$ are in W' , and some reals $\Delta, D, \Lambda, \lambda \geq 1$ such that*

- *For distinct $u, v \in W'$, $\deg_H(v) \leq \Delta$ and $\deg_H(u, v) \leq D$.*
- *For every $e \in \mathcal{B}$, the number of edges $f \in \mathcal{B}$ with $e \setminus W = f \setminus W$ is at most $\Lambda\lambda$.*

- The multiplicity of each $e \in \mathcal{B}$ is at least $\lambda/2$ and at most λ .
- $|\mathcal{B}| = \Omega(\frac{m}{\log m})$ and for every edge $e \in \mathcal{B}$, we have $|e| = p+2$ and $|e \cap W| = 2$.

Let $\pi : W^c \rightarrow \{1, \dots, p+2\}$ be a random function such that $\pi(v)$ is sampled uniformly at random from $\{3, \dots, p+2\}$ if $v \in W' \setminus W$, and uniformly at random from $\{1, \dots, p+2\}$ if $v \in V \setminus W'$, independently for each vertex $v \in W^c$. Let H_π be the multihypergraph with vertex set W and edge set $R := \{e \cap W : e \in E(H), \{3, \dots, p+2\} \subseteq \pi(e \cap W^c), \{1, \dots, p+2\} \not\subseteq \pi(e \cap W^c)\}$, where two copies of $e \cap W$ are added to R if $\pi(e \cap W^c) \cap \{1, \dots, p+2\} = \{3, \dots, p+2\}$. Then H_π is a random $(m, 2\Delta, 2D, \Lambda, 2\lambda, \kappa)$ -multihypergraph of general type.

Proof. It suffices to verify that all conditions in Definition 7.1 are satisfied.

- (i) We let $Y = \{e \cap W : e \in \mathcal{B}\}$, each appearing with multiplicity two, and we let $X = \{e \cap W : e \in E \setminus \mathcal{B}\}$, where each element also appears with multiplicity 2. For every hyperedge $e \in E(H)$ with representatives $e_W^1, e_W^2 = e \cap W \in X \cup Y$, we let $S_{e_W^1}, S_{e_W^2} = e \setminus W$, and for every vertex $a \in W^c$, we let $Z_a = \pi(a)$. Then we let $e_W^1 \in R = E(H_\pi)$ if and only if the set $\{\pi(v) : v \in S_{e_W^1}\}$ contains the set $\{3, \dots, p+2\}$, but does not contain the set $\{1, \dots, p+2\}$, and we let $e_W^2 \in R = E(H_\pi)$ if and only if the set $\{\pi(v) : v \in S_{e_W^2}\}$ is equal to the set $\{3, \dots, p+2\}$ (possibly with multiplicity). (Note that if e_W^1 and e_W^2 are the representatives of a hyperedge $e \in Y$, then the events $e_W^1 \in R$ and $e_W^2 \in R$ coincide because $S_{e_W^1} = S_{e_W^2} \subset W'$ in this case.) Observe that this defines the same random set R that was given in the statement of the lemma.
- (ii) $|X \cup Y| \leq 2e(H) = O(m)$ and, by assumption, we have $|Y| = 2|\mathcal{B}| = \Omega(m/\log m)$.
- (iii) We let T consist of all $e \in Y$ for which $\sum_{v \in S_e} \pi(v)$ is congruent to $\sum_{i=3}^{p+2} i = \frac{p(p+1)}{2}$ modulo p .
- (iv) Immediate from the definition of T .
- (v) Immediate from the definition of T .
- (vi) If $S_f \not\subseteq S_e$, then we have $\mathbb{P}(e \in R, f \in T) = \mathbb{P}(e \in R)\mathbb{P}(f \in T)$. Assume therefore that $S_f \subseteq S_e$. Since $f \in Y$, we have that $|S_f| = p$, and therefore $|S_e \setminus S_f| \leq \kappa - p \leq q+2$. Moreover, for $f \in Y$, since all vertices of f are in W' , we have that each $\pi(v)$ for $v \in S_f$ is uniformly random in $\{3, \dots, p+2\}$, independently of the other vertices, and therefore $\mathbb{P}(f \in T) = 1/p$.

We first deal with the case where $e \in X \cup Y$ is a first representative of an edge $e' \in E(H)$ in the definition of X and Y , i.e. $e \in R$ if and only if $\{\pi(v) : v \in S_e\}$ contains $\{3, \dots, p+2\}$ but does not contain $\{1, \dots, p+2\}$. We claim that

$$\begin{aligned} \mathbb{P}(f \in T, e \in R, \pi^{-1}(1) \cap S_e = A_1, \pi^{-1}(2) \cap S_e = A_2) \\ \geq \mathbb{P}(f \in T)\mathbb{P}(e \in R, \pi^{-1}(1) \cap S_e = A_1, \pi^{-1}(2) \cap S_e = A_2) \end{aligned} \quad (27)$$

for all disjoint pairs $A_1, A_2 \subseteq S_e \setminus W' \subseteq S_e \setminus S_f$. Summing over all such A_1 and A_2 then gives the desired inequality $\mathbb{P}(f \in T, e \in R) \geq \mathbb{P}(f \in T)\mathbb{P}(e \in R)$, so it suffices to establish (27). Note that we may assume that at least one of A_1 and A_2 is empty, as otherwise both sides of (27) are 0.

Now observe that under these assumptions on A_1 and A_2 , the event

$$\{e \in R, \pi^{-1}(1) \cap S_e = A_1, \pi^{-1}(2) \cap S_e = A_2\}$$

is equal to the event

$$\mathcal{E}_{A_1, A_2} := \{\pi(S_e \setminus (A_1 \cup A_2)) = \{3, \dots, p+2\}, \pi^{-1}(1) \cap S_e = A_1, \pi^{-1}(2) \cap S_e = A_2\},$$

so it suffices to prove that $\mathbb{P}(f \in T, \mathcal{E}_{A_1, A_2}) \geq \mathbb{P}(f \in T)\mathbb{P}(\mathcal{E}_{A_1, A_2})$, or, equivalently, that

$$\mathbb{P}(f \in T \mid \mathcal{E}_{A_1, A_2}) \geq \mathbb{P}(f \in T).$$

Clearly, $\mathbb{P}(f \in T \mid \mathcal{E}_{A_1, A_2})$ is equal to the probability that $\sum_{v \in S_f} \pi(v) \equiv p(p+1)/2$ modulo p , conditional on π being surjective from $S_e \setminus (A_1 \cup A_2)$ to $\{3, \dots, p+2\}$. By

Lemma 7.11 (applied with $r = p$ and $t = |S_e \setminus (S_f \cup A_1 \cup A_2)|$), there exists some $p_0 = p_0(t)$ such that for all $p \geq p_0$, the conditional probability above is at least $1/p$. Thus, since $\mathbb{P}(f \in T) = 1/p$, we indeed have $\mathbb{P}(f \in T \mid \mathcal{E}_{A_1, A_2}) \geq \mathbb{P}(f \in T)$. Note that p_0 only depends on $|S_e \setminus (S_f \cup A_1 \cup A_2)|$, which is bounded from above by $q + 2$. Therefore, there is $p_0 = p_0(q)$ such that for all $p \geq p_0$, the inequality (27) holds simultaneously for all A_1, A_2 .

The case where $e \in X \cup Y$ is a second representative of an edge $e' \in E(H)$ in the definition of X and Y (i.e. where $e \in R$ if and only if $\{\pi(v) : v \in S_e\}$ is equal to $\{3, \dots, p + 2\}$) can be treated in a similar way.

- (vii) As $\{e \in R\} \subseteq \{e \in T\} = \{f \in T\}$ (since $3 + \dots + (p + 2) \equiv \frac{p(p+1)}{2} \pmod{p}$), we are left to prove that $\mathbb{P}(e \in R)(1 - \mathbb{P}(e \in T)) = \Omega(1)$. Since it is easy to check that $\mathbb{P}(e \in T) = 1 - \Omega(1)$ and $\mathbb{P}(e \in R) = \Omega(1)$, we get the desired property.
- (viii) It is immediate from the definition of Y and the assumptions of the lemma that we have $|e| = 2$ and $|S_e| = p$ for each $e \in Y$.
- (ix) Immediate from the degree condition in the lemma.
- (x) Immediate from the codegree condition in the lemma, using the assumption that edges in $\mathcal{B}$ are subsets of W' .
- (xi) Immediate from the assumptions of the lemma.
- (xii) Immediate from the assumptions of the lemma.
- (xiii) Let u, v be distinct vertices in W . We let $H_{u,v}$ be the p -uniform multihypergraph on vertex set $W' \setminus W$, where we add the hyperedge S_e for every $e \in Y$ with vertex set $\{u, v\}$. Note that $e_g(u, v) = |E(H_{u,v})|$ and that the multiplicity of every hyperedge in $H_{u,v}$ is at most 2λ by the previous item. Note that π induces a uniformly random colouring of the vertex set of $H_{u,v}$ with colours $\{3, 4, \dots, p + 2\}$, and that $B(u, v)$ corresponds to the number of hyperedges in $H_{u,v}$ in which the sum of the colours of the vertices is congruent to $p(p + 1)/2$ modulo p . Ignoring isolated vertices, the number of vertices in $H_{u,v}$ is $O(m)$. Thus, by applying Lemma 7.6 for $H_{u,v}$ with $a = p(p + 1)/2$, $b = p$, $k = p$ and $C_1 = 13$, there exists a constant $C = C(p, b, C_1) > 0$ such that, with probability at least $1 - \Omega(m^{-13})$, we have

$$|B(u, v) - \mathbb{E}[B(u, v)]| \leq C(\log O(m))^{p/2} \lambda^{1/2} e_g(u, v)^{1/2} \leq C(\log m)^{\kappa/2} \lambda^{1/2} e_g(u, v)^{1/2}.$$

Taking a union bound over all distinct $u, v \in W$ (for which there are $O(m^2)$ choices, ignoring isolated vertices in H) gives the desired result.

This finishes the proof. $\square$

We are now ready to prove Theorem 7.10.

Proof of Theorem 7.10. Let $\varepsilon > 0$, let q be a positive integer and let $k - q \leq r \leq k$, where k is sufficiently large as a function of ε and q . Let H be a k -uniform hypergraph with m edges. By Lemma 3.7, H has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$ unless there exists $U \subseteq V(H)$ such that

- $H[U]$ has $\Omega(m)$ edges,
- for every distinct $u, v \in U$, $\deg_H(v) \leq \Delta = \Theta(m^{2/3-\varepsilon})$ and $\deg_H(u, v) \leq D = O(m^{1/3-2\varepsilon})$.

We now let p be the largest prime such that $p \leq r - 3$. Note that $p \geq (r - 3)/2 \geq k/4$ by Bertrand's postulate.

Our strategy for the rest of the proof is relatively simple, although the analysis itself is rather technical. We first take a uniform random assignment of colours $\{1, \dots, r\}$ for $V(H)$ and reveal each colour class i with $i \geq p + 3$. By Lemma 3.8 this reduces the problem to finding a $(p + 2)$ -cut of large surplus (in expectation) in a naturally defined auxiliary hypergraph. In this auxiliary hypergraph, we further reveal the colour class of every vertex either outside of U , or belonging to a colour class $i \geq 3$. By Lemma 3.9 this further reduces the problem to finding a 2-cut of large surplus in another naturally defined auxiliary hypergraph. With a suitable choice of the set $\mathcal{B}$, Lemma 7.12 shows that this hypergraph is a random multihypergraph of general type as

defined in Definition 7.1. Hence, we can use Theorem 7.4 to prove a suitable lower bound on the expected surplus of this auxiliary hypergraph (of course, the desired 2-cut in this hypergraph is not uniformly random, so here we deviate from the random assignment we considered at the beginning of this paragraph). We now proceed with the detailed proof.

We take a random $\eta : V(H) \rightarrow \{p+3, \dots, r, *\}$, where, for each $p+3 \leq i \leq r$, $\eta(v) = i$ with probability $1/r$, and $\eta(v) = *$ with probability $\frac{p+2}{r}$, independently for all $v \in V(H)$. Let H_η be the hypergraph on vertex set $\eta^{-1}(*)$ and edge set $\{e \cap \eta^{-1}(*) : e \in E(H), \forall j \in [p+3, r], |e \cap \eta^{-1}(j)| \geq 1\}$. Then, by Lemma 3.8, we have that

$$\text{sp}_r(H) \geq \mathbb{E}_\eta[\text{sp}_{p+2}(H_\eta)]. \quad (28)$$

Our goal is now to give a suitable lower bound for $\mathbb{E}_\eta[\text{sp}_{p+2}(H_\eta)]$.

We let $V_\eta = \eta^{-1}(*)$, $U_\eta = U \cap V_\eta$ and $U_\eta^c = U^c \cap V_\eta$. Let F be the hypergraph consisting of the edges of $H_\eta[U_\eta]$ of size exactly $p+2$, and note that (as $H[U]$ has $\Omega(m)$ edges) $f = e(F)$ satisfies $\mathbb{E}[f] \geq \gamma m$ for some constant $\gamma = \gamma(k) > 0$. Applying Corollary 7.8 to F , we obtain that there exist a set $B \subseteq E(F)$ of hyperedges such that $|B| = \Omega\left(\frac{f}{\log f}\right)$, positive integers $\lambda, \Lambda \geq 1$ with $\Lambda = O(m^{\frac{1}{\lceil (p+2)/2 \rceil}}) = O(m^{\frac{1}{\lceil k/4 \rceil}})$ and a subset $S(e) \subset e$ of size p for every $e \in B$ such that

- the multiplicity of every edge $e \in B$ is at least $\lambda/2$ and at most λ ,
- for every edge $e \in B$, the number of edges $f \in B$ with $S(e) = S(f)$ is at most $\lambda\Lambda$.

Let $\omega : V_\eta \rightarrow [p+2]$ be a uniform random $(p+2)$ -cut of V_η , and we let $P\omega$ be the function from V_η to $\{*, 1, \dots, p+2\}$ such that

$$P\omega(v) = \begin{cases} \omega(v) & \text{if } v \in U_\eta^c \text{ or } \omega(v) \in \{3, \dots, p+2\}, \\ * & \text{otherwise.} \end{cases}$$

Let $H_{\eta, P\omega}$ be the multihypergraph on vertex set $(P\omega)^{-1}(*)$ and edge set $R := \{e \cap (P\omega)^{-1}(*) : e \in E(H_\eta), \{3, \dots, p+2\} \subseteq P\omega(e), \{1, \dots, p+2\} \not\subseteq P\omega(e)\}$, where two copies of $e \cap (P\omega)^{-1}(*)$ are added to R if $P\omega(e) \cap \{1, \dots, p+2\} = \{3, \dots, p+2\}$.

Then, by Lemma 3.9, we have that

$$\text{sp}_{p+2}(H_\eta) \geq \mathbb{E}_\omega[\text{sp}(H_{\eta, P\omega})]/2 \quad (29)$$

Therefore, our goal now is to prove a suitable lower bound for $\mathbb{E}_{\eta, \omega}[\text{sp}(H_{\eta, P\omega})]$. Let $W = (P\omega)^{-1}(*)$, let $\mathcal{B} = \{e \in B : e \cap W = e \setminus S(e)\}$, and note that $\mathbb{E}_\omega[|\mathcal{B}|] \geq \delta|B|$ for some constant $\delta = \delta(k) > 0$. We now condition on $W = (P\omega)^{-1}(*)$, and let σ have the same distribution as $P\omega$ conditioned on $W = (P\omega)^{-1}(*)$. Note that the distribution of σ is simply that if $v \in W$, then $\sigma(v) = *$, and for every vertex $v \in W^c$, if $v \in U_\eta$, then $\sigma(v)$ is uniform random in $\{3, \dots, p+2\}$, and if $v \in U_\eta^c$, then $\sigma(v)$ is uniform random in $\{1, \dots, p+2\}$, independently for each vertex $v \in W^c$.

Note that every edge of H_η is the representative of an edge of H , from which we have removed at least $r - p - 2$ vertices (vertices mapped to $p+3, \dots, r$ by η), so that every edge of H_η has size at most $k - r + p + 2 \leq p + q + 2$. If $f \geq \gamma m/10$ and $|\mathcal{B}| \geq \delta|B|/10$, then $|\mathcal{B}| = \Omega(m/\log m)$, and therefore by Lemma 7.12 (applied with $H = H_\eta$ and $W' = U_\eta$), conditional on some η with $f \geq \gamma m/10$ and some W with $|\mathcal{B}| \geq \delta|B|/10$, we have that $H_{\eta, \sigma}$ is a random $(m, 2\Delta, 2D, \Lambda, 2\lambda, \kappa)$ -multihypergraph of general type for $\kappa = p + q + 2$. We now apply Theorem 7.4 and get that

$$\mathbb{E}_{\eta, \omega}[\text{sp}(H_{\eta, P\omega}) \mid |\mathcal{B}| \geq \delta|B|/10, f \geq \gamma m/10] = \Omega\left(\frac{m\lambda}{\beta \log m}\right) - O\left(\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4}\right),$$

where $\beta = (\log m)^{(\kappa+1)/2} \sqrt{4\Delta\Lambda}$. Recalling that $\Delta = \Theta(m^{2/3-\varepsilon})$, $D = O(m^{1/3-2\varepsilon})$ and $\Lambda = O(m^{\frac{1}{\lfloor k/4 \rfloor}})$, we get that

$$\begin{aligned} \frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4} &= \frac{m\lambda}{\beta \log m} \left(\frac{\Lambda D \log m}{\beta} + \frac{D^3 \log m}{\beta^3} \right) \\ &= \frac{m\lambda}{\beta \log m} \cdot O\left(\frac{\Lambda D}{\sqrt{\Delta}} + \frac{D^3}{\Delta^{3/2}} \right) \\ &= \frac{m\lambda}{\beta \log m} \cdot O\left(m^{\frac{1}{\lfloor k/4 \rfloor} - 3\varepsilon/2} + m^{-9\varepsilon/2} \right). \end{aligned}$$

For k sufficiently large, we have $m^{\frac{1}{\lfloor k/4 \rfloor} - 3\varepsilon/2} + m^{-9\varepsilon/2} = o(1)$, which shows that

$$\frac{m\Lambda\lambda D}{\beta^2} + \frac{mD^3\lambda}{\beta^4} = o\left(\frac{m\lambda}{\beta \log m} \right),$$

and therefore

$$\mathbb{E}_{\eta, \omega}[\text{sp}(H_{\eta, P\omega}) \mid |\mathcal{B}| \geq \delta|B|/10, f \geq \gamma m/10] = \Omega\left(\frac{m\lambda}{\beta \log m} \right) = \Omega(m^{2/3}).$$

This implies that

$$\mathbb{E}_{\eta, \omega}[\text{sp}(H_{\eta, P\omega})] = \Omega(\mathbb{P}(|\mathcal{B}| \geq \delta|B|/10, f \geq \gamma m/10) \cdot m^{2/3}) = \Omega(m^{2/3}).$$

Using (28) and (29), we obtain $\text{sp}_r(H) = \Omega(m^{2/3})$. $\square$

7.3. The proof of Lemma 7.11. We first recall the following classical estimate on Stirling numbers of the second kind (see for instance [23]).

Lemma 7.13. *Let t be a fixed nonnegative integer. Then, as r tends to infinity we have*

$$S(r+t, r) \sim \frac{r^{2t}}{2^t t!}.$$

We will also need the following lemma. Here and in the rest of this section, we will think of t as a constant, and so the $O(\cdot)$ terms are allowed to hide a dependence on t .

Lemma 7.14. *Fix a nonnegative integer t . Let $f : [r+t] \rightarrow [r]$ be a uniformly random surjective function. Then the probability that there exists some $i \in [r]$ such that $|f^{-1}(i)| \geq 3$ is $O(r^{-1})$.*

Proof. By a union bound it suffices to show that the probability that $|f^{-1}(1)| \geq 3$ is $O(r^{-2})$. By a simple counting argument, and applying Lemma 7.13, we have

$$\mathbb{P}(|f^{-1}(1)| = s) = \frac{\binom{r+t}{s} (r-1)! S(r+t-s, r-1)}{r! S(r+t, r)} = O\left(\frac{r^s r^{2(t-s+1)}}{r r^{2t}} \right) = O(r^{1-s}).$$

Summing over $s \geq 3$ finishes the proof. $\square$

Given a prime r , and integers $b_1, \dots, b_\ell$, we denote by $N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0)$ the number of solutions of the equation $\sum_{i=1}^\ell b_i x_i \equiv 0$ modulo r , where all the x_i for $i \in [\ell]$ are distinct elements of $\{0, 1, \dots, r-1\}$. The next result is a consequence of Theorem 1.5 in [28], and gives us a formula for $N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0)$. Here and below we use $(r)_\ell$ to denote $\ell! \binom{r}{\ell}$.

Theorem 7.15. *Let r be prime, and let $b_1, \dots, b_\ell$ be integers. Let $\zeta(b_1, \dots, b_\ell; \ell, i)$ be the number of permutations in S_ℓ composed of exactly i cycles, such that for every such cycle $\mathcal{C}$, we have $\sum_{j \in \mathcal{C}} b_j \equiv 0$ modulo r . Then*

$$N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0) = \frac{(r)_\ell}{r} + \frac{r-1}{r} \sum_{i=1}^\ell (-1)^{\ell-i} \zeta(b_1, \dots, b_\ell; \ell, i) r^i.$$

We now introduce some definitions. Fix a nonnegative integer t and a (large) prime r for the rest of this section. The *profile* of a surjective function $f : [r+t] \rightarrow [r]$ is the multiset $P_f = \{a_1, \dots, a_r\}$, where $a_i = |f^{-1}(i) \cap [r]| - 1$ for every $i \in [r]$. We say that a multiset P is a profile if there exists a surjective function $f : [r+t] \rightarrow [r]$ such that $P = P_f$. For a profile $P = \{a_1, \dots, a_r\}$, we let $p(P)$ be the probability that $\sum_{i=1}^r a_i \lambda_i = 0$ in $\mathbb{F}_r$ for $\lambda = (\lambda_1, \dots, \lambda_r)$ a uniform random permutation of $(0, 1, \dots, r-1)$.

The reason we are interested in $p(P)$ is the following simple observation.

Lemma 7.16. *For any profile P , $p(P)$ is the probability that $\sum_{k=1}^r f(k)$ is congruent to $r(r+1)/2$ modulo r for a uniformly random surjective function $f : [r+t] \rightarrow [r]$, conditionally on $P_f = P$.*

Proof. Let $P = \{b_1, \dots, b_r\}$. Conditional on $P_f = P$, the sum $\sum_{k=1}^r f(k) = \sum_{i=1}^r |f^{-1}(i) \cap [r]| \cdot i$ is distributed as $\sum_{i=1}^r (b_i + 1) \lambda_i$, where $(\lambda_1, \dots, \lambda_r)$ is a uniform random permutation of $(0, 1, \dots, r-1)$, since the tuple $(|f^{-1}(1) \cap [r]| - 1, \dots, |f^{-1}(r) \cap [r]| - 1)$ is a uniform random permutation of $(b_1, \dots, b_r)$. But $\sum_{i=1}^r (b_i + 1) \lambda_i \equiv r(r+1)/2 + \sum_{i=1}^r b_i \lambda_i$ modulo r , from which the result follows. $\square$

We will estimate $p(P)$ using the following result.

Lemma 7.17. *Let $P = \{b_1, \dots, b_r\}$ be a profile, and let $1 \leq \ell \leq r$ be such that $b_i = 0$ for all $i > \ell$. Then we have*

$$p(P) = \frac{1}{r} + \frac{r-1}{r \cdot (r)_\ell} \sum_{i=1}^{\ell} (-1)^{\ell-i} \zeta(b_1, \dots, b_\ell; \ell, i) r^i.$$

Proof. Take a uniform random permutation $\lambda = (\lambda_1, \dots, \lambda_r)$ of $(0, 1, \dots, r-1)$. Then $p(P)$ is the probability that $\sum_{i=1}^{\ell} b_i \lambda_i \equiv 0$ modulo r . The number of ordered tuples $(x_1, \dots, x_\ell)$ consisting of distinct elements of $\{0, 1, \dots, r-1\}$ such that $\sum_{i=1}^{\ell} b_i x_i \equiv 0$ modulo r is equal to $N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0)$ by definition. The number of ways to extend such a fixed tuple of distinct elements $(x_1, \dots, x_\ell)$ into a permutation $(x_1, \dots, x_r)$ of $(0, 1, \dots, r-1)$ is $(r-\ell)!$. Thus we have that

$$p(P) = \frac{N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0)(r-\ell)!}{r!} = \frac{N_{\mathbb{F}_r}(b_1, \dots, b_\ell; 0)}{(r)_\ell}.$$

Applying Theorem 7.15, we get

$$p(P) = \frac{1}{r} + \frac{r-1}{r \cdot (r)_\ell} \sum_{i=1}^{\ell} (-1)^{\ell-i} \zeta(b_1, \dots, b_\ell; \ell, i) r^i,$$

as desired. $\square$

For $0 \leq j \leq t$, we let P^j be the profile that has exactly j elements equal to 1 and j elements equal to -1 , and the rest of the elements are 0. Our next result gives an asymptotic formula for $p(P^j)$.

Lemma 7.18. *For every $0 \leq j \leq t$, we have*

$$p(P^j) - \frac{1}{r} = j!(-1)^j r^{-j} + O(r^{-j-1}).$$

Proof. We have $p(P^0) = 1$, so the statement is true for $j = 0$. Assume now that $j \geq 1$. We apply Lemma 7.17 with $\ell = 2j$ and get that

$$p(P^j) - \frac{1}{r} = \frac{r-1}{r \cdot (r)_{2j}} \sum_{i=1}^{2j} (-1)^{2j-i} \zeta(b_1, \dots, b_{2j}; 2j, i) r^i,$$

where $b_i = 1$ for $i \in [j]$ and $b_i = -1$ for $i \in [j+1, 2j]$. Now note that $\zeta(b_1, \dots, b_{2j}; 2j, i) = 0$ for $i \geq j+1$ (as otherwise one of the cycles would consist of exactly one non-zero element), and

that $\zeta(b_1, \dots, b_{2j}; 2j, j) = j!$ (since each cycle in the permutation must consist of exactly one element b_i for $i \in [j]$ and exactly one element b_i for $i \in [j+1, 2j]$). Therefore, we obtain

$$p(P^j) - \frac{1}{r} = j!(-1)^j r^{-j} + O(r^{-j-1}),$$

as wanted. $\square$

We now prove an upper bound on $|p(P) - \frac{1}{r}|$ for any profile P .

Lemma 7.19. *For any profile P with j negative elements, we have*

$$\left| p(P) - \frac{1}{r} \right| = O(r^{-j}).$$

Proof. Let $P = \{b_1, \dots, b_r\}$, where $b_i < 0$ for $i \in [j]$, $b_i > 0$ for $i \in [j+1, j+y]$ and $b_i = 0$ for $i > j+y$. By Lemma 7.17 with $\ell = j+y$, we have

$$\left| p(P) - \frac{1}{r} \right| = O(r^{x-(j+y)}),$$

where x is the maximal number of sets that partition $\{b_1, b_2, \dots, b_{j+y}\}$, such that for every such set $\mathcal{C}$ in the partition, we have $\sum_{b_i \in \mathcal{C}} b_i \equiv 0$ modulo r . Since r is sufficiently large compared to t , we must have $\sum_{b_i \in \mathcal{C}} b_i = 0$ for all such $\mathcal{C}$. Hence, every such $\mathcal{C}$ in the partition must contain at least one positive element, and therefore we have $x \leq y$, which implies that

$$\left| p(P) - \frac{1}{r} \right| = O(r^{-j}),$$

as wanted. $\square$

We slightly abuse notation and for a function $g : [r] \rightarrow [r]$, we will say that the profile of g is the multiset $\{a_1, \dots, a_r\}$ where $a_i = |g^{-1}(i)| - 1$ for every $i \in [r]$. For two nonnegative integers m, n , we let $\text{Surj}_{m,n}$ be the number of surjections from $[m]$ to $[n]$. We now prove the following result.

Lemma 7.20. *Fix $0 \leq j \leq t$. The probability that a uniform random surjection $f : [r+t] \rightarrow [r]$ has profile P^j is*

$$(1 + o(1)) \binom{t}{j} 2^{t-j} \frac{t!}{j!} r^{j-t}$$

as $r \rightarrow \infty$.

Proof. We first count the number of functions $g : [r] \rightarrow [r]$ with profile P^j . We claim that this is $\binom{r}{j} \text{Surj}_{r, r-j}(1 - O(r^{-1}))$. Indeed we first choose $Z := g([r])$ to be a subset of $[r]$ of size $r-j$, for which we have $\binom{r}{j}$ possibilities. Then we choose the map $g' : [r] \rightarrow Z$ such that g' is surjective and there is no $i \in Z$ with $|g'^{-1}(i)| \geq 3$. By Lemma 7.14, the number of possibilities is $\text{Surj}_{r, r-j}(1 - O(r^{-1}))$.

We now count the number of ways we can extend such a function g to a surjective function $f : [r+t] \rightarrow [r]$. There are exactly j elements in $[r]$ which do not have an image through g . Let $X = \{x_1, \dots, x_j\}$ be the set of such elements. Therefore, we are counting the number of functions $h : [t] \rightarrow [r]$ such that $X \subseteq h([t])$. We first choose the size of $h^{-1}(X)$, which we denote by ℓ . Therefore we have that the number of such functions h is

$$\sum_{\ell=j}^t \binom{t}{\ell} \text{Surj}_{\ell, j}(r-j)^{t-\ell} = \binom{t}{j} j! r^{t-j} + O(r^{t-j-1}),$$

as we have $\binom{t}{\ell}$ choices for $H := h^{-1}(X)$, $\text{Surj}_{\ell, j}$ choices for the map $h|_H : H \rightarrow X$, and $r-j$ choices for $h(a)$ for each $a \notin H$.

Putting the above two paragraphs together, we find that the number of surjections $f : [r+t] \rightarrow [r]$ that have profile P^j is

$$\binom{r}{j} \text{Surj}_{r,r-j} \binom{t}{j} j! r^{t-j} (1 + o(1)).$$

Dividing by the number of surjections from $[r+t]$ to $[r]$, we obtain by Lemma 7.13 that the desired probability is

$$\begin{aligned} (1 + o(1)) \frac{\binom{r}{j} \text{Surj}_{r,r-j} \binom{t}{j} j! r^{t-j}}{\text{Surj}_{r+t,r}} &= (1 + o(1)) \frac{\binom{r}{j} (r-j)! S(r, r-j) \binom{t}{j} j! r^{t-j}}{r! S(r+t, r)} \\ &= (1 + o(1)) \frac{\frac{r^{2j}}{2^j j!} \binom{t}{j} r^{t-j}}{\frac{r^{2t}}{2^t t!}} \\ &= (1 + o(1)) \binom{t}{j} 2^{t-j} \frac{t!}{j!} r^{j-t}, \end{aligned}$$

as desired. $\square$

We say that a profile is *good* if it contains exactly j elements 1 and j elements -1 , and the rest of the elements are 0, for some integer j , and *bad* otherwise. In other words, the only profiles which are not bad are the profiles P_j for $0 \leq j \leq t$. The following lemma shows that profiles with exactly j negative elements that are not P_j are unlikely to occur.

Lemma 7.21. *Fix $0 \leq j \leq t$, and let P be a bad profile with exactly j negative elements. The probability that a uniform random surjection $f : [r+t] \rightarrow [r]$ has profile P is $O(r^{j-t-1})$.*

Proof. The proof is similar to the proof of Lemma 7.20. We first claim that there are at most $O(\binom{r}{j} \text{Surj}_{r,r-j} r^{-1})$ functions $g : [r] \rightarrow [r]$ with profile P . Indeed we first choose $Z := g([r])$ to be a subset of $[r]$ of size $r-j$, for which we have $\binom{r}{j}$ possibilities. Then choose the map $g' : [r] \rightarrow Z$, and note that g' is surjective and there is some $i \in Z$ such that $|g'^{-1}(i)| \geq 3$ because P is a bad profile. By Lemma 7.14, the number of possibilities is $O(\text{Surj}_{r,r-j} r^{-1})$, thus giving in total $O(\binom{r}{j} \text{Surj}_{r,r-j} r^{-1})$ choices for the function $g : [r] \rightarrow [r]$.

Then the second part of the proof is completely identical to the one of Lemma 7.20, giving the desired bound. $\square$

We are now ready to prove Lemma 7.11, which we recall for the reader's convenience.

Lemma 7.11. *For every nonnegative integer t , there exists $r_0 \in \mathbb{N}$ such that for every prime $r \geq r_0$, the following holds. We let $f : [r+t] \rightarrow [r]$ be a uniformly random surjective function. Then the probability that $\sum_{i=1}^r f(i)$ is congruent to $r(r+1)/2$ modulo r is greater than $1/r$.*

Proof. By Lemma 7.16, the probability of interest can be written as

$$\sum_P \mathbb{P}(P_f = P) p(P),$$

where f is a uniform random surjection $f : [r+t] \rightarrow [r]$, and the sum is over every possible profile P . Therefore it is equivalent to show that

$$\sum_P \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right) > 0.$$

We now split the sum of interest into two:

$$\sum_P \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right) = \sum_{P \text{ good}} \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right) + \sum_{P \text{ bad}} \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right). \quad (30)$$

We will show that the first term gives a main positive contribution, while the second term is of lower order. By Lemmas 7.18 and 7.20, we have

$$\begin{aligned}
\sum_{P \text{ good}} \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right) &= \sum_{j=0}^t \mathbb{P}(P_f = P^j) \left(p(P^j) - \frac{1}{r} \right) \\
&= \sum_{j=0}^t (1 + o(1)) \binom{t}{j} 2^{t-j} \frac{t!}{j!} r^{j-t} \cdot j! (-1)^j r^{-j} \\
&= (1 + o(1)) t! r^{-t} \sum_{j=0}^t \binom{t}{j} 2^{t-j} (-1)^j \\
&= (1 + o(1)) t! r^{-t}.
\end{aligned} \tag{31}$$

We now estimate the second term. We let P_{bad}^j be the set of bad profiles with exactly j negative elements. By Lemmas 7.19 and 7.21, we have

$$\begin{aligned}
\sum_{P \text{ bad}} \mathbb{P}(P_f = P) \left| p(P) - \frac{1}{r} \right| &= \sum_{j=0}^t \sum_{P \in P_{bad}^j} \mathbb{P}(P_f = P) \left| p(P) - \frac{1}{r} \right| \\
&= O \left(\sum_{j=0}^t \sum_{P \in P_{bad}^j} r^{j-t-1} \cdot r^{-j} \right)
\end{aligned}$$

Since the number of profiles is bounded as a function of t (a profile has at most $2t$ non-zero elements, each of which is at most t in absolute value), we obtain

$$\sum_{P \text{ bad}} \mathbb{P}(P_f = P) \left| p(P) - \frac{1}{r} \right| = O(r^{-t-1}). \tag{32}$$

Plugging (31) and (32) into (30), we get

$$\sum_P \mathbb{P}(P_f = P) \left(p(P) - \frac{1}{r} \right) = (1 + o(1)) t! r^{-t} > 0,$$

which finishes the proof. $\square$

8. A NEW CONSTRUCTION FOR 2-CUT IN 4-GRAPHS

In this short section we prove Proposition 1.7. We recall the statement for the reader's convenience.

Proposition 1.7. *Let $s \geq 4$ and let $S_1, \dots, S_t$ be the blocks of an $(n, s, 2)$ -Steiner system on ground set V . Let G be the 4-uniform hypergraph on vertex set V obtained by placing a complete 4-graph in each S_i . Then $e(G) = \Theta(n^2 s^2)$ and $\text{sp}(G) = O(n^2 + ns^2)$. In particular, when $s = \Theta(\sqrt{n})$ (which holds for projective planes), we have $\text{sp}(G) = O(e(G)^{2/3})$.*

Proof. Clearly, $e(G) = \Theta(ts^4)$. Since $t = \Theta(n^2/s^2)$, we get $e(G) = \Theta(n^2 s^2)$.

We now turn to upper bounding $\text{sp}(G)$. Consider a cut (A, B) of V . There exist $x_1, \dots, x_t$ such that $|A \cap S_i| = s/2 + x_i$ and $|B \cap S_i| = s/2 - x_i$.

Now the surplus of this cut is

$$\begin{aligned}
\text{sp}_G(A, B) &= \sum_{i=1}^t \left(\frac{1}{8} \binom{s}{4} - \binom{s/2 + x_i}{4} - \binom{s/2 - x_i}{4} \right) \\
&= \sum_{i=1}^t \left(\frac{1}{32} s^3 - \frac{1}{8} s^2 x_i^2 - \frac{11}{64} s^2 + \frac{3}{4} s x_i^2 + \frac{7}{32} s - \frac{1}{12} x_i^4 - \frac{11}{12} x_i^2 \right) \\
&\leq \sum_{i=1}^t \left(\frac{1}{32} s^3 - \frac{11}{64} s^2 + \frac{7}{32} s - \left(\frac{1}{8} s^2 - \frac{3}{4} s + \frac{11}{12} \right) x_i^2 \right). \tag{33}
\end{aligned}$$

On the other hand, in S_i , the difference between the number of pairs of vertices cut by (A, B) and the number of pairs of vertices not cut by (A, B) is

$$\binom{s}{2} - 2 \left(\binom{s/2 + x_i}{2} + \binom{s/2 - x_i}{2} \right) = s/2 - 2x_i^2.$$

In V , the number of pairs cut by (A, B) is at most an $O(n)$ additive term greater than the number of pairs not cut by (A, B) . Since each pair of vertices in V belongs to precisely one block S_i , we obtain

$$\sum_{i=1}^t (s/2 - 2x_i^2) \leq O(n),$$

and hence

$$\sum_{i=1}^t -x_i^2 \leq O(n) - \sum_{i=1}^t s/4.$$

Plugging this into (33) and using $\frac{1}{8}s^2 - \frac{3}{4}s + \frac{11}{12} \geq 0$ (which holds for $s \geq 5$; note that the proposition is trivial for $s = 4$), we get

$$\text{sp}_G(A, B) \leq O(s^2 n) + \sum_{i=1}^t \left(\frac{1}{32} s^3 - \frac{11}{64} s^2 + \frac{7}{32} s - \left(\frac{1}{8} s^2 - \frac{3}{4} s + \frac{11}{12} \right) \frac{s}{4} \right) \leq O(s^2 n) + O(ts^2).$$

Using $t = \Theta(n^2/s^2)$, we conclude that $\text{sp}_G(A, B) \leq O(s^2 n) + O(n^2)$, as desired. $\square$

9. CONCLUDING REMARKS

MaxCut for hypergraphs. In this paper we proved that a k -uniform hypergraph with m edges has an r -cut of surplus $\Omega(m^{2/3-\varepsilon_k})$, where $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$. The way our proof is written⁴ gives an ε_k that depends polynomially on k . On the other hand, we believe that our methods can be pushed further to give an ε_k that is exponentially small in k . We omit the details.

Our method can also be used (with suitable modifications) to give polynomially improved lower bounds for all values of (r, k) with the sole exception of $(r, k) = (3, 3)$ (where we can match the best known bound). For $r = k$, our proof gives the asymptotically tight $m^{2/3}/\text{polylog}(m)$ lower bound under the very mild assumption on the hypergraph that no two hyperedges intersect in at least $k - 1$ vertices. The same holds for $r = 2$ if k is even. When $r = 2$ and k is odd, we get the asymptotically tight $m^{2/3}/\text{polylog}(m)$ lower bound under the assumption that no two hyperedges intersect in at least $k - 2$ vertices.

⁴Our proof gives a linear dependence in the most important cases $r = 2$ and $r = k$. In the general case, to get a polynomial bound, one can quantify the dependence of the implicit constant on t in Section 7.3.

Bisection width and positive discrepancy. An equipartition of a finite set is a partition into two parts, whose sizes differ by at most one. A *bisection* of a hypergraph is an equipartition of its vertex set, together with all edges that intersect each part in at least one vertex. The *bisection width* of a hypergraph G , denoted as $\text{bw}(G)$, is the smallest number of edges in a bisection of G . In other words, $\text{bw}(G)$ is the smallest 2-cut of G over all balanced bipartitions.

This parameter has been extensively studied in both Theoretical Computer Science (see, e.g. [17]) and Extremal and Probabilistic Combinatorics. A celebrated result of Alon [2] states that if G is an n -vertex d -regular graph with $d = O(n^{1/9})$, then $\text{bw}(G) \leq e(G)(\frac{1}{2} - \Omega(\frac{1}{\sqrt{d}}))$. While the bound here is tight (e.g. for random d -regular graphs, as proven by Bollobás [8]), the assumption on d being small is not sharp. Indeed, recently the same bound $\text{bw}(G) \leq e(G)(\frac{1}{2} - \Omega(\frac{1}{\sqrt{d}}))$ was established by Raty, Sudakov and Tomon [32] under the weaker assumption $d = O(n^{2/3})$. They also showed that this bound no longer holds for $d = \omega(n^{2/3})$.

Recently, Raty and Tomon [33] obtained a hypergraph analogue of Alon’s result by showing that if G is an n -vertex d -regular k -uniform hypergraph with $d = o(n^{1/2})$, then $\text{bw}(G) \leq e(G)(1 - \frac{1}{2^{k-1}} - \Omega(\frac{1}{\sqrt{d}}))$. While this bound is tight for the range of parameters the result applies to, the assumption that d is small is not optimal. In upcoming work [25], we adapt the methods we developed in this paper to prove the above (tight) upper bound for the bisection width of hypergraphs under a nearly optimal condition on d . Furthermore, for linear hypergraphs, we obtain the optimal condition without any error. Our results reveal a difference between the graph and the hypergraph case that was not demonstrated in previous works.

Finally, we obtain analogous improvements for the closely related notion of positive discrepancy. For details about this hypergraph parameter and its interesting connections to eigenvalues of the adjacency tensor, we refer the reader to [33] and to our upcoming paper [25].

REFERENCES

- [1] N. Alon. Bipartite subgraphs. *Combinatorica*, 16(3):301–311, 1996.
- [2] N. Alon. On the edge-expansion of graphs. *Combinatorics, Probability and Computing*, 6(2):145–152, 1997.
- [3] N. Alon, B. Bollobás, M. Krivelevich, and B. Sudakov. Maximum cuts and judicious partitions in graphs without short cycles. *Journal of Combinatorial Theory, Series B*, 88(2):329–346, 2003.
- [4] N. Alon, M. Krivelevich, and B. Sudakov. MaxCut in H -free graphs. *Combinatorics, Probability and Computing*, 14(5-6):629–647, 2005.
- [5] N. Alon, K. Makarychev, Y. Makarychev, and A. Naor. Quadratic forms on graphs. *Inventiones mathematicae*, 163(3):499–522, 2006.
- [6] I. Balla, L. Hambardzumyan, and I. Tomon. Factorization norms and an inverse theorem for MaxCut. *arXiv preprint arXiv:2506.23989*, 2025.
- [7] I. Balla, O. Janzer, and B. Sudakov. On MaxCut and the Lovasz theta function. *Proceedings of the American Mathematical Society*, 152(05):1871–1879, 2024.
- [8] B. Bollobás. The isoperimetric number of random regular graphs. *European Journal of Combinatorics*, 9(3):241–244, 1988.
- [9] B. Bollobás and A. D. Scott. Discrepancy in graphs and hypergraphs. In *More Sets, Graphs and Numbers: A Salute to Vera Sos and Andras Hajnal*, pages 33–56. Springer, 2006.
- [10] C. Carlson, A. Kolla, R. Li, N. Mani, B. Sudakov, and L. Trevisan. Lower bounds for Max-Cut in H -free graphs via semidefinite programming. *SIAM Journal on Discrete Mathematics*, 35(3):1557–1568, 2021.
- [11] M. Charikar and A. Wirth. Maximizing quadratic programs: Extending Grothendieck’s inequality. In *45th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 54–60. IEEE, 2004.
- [12] D. Conlon, J. Fox, M. Kwan, and B. Sudakov. Hypergraph cuts above the average. *Israel Journal of Mathematics*, 233:67–111, 2019.

- [13] C. S. Edwards. Some extremal properties of bipartite subgraphs. *Canadian Journal of Mathematics*, 25(3):475–485, 1973.
- [14] C. S. Edwards. An improved lower bound for the number of edges in a largest bipartite subgraph. In *Proc. 2nd Czechoslovak Symposium on Graph Theory, Prague*, pages 167–181, 1975.
- [15] P. Erdős. Problems and results in graph theory and combinatorial analysis. *Proc. British Combinatorial Conf., 5th*, pages 169–192, 1975.
- [16] P. Erdős and D. J. Kleitman. On coloring graphs to maximize the proportion of multicolored k -edges. *Journal of Combinatorial Theory*, 5(2):164–169, 1968.
- [17] U. Feige and R. Krauthgamer. A polylogarithmic approximation of the minimum bisection. *SIAM Journal on Computing*, 31(4):1090–1118, 2002.
- [18] G. B. Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 1999.
- [19] S. Glock, O. Janzer, and B. Sudakov. New results for MaxCut in H -free graphs. *Journal of the London Mathematical Society*, 108(2):441–481, 2023.
- [20] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.
- [21] V. Guruswami. Inapproximability results for set splitting and satisfiability problems with no mixed clauses. *Algorithmica*, 38(3):451–469, 2004.
- [22] J. Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001.
- [23] L. C. Hsu. Note on an asymptotic expansion of the n th difference of zero. *The Annals of Mathematical Statistics*, 19(2):273–277, 1948.
- [24] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. John Wiley & Sons, 2011.
- [25] O. Janzer and J. Portier. Improved bounds for the bisection width and positive discrepancy of hypergraphs. In preparation.
- [26] Z. Jin, A. Milojević, I. Tomon, and S. Zhang. From small eigenvalues to large cuts, and Chowla’s cosine problem. *arXiv preprint arXiv:2509.03490*, 2025.
- [27] R. M. Karp. Reducibility among combinatorial problems. *Complexity of Computer Computations*, pages 85–103, 1972. (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972).
- [28] J. Li and X. Yu. Distinct coordinate solutions of linear equations over finite fields. *Finite Fields and Their Applications*, 61:101602, 2020.
- [29] L. Lovász. Coverings and colorings of hypergraphs. In *Proc. 4th Southeastern Conference of Combinatorics, Graph Theory, and Computing*, pages 3–12. Utilitas Mathematica Publishing, 1973.
- [30] A. Megretski. Relaxations of quadratic programs in operator theory and system analysis. In *Systems, Approximation, Singular Integral Operators, and Related Topics: International Workshop on Operator Theory and Applications, IWOTA 2000*, pages 365–392. Springer, 2001.
- [31] M. E. Muller. A note on a method for generating points uniformly on n -dimensional spheres. *Communications of the ACM*, 2(4):19–20, 1959.
- [32] E. Rätty, B. Sudakov, and I. Tomon. Positive discrepancy, MaxCut, and eigenvalues of graphs. *Transactions of the AMS*, to appear.
- [33] E. Rätty and I. Tomon. Bisection width, discrepancy, and eigenvalues of hypergraphs. *arXiv preprint arXiv:2409.15140*, 2024.
- [34] E. Rätty and I. Tomon. Large cuts in hypergraphs via energy. In *Mathematical Proceedings of the Cambridge Philosophical Society*, pages 1–17. Cambridge University Press, 2025.
- [35] A. D. Scott. Judicious partitions and related problems. In *Surveys in Combinatorics 2005*, pages 95–117. London Math. Soc. Lecture Note Ser., 327, Cambridge Univ. Press, Cambridge, 2005.

APPENDIX A. PROOF OF LEMMA 3.8

Proof of Lemma 3.8. Let $\omega : V \rightarrow \{1, \dots, r\}$ be a uniform random r -cut of V . We define the “information restriction” operator P such that for any r -cut $\omega : V \rightarrow \{1, \dots, r\}$, $P\omega$ is the function from V to $\{*, q+1, \dots, r\}$ such that

$$P\omega(v) = \begin{cases} \omega(v) & \text{if } \omega(v) \in \{q+1, \dots, r\}, \\ * & \text{otherwise.} \end{cases}$$

Then it is clear that $P\omega$ has the same distribution as σ , and that a q -cut of H_σ corresponds to an r -cut ω of H of the same size, with $P\omega = \sigma$. Let Z be the random variable associated to the size of the cut ω in H . It is straightforward to see that $\mathbb{E}[Z|P\omega]$ is the average size of a uniform q -cut of $H_{P\omega}$. Therefore, by definition, $H_{P\omega}$ has a q -cut of size $\mathbb{E}[Z|P\omega] + \text{sp}_q(H_{P\omega})$ and the result follows by taking expectation over ω . $\square$

APPENDIX B. PROOF OF LEMMA 3.9

Proof of Lemma 3.9. Let W be a subset of U and consider any $\sigma : W^c \rightarrow \{1, \dots, r\}$ such that $\sigma(U \cap W^c) \subseteq \{3, \dots, r\}$. We treat W and σ as fixed for the moment, though ultimately we will let $W = (P\omega)^{-1}(*)$ and let σ be the restriction of $P\omega$ to W^c . With a slight abuse of notation, we let H_σ be the multihypergraph on vertex set W and edge set $\{e \cap W : e \in E, \{3, \dots, r\} \subseteq \sigma(e \setminus W), \{1, \dots, r\} \not\subseteq \sigma(e \setminus W)\}$, where two copies of $e \cap W$ are added to H_σ if $\sigma(e \setminus W) \cap \{1, \dots, r\} = \{3, \dots, r\}$. We let $\tilde{H}_\sigma \subseteq H$ consist of the set of edges e of H such that $|\sigma(e \setminus W) \cap \{1, 2\}| = 1$, $\{3, \dots, r\} \subseteq \sigma(e \setminus W)$ and $|e \cap W| > 0$. Let H_σ^* be the multihypergraph on vertex set W with edge set $\{e \cap W : e \in E, \sigma(e \setminus W) \cap \{1, \dots, r\} = \{3, \dots, r\}\}$.

Let $\phi : W \rightarrow \{1, 2\}$. Let z_1 be the size of ϕ as a 2-cut of H_σ^* , and let z_2 be the number of edges in $\tilde{H}_\sigma$ that have both a vertex mapped to 1 and a vertex mapped to 2 by ϕ . Then $z = 2z_1 + z_2$ is the size of ϕ as a 2-cut of H_σ since every edge $e \in H_\sigma^*$ was included in H_σ with multiplicity two. Let N_σ^{cut} be the number of edges $e \in E(H)$ such that $|\sigma(e \setminus W)| = r$, i.e. the number of edges of H that can already be seen as cut by σ alone. Define $\psi_{\text{swap}}(\phi)$ to be the mapping $\psi_{\text{swap}}(\phi) : W \rightarrow \{1, 2\}$ satisfying $\psi_{\text{swap}}(\phi)(a) = 3 - \phi(a)$ for every vertex $a \in W$. Let $\Phi : V \rightarrow \{1, \dots, r\}$ be the r -cut equal to ϕ on W and to σ on W^c , and let $\Phi' : V \rightarrow \{1, \dots, r\}$ be the r -cut equal to $\psi_{\text{swap}}(\phi)$ on W and to σ on W^c . Note that for every edge $e \in E(\tilde{H}_\sigma)$ whose corresponding edge in H_σ is not already cut by ϕ , the edge e is cut by exactly one of Φ and Φ' . Therefore, the average size of Φ and Φ' as r -cuts of H is

$$z_1 + z_2 + (e(\tilde{H}_\sigma) - z_2)/2 + N_\sigma^{\text{cut}} = z/2 + e(\tilde{H}_\sigma)/2 + N_\sigma^{\text{cut}}. \quad (34)$$

(Indeed, the total contribution of the edges $e \in E(H)$, $\{3, \dots, r\} \subseteq \sigma(e \setminus W)$ with $|\sigma(e \setminus W) \cap \{1, 2\}| = 0$ is z_1 , the total contribution of the edges with $|\sigma(e \setminus W) \cap \{1, 2\}| = 1$ is $z_2 + (e(\tilde{H}_\sigma) - z_2)/2$, and the total contribution of edges with $|\sigma(e \setminus W) \cap \{1, 2\}| = 2$ is N_σ^{cut} .)

Now let $\omega : V \rightarrow \{1, \dots, r\}$ be a uniformly random r -cut, as in the statement of the lemma. Let Z be the random variable associated to the size of ω as an r -cut of H , and let Z^{part} be the random variable associated to the size of ω restricted to $W := (P\omega)^{-1}(*)$ as a 2-cut of $H_{P\omega}$. Let $P\omega_{\text{rest}}$ be the restriction of $P\omega$ to W^c . It follows from (34) by taking conditional expectation on the outcome $P\omega_{\text{rest}} = \sigma$ that

$$\mathbb{E}_\omega[Z|P\omega_{\text{rest}} = \sigma] = \mathbb{E}_\omega[Z^{\text{part}}|P\omega_{\text{rest}} = \sigma]/2 + e(\tilde{H}_\sigma)/2 + N_\sigma^{\text{cut}}. \quad (35)$$

It is clear that, conditional on $P\omega_{\text{rest}} = \sigma$, the restriction of ω to W is a uniformly random 2-cut of H_σ , and thus $\mathbb{E}_\omega[Z^{\text{part}}|P\omega_{\text{rest}} = \sigma]$ is the average size of a 2-cut of H_σ . Hence, H_σ has a 2-cut of size $\text{sp}(H_\sigma) + \mathbb{E}_\omega[Z^{\text{part}}|P\omega_{\text{rest}} = \sigma]$, and so it follows by (34) that H has an r -cut with surplus

$$(\text{sp}(H_\sigma) + \mathbb{E}_\omega[Z^{\text{part}}|P\omega_{\text{rest}} = \sigma])/2 + e(\tilde{H}_\sigma)/2 + N_\sigma^{\text{cut}} - \mathbb{E}_\omega[Z],$$

which, by (35), is equal to

$$\text{sp}(H_\sigma)/2 + \mathbb{E}_\omega[Z|P\omega_{rest} = \sigma] - \mathbb{E}_\omega[Z].$$

Thus, H has an r -cut with surplus

$$\text{sp}(H_{P\omega})/2 + \mathbb{E}_\omega[Z|P\omega] - \mathbb{E}_\omega[Z].$$

Taking expectation gives the desired result. $\square$

APPENDIX C. PROOF OF LEMMA 3.11

Proof of Lemma 3.11. Let $H = (V, E)$ be an m -edge k -multigraph and let $2 \leq r \leq k - q$. By applying Lemma 3.7, H has an r -cut of surplus $\Omega(m^{2/3-\varepsilon})$ (and so we are done) unless there exists a vertex subset $U \subseteq V$ such that

- $H[U]$ has $\Omega(m)$ edges,
- for every distinct $u, v \in U$, $\deg_H(v) \leq m^{2/3-\varepsilon}$ and $\deg_H(u, v) \leq m^{1/3-2\varepsilon}$.

Let $\omega : V \rightarrow [r]$ be a uniform random r -cut of V , and let $P\omega$ be the function from V to $\{*, 1, \dots, r\}$ such that

$$P\omega(v) = \begin{cases} \omega(v) & \text{if } v \in U^c \text{ or } \omega(v) \in \{3, \dots, r\}, \\ * & \text{otherwise.} \end{cases}$$

Let $H_{P\omega}$ be the graph on vertex set $(P\omega)^{-1}(*)$ and edge set $\{e \cap (P\omega)^{-1}(*) : e \in E, \{3, \dots, r\} \subseteq P\omega(e), \{1, \dots, r\} \not\subseteq P\omega(e)\}$, where an edge e is added twice if $P\omega(e) \cap \{1, \dots, r\} = \{3, \dots, r\}$. By Lemma 3.9, we have

$$\text{sp}_r(H) = \Omega(\mathbb{E}_\omega[\text{sp}(H_{P\omega})]). \quad (36)$$

It is clear by the choice of U that $H_{P\omega}$ has maximum degree $O(m^{2/3-\varepsilon})$, maximum codegree $O(m^{1/3-2\varepsilon})$ and $O(m)$ hyperedges in total. For every edge $e \in E(H[U])$, we have that the probability that $|e \cap (P\omega)^{-1}(*)| = q + 2$ and $P\omega(e) \cap \{1, \dots, r\} = \{3, \dots, r\}$ is $\Omega(1)$. Indeed, letting $v_1, \dots, v_k$ be the vertices that belong to e , the above event occurs if $\omega(v_i) = i + 2$ for $1 \leq i \leq r - 2$, $\omega(v_i) = *$ for $r - 1 \leq i \leq r + q$, and $\omega(v_i) = r$ if $r + q + 1 \leq i \leq k$. Therefore, letting $e_{q+2}(H_{P\omega})$ be the number of edges of $H_{P\omega}$ of size $q + 2$, we have $\mathbb{E}_\omega[e_{q+2}(H_{P\omega})] = \Omega(m)$. As $e_{q+2}(H_{P\omega}) = O(m)$ with probability 1, there exists a constant $\delta = \delta(k) > 0$ such that $\mathbb{P}(e_{q+2}(H_{P\omega}) \geq \delta m) \geq \delta$. By assumption of the lemma, we have

$$\text{sp}(H_{P\omega}) \geq \Omega(m^{2/3-\varepsilon}) \mathbf{1}_{e_{q+2}(H_{P\omega}) \geq \delta m},$$

and by taking expectation over ω and plugging this in (36), we obtain

$$\text{sp}_r(H) = \Omega(\mathbb{E}_\omega[\text{sp}(H_{P\omega})]) = \Omega(m^{2/3-\varepsilon} \mathbb{P}(e_{q+2}(H_{P\omega}) \geq \delta m)) = \Omega(m^{2/3-\varepsilon}),$$

as desired. $\square$

APPENDIX D. PROOF OF LEMMA 5.5

Proof of Lemma 5.5. Let G be the underlying multi-graph of H . For every hyperedge $\{u, v, w\}$ in $E(H)$, we colour the copy of $\{u, v\} \in E(G)$ which comes from the hyperedge $\{u, v, w\}$ with colour w . Let $S_0 \subseteq V(H)$ be such that $H[S_0]$ has m edges, and each vertex in S_0 has degree at most Δ in H . If B is a random subset of S_0 containing each vertex independently with probability $1/3$, then the expected number of edges in $H[S_0]$ with exactly one vertex in B is $4m/9$. Therefore there exists a set $Q \subseteq S_0$ such that at least $4m/9$ edges of $H[S_0]$ contain exactly one vertex of Q . Let $S = S_0 \setminus Q$.

We now sample the vertices of H independently with probability $p = 1/3$, and we let T be the set of sampled vertices. We first condition on $T \setminus Q$, in other words, we reveal T outside of Q , and we treat $T \cap Q$ as the source of randomness. We introduce the following notation.

- $V = S \setminus T$,

- G_0 is the graph on vertex set V which consists of the edges of $G[V]$ whose colour is in Q ,
- G_0^* is the subgraph of G_0 in which we only keep the edges whose colour is in $T \cap Q$,
- G_1 is the graph on vertex set V which consists of the edges of $G[V]$ whose colour is in $V(H) \setminus Q$,
- G_1^* is the subgraph of G_1 in which we only keep the edges whose colour is in $T \setminus Q$,
- G^* is the subgraph of $G[T^c]$ in which we only keep the edges whose colour is in T (where $T^c = V(H) \setminus T$),
- $\tilde{G} = G^*[V] = G_0^* \cup G_1^*$.

Now we have that $\text{sp}(G^*) \geq \text{sp}(\tilde{G})$ by Lemma 3.2. Conditionally on $T \setminus Q$, we have that $\tilde{G}$ is a random $(m, 2\Delta, 2)$ -multihypergraph of linear type, provided that $f = e(G_0)$ satisfies $f \geq m/10$. Indeed, it suffices to verify that all conditions in Definition 5.1 are satisfied.

- (i) Let $X = E(G_0) \cup E(G_1)$ and let $R = E(G_0^*) \cup E(G_1^*)$. For every hyperedge $e \in E(G_0)$ of colour w , we let $S_e = \{w\}$; we then have that $e \in E(G_0^*)$ if and only if $Z_w = 1$, where Z_w is the indicator of the event $w \in T$. For every hyperedge $e \in E(G_1)$, we let $S_e = \emptyset$, and whether $e \in E(G_1^*)$ is already known from $T \setminus Q$. As $\tilde{G} = G_0^* \cup G_1^*$, this concludes this item.
- (ii) This item is trivial.
- (iii) Letting Y be the set of hyperedges $e \in E(G_0)$, then we have $|Y| = f \geq m/10$, and for every $e \in Y$, we have $|e| = 2$ and $\mathbb{P}(e \in R) = 1/3$.
- (iv) This item is trivial.
- (v) This item is trivial.

Thus, by Theorem 5.4, we have

$$\mathbb{E}_T[\text{sp}(G^*)|T \setminus Q] \geq \mathbb{E}_T[\text{sp}(\tilde{G})|T \setminus Q] = \Omega\left(\frac{m \mathbf{1}_{f \geq m/10}}{\sqrt{\Delta}}\right).$$

Taking expectation, we get

$$\mathbb{E}_T[\text{sp}(G^*)] = \Omega\left(\frac{m \cdot \mathbb{P}(f \geq m/10)}{\sqrt{\Delta}}\right).$$

Recall that f is the number of edges in $G[S \setminus T]$ whose colour is in Q . Therefore, every edge of $G[S]$ whose colour is in Q is counted with probability $(1-p)^2 = 4/9$, and thus $\mathbb{E}_T[f] \geq 16m/81$. Since $f = O(m)$ almost surely, it follows that $\mathbb{P}(f \geq m/10) = \Omega(1)$, and plugging this in the previous display gives

$$\mathbb{E}_T[\text{sp}(G^*)] = \Omega\left(\frac{m}{\sqrt{\Delta}}\right).$$

Let W_1, W_2 be a partition of T^c such that $e_{G^*}(W_1, W_2) = \frac{1}{2}e(G^*) + \text{sp}(G^*)$. Since $e_H(T, W_1, W_2) = e_{G^*}(W_1, W_2)$, we obtain

$$\mathbb{E}_T[e_H(T, W_1, W_2)] = \mathbb{E}_T[e_{G^*}(W_1, W_2)] = \mathbb{E}_T\left[\frac{1}{2}e(G^*) + \text{sp}(G^*)\right] = \frac{2}{9}e(H) + \Omega\left(\frac{m}{\sqrt{\Delta}}\right),$$

so H has a 3-cut of surplus $\Omega(\frac{m}{\sqrt{\Delta}})$. $\square$

APPENDIX E. PROOF OF LEMMA 5.6

Proof of Lemma 5.6. Let S be a subset of $V(H)$ such that there are at least m edges $e \in E(H)$ with $|e| \geq 4$ and $|e \cap S| \geq 2$, and each vertex in S has degree at most Δ in H . We start by taking a random subset W of S by selecting each vertex with probability $1/2$ independently, and let $W^c = V(H) \setminus W$. For every hyperedge $e \in E(H)$ such that $|e| \geq 4$ and $|e \cap S| \geq 2$, the event $\{|e \cap W| \geq 2, |e \cap W^c| \geq 2\}$ happens with probability $\Omega(1)$. Therefore we may fix for the rest of the proof an outcome of W such that the number of hyperedges $e \in E(H)$ satisfying $|e \cap W| \geq 2$ and $|e \cap W^c| \geq 2$ is $\Omega(m)$. We put each vertex $v \notin W$ in part 1 or part 2 of the cut uniformly

at random and independently, and we let σ be this assignment. We now construct the following auxiliary multihypergraph H_σ on vertex set W : for every edge $e \in E(H)$, if all vertices of e are in W then we add two copies of e , otherwise, if $|\sigma(e \setminus W)| = 1$ and $|e \cap W| > 0$, we add one copy of $e \cap W$. By Corollary 3.10, we have

$$\text{sp}(H) \geq \mathbb{E}_\sigma[\text{sp}(H_\sigma)]/2 \quad (37)$$

We note that the graph H_σ is a random $(m, 2\Delta, k)$ -multihypergraph of linear type. Indeed, it suffices to verify that all conditions in Definition 5.1 are satisfied.

- (i) We let $X = \{e \cap W : e \in E(H)\}$, where, if $e \subset W$, the element $e \cap W = e$ appears with multiplicity 2. For every hyperedge $e \in E(H)$ with representative $e_W = e \cap W \in X$, we set $S_{e_W} = e \setminus W$, and for every vertex $a \in W^c$, we let Z_a be the random variable associated to which part of the cut a has been assigned by σ . Then it is clear that the event $\{e_W \in E(H_\sigma)\}$ is a function of $\{Z_i : i \in S_{e_W}\}$.
- (ii) For two edges $e, f \in X$, we have $|S_e \cap S_f| \leq 1$ since H is a linear hypergraph. It is then clear that we have $\mathbb{P}(e \in R, f \in R) = \mathbb{P}(e \in R)\mathbb{P}(f \in R)$.
- (iii) Let $Y = \{e \cap W : e \in E(H), |e \cap W| \geq 2, |e \cap W^c| \geq 2\}$. We have by choice of W that $|Y| = \Omega(m)$, and it is easy to check that for every $e \in Y$, we have $|e| \geq 2$ and $\Omega(1) = \mathbb{P}(e \in E(H_\sigma)) = 1 - \Omega(1)$.
- (iv) This item is trivial.
- (v) This item is trivial.

By Theorem 5.4, we have

$$\mathbb{E}_\sigma[\text{sp}(H_\sigma)] = \Omega\left(\frac{m}{\sqrt{\Delta}}\right),$$

which, together with (37), finishes the proof. $\square$