THE EXISTENCE OF DESIGNS

PETER KEEVASH

Abstract. We prove the existence conjecture for combinatorial designs, answering a question of Steiner from 1853. More generally, we show that the natural divisibility conditions are sufficient for clique decompositions of simplicial complexes that satisfy a certain pseudorandomness condition.

1. Introduction

A *Steiner system* with parameters (n, q, r) is a set S of q-subsets of an n-set[1] X, such that every r-subset of X belongs to exactly one element of S. The question of whether there is a Steiner system with given parameters is one of the oldest problems in combinatorics, dating back to work of Plücker (1835), Kirkman (1846) and Steiner (1853); see [42] for a historical account.

More generally, we say that a set S of q-subsets of an n-set X is a *design* with parameters (n, q, r, λ) if every r-subset of X belongs to exactly λ elements of S. There are some obvious necessary 'divisibility conditions' for the existence of such S, namely that $\binom{q-i}{r-i}$ divides $\lambda\binom{n-i}{r-i}$ for every $0 \leq i \leq r-1$ (fix any i-subset I of X and consider the sets in S that contain I). It is not known who first advanced the 'Existence Conjecture' that the divisibility conditions are also sufficient, apart from a finite number of exceptional n given fixed q, r and λ.

The case $r = 2$ has received particular attention because of its connections to statistics, under the name of 'balanced incomplete block designs'. We refer the reader to [5] for a summary of the large literature and applications of this field. The Existence Conjecture for $r = 2$ was a long-standing open problem, eventually resolved by Wilson [43, 44, 45] in a series of papers that revolutionised Design Theory, and had a major impact in Combinatorics. In this paper, we prove the Existence Conjecture in general, via a new method, which we will refer to as Randomised Algebraic Constructions.

1.1. Results.

The Existence Conjecture will follow from a more general result on clique decompositions of hypergraphs that satisfy a certain pseudorandomness condition. To describe this we make the following definitions.

Definition 1.1. A *hypergraph* G consists of a vertex set $V(G)$ and an edge set $E(G)$, where each $e \in E(G)$ is a subset of $V(G)$. We identify G with $E(G)$. If every edge has size r we say that G is an *r-graph.* For $S \subseteq V(G)$, the *neighbourhood* $G(S)$ is the $(r - |S|)$-graph $\{f \subseteq V(G) \setminus S : f \cup S \in G\}$. For an r-graph H, an *H-decomposition* of G is a partition of $E(G)$ into subgraphs isomorphic to H. Let K^r_q be the complete r-graph on q vertices.

Date: January 15, 2014.

Mathematical Institute, University of Oxford, Oxford, UK. Email keevash@maths.ox.ac.uk. Research supported in part by ERC grant 239696 and EPSRC grant EP/G056730/1.

[1] i.e. $|X| = n$ and S consists of subsets of X each having size q

Note that a Steiner system with parameters (n, q, r) is equivalent to a K_q^r-decomposition of K_n^r. It is also equivalent to a perfect matching (a set of edges covering every vertex exactly once) in the auxiliary $\binom{q}{r}$-graph on $\binom{[n]}{r}$ (the r-subsets of $[n]$) with edge set $\{\binom{Q}{r} : Q \in \binom{[n]}{q}\}$. The next definition generalises the necessary divisibility conditions described above.

Definition 1.2. Suppose G is an r-graph. We say that G is K_q^r-*divisible* if $\binom{q-i}{r-i}$ divides $|G(e)|$ for any i-set $e \subseteq V(G)$, for all $0 \leq i \leq r$.

Next we formulate a simplified version of our quasirandomness condition. It is easy to see that it holds whp if $G = G^r(n,p)$ is the usual random r-graph and n is large given p, c and h.

Definition 1.3. Suppose G is an r-graph on n vertices. We say that G is (c,h)-*typical* if there is some $p > 0$ such that for any set A of $(r-1)$-subsets of $V(G)$ with $|A| \leq h$ we have $|\cap_{S \in A} G(S)| = (1 \pm c)p^{|A|}n$.

Now we can state a simplified version of our main theorem.

Theorem 1.4. *Let $1/n \ll c \ll d, 1/h \ll 1/q \leq 1/r$. Suppose G is a K_q^r-divisible (c,h)-typical r-graph on n vertices with $|G| > dn^r$. Then G has a K_q^r-decomposition.*

Applying this with $G = K_n^r$, we deduce that for large n the divisibility conditions are sufficient for the existence of Steiner systems. That they suffice for the existence of designs requires a generalisation to multicomplexes that we will explain later. Our method gives a randomised algorithm for constructing designs and also an estimate on the number of designs with given parameters (see Theorem 6.8).

Theorem 1.4 gives new results even in the graph case ($r = 2$); for example, it is easy to deduce that the standard random graph model $G(n, 1/2)$ whp has a partial triangle decomposition that covers all but $(1 + o(1))n/4$ edges: deleting a perfect matching on the set of vertices of odd degree and then at most two 4-cycles gives a graph satisfying the hypotheses of the theorem. This is the asymptotically best possible leave, as whp there are $(1+o(1))n/2$ vertices of odd degree and any partial triangle decomposition must leave at least one edge uncovered at each vertex of odd degree.

We also note that if an r-graph G on n vertices satisfies $|G(S)| \geq (1-c)n$ for every $(r-1)$-subset S of $V(G)$ then it is (hc, h)-typical (with $p = 1$), so we also deduce a minimum $(r-1)$-degree version of the theorem, generalising Gustavsson's minimum degree version [15] of Wilson's theorem.

1.2. **Related work.** As a weaker version of the Existence Conjecture, Erdős and Hanani [7] posed the question of the existence of asymptotically optimal Steiner systems; equivalently, finding $(1 - o(1))\binom{q}{r}^{-1}\binom{n}{r}$ edge-disjoint K_q^r's in K_n^r. This was proved by Rödl [33], by developing a new semi-random construction method known as the 'nibble', which has since had a great impact on Combinatorics (see e.g. [9, 12, 19, 24, 25, 28, 32, 38, 41] for related results and improved bounds). It will also play an important role in this paper.

Regarding exact results, we have already mentioned Wilson's theorem, and Gustavsson's minimum degree generalisation thereof. We should also note the seminal work of Hanani [16, 17], which (inter alia) answers Steiner's problem for $(q, r) \in \{(4, 2), (4, 3), (5, 2)\}$ and all n (the case $(q, r) = (3, 2)$ was solved by Kirkman, before Steiner posed the problem). Besides these, we again refer to [5] as an introduction to the huge literature on the construction of designs. One should note that before the results of the current paper, there were only finitely many known Steiner systems with $r \geq 4$, and it was not known if there were any Steiner systems with $r \geq 6$.

Even the existence of designs with $r \geq 7$ and any 'non-trivial' λ was open before the breakthrough result of Teirlinck [39] confirming this. An improved bound on λ and a probabilistic method (a local limit theorem for certain random walks in high dimensions) for constructing many other rigid combinatorial structures was recently given by Kuperberg, Lovett and Peled [27]. Their result for designs is somewhat complementary to ours, in that they can allow the parameters q and r to grow with n, whereas we require them to be (essentially) constant. They also obtain much more precise estimates than we do for the number of designs (within their range of parameters).

Another recent result, due to Ferber, Hod, Krivelevich and Sudakov [8] gives a short probabilistic construction of 'almost Steiner systems', in which every r-subset is covered by either one or two q-subsets.

Another relaxation of the conjecture, which will play an important role in this paper, is obtained by considering 'integral designs', in which one assigns integers to the copies of K^r_q in K^r_n such that for every edge e the sum of the integers assigned to the copies of K^r_q containing e is a constant independent of e. Graver and Jurkat [14] and Wilson [46] showed that the divisibility conditions suffice for the existence of integral designs (this is used in [46] to show the existence for large λ of integral designs with non-negative coefficients). Wilson [47] also characterised the existence of integral H-decompositions for any r-graph H.

1.3. **Randomised Algebraic Constructions.** We make some brief comments here on our new method of Randomised Algebraic Constructions. The idea is to start by taking a random subset of a model for the problem that is algebraically defined. For the problem of this paper, the result is a partial K^r_q-decomposition that covers a constant fraction of the edge set, and also carries a rich structure of possible local modifications. We treat this partial decomposition as a template for the final decomposition. By various applications of the nibble and greedy algorithms, we can choose another partial K^r_q-decomposition that covers all edges not in the template, which also spills over slightly into the template, so that some edges are covered twice.

To handle the 'spill', we express as it as an integral design that is the difference $D^+ - D^-$ of two partial K^r_q-decompositions of the underlying r-graph of the template. We apply local modifications to the template decomposition so that it includes D^+, then delete D^+ and replace it by D^-. The resulting partial decomposition of the template has a hole that exactly matches the spill, so fits together with the previous partial decomposition to provide a complete decomposition.

At this level of generality, our local modification method sounds somewhat similar to that of Gustavsson, and also to the 'absorbing method' of Rödl, Ruciński and Szemerédi [35]. However, Gustavsson's method is inherently an argument for graphs rather than hypergraphs, and the absorbing method relies on a dense set of 'absorbing configurations' that cannot be guaranteed in the sparse setting of the decomposition problem.

We also remark that our argument is by induction on r, and so we will need to consider the general setting of typical simplicial complexes (to be defined later): even if we just want to prove the existence of Steiner (n, q, r)-systems, our proof uses the existence of K_q^s-decompositions of typical complexes with $s < r$. A final comment is that although our quasirandomness condition resembles those used in Hypergraph Regularity Theory, we make no use of this theory in our proof, so our constants could in principle be moderately effective (but we do not attempt to calculate them).

1.4. **Notation and terminology.** Here we gather some notation and terminology that is used throughout the paper. We write $[n] = \{1, \dots, n\}$ and $[m, n] = [n] \setminus [m-1]$. For a set S, we write $\binom{S}{r}$ for the set of r-subsets of S, $\binom{S}{\le r} = \cup_{i \le r} \binom{S}{i}$ and $\binom{S}{<r} = \cup_{i<r} \binom{S}{i}$. We say that an event E holds with high probability (whp) if $\mathbb{P}(E) = 1 - e^{-\Omega(n^c)}$ for some $c > 0$ as $n \to \infty$; note that when n is sufficiently large, by union bounds we can assume that any specified polynomial number of such events all occur. We write $x \ll y$ to mean for any $y \ge 0$ there exists $x_0 \ge 0$ such that for any $x \le x_0$ the following statement holds. Similar statements with more constants are defined similarly. We write $a = b \pm c$ to mean $b - c \le a \le b + c$. Throughout the paper we omit floor and ceiling symbols where they do not affect the argument.

Suppose X and Y are sets. We write Y^X for the set of vectors with entries in Y and coordinates indexed by X. We identify $v \in \{0,1\}^X$ with the set $\{x \in X : v_x = 1\}$. We identify $v \in \mathbb{N}^X$ with the multiset in X in which x has multiplicity v_x (for our purposes $0 \in \mathbb{N}$). We adopt the convention that 'for each $x \in v$' means that x is considered v_x times in any statement or algorithm. We also apply similar notation and terminology as for multisets to vectors $v \in \mathbb{Z}^X$ (which one might call 'intsets'). Here our convention is that 'for each $x \in v$' means that x is considered $|v_x|$ times in any statement or algorithm, and has a sign attached to it (the same as that of v_x); we also refer to x as a 'signed element' of v. Arithmetic on vectors in $\mathbb{Z}^X$ is to be understood pointwise, i.e. $(v + v')_x = v_x + v'_x$ and $(vv')_x = v_x v'_x$ for $x \in X$. For $v \in \mathbb{Z}^X$ we write $|v| = \sum_{x \in X} |v_x|$. We also write $v = v^+ - v^-$, where $v_x^+ = \max\{v_x, 0\}$ and $v_x^- = \max\{-v_x, 0\}$ for $x \in X$. For $X' \subseteq X$ we define $v[X'] \in \mathbb{Z}^{X'}$ by $v[X']_x = v_x$ for $x \in X'$. If G is a hypergraph, $v \in \mathbb{Z}^G$ and $e \in G$ we define $v(e) \in \mathbb{Z}^{G(e)}$ by $v(e)_f = v_{e \cup f}$ for $f \in G(e)$.

1.5. **Organisation of the paper.** In the next section we review some basic probabilistic methods and state a version of the nibble. Section 3 concerns basic properties of typicality for simplicial complexes. In section 4 we generalise the results of Graver and Jurkat and of Wilson on integral designs to the setting of typical complexes. Section 5 develops the algebraic ingredients of our construction. We prove our main results in section 6. The final section contains some concluding remarks and open problems.

2. Probabilistic methods

In this section we gather some results that will be used throughout the paper, mostly concerning concentration of probability, and ending with a version of the nibble. We start by considering sums of independent $\{0,1\}$-valued random variables.

Definition 2.1. A random variable B is *p-Bernoulli* if $\mathbb{P}(B=1)=p$ and $\mathbb{P}(B=0)=1-p$. We say B is *Bernoulli* if it is p-Bernoulli for some $p=\mathbb{E}B$. We say that a random variable X is *pseudobinomial* if we can write $X=\sum_{i=1}^n X_i$, where $X_1,\dots,X_n$ are independent Bernoulli variables. We do not assume that $\mathbb{E}X_1,\dots,\mathbb{E}X_n$ are equal; if they are we say that X is *binomial.* We also say that a random subset S of a set V is *p-binomial* if each element of V belongs to S independently with probability p.

The following estimate is commonly known as the Chernoff bound, see e.g. [18] Theorems 2.1 and 2.8.

Lemma 2.2. *If X is pseudobinomial and $a\ge 0$ then $\mathbb{P}(X<\mathbb{E}X-a)<e^{-a^2/2\mathbb{E}X}$ and $\mathbb{P}(X>\mathbb{E}X+a)<e^{-a^2/2(\mathbb{E}X+a/3)}$.*

We remark that Lemma 2.2 can be applied to hypergeometric random variables, i.e. random variables of the form $|A\cap B|$ where $B\subseteq V$ are fixed sets and $A\subseteq V$ is uniformly random among a-subsets of V, for some a. This follows by Lemma 1 of [40] (see also Remark 2.11 of [18]).

Next we need an estimate for the expected deviation of a pseudobinomial variable from its mean. We remark that for binomial variables, there is an exact formula due to De Moivre (see [6] for an entertaining exposition and some generalisations).

Lemma 2.3. *There is $C>0$ such that for any pseudobinomial X with mean μ we have $\mathbb{E}|X-\mu|\le C\sqrt{\mu}$.*

Proof. Write $\mathbb{E}|X-\mu|=\sum_{t\ge 0}|t-\mu|\mathbb{P}(X=t)=E_0+E_1$, where E_i is the sum of $|t-\mu|\mathbb{P}(X=t)$ over $|t-\mu|>\frac{1}{2}C\sqrt{\mu}$ for $i=0$ or $|t-\mu|\le\frac{1}{2}C\sqrt{\mu}$ for $i=1$. Clearly, $E_1\le\frac{1}{2}C\sqrt{\mu}$, and by Chernoff bounds, for C large,

$$E_0\le\sum_{a>\frac{1}{2}C\sqrt{\mu}} a(e^{-a^2/2\mu}+e^{-a^2/2(\mu+a/3)})\le\tfrac{1}{2}C\sqrt{\mu}. \qquad \square$$

Now we describe some more general results on concentration of probability. We make the following standard definitions. (In this paper all probability spaces are finite, and will only be referred to implicitly via random variables.)

Definition 2.4. Let Ω be a (finite) probability space. An *algebra* (on Ω) is a set $\mathcal{F}$ of subsets of Ω that includes Ω and is closed under intersections and taking complements. A *filtration* (on Ω) is a sequence $\mathcal{F} = (\mathcal{F}_i)_{i\geq 0}$ of algebras such that $\mathcal{F}_i \subseteq \mathcal{F}_{i+1}$ for $i \geq 0$. We say $A = (A_i)_{i\geq 0}$ is a *supermartingale* (wrt $\mathcal{F}$) if $\mathbb{E}(X_{i+1}|\mathcal{F}_i) \leq X_i$ for $i \geq 0$.

Now we can state a general result of Freedman [10] that essentially implies all of the bounds we will use (perhaps with slightly weaker constants).

Lemma 2.5. *Let $(A_i)_{i\geq 0}$ be a supermartingale wrt a filtration $\mathcal{F} = (\mathcal{F}_i)_{i\geq 0}$. Suppose that $A_{i+1} - A_i \leq b$ for all $i \geq 0$, and let E be the 'bad' event that there exists $j \geq 0$ with $A_j \geq A_0 + a$ and $\sum_{i=1}^{j} Var[A_i \mid \mathcal{F}_{i-1}] \leq v$. Then*

$$\mathbb{P}(E) \leq \exp\left(-\tfrac{a^2}{2(v+ab)}\right).$$

We proceed to give some useful consequences of Lemma 2.5. First we make another definition.

Definition 2.6. Suppose Y is a random variable and $\mathcal{F} = (\mathcal{F}_0, \dots, \mathcal{F}_n)$ is a filtration. We say that Y is *(C,μ)-dominated (wrt $\mathcal{F}$)* if we can write $Y = \sum_{i=1}^{n} Y_i$, where Y_i is $\mathcal{F}_i$-measurable, $|Y_i| \leq C$ and $\mathbb{E}[|Y_i| \mid \mathcal{F}_{i-1}] < \mu_i$ for $i \in [n]$, where $\sum_{i=1}^{n} \mu_i < \mu$.

Lemma 2.7. *If Y is (C,μ)-dominated then*

$$\mathbb{P}(|Y| > (1+c)\mu) < 2e^{-\mu c^2/2(1+2c)C}.$$

Proof. Let $A_i = \sum_{j<i}(Y_j - \mu_j)$ for $i \geq 0$; then $(A_i)_{i\geq 0}$ is a supermartingale and

$$\begin{aligned} Var[A_i \mid \mathcal{F}_{i-1}] &= Var[Y_i \mid \mathcal{F}_{i-1}] \\ &\leq \mathbb{E}[Y_i^2 \mid \mathcal{F}_{i-1}] \leq C\mathbb{E}[|Y_i| \mid \mathcal{F}_{i-1}] \leq C\mu_i. \end{aligned}$$

By Lemma 2.5 applied with $a = c\mu$, $b = 2C$ and $v = C\mu$ we obtain

$$\mathbb{P}(Y > (1+c)\mu) < e^{-\mu c^2/2(1+2c)C}.$$

Similarly, considering $A_i = -\sum_{j<i}(Y_j + \mu_j)$ gives the same estimate for $\mathbb{P}(Y < -(1+c)\mu)$. □

Next we state the well-known inequality of Azuma [3]. (A slightly weaker version follows from Lemma 2.5.)

Lemma 2.8. *(Azuma) Let $(A_i)_{i=0}^{n}$ be a supermartingale wrt a filtration $\mathcal{F} = (\mathcal{F}_i)_{i\geq 0}$. Suppose that $|A_{i+1} - A_i| \leq b_i$ for $i \geq 0$. Then*

$$\mathbb{P}(A_n - A_0 \geq a) \leq \exp\left(-\frac{a^2}{2\sum_{i=1}^{n} b_i^2}\right).$$

We record several consequences of Azuma's inequality (see e.g. [30]).

Definition 2.9. Suppose $f : S \to \mathbb{R}$ where $S = \prod_{i=1}^{n} S_i$ and $b = (b_1, \dots, b_n)$ with $b_i \geq 0$ for $i \in [n]$. We say that f is *b-Lipschitz* if for any $s, s' \in S$ that differ only in the ith coordinate we have $|f(s) - f(s')| \leq b_i$. We also say that f is *B-varying* where $B = \sum_{i=1}^{n} b_i^2$.

Lemma 2.10. *Suppose $Z = (Z_1, \dots, Z_n)$ is a sequence of independent random variables, and $X = f(Z)$, where f is a B-varying function. Then*

$$\mathbb{P}(|X - \mathbb{E}X| > a) \le 2e^{-a^2/2B}.$$

Let S_X be the symmetric group on X.

Definition 2.11. Suppose $f : S_{[n]} \to \mathbb{R}$ and $b \ge 0$. We say that f is *b-Lipschitz* if for any $\sigma, \sigma' \in S$ such that $\sigma = \tau \circ \sigma'$ for some transposition $\tau \in S_{[n]}$ we have $|f(\sigma) - f(\sigma')| \le b$.

Lemma 2.12. *Suppose $f : S_{[n]} \to \mathbb{R}$ is b-Lipschitz, $\sigma \in S_{[n]}$ is uniformly random and $X = f(\sigma)$. Then*

$$\mathbb{P}(|X - \mathbb{E}X| > a) \le 2e^{-a^2/2nb^2}.$$

We will also need a common generalisation of Lemmas 2.10 and 2.12, which perhaps has not appeared before, but is proved in the same way. It considers functions in which the input consists of n independent random injections $\pi_i : [a'_i] \to [a_i]$: if $a'_i = 1$ this is a random element of $[a_i]$; if $a'_i = a_i$ this is a random permutation of $[a_i]$.

Definition 2.13. Let $a = (a_1, \dots, a_n)$ and $a' = (a'_1, \dots, a'_n)$, where $a_i \in \mathbb{N}$ and $a'_i \in [a_i]$ for $i \in [n]$, and $\Pi(a, a')$ be the set of $\pi = (\pi_1, \dots, \pi_n)$ where $\pi_i : [a'_i] \to [a_i]$ is injective. Suppose $f : \Pi(a, a') \to \mathbb{R}$ and $b = (b_1, \dots, b_n)$ with $b_i \ge 0$ for $i \in [n]$. We say that f is *b-Lipschitz* if for any $i \in [n]$ and $\pi, \pi' \in \Pi(a, a')$ such that $\pi_j = \pi'_j$ for $j \ne i$ and $\pi_i = \tau \circ \pi'_i$ for some transposition $\tau \in S_{[a_i]}$ we have $|f(s) - f(s')| \le b_i$. We also say that f is *B-varying* where $B = \sum_{i=1}^n a'_i b_i^2$.

Lemma 2.14. *Suppose $f : \Pi(a, a') \to \mathbb{R}$ is B-varying, $\pi \in \Pi(a, a')$ is uniformly random and $X = f(\pi)$. Then*

$$\mathbb{P}(|X - \mathbb{E}X| > t) \le 2e^{-t^2/2B}.$$

We conclude with a version of the nibble (which was discussed in the introduction). The following theorem of Pippenger (unpublished) generalises the result of Rödl to give a nearly perfect matching in any r-graph that is approximately regular and has small codegrees; see also [32] for a powerful generalisation of this result.

Theorem 2.15. *For any integer $r \ge 2$ and real $a > 0$ there is $b > 0$ so that if G is an r-graph such that there is some D for which $|G(x)| = (1 \pm b)D$ for every vertex x and $|G(xy)| < bD$ for every pair of vertices x, y, then G has a matching covering all but at most an vertices.*

Our applications of this result take the following form.

Definition 2.16. Suppose J is an r-graph on $[n]$ and $\theta > 0$. We say that J is *θ-bounded* if $|J(e)| < \theta n$ for all $e \in \binom{[n]}{r-1}$.

Lemma 2.17. *Let $1/n \ll b, d_0 \ll a \ll 1/r$ and $d \ge d_0$. Suppose G is an r-graph on $[n]$ and $\mathcal{Q} \subseteq K_q^r(G)$. Let A be the auxiliary $\binom{q}{r}$-graph where $V(A) = E(G)$ and $E(A) = \{\binom{Q}{r} : Q \in \mathcal{Q}\}$. Suppose also that for any $e \in E(G) = V(A)$ we have $|A(e)| = (1 \pm b)dn^{q-r}$. Then there is a matching M in A such the r-graph $G' = V(A) \setminus V(M)$ is a-bounded.*

Note that the weaker conclusion $|G'| < a|G|$ is immediate from Theorem 2.15, as for for any $\{e, e'\} \subseteq E(G)$ we have $|A(ee')| < n^{q-r-1} < bdn^{q-r}$. The stronger conclusion that G' is a-bounded follows from the proof: applying large deviation inequalities in each bite of the nibble gives whp $|G'(e)| < an$ for all $e \in \binom{[n]}{r-1}$.

3. Typicality

In this section we define typical complexes, which are the quasirandom structures used throughout the paper, and prove several lemmas regarding their basic properties. We start with a series of definitions that are reminiscent of those in Hypergraph Regularity Theory (see e.g. [11, 36, 37] but note that we do not use this theory here).

Definition 3.1. An *complex* G is a hypergraph such that whenever $e' \subseteq e \in G$ we have $e' \in G$. We say G is an *q-complex* if $|e| \leq q$ for all $e \in G$. We write G_i, $G_{<i}$, $G_{\leq i}$ respectively for the set of $e \in G$ such that $|e| = i$, $|e| < i$, $|e| \leq i$. For any hypergraph H we write $H^{\leq}$ for the complex *generated* by H, i.e. the set of $e \subseteq V(H)$ such that $e \subseteq f$ for some $f \in H$.

Definition 3.2. Suppose G is a complex and $e \in G$. The *neighbourhood* of e is the complex $G(e)$ consisting of all sets f such that $e \cap f = \emptyset$ and $e \cup f \in G$.

Definition 3.3. Suppose that G is a q-complex and G' is a q'-complex. We define the *restriction* of G to G' as the q-complex $G[G']$ consisting of all $e \in G$ such that $e' \in G'$ for all $e' \subseteq e$ with $|e'| \leq q'$. Typically we have $G' \subseteq G$, although we do not require this. Also, if $J \subseteq G_s$ for some $s \in [q]$ we define $G[J] = G[J \cup G_{<s}]$.

To understand the next definition, it is perhaps helpful to focus at first on the case $i = 3$: in words, the 3-density of G is the proportion of triangles in G_2 that are also triples in G_3.

Definition 3.4. Let G be a complex and K be the complete complex on $V(G)$. The *(relative) i-density* of G is

$$d_i(G) = |G_i|/|K[G_{<i}]_i|.$$

If $d_i(G)$ is undefined then we adopt the convention that $d_i(G) = 1$.

When unspecified, K will always denote the complete complex on $V(G)$. Next we introduce some general notation for extension variables.

Definition 3.5. Let H and G be complexes.

(i) We say that $\phi : V(H) \to V(G)$ is an *embedding* of H in G if ϕ is injective and $\phi(e) \in G$ for all $e \in H$.
(ii) Suppose $F \subseteq V(H)$ and ϕ is an embedding of $H[F]$ in G. We call $E = (\phi, F, H)$ a *rooted extension (in G)* with *base* F and *size* $|V(H)|$.
(iii) We say that E is *simple* if $|F| = |V(H)| - 1$.
(iv) We write $X_E(G)$ for the set or number of embeddings ϕ^* from H to G such that $\phi^*|_F = \phi$.
(v) If $H[F] \subseteq G$ we write $\iota : F \to V(G)$ for the identity embedding.
(vi) If $F = \emptyset$ we use H to denote $E = (\iota, \emptyset, H)$.

(vii) We write $X_q(G) = X_{[q]^\leq}(G)$.

Finally we can define typicality.

Definition 3.6. Let G be a complex, $E = (\phi, F, H)$ be a rooted extension in G and

$$\pi_E(G) = \prod_{e \in H \setminus H[F]} d_{|e|}(G).$$

We say that E is *c-typical (in G)* if

$$X_E(G)/X_E(K) = (1 \pm c)\pi_E(G).$$

We say that G is *(c,h)-typical* if every simple rooted extension in G of size at most h is c-typical.

As a simple example, we note that $\pi_E(K) = 1$ for any E so K is (c,h)-typical for any c and h. More generally, suppose $d = (d_i : i \in [q])$ with $d_i \in [0,1]$ and let $G(n,d)$ be the random q-complex on $[n]$ generated by the following iterative procedure: for $i \in [q]$, given G_j for $j < i$, include each $e \in K[G_{<i}]_i$ in G_i independently with probability d_i. It is not hard to see that for fixed $h \in \mathbb{N}$, $c > 0$, $d_i > 0$ for $i \in [q]$ whp $G(n,d)$ is (c,h)-typical (for large n).

We also remark that if G is (c,h)-typical and $E = (\phi, F, H)$ is a rooted extension in G of size at most h (not necessarily simple) then $X_E(G)/X_E(K) = (1 \pm c)^h \pi_E(G)$; since E can be constructed as a sequence of simple extensions this is immediate from the definition.

Next we generalise the above definitions to define typicality of a pair of complexes.

Definition 3.7. We say that (G, G') is a *complex pair* if G' is a subcomplex of G. Suppose that (G, G') and (H, H') are complex pairs. We say that an injection $\phi : V(H) \to V(G)$ is an *embedding* of (H, H') in (G, G') if $\phi(e) \in G$ for all $e \in H$ and $\phi(e) \in G'$ for all $e \in H'$. Suppose $F \subseteq V(H)$ and ϕ is an embedding of $(H[F], H'[F])$ in (G, G'). We call $E = (\phi, F, H, H')$ a *rooted extension (in (G, G'))* with *root* ϕ, *base* F and *size* $|V(H)|$. We say that E is *simple* if $|F| = |V(H)| - 1$. We write $X_E(G, G')$ for the set or number of embeddings ϕ^* of (H, H') in (G, G') such that $\phi^*|_F = \phi$. Let

$$\pi_E(G, G') = \prod_{e \in H' \setminus H'[F]} d_{|e|}(G') \prod_{e \in H \setminus (H' \cup H[F])} d_{|e|}(G).$$

We say that E is *c-typical* if

$$X_E(G, G')/X_E(K, K) = (1 \pm c)\pi_E(G, G').$$

We say that (G, G') is *(c,h)-typical* if every simple rooted extension in (G, G') of size at most h is c-typical.

Note that if G is (c,h)-typical then so is (G, G), as if $E = (\phi, F, H, H')$ and $E^- = (\phi, F, H)$ we have $X_E(G, G) = X_{E^-}(G)$. Note also that if (G, G') is (c,h)-typical then G and G' are both (c,h)-typical (the converse does not hold), as if $E = (\phi, F, H)$, $E^1 = (\phi, F, H, \emptyset)$, $E^2 = (\phi, F, H, H)$ we have $X_E(G) = X_{E^1}(G, G')$ and $X_E(G') = X_{E^2}(G, G')$.

The next lemma provides a useful criterion for typicality that will be used in several other lemmas.

Lemma 3.8. *Let $0 < c \ll 1/h \ll 1/q$, let $d_i, d'_i > 0$ for $i \in [q]$, and (G, G') be a q-complex pair. Suppose that for any simple rooted extension $E = (\phi, F, H, H')$ in (G, G') of size at most h we have*

$$X_E(G, G')/X_E(K, K) = (1 \pm c)\pi_E,$$

where

$$\pi_E = \prod_{e \in H' \setminus H'[F]} d'_{|e|} \prod_{e \in H \setminus (H' \cup H[F])} d_{|e|}.$$

Then (G, G') is $(3^h c, h)$-typical, with

$$d_i(G) = (1 \pm 3ic)d_i \quad \text{and} \quad d_i(G') = (1 \pm 3ic)d'_i \quad \text{for} \quad i \in [q].$$

Proof. For any extension E we can construct E as a sequence of simple extensions and apply the hypothesis to each step. In particular, for $i \in [q]$, letting $E = (\iota, \emptyset, [i]^{\leq}, \emptyset)$, we obtain

$$|G_i| = (1 \pm c)^i n^i \prod_{j \in [i]} d_j^{\binom{i}{j}}.$$

Similarly, letting $E = (\iota, \emptyset, [i]^{<}, \emptyset)$, we obtain

$$|K[G_{<i}]_i| = (1 \pm c)^i n^i \prod_{j \in [i-1]} d_j^{\binom{i}{j}}.$$

Thus $d_i(G) = |G_i|/|K[G_{<i}]_i| = (1 \pm 3ic)d_i$. Similarly, $d_i(G') = (1 \pm 3ic)d'_i$. Now for any simple rooted extension E of size at most h, substituting these estimates in $X_E(G, G')/X_E(K, K) = (1 \pm c) \prod \pi_E$, we see that (G, G') is $(3^h c, h)$-typical. □

Next we show that typicality is preserved under restrictions.

Lemma 3.9. *Let $0 < c \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair with $d_i(G') > 0$ for $i \in [q]$. Let $s \in [q]$ and $G'' = G[G'_{\leq s}]$. Then (G, G'') and (G'', G') are $(3^h c, h)$-typical with $d_i(G'') = (1 \pm 3ic)d_i(G)$ for $i > s$.*

Proof. Consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G'') of size at most h. Writing $E' = (\phi, F, H, H'_{\leq s})$, we have

$$\begin{aligned} X_E(G, G'')/X_E(K, K) &= X_{E'}(G, G')/X_{E'}(K, K) = (1 \pm c)\pi_{E'}(G, G') \\ &= (1 \pm c) \prod_{e \in H'_{\leq s} \setminus H'[F]} d_{|e|}(G') \prod_{e \in H \setminus (H'_{\leq s} \cup H[F])} d_{|e|}(G). \end{aligned}$$

Thus the required typicality of (G, G'') follows from Lemma 3.8 applied with $d_i = d_i(G)$ for $i \in [q]$, $d'_i = d_i(G')$ for $i \in [s]$ and $d'_i = d_i(G)$ for $i \in [q] \setminus [s]$.

Similarly, consider any simple rooted extension $E = (\phi, F, H, H')$ in (G'', G') of size at most h. Writing $E' = (\phi, F, H, H' \cup H_{\leq s})$, we have

$$\begin{aligned} X_E(G'', G')/X_E(K, K) &= X_{E'}(G, G')/X_{E'}(K, K) = (1 \pm c)\pi_{E'}(G, G') \\ &= (1 \pm c) \prod_{e \in (H' \cup H_{\leq s}) \setminus H[F]} d_{|e|}(G') \prod_{e \in H \setminus (H' \cup H_{\leq s} \cup H[F])} d_{|e|}(G). \end{aligned}$$

Thus the required typicality of (G'', G') follows from Lemma 3.8 applied with $d'_i = d_i(G')$ for $i \in [q]$, $d_i = d_i(G')$ for $i \in [s]$ and $d_i = d_i(G)$ for $i \in [q] \setminus [s]$. $\square$

Now we show that typicality is preserved by taking neighbourhoods.

Lemma 3.10. *Let $0 < c \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair with $d_i(G') > 0$ for $i \in [q]$.*

(i) Suppose $e \in G'$ and let $G^* = G'[G'(e)]$. Then (G, G^*) is $(3^h c, h - |e|)$-typical with $d_i(G^*) = (1 \pm 3hc) \prod_{j \in [q]} d_j(G')^{\binom{|e|}{j-i}}$ for $i \in [q]$.

(ii) Suppose $e \in G$ and let $G^* = G'[G(e)]$. Then (G, G^*) is $(3^h c, h - |e|)$-typical with $d_i(G^*) = (1 \pm 3hc) d_i(G') \prod_{j=i+1}^{q} d_j(G)^{\binom{|e|}{j-i}}$ for $i \in [q]$.

Proof. For (i), consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G^*) of size at most $h - |e|$. We define $E^+ = (\phi^+, F^+, H^+, H^{+\prime})$ in (G, G'), where $V(H^+)$ is obtained from $V(H)$ by adding a set S of $|e|$ new vertices, H^+ and $H^{+\prime}$ are respectively obtained from H and H' by adding all sets $f \cup S'$ of size at most q where $S' \subseteq S$ and $f \in H'$, we let $F^+ = F \cup S$ and we extend ϕ to ϕ^+ by mapping S to e. Then $X_E(G, G^*) = X_{E^+}(G, G')$, and by typicality (recalling that $d_i(G) = 1$ if G_i is undefined)

$$\begin{aligned}
&X_{E^+}(G, G^*)/X_{E^+}(K, K) = (1 \pm c)\pi_{E^+}(G, G') \\
&= (1 \pm c) \prod_{\substack{f \in H' \setminus H'[F] \\ S' \subseteq S}} d_{|f \cup S'|}(G') \prod_{f \in H \setminus (H' \cup H[F])} d_{|f|}(G) \\
&= (1 \pm c) \prod_{\substack{f \in H' \setminus H'[F] \\ j \in [q]}} d_j(G')^{\binom{|e|}{j-|f|}} \prod_{f \in H \setminus (H' \cup H[F])} d_{|f|}(G).
\end{aligned}$$

Thus the required typicality of (G, G^*) follows from Lemma 3.8 applied with $d_i = d_i(G)$ and $d'_i = \prod_{j \in [q]} d_j(G')^{\binom{|e|}{j-i}}$ for $i \in [q]$.

Similarly, for (ii), consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G^*) of size at most $h - |e|$. We define $E^+ = (\phi^+, F^+, H^+, H')$ in (G, G'), where $V(H^+)$ is obtained from $V(H)$ by adding a set S of $|e|$ new vertices, H^+ is obtained from H by adding all sets $f \cup S'$ of size at most q where $S' \subseteq S$ and $f \in H'$, we let $F^+ = F \cup S$ and we extend ϕ to ϕ^+ by mapping S to e. Then $X_E(G, G^*) = X_{E^+}(G, G')$, and by typicality

$$
\begin{aligned}
&X_{E^+}(G, G^*)/X_{E^+}(K, K) = (1 \pm c)\pi_{E^+}(G, G') \\
&= (1 \pm c) \prod_{f \in H' \setminus H'[F]} d_{|f|}(G') d_{|f|}(G)^{-1} \prod_{\substack{f \in H' \setminus H'[F] \\ S' \subseteq S}} d_{|f \cup S'|}(G) \\
&\quad \times \prod_{f \in H \setminus (H' \cup H[F])} d_{|f|}(G) \\
&= (1 \pm c) \prod_{f \in H' \setminus H'[F]} d_{|f|}(G') d_{|f|}(G)^{-1} \left(\prod_{j \in [q]} d_j(G)^{\binom{|e|}{j-|f|}} \right) \\
&\quad \times \prod_{f \in H \setminus (H' \cup H[F])} d_{|f|}(G).
\end{aligned}
$$

Thus the required typicality of (G, G^*) follows from Lemma 3.8 applied with $d_i = d_i(G)$ and $d'_i = d_i(G') \prod_{j=i+1}^{q} d_j(G')^{\binom{|e|}{j-i}}$ for $i \in [q]$. □

The next lemma is an immediate corollary of the previous one.

Lemma 3.11. *Let $0 < c \ll 1/h \ll 1/q$. Suppose G is a (c, h)-typical q-complex with $d_i(G) > 0$ for $i \in [q]$. Let $e \in G$.*

Then $G(e)$ is a $(3^h c, h - |e|)$-typical $(q - |e|)$-complex with $d_i(G(e)) = (1 \pm 3hc) \prod_{j \in [q]} d_j(G)^{\binom{|e|}{j-i}}$ for $i \in [q - |e|]$.

Now we need another general criterion for typicality that will be used in several further lemmas below.

Lemma 3.12. *Let $0 < \theta, c, p_0 \ll 1/h \ll 1/q$. Let (G, G') be a (c, h)-typical q-complex pair and G'' be a subcomplex of G' with $d_i(G'') > 0$ for $i \in [q]$. Suppose that, for some $p \geq p_0$ and $s \in [q]$, for any simple rooted extension $E = (\phi, F, H, H')$ in (G, G'') of size at most h we have*

$$X_E(G, G'') = (1 \pm \theta) p^{|H'_s \setminus H'[F]|} X_E(G, G').$$

Then (G, G'') is $(c + 3^h \theta, h)$-typical, $d_i(G'') = d_i(G')$ for $i < s$, $d_s(G'') = (1 \pm 3q\theta) d_s(G') p$ and $d_i(G'') = (1 \pm 3q\theta) d_i(G')$ for $i > s$.

Proof. Letting $E = (\iota, [s-1], [s]^{\leq}, [s]^{\leq})$ we see that every set in G'_{s-1} is contained in at least one set of G''_s, so $G''_i = G'_i$ for $i < s$, giving $d_i(G'') = d_i(G')$ for $i < s$. Also, for any extension E we can construct E as a sequence of simple extensions and apply the hypothesis to each step. In particular, for $i \geq s$, letting $E = (\iota, \emptyset, [i]^{\leq}, [i]^{\leq})$, we obtain $|G''_i| = (1 \pm \theta)^i p^{\binom{i}{s}} |G'_i|$. Similarly, letting $E = (\iota, \emptyset, [i]^{<}, [i]^{<})$, where $i > s$, we obtain $|K[G''_{<i}]_i| = (1 \pm \theta)^i p^{\binom{i}{s}} |K[G'_{<i}]_i|$.

Now $d_s(G'') = |G''_s|/|K[G''_{<s}]_s| = (1\pm\theta)^s p|G'_s|/|K[G'_{<s}]_s| = (1\pm\theta)^s p d_s(G')$. Also, for $i > s$, we have

$$\begin{aligned} d_i(G'') &= |G''_i|/|K[G''_{<i}]_i| \\ &= (1\pm\theta)^i p^{\binom{i}{s}} |G'_i|/(1\pm\theta)^i p^{\binom{i}{s}} |K[G'_{<i}]_i| \\ &= (1 \pm 3i\theta) d_i(G'). \end{aligned}$$

Finally, for any simple rooted extension E of size at most h, substituting these density estimates in the hypothesis of the lemma gives

$$\begin{aligned} X_E(G,G'')/X_E(K,K) &= (1\pm\theta)p^{|H'_s \setminus H'[F]|}(1\pm c)\pi_E(G,G') \\ &= (1 \pm c \pm 3^h\theta)\pi_E(G,G''). \end{aligned}$$ □

The statement and proof of the next lemma are very similar to that of the previous one, so we omit the proof.

Lemma 3.13. *Let $0 < \theta, c, p_0 \ll 1/h \ll 1/q$. Suppose (G,G^*) and (G^*,G') are q-complex pairs with $d_i(G') > 0$ for $i \in [q]$ and (G,G') is (c,h)-typical. Suppose that, for some $p \geq p_0$ and $s \in [q]$, for any simple rooted extension $E = (\phi, F, H, H')$ in (G^*,G') of size at most h we have*

$$X_E(G^*,G') = (1\pm\theta)p^{|H_s \setminus H[F]|} X_E(G,G').$$

Then (G^*,G') is $(c+3^h\theta, h)$-typical, $d_i(G^*) = d_i(G)$ for $i < s$, $d_s(G^*) = (1 \pm 3q\theta)d_s(G)p$ and $d_i(G^*) = (1\pm 3q\theta)d_i(G)$ for $i > s$.

Next we show that typicality is preserved under deletion of bounded subgraphs. First we reformulate Definition 2.16 for convenient use with subgraphs of complexes (it is equivalent, as $|J(e)| = 0$ for $e \in \binom{[n]}{i-1} \setminus G_{i-1}$).

Definition 3.14. Suppose G is a q-complex on $[n]$, $i \in [q]$, $J \subseteq G_i$ and $\theta > 0$. We say that J is *θ-bounded (wrt G)* if $|J(e)| < \theta n$ for all $e \in G_{i-1}$.

Lemma 3.15. *Let $0 < \theta \ll \theta' \ll c, d \ll 1/h \ll 1/q$ and $s \in [q]$. Suppose (G,G') is a (c,h)-typical q-complex pair with $d_i(G') \geq d$ for $i \in [q]$.*

(i) Let $J \subseteq G'_s$ be θ-bounded wrt G' and $G'' = G'[G'_s \setminus J]$. Then (G,G'') is $(c+\theta', h)$-typical with $d_i(G'') = (1\pm\theta')d_i(G')$ for $i \in [q]$.

(ii) Let $J \subseteq G_s \setminus G'_s$ be θ-bounded wrt G and $G^1 = G[G_s \setminus J]$. Then (G^1,G') is $(c+\theta', h)$-typical with $d_i(G^1) = (1\pm\theta')d_i(G)$ for $i \in [q]$.

Proof. For (i), consider any simple rooted extension $E = (\phi, F, H, H')$ in (G,G'') of size at most h. Since J is θ-bounded wrt G' we have

$$X_E(G,G') \geq X_E(G,G'') \geq X_E(G,G') - 2^h\theta|V(G)|,$$

so $X_E(G,G'') = (1\pm\theta_0)X_E(G,G')$, where $\theta \ll \theta_0 \ll \theta'$. Typicality now follows from Lemma 3.12 (with $p = 1$). Similarly, for (ii), consider any simple rooted extension $E = (\phi, F, H, H')$ in (G^1,G') of size at most h. As in (i),

$$X_E(G,G') \geq X_E(G^1,G') \geq X_E(G,G') - 2^h\theta|V(G)|,$$

so $X_E(G^1,G') = (1\pm\theta_0)X_E(G,G')$. Typicality follows from Lemma 3.13. □

In the next lemma we apply inclusion-exclusion to estimate the number of extensions in which we do not allow the images of certain sets to lie in some subcomplex.

Lemma 3.16. *Let $0 < c \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair with $d_i(G') > 0$ for $i \in [q]$. Let (ϕ, F, H, H') be a rooted extension in (G, G') with $|V(H)| \leq h$ and $A \subseteq H \setminus H'$ such that $e' \in H$ for all $e' \subsetneq e \in A$. Let $X^A_E(G, G')$ be the number of $\phi \in X_E(G, G')$ such that $\phi(f) \in G \setminus G'$ for all $f \in A$. Suppose that $d_{|f|}(G) \geq 2d_{|f|}(G')$ for all $f \in A$. Then*

$$X^A_E(G, G') = (1 \pm 2^{h!}c) X_E(K, K)\pi_E(G, G') \prod_{f \in A} (1 - d_{|f|}(G')/d_{|f|}(G)).$$

Proof. For $I \subseteq A$ let N_I be the number of $\phi \in X_E(G, G')$ such that $\phi(f) \in G'$ for all $f \in I$. By inclusion-exclusion, $X^A_E(G, G') = \sum_{I \subseteq A} (-1)^{|I|} N_I$. By typicality applied to $(\phi, F, H, H' \cup I)$,

$$N_I = (1 \pm c)^h X_E(K, K)\pi_E(G, G') \prod_{f \in I} (d_{|f|}(G')/d_{|f|}(G)).$$

We also note that

$$\prod_{f \in A} (1 - d_{|f|}(G')/d_{|f|}(G)) = \sum_{I \subseteq A} (-1)^{|I|} \prod_{f \in I} (d_{|f|}(G')/d_{|f|}(G)).$$

Thus the main terms in the inclusion-exclusion formula for $X^A_E(G, G')$ combine to give $X_E(K, K)\pi_E(G, G') \prod_{f \in A}(1 - d_{|f|}(G')/d_{|f|}(G))$. For each $I \subseteq A$, the error term can be bounded by $2hc\pi_E(G, G')X_E(K, K)$, and there are $2^{|A|} < 2^{h^q}$ such terms, so this easily implies the stated estimate. □

We deduce the following 'complementation' property of typicality.

Lemma 3.17. *Let $0 < c \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair with $d_i(G') > 0$ for $i \in [q]$ and $G'_i = G_i$ for $i < r$. Let $r \in [q]$, suppose that $d_r(G) \geq 2d_r(G')$, and let $G'' = G[G_r \setminus G'_r]$. Then (G, G'') is $(4^{h!}c, h)$-typical with $d_r(G'') = (1 \pm 4^{h!}c)(d_r(G) - d_r(G'))$ and $d_i(G'') = (1 \pm 4^{h!}c)d_i(G)$ for $i \in [q] \setminus \{r\}$.*

Proof. Consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G'') of size at most h. Writing $E' = (\phi, F, H, H_{<r})$ and $A = H'_r \setminus H'[F]$, by Lemma 3.16 and typicality of (G, G') we have

$$X_E(G, G'') = X^A_{E'}(G, G') = (1 \pm 3^{h!}c)X_{E'}(G, G')(1 - d_r(G')/d_r(G))^{|A|}.$$

The lemma now follows from Lemma 3.12, applied with $s = r$ and $p = 1 - d_r(G')/d_r(G)$. □

We conclude our analysis of typicality with two lemmas showing that it is preserved under random reductions. First we consider restricting to a random subset of fixed size.

Lemma 3.18. *Let $0 < 1/n \ll \varepsilon, c, d \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair on n vertices with $d_i(G') \geq d$ for $i \in [q]$. Fix $N > n^\varepsilon$, suppose J is a uniformly random subset of G'_1 of size N, and let $G'' = G'[J]$. Then whp (G, G'') is $(c + N^{-1/3}, h)$-typical with $d_i(G'') = (1 \pm N^{-1/3})d_i(G')$ for $i > 1$.*

Proof. Consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G'') of size at most h. Write $V(H) \setminus F = \{x\}$. If $\{x\} \notin H'$ then $X_E(G, G'') = X_E(G, G')$. Otherwise, $\mathbb{E}X_E(G, G'') = pX_E(G, G') = \Omega(N)$, where $p = N/|G'_1|$, so by Lemma 2.2 whp $X_E(G, G'') = (1 \pm N^{-0.4})pX_E(G, G')$. The lemma now follows from Lemma 3.12. □

Our second random reduction is that of restricting to a binomial random subset of any level of the complex; it will also be convenient in later applications to allow replacement of a bounded subgraph of the deleted sets.

Lemma 3.19. *Let $0 < 1/n \ll \theta \ll \theta' \ll c, d, p_0 \ll 1/h \ll 1/q$. Suppose (G, G') is a (c, h)-typical q-complex pair on n vertices with $d_i(G') \geq d$ for $i \in [q]$. Suppose J is p-binomial in G'_s for some $s \in [q]$ and $p \geq p_0$. Let $J' \subseteq G'_s$ be θ-bounded wrt G' and $G'' = G'[J \cup J']$. Then whp (G, G'') is $(c+\theta', h)$-typical, $d_i(G'') = d_i(G')$ for $i < s$, $d_s(G'') = (1 \pm \theta')d_s(G')\mathbb{E}|J|/|G'_s|$ and $d_i(G'') = (1 \pm \theta')d_i(G')$ for $i > s$.*

Proof. Consider any simple rooted extension $E = (\phi, F, H, H')$ in (G, G'') of size at most h. Let $p = \mathbb{E}|J|/|G'_s|$ and $\theta \ll \theta_0 \ll \theta'$. Then $\mathbb{E}X_E(G, G'') = p^{|H'_s \setminus H'[F]|}X_E(G, G') \pm 2^h\theta n$, as J' is θ-bounded wrt G', so by Lemma 2.2 whp $X_E(G, G'') = (1 \pm \theta_0)p^{|H'_s \setminus H'[F]|}X_E(G, G')$. The lemma now follows from Lemma 3.12. □

4. Integral designs

In this section we generalise to the context of typical complexes the results of Graver and Jurkat and of Wilson on the module structure of integral designs.

4.1. Null designs.

We start by discussing the original result. Suppose $0 \leq r \leq q \leq n$. We write $M = M(n, q, r)$ for the incidence matrix with rows indexed by $\binom{[n]}{q}$ and columns indexed by $\binom{[n]}{r}$, where M_{ef} is 1 if $f \subseteq e$ or 0 otherwise. A *null design* with parameters (n, q, r) is a vector $v \in \mathbb{Z}^{\binom{[n]}{q}}$ with $vM = 0$, i.e. an assignment of integers to the q-subsets of $[n]$ such that for every r-subset e the sum of integers over all q-subsets containing e is zero. One motivation for considering null designs is that the difference of any two designs with the same parameters is a null design.

Let $W = W(n, q, r)$ denote the set of null designs with parameters (n, q, r); it is an abelian group under addition. The rank of W (as a free $\mathbb{Z}$-module) is $\binom{n}{q} - \binom{n}{r}$ for $r \leq q \leq n - r$ or 0 otherwise (see [14, 13]). An integral basis is provided by a simple construction which we now describe, specialising to $W = W(n, r, r-1)$, which is the only case that we need.

Definition 4.1. The *r-octahedron* $O(r)$ is the r-complex on $\{(j, x) : j \in [r], x \in \{0, 1\}\}$ generated by all $e_x = \{(i, x_i) : i \in [r]\}$ with $x \in \{0, 1\}^r$. We write $O = O(r)$ when r is clear from the context. The *sign* of e_x is $s(e_x) = (-1)^{\sum_{i \in [r]} x_i}$. We write $e_0 = [r] \times \{0\}$ and $e_1 = [r] \times \{1\}$.

To any embedding ϕ of $O(r)$ in $\binom{[n]}{\leq r}$ we can associate $v \in \mathbb{Z}^{\binom{[n]}{q}}$ supported on $\phi(O(r)_r)$, in which the entry for $\phi(e_x)$ is $s(e_x)$. Graver and Jurkat [14] show that one can choose a subset of these embeddings that forms an integral basis of W. In particular, the octahedra generate W; we will show that this also holds more generally in typical complexes.

First we introduce some general notation that will be used throughout the remainder of the paper; In the following definition we will often take $q = r$ and $H = O(r)$, with signs $s(e)$ as in Definition 4.1 for $|e| = r$ and $s(e) = 0$ otherwise.

Definition 4.2. Let H be a q-complex and $s : H \to \{-1, 0, 1\}$. We call (H, s) a *signed q-complex*. Often we suppress s and refer to H as a signed complex. Given a q-complex G, $\Phi \in \mathbb{Z}^{X_H(G)}$ and $H' \subseteq H$ we define the *H'-boundary* of Φ by

$$\partial_{H'}\Phi = \sum_{\phi \in X_H(G)} \Phi(\phi) \sum_{e \in H'} s(e)\phi(e) \in \mathbb{Z}^G.$$

We write $\partial\Phi = \partial_{H'}\Phi$ if H' is clear from the context. If for some $r \leq q$ we have $s(e) = 0$ whenever $|e| \neq r$ then we also view $\partial_{H'}\Phi$ as an element of $\mathbb{Z}^{G_r}$. We let

$$H^+ = \{e \in H : s(e) = 1\} \quad \text{and} \quad H^- = \{e \in H : s(e) = -1\}.$$

We define $\partial_+\Phi = \partial_{H^+}\Phi^+ - \partial_{H^-}\Phi^-$ and $\partial_-\Phi = \partial_{H^+}\Phi^- - \partial_{H^-}\Phi^+$.

Note that $\partial\Phi = \partial_+\Phi - \partial_-\Phi$, with all positive contributions to edges recorded by $\partial_+\Phi$ and all negative contributions by $\partial_-\Phi$. When comparing this with the identity $\partial\Phi = (\partial\Phi)^+ - (\partial\Phi)^-$, one should note that there may be cancellations between $\partial_+\Phi$ and $\partial_-\Phi$. We also note that $\partial_{H'} : \mathbb{Z}^{X_H(G)} \to \mathbb{Z}^G$ is $\mathbb{Z}$-linear for any $H' \subseteq H$. Next we introduce some more convenient terminology for null designs.

Definition 4.3. Suppose G is a q-complex and $0 \leq s \leq r \leq q$. We say that $J \in \mathbb{Z}^{G_r}$ is *s-null* if $|J^+(e)| = |J^-(e)|$ for all $e \in G_s$. We say that J is *null* if it is $(r-1)$-null.

Equivalently, J is s-null if $\partial_{K_r^s} J = 0$. Here we think of K_r^s as the signed r-complex $([r]^{\leq}, s)$, where $s(e) = 1$ if $|e| = s$ and $s(e) = 0$ otherwise. For future reference, we remark that for any q-complex G and $\Phi \in \mathbb{Z}^{X_q(G)}$ we have $\partial_+\Phi = \partial\Phi^+$ and $\partial_-\Phi = \partial\Phi^-$. We note some further simple properties of the definition.

Proposition 4.4.

(i) If $\Phi \in \mathbb{Z}^{X_O(G)}$ then $\partial\Phi$ is null.

(ii) If $J \in \mathbb{Z}^{G_r}$ is s-null and $e \in G_{\leq s}$ then $J(e) \in \mathbb{Z}^{G(e)_{r-|e|}}$ is $(s - |e|)$-null.

Proof. Any $e \in O(r)_{r-1}$ is contained in one edge of sign $+1$ and one edge of sign -1, so (i) holds by linearity of ∂. Also, for any $f \in G(e)_{s-|e|}$ we have $|J(e)^+(f)| = |J^+(e \cup f)| = |J^-(e \cup f)| = |J(e)^-(f)|$, so (ii) holds. $\square$

One can reformulate Proposition 4.4(i) in terms of boundary operators: we have $\partial_{K_r^{r-1}}\partial_{O(r)} = 0$. Next we reformulate the result of Graver and Jurkat, adding a boundedness condition that we need later. It can be viewed as an exactness statement for boundary operators: we have $\mathrm{Im}(\partial_{O(r)}) = \mathrm{Ker}(\partial_{K_r^{r-1}})$.

Lemma 4.5. *Let G be a complete r-complex on $R \geq 2r$ vertices, let $h \gg r$ and suppose $J \in \mathbb{Z}^{G_r}$ is null. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ with $|\Phi| \leq h|J^+|$ such that $\partial\Phi = J$.*

Proof. Recall that the module of such J has a basis consisting of octahedra. Writing B for the matrix with columns equal to the basis vectors, the required Φ is the solution of $B\Phi = J$ (which is integral). Let B_0 be a square submatrix of B with full rank and let $H \subseteq G'_r$ index the rows of B_0. Then $\Phi = B_0^{-1}J[H]$, so we deduce $|\Phi| \leq h|J^+|$. □

4.2. **Octahedral decomposition.** In this subsection we prove the main lemma of the section, which generalises the results of Graver and Jurkat and of Wilson to typical complexes, and has several additional features that will be useful when we apply it later: informally, we can control boundedness if J is bounded, ensure simplicity if J is simple, and use bad sets only when forced to do so. We formalise these in the following definitions.

Definition 4.6. We say $J \in \mathbb{Z}^{G_r}$ is *N-simple* if $|J(e)| \leq N$ for all $e \in G_r$. We say $\Phi \in \mathbb{Z}^{X_H(G)}$ is *N-simple (for H)* if $\partial_+\Phi$ and $\partial_-\Phi$ are N-simple. We say that J or Φ is *simple* if it is 1-simple.

Definition 4.7. Suppose G is a q-complex and $r \in [q]$. We say $J \in \mathbb{Z}^{G_r}$ is *θ-bounded (wrt G)* if J^+ and J^- are θ-bounded wrt G.

Definition 4.8. Suppose G is an r-complex, $B \subseteq G_r$ and $\Phi \in \mathbb{Z}^{X_O(G)}$. We say that Φ is *B-avoiding* if there is no $e \in B$ with $\partial_+\Phi(e) \neq 0$ and $\partial_-\Phi(e) \neq 0$.

Lemma 4.9. *Let $1/n \ll \theta \ll \theta' \ll c, d \ll 1/h \ll 1/r$ and $\eta \ll d$.*

Suppose G is a (c,h)-typical r-complex on n vertices with $d_i(G) > d$ for $i \in [r]$, and $J \in \mathbb{Z}^{G_r}$ is null and θN-bounded, for some $N \geq 1$. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $\partial\Phi = J$ and $\partial_\pm\Phi$ are $\theta' N$-bounded.

Furthermore, if $B \subseteq G_r$ is η-bounded then Φ can made B-avoiding, and if J is N-simple then Φ can made N-simple.

The following informal terminology will occasionally be helpful. Suppose $\partial\Phi = J$ and $e \in G_r$. We think of $|J(e)|$ uses of e by octahedra contributing to e with the same sign as $J(e)$ as being 'forced' uses. We think of the remaining uses of e as 'unforced': these cancel, as there are an equal number of uses with each sign. Thus Φ is B-avoiding if there are no unforced uses of edges in B.

The idea of the proof of Lemma 4.9 is to iteratively reduce J by subtracting null designs $\partial\Phi$, taking care to control boundedness until the support is so small that one can afford to apply Lemma 4.5. We start by introducing a general random extension process that will be used in the proof and throughout the remainder of the paper.

Definition 4.10. Suppose (G, G') and (H, H') are q-complex pairs, $F \subseteq V(H)$, $N \geq 1$, $\mathcal{B} = (B^i : i \in [t])$ with $B^i \subseteq G$, and $\mathcal{E} = (E_i : i \in [t])$ is a sequence of rooted extensions $E_i = (\phi_i, F, H, H')$ in (G, G').

The *$(\mathcal{E}, N, r, \mathcal{B})$-process* is the random sequence $\Phi^* = (\phi_i^* : i \in [t])$, where ϕ_i^* is an embedding of (H, H') in (G, G') such that $(\phi_i^*)|_F = \phi_i$, and letting

C^i be the set of $e \in G_r$ such that $e \in \phi_j^*(H_r \setminus H[F])$ for N choices of $j < r$, we choose ϕ_i^* uniformly at random subject to $\phi_i^*(f) \notin B^i \cup C^i$ for $f \in H \setminus H[F]$. If there is no such choice of ϕ_i^* then the process aborts.

For $f \in H$, the *f-boundary* of Φ^* is $\partial_f \Phi^* = \sum_{i\in[t]} \phi_i^*(f) \in \mathbb{N}^{G_{|f|}}$. We also write $\partial_f^j \Phi^* = \sum_{i\in[j]} \phi_i^*(f)$ for $j \in [t]$.

We often specify $\mathcal{E}$ in advance and fix $B^i = B$ for $i \in [t]$, but sometimes take ϕ_i and B^i to be random, depending on ϕ_j^* for $j < i$. If $H = H'$ and $G = G'$ then we simplify notation by referring to extensions of complexes rather than complex pairs.

Lemma 4.11. *Let $0 < 1/n \ll \theta \ll \theta' \ll c, d \ll 1/h \ll 1/q \le 1/r \le 1/s$ and $\eta \ll d$. Suppose (G, G') is a (c, h)-typical q-complex pair on n vertices with $d_i(G) > d$ for $i \in [q]$. Let (H, H') be a q-complex pair with $|V(H)| \le h$ and $F \subseteq V(H)$, such that for any $g \in H_r \setminus H[F]$ there is $f \in H_s[F]$ with $f \setminus g \neq \emptyset$. Suppose $\mathcal{E} = (E_i : i \in [t])$ is a (possibly random) sequence of rooted extensions $E_i = (\phi_i, F, H, H')$ in (G, G'), and $\mathcal{B} = (B^i : i \in [t])$ is a (possibly random) sequence with $B^i \subseteq G$. Let $N \ge 1$ and Φ^* be the $(\mathcal{E}, N, r, \mathcal{B})$-process. For $i \in [t]$ let $\mathcal{B}^i$ be the 'bad' event that*

(i) $\partial_f^i \Phi^*$ is $\theta N n^{r-s}$-bounded for all $f \in H_s[F]$ and each $B_j^{i'}$, $i' \le i$, $j \in [q]$ is η-bounded, but
(ii) $\partial_g^i \Phi^*$ is not $\theta' N$-bounded for some $g \in H_r \setminus H[F]$ or the process aborts before step i.

Then whp $\mathcal{B}^t$ does not hold.

Proof. Define a stopping time τ to be the first i for which $\mathcal{B}^i$ holds, or ∞ if there is no such i. We claim whp $\tau = \infty$. We estimate $\mathbb{P}(\tau = i')$ for some $i' < \infty$ as follows. For $i \le i'$ we can assume that $\partial_f^i \Phi^*$ is $\theta N n^{r-s}$-bounded for all $f \in H_s[F]$ and each B_j^i, $j \in [q]$ is η-bounded (otherwise $\mathcal{B}^{i'}$ cannot hold).

Fix $i < i'$ and consider the choice of ϕ_i^*. Let $\theta', \eta \ll \delta \ll d$. Since (G, G') is (c, h)-typical, there are at least $2\delta n^{|V(H)\setminus F|}$ embeddings of (H, H') in (G, G') that restrict to ϕ_i on F. Of these, we claim that at most $\delta n^{|V(H)\setminus F|}$ have $\phi_i^*(g) \in B^i \cup C^i$ for some $g \in H \setminus H[F]$. To see this, note that $\mathcal{B}^i$ does not hold, so $\partial_g^i \Phi^*$ is $\theta' N$-bounded for all $g \in H_r \setminus H[F]$ (so C^i is $2^h\theta'$-bounded). Thus for any $g \in H \setminus H[F]$ there are at most $(2^h\theta' + \eta)n^{|g\setminus F|}$ embeddings ϕ of $(H[F \cup g], H'[F \cup g])$ in (G, G') with $\phi|_F = \phi_i$ and $\phi(g) \in B^i \cup C^i$. Each of these extends to at most $n^{|V(H)\setminus(F\cup g)|}$ choices of ϕ^*. There are at most 2^h choices of g, so the claimed bound follows. Thus we choose ϕ_i^* randomly from at least $\delta n^{|V(H)\setminus F|}$ options.

Next we fix $g \in H_r \setminus H[F]$, $e \in G_r$ and estimate $\partial_g^{i'} \Phi^*(e)$. By assumption there is $f \in H_s[F]$ with $f \setminus g \neq \emptyset$. Let $j = |f \setminus g|$. Since $\partial_f^i \Phi^*$ is $\theta N n^{r-s}$-bounded for $i < i'$, the number of $i < i'$ with $|\phi_i(f) \setminus e| = j$ is at most $2^r n^{j-1} \cdot \theta N n^{r-s+1}$. For each such i, since $|e \setminus \phi_i(f)| = r - s + j$, there are at most $n^{|V(H)|-|F|-(r-s+j)}$ choices of ϕ_i^* with $\phi_i^*(g) = e$. Writing $\mathcal{F}$ for the algebra generated by the previous choices, we have $\mathbb{P}[\phi_i^*(g) = e \mid \mathcal{F}] <$

$2^r\delta^{-1}n^{-(r-s+j)}$. Thus

$$\mathbb{E}[\partial_g^{i'}\Phi^*(e) \mid \mathcal{F}] < 2^r n^{j-1} \cdot \theta N n^{r-s+1} \cdot 2^r\delta^{-1}n^{-(r-s+j)} = 2^{2r}\delta^{-1}\theta N.$$

Now for any $e' \in G_{r-1}$, $|\partial_g^{i'}\Phi^*(e')|$ is $(2^h, \mu)$-dominated with $\mu = 2^{2r}\delta^{-1}\theta N n$. By Lemma 2.7 whp $|\partial_g^{i'}\Phi^*(e')| < \theta' N n$, as required. □

The proof of Lemma 4.9 uses induction on r. In some lemmas below we will use the induction hypothesis. Throughout we use the following hierarchy of constants:

$$1/n \ll \gamma \ll \gamma' \ll \theta \ll \theta' \ll c, d \ll 1/h \ll 1/r \text{ and } \theta', \eta \ll \delta \ll d.$$

We also assume throughout that $N \geq 1$ and G is a (c, h)-typical r-complex on n vertices with $d_i(G) > d$ for $i \in [r]$. Our first application of extension processes is to the following 'simplification' lemma.

Lemma 4.12. *Suppose $J \in \mathbb{Z}^{G_r}$ is θN-bounded. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $J_s = J - \partial\Phi$ is N-simple and $\partial_\pm\Phi$ are $\theta' N$-bounded.*

Proof. We let $H = O(r)$, $F = e_0$ and choose $E_i = (\phi_i, F, H)$ and $s_i \in \{-1, 1\}$ for $i \in [t]$ such that $\sum_{i\in[t]} s_i\phi_i(e_0) = J$. We also let $\mathcal{E} = (E_i : i \in [t])$ and $\Phi^* = (\phi_i^* : i \in [t])$ be the $(\mathcal{E}, N, r, \emptyset)$-process. Then we set $\Phi = \sum_{i\in[t]} s_i\phi_i^*$. Then $J_s = J - \partial\Phi$ is N-simple, and each of $\partial_\pm\Phi$ is dominated by $\sum_{f\in H}\partial_f\Phi^*$, so is whp $\theta' N$-bounded by Lemma 4.11. □

The next lemma allows us to focus the support of J inside a subcomplex G' that is not too sparse compared with G.

Lemma 4.13. *Suppose $J \in \mathbb{Z}^{G_r}$ is null and θN-bounded wrt G and (G, G') is (c, h)-typical with $d_i(G') > \theta' d_i(G)$ for $i \in [r]$. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $\partial_\pm\Phi$ are $\theta' N$-bounded wrt G, $J' = J - \partial\Phi \in \mathbb{Z}^{G'_r}$ and $J'^{\pm}$ are $\theta' N$-bounded wrt G'.*

Proof. We introduce additional constants θ_{ijk} for $0 \leq i, j \leq r$, $k \in [3]$ such that $\theta_{i(j-1)3} \ll \theta_{ij1} \ll \theta_{ij2} \ll \theta_{ij3}$ for $i, j \in [r]$, and $\theta_{001} = \theta$, $\theta_{rr3} = \theta'$ and $\theta_{i0k} = \theta_{(i-1)rk}$ for $i \in [r]$, $k \in [3]$ (for convenient notation). We start by applying Lemma 4.12 to obtain $\Phi^s \in \mathbb{Z}^{X_O(G)}$ such that $J^s = J - \partial\Phi^s$ is N-simple and $\partial_\pm\Phi^s$ are $\theta_{003}N$-bounded wrt G.

We will refer to sets in G' as 'good' and sets in $G \setminus G'$ as 'bad'. We let $G^i = G[G'_{\leq i}]$ be the subcomplex of G consisting of all sets f such that every i-subset of f is good. We let G^{ij} be the set of $f \in G^{i-1}$ such that the union f' of all bad i-sets in f has size at most $r - j$. Note that $G^0 = G$, $G^r = G'$, $G^{i0} = G^{i-1}$ and $G^{ir} = G^i$. By Lemma 3.9 (applied twice) each (G^{i-1}, G^i) is $(9^h c, h)$-typical with $d_j(G^i) = d_j(G')$ for $j \leq i$ and $d_j(G^i) = (1 \pm 3hc)d_j(G)$ for $j > i$.

Now we construct $\Phi' = \sum_{i,j\in[r]} \Phi'^{ij}$, such that defining $J^{ij} = J^{i(j-1)} - \partial\Phi'^{ij}$ for $i, j \in [r]$, with $J^{10} = J^s$ and $J^{i0} = J^{(i-1)r}$ for $2 \leq i \leq r$, we have $J^{ij} \in \mathbb{Z}^{G_r^{ij}}$ for $i, j \in [r]$ that is N-simple and $\theta_{ij1}N$-bounded wrt G. By Proposition 4.4 each such J^{ij} is null.

We construct Φ'^{ij} by the following iterative procedure. Initially we set $\Phi'^{ij} = 0$ and $J^{ij} = J^{i(j-1)} - \partial\Phi'^{ij} = J^{i(j-1)}$. Suppose at some round we have constructed $J^{ij} = J^{i(j-1)} - \partial\Phi'^{ij}$ and we have $J^{ij}(f) \neq 0$ for some $f \in G_r^{i(j-1)} \setminus G_r^{ij}$. Then every subset of f of size less than i is good, and the union f' of the bad i-sets in f has size $r-j+1$. Note that for any $f' \subseteq e \in J^{ij}$, since $e \in G^{i(j-1)}$, the union of the bad i-sets in e is also equal to f', so there are no bad i-sets in e that are not contained in f'. By Lemma 3.11 $G^* := G^{i-1}(f')$ is a $(27^h c, h-r)$-typical $(j-1)$-complex with $d_i(G^*) > d^{2^r}$ for $i \in [j-1]$. Each round will have the property that no new sets in $G_r^{i(j-1)} \setminus G_r^{ij}$ are added to J^{ij}. Thus $J^* := J^{ij}(f') = J^{i(j-1)}(f') \in \mathbb{Z}^{G^*_{j-1}}$ is N-simple and $\theta_{i(j-1)1}N$-bounded wrt G^*.

Below we will choose some $\Phi^{f'} \in \mathbb{Z}^{X_O(G^*)}$ such that $\partial_\pm \Phi^{f'}$ are $\theta_{i(j-1)3}N$-bounded wrt G. We let B be the set of $e \in G^*_{j-1}$ such that there is some $(r-j)$-set $g \subseteq f'$ such that e was in $\partial_\pm \Phi^{f^*}$ for at least $\sqrt{\theta_{i(j-1)3}}Nn$ previous choices of f^* playing the role of f' with $g \subseteq f^*$. Then $|B(e')| < r\sqrt{\theta_{i(j-1)3}}n$ for any $e' \in G^*_{j-2}$, otherwise we would obtain an $(r-j)$-set g and at least $\theta_{i(j-1)3}Nn^2$ pairs (f^*, e) with $g \subseteq f^*$, $e' \subseteq e$ and $e \in \partial_\pm \Phi^{f^*}$, contradicting $\sum_{f^*: g \subseteq f^*} |\partial_\pm \Phi^{f^*}(e')| < \theta_{i(j-1)3}Nn^2$. It follows that B is θ_{ij1}-bounded wrt G'. By the induction hypothesis of Lemma 4.9 there is an N-simple $\Phi^{f'} \in \mathbb{Z}^{X_O(G^*)}$ such that $\Phi^{f'}$ is B-avoiding, $J^* = \partial \Phi^{f'}$ and $\partial_\pm \Phi^{f'}$ are $\theta_{i(j-1)3}N$-bounded wrt G^*.

Now for each $\psi \in \Phi^{f'}$ we choose a random $\phi \in X_O(G)$ such that:

(i) $f' = [j,r] \times \{0\}$,
(ii) $\phi(k,x) = \psi(k,x)$ for $k \in [j-1]$, $x \in \{0,1\}$,
(iii) $\phi(I)$ is good for every $I \in O_i$ with $\phi(I) \not\subseteq f'$, and
(iv) $\phi(e_x) \notin C^{ij}$ for all $x \in \{0,1\}^r$ with $x[j,r] \neq 0$,

where

$$C^{ij} = \{e : |J^{ij}(e)| = N\}.$$

(We will show below that whp there are many such ϕ.) Then we add ϕ to Φ'^{ij} with the same sign as ψ. (Recall our convention that 'each $\psi \in \Phi^{f'}$' means that ψ is considered $|\Phi^{f'}(\psi)|$ times by the algorithm.) By construction, the new J^{ij} is zero on sets containing f', is N-simple, and for every new set $e \in J^{ij}$ the union of the bad i-sets in e is a strict subset of f'. Thus if successful the procedure terminates with an N-simple $J^{ij} \in \mathbb{Z}^{G_r^{ij}}$.

To analyse the procedure, let $\mathcal{B}$ be the 'bad' event that $J^{ij\pm}$ are not $\theta_{ij1}N$-bounded. Define a stopping time τ to be the first round after which $\mathcal{B}$ holds, or ∞ if there is no such round. We will show whp $\tau = \infty$. Consider a round of the procedure where we consider f' as above and fix $\psi \in \Phi^{f'}$. We claim that the number of choices of ϕ satisfying (i–iii) above is equal to $X_E(G^{i-1}, G^i)$, where $E = (\phi_0, F, H, H')$ is defined by $H = O(r)$, $F = F_1 \cup F_2$, where $F_1 = [j,r] \times \{0\}$ and $F_2 = [j-1] \times \{0,1\}$, $H' = O(r)_{\le i} \setminus \binom{F_1}{i}$, $\phi_0(F_1) = f'$ and $\phi_0|_{F_2} = \psi$. To see this, note that $J^{ij} \in \mathbb{Z}^{G_r^{i(j-1)}}$, so $\phi_0(H'[F_1]) \subseteq G^{i-1}_{<i} = G^i_{<i}$. Then, since all bad i-sets in $\phi_0(H'[F])$ are contained in f', we have $\phi_0(H'[F]) \subseteq G^i$. Thus $X_E(G^{i-1}, G^i)$ is well-defined, and the claim is now clear. Recalling that (G^{i-1}, G^i) is $(9^h c, h)$-typical with each $d_k(G^i) > d_k(G')/2$, there are at least $2\delta n^{r-j+1}$ choices for ϕ satisfying (i-iii). To satisfy (iv), since $\mathcal{B}$ does not hold at the start of the round, C^{ij} is

$2\theta_{ij1}$-bounded, so we exclude at most $2|H|\theta_{ij1}n^{r-j+1}$ choices for ϕ, leaving at least δn^{r-j+1} choices.

Now we fix $e \in G_{r-1}$ and analyse the contributions to $J^{ij\pm}(e)$ as follows. We can bound the number of forced uses by $|J^{i(j-1)+}(e)| + |J^{i(j-1)-}(e)| < 2\theta_{i(j-1)1}Nn$. For the unforced uses, fix k and k' and consider those f' with $|e \cap f'| = k$ and those $\psi \in \Phi^{f'}$ with $|e \cap Im(\psi)| = k'$. We can choose $e \cap Im(\psi)$ and $e \cap f'$ in fewer than 2^{2r} ways.

Now consider the number of ways of extending to f' and $e' = \psi(e_x)$ containing $e \cap Im(\psi)$ for some $x \in \{0,1\}^{j-1}$. There are at most 2^{2j} choices for x and $\psi^{-1}|_{e'}$. If $f' \subseteq e$ then f' is given, and there are at most $\theta_{i(j-1)3}n^{r-k-k'}$ choices for e', as $|\partial_\pm \Phi^{f'}(f)| < \theta_{i(j-1)3}n$ for every $f \in G^*_{j-2}$. Otherwise, there are at most $n^{j-1-k'}$ choices for e', and by choice of B at most $\sqrt{\theta_{i(j-1)3}}n^{r-j+1-k}$ choices for f' such that $e' \notin B$, as we are only considering unforced uses of e', and each $\Phi^{f'}$ is B-avoiding.

Since $\Phi^{f'}$ is N-simple there are at most N choices for ψ given e', so the number of choices for such f' and ψ is at most $2^{2r}\sqrt{\theta_{i(j-1)3}}Nn^{r-k-k'}$. Then since there are $r-1-k-k'$ more vertices of e to cover, there are at most $n^{k+k'-j+2}$ choices for the remaining $r-j+1$ vertices of ϕ.

The probability of such a choice (conditional on the history of the procedure) is at most $n^{k+k'-j+2}/\delta n^{r-j+1}$. Summing over k and k', we see that $|J^{ij\pm}(e)|$ is $(1,\mu)$-dominated with $\mu = r^2 2^{6r}\delta^{-1}\sqrt{\theta_{i(j-1)3}}Nn$. By Lemma 2.7 whp $|J^{ij\pm}(e)| < \theta_{ij1}Nn$, so $\tau = \infty$, as claimed. Thus we can construct Φ' with the properties stated above. Setting $\Phi = \Phi^s + \Phi'$ and $J' = J - \partial\Phi = J^{rr}$ we have $J^{rr} \in \mathbb{Z}^{G'_r}$ with $|J^{rr\pm}(e)| < \theta_{rr1}Nn < \theta' N|V(G')|$ for all $e \in G'_{r-1}$. The above estimates also give $|\partial_\pm \Phi(e)| < \theta' Nn$ for every $e \in G_{r-1}$. □

The next lemma will be used at the end of the reduction when the support of J is small enough that we can afford a crude boundedness estimate; it will also play a key role in the cancellation procedure of Lemma 4.15.

Lemma 4.14. *Suppose $J \in \mathbb{Z}^{G_r}$ is null. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $\partial\Phi = J$. Furthermore, if J is θN-bounded then for any $2r$-set X such that $G_r[X]$ is complete we can choose Φ such that $|\partial_\pm\Phi(e)| < \theta' Nn^{|e\cap X|+1}$ for any $e \in G_{r-1}$.*

Proof. Choose $\theta \ll \theta_0 \ll \theta'$. Fix any $2r$-set X such that $G_r[X]$ is complete. Let G' be the set of $f \in G$ such that $e = f' \cup e' \in G$ for any $f' \subseteq f$ and $e' \subseteq X$ such that $|e| \le r$. By Lemma 3.10(ii) repeatedly applied, (G, G') is $(2^{h^2}c, h - 2^{2r})$-typical with $d_i(G') > d^{10^r}$ for $i \in [r]$. By Lemma 4.13 there is $\Phi_0 \in \mathbb{Z}^{X_O(G)}$ such that $\partial_\pm\Phi_0$ are $\theta_0 N$-bounded wrt G, $J^0 = J - \partial\Phi_0 \in \mathbb{Z}^{G'_r}$ and $J^{0\pm}$ are $\theta_0 N$-bounded wrt G'.

Next we construct $\Phi_1 = \sum_{i\in[r+1]}\Phi_1^i$ such that defining $J^i = J^{i-1} - \partial\Phi_1^i$ we have $J^i(f) = 0$ for all $f \in G_r$ with $|f \cap X| < i$. By Proposition 4.4 each such J^i is null. We construct Φ_1^i by the following procedure. Consider any $f \in G_r$ with $J^i(f) \neq 0$ and $|f \cap X| < i$. Then $f' = f \setminus X$ has size $r - i + 1$ and $J' = J^i(f')$ is supported on X by construction of J^{i-1}. By Lemma 4.5 there is $\Phi^{f'} \in \mathbb{Z}^{X_O(G[X])}$ with $|\Phi^{f'}| \le h|J'^+|$ such that $\partial\Phi^{f'} = J'$. For each $\psi \in \Phi^{f'}$ we choose an arbitrary $\phi \in X_O(G)$ such that:

(i) $f' = \{\phi(k,0) : k \in [i,r]\}$,
(ii) $\phi(k,x) = \psi(k,x)$ for $k \in [i-1]$, $x \in \{0,1\}$, and
(iii) $\phi(k,1) \in X$ for $k \in [i,r]$;

any choice of distinct vertices of X is valid in (iii) by definition of G'. Then we add ϕ to Φ^i_1 with the same sign as ψ. Repeating this for all such f we obtain $J^i = J^{i-1} - \partial\Phi^i_1$ with the required property.

Defining $\Phi = \Phi_0 + \Phi_1$ we have $J - \partial\Phi = J^{r+1} = 0$. Also, for any $f' \in G_{r-i+1}$ disjoint from X we have $|J^{i\pm}(f')| \le 2^r|\Phi^{f'}| \le 2^r h|J^{i-1}(f')|$. Then for any $f \in G$ disjoint from X with $|f| \le r-i+1$, summing the previous estimate over $f' \in G_{r-i+1}$ disjoint from X with $f \subseteq f'$ we obtain $|J^{i\pm}(f)| < 2^r h|J^{i-1}(f)|$. Iterating, for any $e \in G_{r-1}$ we have $|\partial_\pm\Phi_1(e)| < (2^r h)^r|J^0(e \setminus X)|$. We deduce that $|\partial_\pm\Phi(e)| < \theta' N n^{|e\cap X|+1}$. □

The next lemma is the engine of the proof: it gives a cancellation procedure for improving the boundedness of J, via randomised rounding of a fractional relaxation.

Lemma 4.15. *Suppose $J \in \mathbb{Z}^{G_r}$ is null and θN-bounded. Then there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $J' = J - \partial\Phi$ is γN-bounded and $\partial_\pm\Phi$ are $\theta' N$-bounded.*

Proof. Let $1/n \ll \nu \ll \gamma$ and $\theta \ll \theta_1 \ll \theta_2 \ll \theta_3 \ll \theta_4 \ll \theta'$. Consider the random r-complex G' where $G'_i = G_i$ for $i < r$ and G'_r is ν-binomial in G. We form $\Phi^1 \in \mathbb{Z}^{X_O(G)}$ by choosing for each $e \in J$ a uniformly random $\phi \in X_O(G)$ with $\phi(e_0) = e$ and $\phi(e_x) \in G'_r$ for all nonzero $x \in \{0,1\}^r$; we add ϕ to Φ^1 with the same sign as e. Then $J^1 = J - \partial\Phi^1 \in \mathbb{Z}^{G'_r}$.

We estimate $|J^1(e')|$ for any $e' \in G_{r-1}$ as follows. Write $e' = \{v_1, \dots, v_{r-1}\}$, fix $I \subseteq [r-1]$, and consider the contribution from $e = \{u_1, \dots, u_r\} \in J$ with $I = \{i \in [r-1] : u_i \ne v_i\}$ and $\phi \in X_O(G)$ with $\phi(i,0) = u_i$ for $i \in [r]$ and $\phi(i,1) = v_i$ for $i \in I$. There are at most $\theta N n^{|I|+1}$ choices for $e \in J$, and by Lemma 2.2 whp at most $2\nu^{2^{|I|}-1}\theta N n^{|I|+1}$ of these have $\{v_i : i \in I'\} \cup \{u_i : i \in [r] \setminus I'\} \in G'_r$ for all $\emptyset \ne I' \subseteq I$. For each such e there are at most $n^{r-|I|}$ choices of ϕ as above with $e' \in \phi(O)$, and by Lemma 2.2 whp at most $2\nu^{2^r - 2^{|I|}} n^{r-|I|}$ of these have $\{v_i : i \in I'\} \cup \{u_i : i \in [r] \setminus I'\} \in G'_r$ for all $I' \not\subseteq I$. By Lemma 3.19 whp G' is $(2c,h)$-typical with $d_r(G') = (1 \pm c)\nu d_r(G)$. Thus the total number of choices for ϕ is at least $\nu^{2^r-1}\delta n^r$, so $\mathbb{P}(e' \in \phi(O)) < 2\delta^{-1}\nu^{1-2^{|I|}} n^{-|I|}$. Then

$$\mathbb{E}|J^1(e')| < 2\nu^{2^{|I|}-1}\theta N n^{|I|+1} \cdot 2\delta^{-1}\nu^{1-2^{|I|}} n^{-|I|} = 4\delta^{-1}\theta N n,$$

so by Lemma 2.2 whp $|J^1(e')| < \theta_1 N n$ for all $e' \in G_{r-1}$.

Next we find a fractional version of the desired construction. Let $\mathcal{X}$ be the set of $2r$-sets X such that $G'_r[X]$ is complete; we have $|\mathcal{X}| > \delta n^{2r}$ by typicality. For each $X \in G'_r$ we apply Lemma 4.14 to $J^1 \in \mathbb{Z}^{G'_r}$ obtaining $\Phi^X \in \mathbb{Z}^{X_O(G')}$ such that $\partial\Phi^X = J^1$ and $|\partial_\pm\Phi^X(e)| < \theta_2 N n^{|e\cap X|+1}$ for any $e \in G'_{r-1}$. Let $\Psi = |\mathcal{X}|^{-1}\sum_{X\in\mathcal{X}}\Phi^X \in \mathbb{Q}^{X_O(G')}$. Then $\partial\Psi = J^1$ (defining ∂ by the same formula as for integer vectors). Also, for any $e' \in G'_{r-1}$ and $0 \le j \le r$ there are at most n^{2r-j} choices of X with $|e' \cap X| = j$, so $|\partial_\pm\Psi(e')| < |\mathcal{X}|^{-1}\sum_{j=0}^r n^{2r-j} \cdot \theta_2 N n^{j+1} < \theta_3 N n$.

We obtain $\Phi^2 \in \mathbb{Z}^{X_O(G)}$ from Ψ by the following randomised rounding procedure. For each $\phi \in X_O(G)$ with $\Psi(\phi) \geq 0$, we write $m_\phi = \lfloor \Psi(\phi) \rfloor$, $p_\phi = \Psi(\phi) - m_\phi$ and let $\Phi^2(\phi) = m_\phi + X_\phi$, where X_ϕ is p_ϕ-Bernoulli. Similarly, for each $\phi \in X_O(G)$ with $\Psi(\phi) < 0$, we write $m_\phi = \lfloor |\Psi(\phi)| \rfloor$, $p_\phi = |\Psi(\phi)| - m_\phi$ and let $\Phi^2(\phi) = -m_\phi - X_\phi$, where X_ϕ is p_ϕ-Bernoulli. We take all the X_ϕ's to be independent. Writing s_ϕ for the sign of $\Psi(\phi)$, we have $\Phi^2(\phi) = s_\phi(m_\phi + X_\phi)$ and $\Psi(\phi) = s_\phi(m_\phi + \mathbb{E}X_\phi)$. Thus $\mathbb{E}\Phi^2(\phi) = \Psi(\phi)$, and so $\mathbb{E}\partial\Phi^2(e) = J^1(e)$ for all $e \in G'_r$.

Let $J_* = \partial\Phi^2 - J^1$. Then for any $e \in G'_r$ we can write

$$J_*(e) = \sum s(e_x)s_\phi(X_\phi - \mathbb{E}X_\phi),$$

where we sum over all (ϕ, x) such that $\phi(e_x) = e$. Thus we can write $|J_*(e)| \leq |Y_1 - \mu_e| + |Y_2 - \mu_e|$ for some μ_e, where Y_1 and Y_2 are pseudobinomial variables, both having mean μ_e. Since $\mathbb{E}X_\phi \leq |\Psi(\phi)|$ for all ϕ we have $\mu_e \leq \partial_+\Psi(e) + \partial_-\Psi(e)$. Then for any $e' \in G_{r-1}$ we have $\sum_{e:e' \subseteq e} \mu_e < 2\theta_3 Nn$.

For each $e \in G'_r$ we have $\mathbb{E}|J_*(e)| < h\sqrt{\mu_e}$ by Lemma 2.3. Also, for each $e' \in G_{r-1}$ we have $|G'_r(e')| < 2\nu n$, as G' is $(2c, h)$-typical with $d_r(G') < 1.1\nu$. Then by the inequality of power means

$$\begin{aligned}\mathbb{E}|J_*(e')| &< \sum_{e:e' \subseteq e} h\sqrt{\mu_e} \leq h|G'_r(e')|(\theta_3 Nn/|G'_r(e')|)^{1/2} \\ &< h(2\nu n \cdot \theta_3 Nn)^{1/2} < \nu^{1/3} Nn.\end{aligned}$$

Also, any rounding decision in Φ^2 affects $|J_*(e')|$ by at most 1, so by Lemma 2.10 whp $|J_*(e')| < 2\nu^{1/3} Nn$ for all $e' \in G_{r-1}$. Similarly, whp $|\partial_\pm \Phi^2(e')| < 2\theta_3 Nn$ for all $e' \in G_{r-1}$. Setting $\Phi = \Phi^1 + \Phi^2$ and $J' = J - \partial\Phi$ we have $|J'(e)| < \gamma Nn$ and $|\partial_\pm \Phi(e)| < \theta' Nn$ for every $e \in G_{r-1}$. □

So far we have not addressed the simplicity and avoiding conditions of Lemma 4.9; this is handled by our last preliminary lemma.

Lemma 4.16. *Suppose $\Phi \in \mathbb{Z}^{X_O(G)}$ such that $J = \partial\Phi$ is N-simple and $\partial_\pm\Phi$ are θN-bounded, and $B \subseteq G_r$ is η-bounded. Then there is $\Psi \in \mathbb{Z}^{X_O(G)}$ such that Ψ is N-simple and B-avoiding, $\partial\Psi = J$ and $\partial_\pm\Psi$ are $\theta' N$-bounded.*

Proof. We apply the following iterative procedure to eliminate certain 'bad' pairs of octahedra. Suppose at some round we have defined some $\Psi \in \mathbb{Z}^{X_O(G)}$ with $\partial\Psi = J$, but we are not done, because Ψ is not N-simple or not B-avoiding. Then we can choose a *bad pair* (ϕ, ϕ'), which consists of two signed elements of Ψ, that both use some edge e with opposite signs, such that $e \in B$ or e appears in more than N octahedra in Ψ. We will now show how to replace $\phi + \phi'$ by some $\Delta \in \mathbb{Z}^{X_O(G)}$ that is equivalent, in that $\partial\Delta = \partial(\phi + \phi')$, such that e does not appear in any octahedron in Δ.

Let H be the r-complex on $[r] \times [4]$ generated by all r-sets of the form $\{(i, x_i) : i \in [r]\}$ for some $x \in [4]^{[r]}$. By identifying $[4]$ with $\{0,1\}^2$ we can decompose H into 2^r octahedra O_x, $x \in \{0,1\}^r$, where O_x is generated by all r-sets of the form $e_{xy} = \{(i, (x_i, y_i))\}$ for some $y \in \{0,1\}^r$. We assign e_{xy} the sign $s(e_{xy}) = (-1)^{\sum x_i + \sum y_i}$, so that for each octahedron O_x we take the usual sign of an edge multiplied by $s(e_x)$. Note that the sign of every r-set is invariant if in any coordinate we apply a transposition to the identification, namely whenever $j \in [4]$ was identified with $(a, b) \in \{0, 1\}^2$ we identify it instead with (b, a).

Now consider $\Phi^H \in \mathbb{Z}^{X_O(H)}$ obtained taking embeddings ϕ_x of O as O_x with sign $s(e_x)$ for all $x \in \{0, 1\}^r$, and also embeddings ϕ'_x of O as O'_x with sign $-s(e_x)$ for all $x \in \{0, 1\}^r$, where the O_x are defined with respect to some fixed identification, and the O'_x are defined with respect to the identification obtained applying transpositions in every coordinate of $[r]$. Note that $\partial\Phi^H = 0$, as each r-set occurs once with each sign. The pair of octahedra containing the edge e_{00} (which is the same in both identifications) will play a special role below: with respect to the first identification, O_0 uses $(0, 0)$ and $(0, 1)$ in every coordinate of $[r]$, whereas O'_0 uses $(0, 0)$ and $(1, 0)$ in every coordinate of $[r]$.

Now we choose $\psi_i \in X_H(G)$ for $i \in [3]$ such that $\psi_1(O_0) = \phi(O)$ with $\psi_1(e_0) = e$, $\psi_2(O_0) = \phi'(O)$ with $\psi_2(e_0) = e$, $\psi_3(O_0) = \psi_1(O'_0)$ and $\psi_3(O'_0) = \psi_2(O'_0)$ with $\psi_3(e_0) = e$, and there are no other identifications of vertices in $Im(\psi_i)$, $i \in [3]$. We choose ψ_1, ψ_2, ψ_3 uniformly at random subject to no edge of $\cup_{i\in[3]}\psi_i(H)$ outside of $\phi(O) \cup \phi'(O)$ being used N times by any octahedron in Φ. Finally, we let

$$\Delta = \phi + \phi' + s_1\psi_1(\Phi^H) + s_2\psi_2(\Phi^H) + s_3\psi_3(\Phi^H),$$

choosing signs $s_1, s_2, s_3 \in \{-1, 1\}$ such that $s_1\psi_1(O_0)$ cancels $\phi(O)$, $s_2\psi_2(O_0)$ cancels $\phi'(O)$ and $s_3(\psi_3(O_0)-\psi_3(O'_0))$ cancels $-s_1\psi_1(O'_0)-s_2\psi_2(O'_0)$; to see that s_3 is well-defined, note that e has opposite signs in $\psi_3(O_0)$ and $-\psi_3(O'_0)$, and by choice of (ϕ, ϕ') also in ϕ and ϕ', so in $s_1\psi_1(O_0)$ and $s_2\psi_2(O_0)$ and so in $s_1\psi_1(O'_0)$ and $s_2\psi_2(O'_0)$. Then $\partial\Delta = \partial(\phi+\phi')$, as $\partial\Phi^H = 0$. Also, e does not appear in any octahedron in Δ, as the coefficients of all such octahedra cancel by construction.

We can describe the choice of ψ_i, $i \in [3]$ by a rooted extension $E = (\phi^*, F^*, H^*)$, where H^* consists of 3 copies of H that have the same identifications as $\psi_i \in X_H(G)$ for $i \in [3]$ and are otherwise disjoint, and F^* is such that ϕ^* identifies $H[F^*]$ with $\phi(O) \cup \phi'(O)$. Note that E depends on how $\phi(O)$ and $\phi'(O)$ intersect outside of e: there are r mutually non-isomorphic possibilities. Thus we can describe the algorithm by r versions of the $(\mathcal{E}, N, r, B)$-process running in parallel, where each $\mathcal{E}$ has the form $\mathcal{E} = (E_i : i \in [t])$, and $E_i = (\phi^*_i, F^*, H^*)$ is such that $\phi^*_i[F^*]$, $i \in [t]$ runs through all bad pairs of octahedra of the isomorphism type corresponding to $\mathcal{E}$. Here we note that the set of edges e for which we choose a bad pair on e is fixed at the start of the algorithm, as we do not create any new such edges, but the set of octahedra present on e when we consider e depends on the history of the algorithm, and there may be many choices for a bad pair.

Each of $\partial_\pm\Phi$ is a sum of some edge-boundaries of the $(\mathcal{E}, N, r, B)$-processes Φ^*, so whp $\theta' N$-bounded by Lemma 4.11; we choose $e_0 \in H[F]$ when applying the lemma, noting that $\partial_{e_0}\Phi^* \leq \partial_+\Psi + \partial_-\Psi$ is $2\theta N$-bounded. By construction, $\partial\Psi = J$ throughout, and when the algorithm terminates, Ψ is N-simple and B-avoiding. □

Finally, we can prove the main lemma of this section.

Proof of Lemma 4.9. The argument is by induction on r, as the result for smaller r was used in some lemmas above. The result is trivial for $r = 0$, so we suppose $r \geq 1$. (In fact, it is not hard to give a direct non-inductive argument for the case $r = 1$, but we omit the details.)

Choose $\theta \ll \theta_0 \ll \theta'$. First we construct Ψ such that $\partial\Psi = J$ and $\partial_\pm\Psi$ are $2\theta_0 N$-bounded. We do so by an iterative process. At the start of round t we will have $U^t \subseteq V(G)$ with $|U^t| = \theta^t n \geq n^{1/3r}$ and $G^t = G[U^t]$ such that (G^t, G^{t+1}) is $(2c, h)$-typical with $d_i(G^t) = (1 \pm n^{-1/9r})d_i(G)$ for $i > 1$. We claim that the required property holds whp if we choose U^t uniformly at random. For by Lemma 3.18 whp (G, G^t) is $(c + n^{-1/9r}, h)$-typical with $d_i(G^t) = (1 \pm n^{-1/9r})d_i(G)$ for $i > 1$. By definition, the same is true of G^t, and so (G^t, G^t). Then again by Lemma 3.18 whp (G^t, G^{t+1}) is $(c + 2n^{-1/9r}, h)$-typical, as claimed.

We will also have $J^t \in \mathbb{Z}^{G^t_r}$ that is null and θN-bounded wrt G^t. We take $J^0 = J$, so this holds for $t = 0$ by assumption. Now suppose that we have J^t for some $t \geq 0$ and $|U^{t+1}| \geq n^{1/3r}$. By Lemma 4.15 there is $\Phi^{t1} \in \mathbb{Z}^{X_O(G^t)}$ such that $J^{t1} = J^t - \partial\Phi^{t1}$ is γN-bounded wrt G^t and $\partial_\pm\Phi^{t1}$ are $\theta_0 N$-bounded wrt G^t. By Lemma 4.13, applied with (γ, θ) in place of (θ, θ'), there is $\Phi^{t2} \in \mathbb{Z}^{X_O(G)}$ such that $\partial_\pm\Phi^{t2}$ are θN-bounded wrt G^t, $J^{t2} = J^{t1} - \partial\Phi^{t2} \in \mathbb{Z}^{G^{t+1}_r}$ and $J^{t2\pm}$ are θN-bounded wrt G^{t+1}. Setting $\Phi^t = \Phi^{t1} + \Phi^{t2}$ and $J^{t+1} = J^t - \partial\Phi^t = J^{t2}$ we have the required properties for the next step.

At the first $t = T$ where we have $|U^{t+1}| < n^{1/3r}$, we instead apply Lemma 4.14 to choose $\Phi^t \in \mathbb{Z}^{X_O(G)}$ such that $\partial\Phi^t = J^t$ and $|\partial_\pm\Phi^t(e)| < Nn^{(r+1)/3r}$ for any $e \in G^t_{r-1}$. Then we set $\Psi = \sum_{t=0}^T \Phi^t$, so $\partial\Psi = J$. Now for any $e \in G_{r-1}$ we have $|\partial_\pm\Psi(e)| = \sum_{t=0}^T |\partial_\pm\Phi^t(e)| \leq Nn^{(r+1)/3r} + \sum_{t=0}^{T-1} \theta_0 N(\theta^t n) < 2\theta_0 Nn$. Finally, by Lemma 4.16 applied to Ψ, there is $\Phi \in \mathbb{Z}^{X_O(G)}$ such that Φ is N-simple and B-avoiding, $\partial\Phi = J$ and $\partial_\pm\Phi$ are $\theta' N$-bounded. □

4.3. **Integral designs.** In the remainder of the section we give an application of Lemma 4.9 to a result on integral designs. We start with a family of constructions that generalise the octahedra described above.

Definition 4.17.

(i) Suppose H is a signed q-complex. The *suspension* SH of H is the signed $(q+1)$-complex obtained from H by adding two new vertices x^+ and x^-, adding all $e^+ = e \cup \{x^+\}$ and $e^- = e \cup \{x^-\}$ for $e \in H$, and defining $s(e^+) = s(e)$, $s(e^-) = -s(e)$ and $s(e') = 0$ if $e' \subseteq V(H)$.

(ii) Suppose G is a q-complex and $\Phi \in \mathbb{Z}^{X_H(G)}$. We define the *suspension* $S\Phi \in \mathbb{Z}^{X_{SH}(SG)}$ of Φ by $S\Phi = \sum_{\phi \in X_H(G)} \Phi(\phi) S\phi$, where for $\phi \in X_H(G)$ we define $S\phi \in X_{SH}(SG)$ by $S\phi|_H = \phi$ and $S\phi(x^\pm) = x^\pm$. We also write $\phi^\pm = S\phi|_{V(SH)\setminus\{x^\mp\}}$.

We note the following properties of the definition.

Proposition 4.18.

(i) $S : \mathbb{Z}^{X_H(G)} \to \mathbb{Z}^{X_{SH}(SG)}$ *is* $\mathbb{Z}$*-linear,*

(ii) $SO(r) = O(r+1)$,

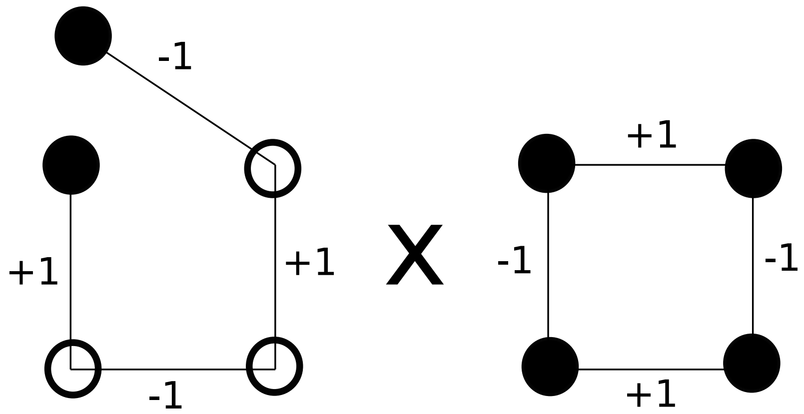

FIGURE 1. The unmodified $(4,3)$-move.

(iii) $\partial S = S\partial$.

Proof. Statement (i) is clear from the definition. For (ii), we can identify $SO(r)$ with $O(r+1)$ by mapping x^+ to $(r+1,0)$, x^- to $(r+1,1)$ and (j,x) to itself for all $j \in [r]$, $x \in \{0,1\}$; by construction each edge has the correct sign. For (iii), by linearity of S and ∂ it is enough to show that $\partial S\phi = S\partial\phi$ for all $\phi \in X_H(G)$. This follows from

$$\partial S\phi = \sum_{e' \in SH} s(e')S\phi(e') = \sum_{e \in H} s(e)(\phi^+(e^+) - \phi^-(e^-)), \quad \text{and}$$

$$S\partial\phi = S\sum_{e \in H} s(e)\phi(e) = \sum_{e \in H} s(e)(\phi(e)^+ - \phi(e)^-). \qquad \square$$

Next we describe a construction that implements an r-octahedron as a signed combination of K^r_q's. The case when $q=4$ and $r=3$ is illustrated in Figure 1. It is a signed 4-complex, in which there are 16 quadruples, obtained by taking all possible unions of a pair on the left with a pair on the right, with sign given by the product of the signs of the pairs. Note that every triple is contained in at most one quadruple of each sign, and the K^3_4 boundary is a signed 3-octahedron supported on the six solid discs.

Definition 4.19. Let G^{q1} be the q-complex on $[q]^2 \cup \{\infty\}$ generated by the 'columns' $c_j = \{(i,j) : i \in [q]\}$ for $j \in [q]$, the 'rows' $r_i = \{(i,j) : j \in [q]\}$ for $i \in [2,q]$, and $r'_1 = \{(1,j) : j \in [2,q]\} \cup \{\infty\}$. We define

$$\Phi_{q1} = \sum_{j \in [q]} c_j - \sum_{i \in [2,q]} r_i - r'_1 \in \mathbb{Z}^{X_q(G^{q1})},$$

where we identify each $e \in G^{q1}_q$ with an arbitrary fixed bijective map from $[q]$ to e.

In general, we define G^{qr} and $\Phi_{qr} \in \mathbb{Z}^{X_q(G^{qr})}$ inductively as follows. We let $G^{qr} = SG^{(q-1)(r-1)}$. For $\phi \in X_{q-1}(G^{(q-1)(r-1)})$ we consider ϕ^+ or ϕ^- as an element of $X_q(G^{qr})$ by identifying x^+ or x^- with the vertex q of $[q]^{\leq}$. For $r > 1$ we obtain $\Phi_{qr} \in \mathbb{Z}^{X_q(G^{qr})}$ from $\Phi_{(q-1)(r-1)}$ by including ϕ^+, ϕ^-

for each $\phi \in \Phi_{(q-1)(r-1)}$, respectively with the same, opposite sign to ϕ. We refer to Φ_{qr} as the *unmodified* (q,r)*-move*.

In the next section we will modify Φ_{qr} to define the general (q,r)-move, which will have the additional property of being simple wrt K_q^r. The following proposition shows that we do need need any modification in the case $q = r+1$.

Proposition 4.20. *$\Phi_{(r+1)r}$ is simple wrt K_{r+1}^r.*

Proof. We argue by induction. For the base case we have $\Phi_{21} = c_1 + c_2 - r_1' - r_2$. Clearly each vertex is covered at most once with each sign, so Φ_{21} is simple wrt K_2^1. For the induction step, suppose that $r \geq 1$ and $\Phi_{(r+1)r}$ is simple wrt K_{r+1}^r. Consider $e' \in G^{(r+2)(r+1)} = SG_{r+1}^{(r+1)r}$. Suppose first that $e' = e^+$ for some $e \in G_r^{(r+1)r}$. Then

$$\partial_\pm \Phi_{(r+2)(r+1)}(e') = \partial \Phi^\pm_{(r+2)(r+1)}(e') = \partial \Phi^\pm_{(r+1)r} e = \partial_\pm \Phi_{(r+1)r} e \leq 1.$$

Similarly, if $e' = e^-$ for some $e \in G^{(r+1)r}$ then

$$\partial_\pm \Phi_{(r+2)(r+1)}(e') = \partial \Phi^\pm_{(r+2)(r+1)}(e') = \partial \Phi^\mp_{(r+1)r} e = \partial_\mp \Phi_{(r+1)r} e \leq 1.$$

Finally, suppose $e' \in G_{r+1}^{(r+1)r}$, and note that there is a unique $\phi \in \Phi_{(r+1)r}$ with $\phi([r+1]) = e'$; this property clearly holds for all r by induction, and is the place where the argument would fail for $q > r+1$. Then ϕ^+ and ϕ^- are the only elements of $\Phi_{(r+2)(r+1)}$ whose images contain e', and they have opposite signs. □

We also note the following key property that holds for any q and r.

Proposition 4.21. *$\partial_{K_q^r} \Phi_{qr} = O(r)_r$.*

Proof. When $r = 1$ we have $\partial \Phi_{q1} = (1,1) - \infty$, which is $O(1)$ under the identification of $(1,1)$ with 0 and ∞ with 1. For $r > 1$, we have $\partial \Phi_{qr} = \partial S \Phi_{(q-1)(r-1)}$, as writing $K = ([q-1]^{\leq}, s)$ for K_{q-1}^{r-1} as a signed complex, for each $\phi \in \Phi_{(q-1)(r-1)}$ we have $\partial_{(SK)_r} S\phi = \partial_{K_q^r} \phi^+ - \partial_{K_q^r} \phi^+$. Then by Proposition 4.18 and induction we have $\partial \Phi_{qr} = S \partial \Phi_{(q-1)(r-1)} = SO(r-1)_r = O(r)_r$. □

We conclude by showing the existence of integral decompositions under the divisibility conditions. Our result is tailored to the intended application later, so we only consider J with multiplicities 0 or 1 (i.e. a subgraph). In this case, we obtain an expression that can be viewed as the difference of two K_q^r-decompositions, which is a key part of the proof strategy discussed in the introduction. We can also maintain boundedness if J is bounded.

In fact, we will not use Lemma 4.22 in the proof of the main theorem, but Lemma 5.28 instead (which is its 'linear version'). However, we include the proof below as a warmup to the later result, as it contains many of the main ideas but avoids some additional complications. For simplicity we only give the proof in the case $q = r+1$; for the general case one can apply the same proof using the general (q,r)-move defined in the next section.

If f is an embedding of a q-complex H in G and $\Phi \in \mathbb{Z}^{X_q(H)}$ we define $f(\Phi) = \sum_{\phi \in X_q(H)} \Phi(\phi)(f \circ \phi) \in \mathbb{Z}^{X_q(G)}$.

Lemma 4.22. *Let $1/n \ll \theta \ll \theta' \ll c, d \ll 1/h \ll 1/q \le 1/r$. Suppose G is a (c,h)-typical q-complex on n vertices with $d_i(G) > d$ for $i \in [q]$ and $J \subseteq G_r$ is θ-bounded and K^r_q-divisible. Then there is a simple $\Phi \in \mathbb{Z}^{X_q(G)}$ such that $\partial_{K^r_q}\Phi = J$ and $\partial_\pm \Phi$ are θ'-bounded.*

Proof assuming $q = r+1$**.** Let $\theta \ll \theta_0 \ll \theta'_0 \ll \cdots \ll \theta_r \ll \theta'_r \ll \theta'$. We construct $\Phi = \sum_{i=0}^r \Phi^i$ such that $\partial_\pm \Phi^i$ are θ_i-bounded, and defining $J^0 = J - \partial\Phi^0$ and $J^i = J^{i-1} - \partial\Phi^i$ for $i \in [r]$, each J^i is simple and i-null. (When unspecified, ∂ means $\partial_{K^r_q}$.) To construct Φ^0 we sum $\binom{q}{r}^{-1}|J|$ elements of $X_q(G)$, chosen sequentially uniformly at random from $X_q(G)$ subject to remaining simple, i.e. whenever we choose some ϕ in future choices we do not allow any ϕ' with $\partial\phi' \cap \partial\phi \ne \emptyset$. Then $J^0 = J - \partial\Phi^0$ is simple and 0-null and whp $\partial_\pm\Phi^0$ are θ_0-bounded (this follows from the proof of Lemma 4.11; we omit the details).

Now we construct Φ^i given J^{i-1} for $i \ge 1$. Note that J^{i-1} is K^r_q-divisible, so $\binom{q-i}{r-i}$ divides $|J^{i-1}(e)|$ for all $e \in G_i$. Then $J' := \binom{q-i}{r-i}^{-1}\partial_{K^i_r} J^{i-1} \in \mathbb{Z}^{G_i}$. For any $e \in G_{i-1}$ we have

$$|J'^{\pm}(e)| = (r-i+1)\binom{q-i}{r-i}^{-1}|J^{(i-1)\pm}(e)|.$$

This implies that J' is null (as J^{i-1} is $(i-1)$-null) and J' is $\theta_{i-1}n^{r-i}$-bounded (as J^{i-1} is θ_{i-1}-bounded). By Lemma 4.9 with $N = n^{r-i}$ there is an n^{r-i}-simple $\Psi \in \mathbb{Z}^{X_{O(i)}(G)}$ such that $\partial_{O(i)}\Psi = J'$ and $\partial_\pm\Psi$ are $\theta'_{i-1}n^{r-i}$-bounded.

Next we let $H = G^{qi}$ be as in Definition 4.19, and $F \subseteq V(H)$ be such that $\partial_{K^i_q}\Phi_{qi}$ is $O(i)_i$ on F (by Proposition 4.21). We choose $E_u = (\phi_u, F, H)$ and $s_u \in \{-1, 1\}$ for $u \in [t]$ such that $\Psi = \sum_u s_u \phi_u$. We also let $\mathcal{E} = (E_u : u \in [t])$ and $\mathcal{B} = (B^u : u \in [t])$, where B^u is the set of e such that $\partial_+\Phi^j(e)$ or $\partial_-\Phi^j(e)$ is non-zero for some $j < u$, and also all $e \in J' = J^{r-1}$ if $i = r$. We let $\Phi^* = (\phi^*_u : u \in [t])$ be the $(\mathcal{E}, 1, r, \mathcal{B})$-process. Then we set $\Phi^i = \sum_{u\in[t]} s_u \phi^*_u(\Phi_{qi}) \in \mathbb{Z}^{X_q(G)}$.

By Lemma 4.11 (with $N = 1$, $s = i$ and $\theta = \theta'_{i-1}$) whp $\partial_\pm\Phi^i$ are θ_i-bounded. Also, $\sum_{j=0}^i \Phi^j$ is simple by Definition 4.10 and Proposition 4.20 (assuming $q = r+1$). Furthermore, for any $e \in G_i$ we have

$$|\partial_{K^r_q}\Phi^i(e)| = \binom{q-i}{r-i}\partial_{K^i_q}\Phi^i(e) = \partial_{K^i_r}J^{i-1}(e) = |J^{i-1}(e)|,$$

since $\partial_{K^i_q}\Phi^i = \partial_{O(i)}\Psi = J'$. Thus $J^i = J^{i-1} - \partial\Phi^i$ is simple and i-null.

Finally, $\Phi = \sum_{i=0}^r \Phi^i$ is simple, $J^r = J - \partial\Phi$ is r-null, so identically zero, and for $e \in G_{r-1}$ we have $\partial_\pm\Phi(e) \le \sum_{i=0}^r \partial_\pm\Phi^i(e) \le \sum_{i=0}^r \theta_i n \le \theta' n$. □

5. Linear designs

This section develops the algebraic ingredients of our Randomised Algebraic Construction.

5.1. Setting and notation.

Throughout we consider a q-complex G on $n = p^a$ vertices, where p is a fixed prime and a is large. We identify $V(G)$ with the field $\mathbb{F}_{p^a}$, which we also view as a vector space over the subfield $\mathbb{F}_p$. We introduce the following notation.

Definition 5.1. For $S \subseteq \mathbb{F}_{p^a}$ we let $\dim(S)$ be the dimension of the $\mathbb{F}_p$-subspace generated by S. Similarly, for $x \in \mathbb{F}_{p^a}^s$ for some s we let $\dim(x)$ be the dimension of the $\mathbb{F}_p$-subspace generated by $\{x_1, \dots, x_s\}$.

In the following notation we suppress the dependence on s, which is to be understood from the context.

Definition 5.2. For any $s \geq 1$ we let e^i denote the ith unit vector in $\mathbb{F}_p^s$, where $e^i_i = 1$ and $e^i_j = 0$ for $j \neq i$. For $I \subseteq [s]$ let e^I be the matrix with rows indexed by I with e^i in row $i \in I$.

We also require some notation for matrices.

Definition 5.3. Let M be an r by c matrix in $\mathbb{F}_{p^a}^{r\times c}$. We write $Row(M) \subseteq \mathbb{F}_{p^a}^c$ for the row space of M and $Col(M) \subseteq \mathbb{F}_{p^a}^r$ for the column space of M.

Note that e^iM is the ith row of M and Me^i is the ith column of M. Similarly, e^IM for $I \subseteq [r]$ is the submatrix of M induced by the rows in I, and Me^I for $I \subseteq [c]$ is the submatrix of M induced by the columns in I.

The spaces of the previous definition are to be understood with respect to the field $\mathbb{F}_{p^a}$, e.g. $Row(M)$ is the set of sums $\sum_{i=1}^r a_i e^i M$ with $a_i \in \mathbb{F}_{p^a}$ for $i \in [r]$; if we want to restrict attention to the base field $\mathbb{F}_p$ then we will say so explicitly.

5.2. **Linear forms and extensions.** In this subsection we present some properties of linear forms and define the linear analogue of extensions. We start with some notation and terminology for linear forms.

Definition 5.4. Let $z = (z_i : i \in [g])$ be indeterminates, and $L = (L_v : v \in X)$, where the L_v are $\mathbb{F}_{p^a}$-linear forms in z.

- (i) Let $M(L)$ be the matrix with rows indexed by X and columns indexed by z, where $M(L)_{vz_i}$ is the coefficient of z_i in L_v. Let $C(L)$ be the column vector indexed by X where $C(L)_v$ is the constant term in L_v. Let $M^+(L)$ be the matrix obtained from $M(L)$ by adding $C(L)_v$ as a column indexed by 1.
- (ii) Write
$$Im(L) = \{M^+(L)y : y \in \mathbb{F}_{p^a}^{g+1}, y_{g+1} = 1\}$$
and $\dim(L) = \text{rank}(M(L))$ for the affine $\mathbb{F}_{p^a}$-dimension of $Im(L)$. For $S \subseteq X$ we write $L_S = (L_v : v \in S)$ and
$$Span(S) = \{v \in X : M(L)^v \in Row(M(L_S))\}.$$
- (iii) We say that L is *simple* if $L_v \neq L_{v'}$ for $v \neq v'$.
- (iv) We say that L is *basic* if $M(L)$ has entries in $\mathbb{F}_p$.

Definition 5.5.

- (i) Let $L = (L_v : v \in X)$ be linear forms in variables $z = (z_i : i \in [g])$ and $L^z = (L^z_i : i \in [g])$ be linear forms in variables $y = (y_i : i \in [g'])$. We define the *specialisation* $L[L^z] = (L[L^z]_v : v \in X)$ of L to L^z by
$$L[L^z]_v(y) = L_v(L^z(y)).$$
- (ii) If $\dim(C(L), C(L^z)) = \dim(C(L)) + g$ we say that L^z is *generic wrt* L, and $L[L^z]$ is a *generic specialisation* of L.

(iii) If $\text{rank}(M(L^z)) = \dim(L)$ we say L^z is a *change of variables* to y.

We note the following properties of specialisations.

Proposition 5.6.

(i) $M(L[L^z]) = M(L)M(L^z)$ *and* $C(L[L^z]) = M(L)C(L^z) + C(L)$.

(ii) If $L[L^z]$ *is obtained by change of variables then* $Im(L[L^z]) = Im(L)$.

(iii) If L *is basic,* $a \in \mathbb{F}_p^X$ *with* $a \cdot L \neq 0$ *and* L^z *is generic wrt* L *then* $a \cdot L[L^z] \neq 0$.

(iv) If L *is basic and simple and* L^z *is generic wrt* L *then* $L[L^z]$ *is simple.*

Proof. To see (i), note that $L[L^z]_v(y) = M(L)_v(M(L^z)y + C(L^z)) + C(L)_v$ for $v \in X$. We deduce (ii), as $\dim(L[L^z]) = \text{rank}(M(L[L^z])) = \text{rank}(M(L)) = \dim(L)$. For (iii), either $a \cdot L$ is constant, in which case $a \cdot L[L^z] = a \cdot L \neq 0$, or $aM(L) \neq 0$, so $a \cdot L[L^z]$ has constant term $aM(L)C(L^z) + a \cdot C(L)$, which is an $\mathbb{F}_p$-linear form in the entries of $C(L)$ and $C(L^z)$ with not all coefficients of $C(L^z)$ equal to zero, so is non-zero by genericity. Now (iv) follows from (iii), where a ranges over all vectors in which one entry is 1, one entry is -1, and the other entries are 0. □

The next lemma shows the existence of specialisations in which we specify the values of some of the forms, and gives a dimension formula for such specialisations.

Lemma 5.7. *Suppose* $L = (L_v : v \in X)$ *are linear forms in variables* $z = (z_i : i \in [g])$. *Let* $S \subseteq X$ *and* $e \in Im(L_S)$. *Then there is a specialisation* $L[L^z]$ *of* L *such that* $Im(L[L^z]) = \{v \in Im(L) : v[S] = e\}$, *which is basic if* L *is basic. For* $S' \subseteq X$ *let*

$$R(L)^{SS'} = Row(M(L_S)) \cap Row(M(L_{S'})).$$

Then

$$\dim(L[L^z]_{S'}) = \dim(L_{S'}) - \dim(R(L)^{SS'}) = \dim(L_{S\cup S'}) - \dim(L_S).$$

Proof. For the first statement, let $(M \mid M')$ be a submatrix of $M(L)$ whose columns form a basis for $Col(M(L))$ such that the columns of $M'[S]$ form a basis of $Im(L_S) - e$. Let (y, y') be the subsequence of z consisting of variables corresponding to columns in $(M \mid M')$, and C^z be the change of variables to (y, y').

Since $Col(M \mid M') = Col(M(L))$, for any $i \in [g]$ we can choose a^i such that $(M \mid M')a^i = M(L)e^i$; then we have $C^z_{z_i} = a^i \cdot (y, y')$. Note that there is a unique choice of $y' = c'$ such that $L[C^z]_S(y, c') = e$ for all y. Let L^z in variables y be obtained by substituting $y' = c'$ in C^z. Then $Im(L[L^z]) = \{v \in Im(L) : v[S] = e\}$. If L is basic then each a^i has coefficients in $\mathbb{F}_p$, so $L[L^z]$ is basic.

For the second statement, fix $I^* \subseteq I^S, I^{S'} \subseteq I^X$ such that the rows of $M(L_{I^*})$ form a basis of $R(L)^{SS'}$ and for $Y \in \{S, S', X\}$ the rows of $M(L_{I^Y})$ form a basis of $Row(M(L_Y))$. Then $Im(L_{I^Y}) = \mathbb{F}_{p^a}^{|I^Y|}$ for each $Y \in \{S, S', X\}$ and $Im(L_{I^Y})$ is in bijection with $Im(L_Y)$. In particular, there is a unique choice of $L_{I^S} \in \mathbb{F}_{p^a}^{|I^S|}$ corresponding to choosing $L_S = e$. This fixes

L_v for $v \in I^*$, so $\dim(L[L^z]_{S'}) = |I^{S'} \setminus I^*| = \dim(L_{S'}) - \dim(R(L)^{SS'}) = \dim(L_{S \cup S'}) - \dim(L_S)$. □

We conclude this subsection by defining the linear analogue of extensions. We remark that the permutation σ in the definition is a convenient device; we will later take it to be the identity without loss of generality.

Definition 5.8.

(i) Suppose G is a q-complex with $V(G) = \mathbb{F}_{p^a}$ and σ is a permutation of $\mathbb{F}_{p^a}$. Let H be a q-complex, $z = (z_i : i \in [g])$ be indeterminates, and $L = (L_v(z) : v \in V(H))$ be distinct linear forms in z. Let F be the set of $v \in V(H)$ such that $L_v(z) = L_v$ is constant, and suppose that $v \mapsto \sigma(L_v)$ is an embedding of $H[F]$ in G. We say that $E = (L, H)$ is a *linear extension* in G (wrt σ) with *base* F and *variables* z. We say that an injection $\phi : V(H) \to V(G)$ is an *L-embedding* of H in G (wrt σ) if $\phi(f) \in G$ for all $f \in H$, and for some $y \in \mathbb{F}_{p^a}^g$ we have $\phi(v) = \sigma(L_v(y))$ for all $v \in V(H)$.

(ii) Now suppose also that G' is a subcomplex of G and H' is a subcomplex of H, such that $v \mapsto \sigma(L_v)$ is an embedding of $H'[F]$ in G'. We say that $E = (L, H, H')$ is a *linear extension* in (G, G') (wrt σ). We say that an L-embedding of H in G (wrt σ) is an *L-embedding* of (H, H') in (G, G') (wrt σ) if we also have $\phi(f) \in G'$ for all $f \in H'$.

(iii) We say that E is *r-generic* if for every $e \in H'_r$ there exists $y \in \mathbb{F}_{p^a}^g$ such that $\dim(L_e(y)) = r$.

(iv) We say that E is *basic* if its linear forms are basic.

For future reference we note some properties of genericity.

Proposition 5.9. *Suppose $E = (L, H, H')$ is a linear extension in (G, G') with base F and variables $z = (z_i : i \in [g])$.*

(i) All but $O(n^{g-1})$ choices of $z = y$ define an L-embedding that is a generic specialisation of E.

(ii) If E is r-generic and basic and $z = y$ defines an L-embedding that is a generic specialisation of E then $\dim(L_e(y)) = r$ for all $e \in H'_r$.

Proof. For (i), we can choose the entries of y sequentially, noting that in each step, all but $O(1)$ choices are not in the linear span of the previous choices and the constant terms in L. For (ii), note that for any $e \in H'_r$ and nonzero $a \in \mathbb{F}_p^e$, we have $a \cdot L_e(z) \neq 0$ as E is r-generic, so $a \cdot L_e(y) \neq 0$ by Proposition 5.6(iii). □

5.3. **Templates.** In this subsection we define the template for our decomposition, and show that is both typical and 'linearly typical' (to be defined). The construction requires a matrix with the following genericity property.

Definition 5.10. Let $M \in \mathbb{F}_p^{q \times r}$ be a q by r matrix. We say that M is *generic* if every square submatrix of M is non-singular.

It is easy to see that if $p > p_0(q, r)$ is large enough then there is a generic matrix M (e.g. consider a random matrix and take a union bound of the probabilities that any square submatrix is singular). Henceforth we assume that M is a fixed generic q by r matrix.

Next we will define the template. The intuition is that the 'model' for the template is the set S of all vectors My with $y \in \mathbb{F}_{p^a}^r$. Indeed, S has the following property that is reminiscent of K_q^r-decompositions: for every r-set e and injective map $\pi_e : e \to [q]$, we can reconstruct the unique y such that $(My)_i = x$ for all $x \in e$, $i = \pi_e(x)$, namely $y = (e^I M)^{-1} v$, where $I = \pi_e(e)$ and $v_i = x$ for $x \in e$, $i = \pi_e(x)$.

However, the model has three significant defects: (i) any e is covered many times by S, corresponding to the choices of π_e, (ii) we are considering a complex that is typical rather than complete, so some of the above sets may not belong to the complex, and (iii) some vectors My do not have distinct coordinates, so do not correspond to q-sets.

However, we can at least obtain a partial K_q^r-decomposition that covers a constant fraction of the r-sets, by randomly permuting the vertex set, only allowing y with $\dim(y) = r$, and making each r-set e randomly 'decide' on some fixed injection π_e. Now we give the definition; it is well-defined by Proposition 5.12, and again, we will later take σ to be the identity.

Definition 5.11. Suppose G is a q-complex with $V(G) = \mathbb{F}_{p^a}$. Let $M \in \mathbb{F}_p^{q \times r}$ (be generic). Let $\pi = (\pi_e : e \in G_r)$, where $\pi_e : e \to [q]$ are uniformly random injections, and σ be a uniformly random permutation of $\mathbb{F}_{p^a}$, all random choices being independent. Let $D = D(G, M, \pi, \sigma)$ be the set of all $\phi \in X_q(G)$ such that

(i) $\phi^{-1}(x) = \pi_e(x)$ for all $x \in e \in \partial_{K_q^r} \phi$, and
(ii) for some $y \in \mathbb{F}_{p^a}^r$ with $\dim(y) = r$ we have $\phi(i) = \sigma(e^i M y)$ for all $i \in [q]$.

The *M-template* of G is the random q-complex $G(M)$ with $G(M)_r = \partial_{K_q^r} D$ and $G(M)_i$ for $i < r$ consisting of all $e \in G_i$ such that $\dim(\sigma^{-1}(e)) = i$, and $G(M)_i$ for $i > r$ is defined by $G(M) = G[G(M)_{\leq r}]$.

For $e \in G(M)_r$ we let $\phi^e \in D$ be such that $e \in \partial_{K_q^r} \phi^e$. We write $M(e) = \phi^e([q])$ and $M(e)_i = \phi^e(i)$ for $i \in [q]$. We let $M(G)$ be the set of all $M(e)$ with $e \in G(M)_r$.

The key property of the template, given in the first part of the following proposition, is that its q-sets form a K_q^r-decomposition of its underlying r-graph; this shows that ϕ^e is unique in Definition 5.11. The second part of the proposition shows that $G(M)$ is a complex.

Proposition 5.12.

(i) $M(G)$ is a K_q^r-decomposition of $G(M)_r$,
(ii) $\dim(\sigma^{-1}(e)) = r$ for all $e \in G(M)_r$.

Proof. For (i), consider any $e \in G(M)_r$. Let $v(e) \in \mathbb{F}_{p^a}^r$ have the elements of e as its coordinates, in the order induced by π_e. Let $I = \pi_e(e) \subseteq [q]$ and $y(e) = (e^I M)^{-1} \sigma^{-1}(v(e)) \in \mathbb{F}_{p^a}^r$. Then the unique $\phi^e \in D$ such that $e \in \partial_{K_q^r} \phi^e$ is given by $\phi^e(i) = \sigma(e^i M y(e))$ for $i \in [q]$. For (ii), note that $\sigma^{-1}(v(e)) = e^I M y(e)$, where $e^I M$ is non-singular and $\dim(y(e)) = r$ by Definition 5.11. Then for any $c \in \mathbb{F}_p^r$ with $c \cdot \sigma^{-1}(v(e)) = 0$ we have $c e^I M y(e) = 0$, so $c e^I M = 0$, so $c = 0$. □

Next we will define typicality of linear extensions with respect to a template. We require the following definition which takes account of the edge correlations inherent in the construction.

Definition 5.13. Suppose G^* is a value of the M-template $G(M)$ given π and σ. Let $E = (L, H, H^*)$ be a linear extension in (G, G^*) with base F and variables z. Suppose $\mathcal{M} \subseteq H^*_q$ with $|f \cap F| < r$ for all $f \in \mathcal{M}$ and for each $f \in \mathcal{M}$ let $\sigma_f : f \to [q]$ be a bijection. We say that E is M*-closed* if each $e \in H^*_r \setminus H[F]$ is contained in a unique $M(e) \in \mathcal{M}$, and for each $f \in \mathcal{M}$ there are linear forms $L^f = (L^f_i : i \in [r])$ in z such that $L_v(z) = e^i M L^f(z)$ for all $v \in f$, $i = \sigma_f(v)$.

Now we can define the linear analogue of typicality.

Definition 5.14. Suppose G^* is a value of the M-template $G(M)$ given π and σ and $E = (L, H, H^*)$ is a linear extension in (G, G^*) with base F. Let $X_E(G, G^*)$ be the set or number of L-embeddings ϕ^* of (H, H^*) in (G, G^*). Let $UE = (\phi, F, H, H^*)$ be the underlying rooted extension of E, where $\phi(v) = \sigma(L_v)$ for $v \in F$. Let

$$z_{qr} = \prod_{i=1}^{r} (q + 1 - i)^{-1} \quad \text{and} \quad Z_E = z_{qr}^{|H^*_r \setminus H^*[F]| \binom{q}{r}}.$$

We say that E is M*-linearly* c*-typical (in* (G, G^*)*)* if

$$X_E(G, G^*) = (1 \pm c) Z_E \pi_{UE}(G, G) n^{\dim(L)}.$$

We say that (G, G^*) is M*-linearly* (c, h, r)*-typical* if every M-closed r-generic linear extension in (G, G^*) of size at most h is c-typical.

Next we show that whp the template is both typical and linearly typical.

Lemma 5.15. *Let* $0 < 1/n \ll c, d \ll 1/h \ll 1/p \ll 1/q \le 1/r$ *with* $n = p^a$*. Suppose that* G *is a* (c, h)*-typical* q*-complex with* $V(G) = \mathbb{F}_{p^a}$ *and* $d_i(G) > d$ *for* $i \in [q]$*. Let* $G(M)$ *be the* M*-template wrt random* π *and* σ*. Then whp* $(G, G(M))$ *is* M*-linearly* $(2hc, h, r)$*-typical and* $(3^h c, h/q)$*-typical with* $d_i(G(M)) = (1 \pm 9qhc) d_i(G)$ *for* $i \in [q] \setminus \{r\}$ *and* $d_r(G(M)) = (1 \pm 9qhc) d^*_r$*, where*

$$d^*_r = d_r(G) \prod_{i=1}^{r} (q + 1 - i)^{1 - \binom{q}{r}} d_i(G)^{\binom{q}{i} - \binom{r}{i}}.$$

Proof. Consider any M-closed r-generic linear extension $E = (L, H, H^*)$ in $(G, G(M))$ of size at most h with base F and variables $z = (z_i : i \in [g])$, where we condition on $\sigma(L_v)$, $v \in F$. By change of variables (Definition 5.5) we can assume $g = \dim(L)$. Fix any $\phi \in X_{UE}(G, G)$ and $y \in \mathbb{F}^g_{p^a}$ such that $z = y$ defines a generic specialisation of E. Note that $v \mapsto L_v(y)$ is injective by Proposition 5.6(iv). In order for ϕ to be realised as an L-embedding $v \mapsto \sigma(L_v(y))$ of (H, H^*) in $(G, G(M))$ we need $\sigma(L_v(y)) = \phi(v)$ for all $v \in V(H)$ and $\pi_e(u) = \sigma_{M(f)}(v)$ for all $v \in f \in H^*_r \setminus H^*[F]$, $u = \phi(v)$, $e = \phi(f)$ (with notation as in Definitions 5.13 and 5.14). Under the random choice of σ and π, this occurs with probability $(1 + O(1/n)) n^{-|V(H) \setminus F|} Z_E$.

Let X' be the number of choices of ϕ and y as above such that $v \mapsto \sigma(L_v(y)) = \phi(v)$ is an L-embedding of (H, H^*) in $(G, G(M))$. There are

$X_{UE}(G,G) = (1 \pm c)^h n^{|V(H)\setminus F|} \pi_{UE}(G,G)$ choices for ϕ and $(1+O(1/n))n^g$ choices for y by Proposition 5.9(i), so

$$\mathbb{E}X' = (1 \pm 1.1hc) n^g Z_E \pi_{UE}(G,G).$$

Next note that any transposition of σ affects X' by $O(n^{g-1})$, as for any $v \in V(H) \setminus F$ and $x \in \mathbb{F}_{p^a}$ there are n^{g-1} choices of y with $\sigma(L_v(y)) = x$. Also, conditional on σ, for any $f \in H^*_r \setminus H^*[F]$, there are at most $n^{\dim(L_f)}$ sets $e = L_f(y)$, and for each such e, by Lemma 5.7 any transposition of π_e affects X' by $O(n^{g-\dim(L_f)})$. Thus by Lemma 2.14 whp

$$X' = (1 \pm 1.2hc) n^g Z_E \pi_{UE}(G,G).$$

Furthermore, we can bound $|X' - X_E(G,G(M))|$ by the number of choices of $\phi \in X_{UE}(G,G)$ and $y \in \mathbb{F}^g_{p^a}$ such that $z = y$ does not define a generic specialisation of E and $\sigma(L_v(y)) = \phi(v)$ for all $v \in V(H)$. There are $O(n^{g-1})$ choices for y, so by Lemma 2.14 whp $|X' - X_E(G,G(M))| = O(n^{g-1})$. Thus whp

$$X_E(G,G(M)) = (1 \pm 2hc) n^g Z_E \pi_{UE}(G,G),$$

so $(G,G(M))$ is M-linearly $(2hc,h,r)$-typical.

Now consider any simple rooted extension $E = (\phi, F, H, H^*)$ in $(G,G(M))$ with $|V(H)| \le h/q$. We condition on $\sigma^{-1}(x)$, $x \in \phi(F)$. Let H', $H^{*\prime}$ be the q-complexes obtained from H, H^* by adding for each $e \in H^*_r \setminus H^*[F]$ a set S^e of $q-r$ new vertices and making $M(e) = e \cup S^e$ a copy of $[q]^{\le}$. Let $\mathcal{M} = \{M(e) : e \in H^*_r \setminus H^*[F]\}$, fix any bijections $\sigma_f : f \to [q]$, $f \in \mathcal{M}$ and let $E' = (L', H', H^{*\prime})$ be the linear extension in $(G,G(M))$ with variables $z = (z_v : v \in V(H) \setminus F)$, where

$$\begin{aligned}
L'_v &= \sigma^{-1}(\phi(v)) && \text{for } v \in F,\\
L'_v &= z_v && \text{for } v \in V(H) \setminus F, \text{ and}\\
L'_v &= e^i M(e^I M)^{-1} L'_e && \text{for } e \in H^*_r \setminus H^*[F],\ v \in M(e),\\
& && I = \sigma_{M(e)}(e),\ i = \sigma_{M(e)}(v).
\end{aligned}$$

Note that for any $\phi' \in X_{E'}(G,G(M))$ we have $\phi'|_{V(H)} \in X_E(G,G(M))$. Also, E' is M-closed by construction, and r-generic as M is generic.

Conversely, for any $\phi \in X_E(G,G(M))$, conditioning on π_f for $f \in \phi(H^*_r \setminus H^*[F])$, for any choice of bijections $\sigma_{f'} : f' \to [q]$, $f' \in \mathcal{M}$ with

$$\sigma_{M(e)}|_e = \pi_{\phi(e)} \circ \phi|_e \ \text{ for all } \ e \in H^*_r \setminus H^*[F],$$

by definition of the M-template we obtain $\phi' \in X_{E'}(G,G(M))$ by defining $\phi'(x) = M(f)_i$ for $x \in e \in H^*_r \setminus H^*[F]$, $f = \phi(e)$, $i = \sigma_{M(e)}(x)$.

There are $z_{qr}^{-|H^*_r \setminus H^*[F]|}$ distinct versions of $X_{E'}(G,G(M))$ that contribute to $X_E(G,G(M))$, corresponding to the choices of $\sigma_{M(f)}|_f$ for $f \in \phi(H^*_r \setminus H^*[F])$, so since $(G,G(M))$ is M-linearly $(2hc,h,r)$-typical we have

$$\begin{aligned}
X_E(G,G(M)) &= (1 \pm 3hc) n^{\dim(L')} z_{qr}^{-|H^*_r \setminus H^*[F]|} Z_{E'} \pi_{UE'}(G,G)\\
&= (1 \pm 3hc) n^{|V(H)\setminus F|} \pi_E(G,G) (d^*_r / d_r(G))^{|H^*_r \setminus H^*[F]|}.
\end{aligned}$$

The lemma now follows from Lemma 3.8. □

5.4. **Linear designs.** For the remainder of the section we fix

$$0 < 1/n \ll \theta \ll c, d \ll 1/h \ll 1/p \ll 1/q \leq 1/r$$

with $n = p^a$ and $\theta \ll \theta', \eta \ll \delta \ll d$, a (c,h)-typical q-complex G with $V(G) = \mathbb{F}_{p^a}$ and $d_i(G) > d$ for $i \in [q]$, and a value G^* of the M-template $G(M)$ satisfying Lemma 5.15; we write $M^* = M(G)$, and assume without loss of generality that σ is the identity.

In this subsection we establish the linear analogue of Lemma 4.22. We start by generalising our previous boundedness definitions to the linear setting.

Definition 5.16. Let A be an affine space over $\mathbb{F}_{p^a}$ and let W be the vector space of which it is a translate. We say that A is *basic* if W has a basis of vectors with all coefficients in $\mathbb{F}_p$. For $v \in \mathbb{F}_{p^a}^s$ and $d \in \mathbb{F}_p^s$ we write $L(v,d)$ for the basic line consisting of all $v + \mu d$ with $\mu \in \mathbb{F}_{p^a}$.

Since p is fixed and $n = p^a$ is large, it is helpful to think of the number of basic affine spaces of any fixed dimension as a constant independent of n.

Definition 5.17.

(i) We say that $J \in \mathbb{N}^{\mathbb{F}_{p^a}^s}$ is *linearly θ-bounded* if $\sum_{e \in L} J(e) < \theta n$ for any basic line L in $\mathbb{F}_{p^a}^s$. We say that $J \in \mathbb{Z}^{\mathbb{F}_{p^a}^s}$ is *linearly θ-bounded* if J^+ and J^- are linearly θ-bounded.

(ii) For $J \in \mathbb{Z}^{G_s}$ we define $v(J) \in \mathbb{Z}^{\mathbb{F}_{p^a}^s}$ by summing for each $e \in J$ (with multiplicity and sign) all vectors whose set of coordinates is e (in any order). We say that J is *linearly θ-bounded* if $v(J)$ is linearly θ-bounded.

We note the following basic properties of linear boundedness.

Proposition 5.18.

(i) If $J \in \mathbb{N}^{G_s}$ is linearly θ-bounded then it is θ-bounded.

(ii) If $J \in \mathbb{N}^{\mathbb{F}_{p^a}^s}$ is linearly θ-bounded and A is a basic affine subspace of $\mathbb{F}_{p^a}^s$ then $\sum_{e \in A} v(J)(e) < \theta|A|$.

Proof. For (i), consider any $e \in G_{s-1}$, order e as $(v_1, \dots, v_{s-1})$ and let $L = L(x,d)$ where $x_i = v_i$ for $i \in [s-1]$, $x_s = 0$, $d_i = 0$ for $i \in [s-1]$, $d_s = 1$. Then $|J(e)| = \sum_{e \in L} v(J)(e) < \theta n$. For (ii), fix any basic direction d such that $A + d = A$, and write A as the disjoint union of a set of basic lines $\mathcal{L}$, each in direction d. Then $\sum_{e \in A} v(J)(e) = \sum_{L \in \mathcal{L}} \sum_{e \in L} v(J)(e) < \sum_{L \in \mathcal{L}} \theta|L| = \theta|A|$. □

Next we require a linear analogue of Lemma 4.11. We will add two further properties from the following definition that will be useful in the proof of our main theorem.

Definition 5.19. Suppose $J \subseteq G_r^*$.

(i) We say that J is *M-simple* if $|J \cap \partial_{K_q^r} M(e)| \leq 1$ for all $e \in G_r^*$.

(ii) We say that J is *M-linearly θ-bounded* if $\cup_{e \in J} \partial M(e)$ is linearly θ-bounded.

To implement these properties, and also a full-dimensionality condition, it is convenient to introduce the following instance of the extension process defined previously.

Definition 5.20. Suppose (H,H') is a q-complex pair, $F \subseteq V(H)$, $N \geq 1$, $\mathcal{B} = (B^i : i \in [t])$ with $B^i \subseteq G$, and $\mathcal{E} = (E_i : i \in [t])$ is a sequence of rooted extensions $E_i = (\phi_i, F, H, H')$ in (G, G^*). The $(\mathcal{E}, N, r, \mathcal{B}, M)$*-process* is the $(\mathcal{E}, N, r, \mathcal{A})$-process, for $\mathcal{A} = (B^i \cup C^i : i \in [t])$ where for $s > r$ we let C^i_s consist of all minimally dependent $e \in G^*_s$, we let C^i_{r+1} consist of all minimally dependent $e \in G^*_{r+1}$ and the union of all $\partial_{K_q^{r+1}} M(e)$, and we let C^i_r be the union of all $\partial_{K_q^r} M(e)$ such that $e \in \phi^*_j(H_r \setminus H[F])$ for N choices of $j < r$.

Note that the C^i of Definition 5.20 contains that of Definition 4.10.

Lemma 5.21. *Let (H,H') be a q-complex pair with $|V(H)| \leq h/q$ and $F \subseteq V(H)$, such that for any $g \in H_r \setminus H[F]$ there is $f \in H_s[F]$ with $f \setminus g \neq \emptyset$. Suppose $\mathcal{E} = (E_i : i \in [t])$ is a (possibly random) sequence of rooted extensions $E_i = (\phi_i, F, H, H')$ in (G, G^*) and $\mathcal{B} = (B^i : i \in [t])$ is a (possibly random) sequence with $B^i \subseteq G$. Let $N \geq 1$ and Φ^* be the $(\mathcal{E}, N, r, \mathcal{B}, M)$-process. For $i \in [t]$ let $\mathcal{B}^i$ be the 'bad' event that*

(i) $\partial^i_f \Phi^*$ is $\theta N n^{r-s}$-bounded for all $f \in H_s[F]$ and each $B^{i'}_j$, $i' \leq i$, $j \in [q]$ is η-bounded, but

(ii) $\partial^i_g \Phi^*$ is not M-linearly $\theta' N$-bounded for some $g \in H_r \setminus H[F]$ or the process aborts before step i.

Then whp $\mathcal{B}^t$ does not hold.

Proof. We define a stopping time τ to be the first i for which $\mathcal{B}^i$ holds, or ∞ if there is no such i. We claim whp $\tau = \infty$. We estimate $\mathbb{P}(\tau = i')$ for some $i' < \infty$ as follows. For $i \leq i'$ we can assume that $\partial^i_f \Phi^*$ is $\theta N n^{r-s}$-bounded for all $f \in H_s[F]$ and each B^i_j, $j \in [q]$ is η-bounded (otherwise $\mathcal{B}^{i'}$ cannot hold).

Fix $i < i'$ and consider the choice of ϕ^*_i. Since (G, G^*) is $(3^h c, h/q)$-typical, there are at least $2\delta n^{|V(H)\setminus F|}$ embeddings of (H,H') in (G,G^*) that restrict to ϕ_i on F. Of these, we claim that at most $\delta n^{|V(H)\setminus F|}$ have $\phi^*_i(g) \in B^i \cup C^i$ for some $g \in H \setminus H[F]$. To see this, first note that for $s > r$, any $e \in G_{s-1}$ is contained in $O(1)$ minimally dependent s-sets. Also, any $e \in G_r$ is contained in at most one $f \in M^*$, so $O(1)$ sets in the union of all $\partial_{K_q^{r+1}} M(e)$. Furthermore, $\mathcal{B}^i$ does not hold, so $\partial^i_g \Phi^*$ is M-linearly $\theta' N$-bounded for all $g \in H_r \setminus H[F]$. Then $B^i \cup C^i$ is $(2^h \theta' + \eta)$-bounded by Definitions 5.19 and 5.20. Thus for any $g \in H \setminus H[F]$ there are at most $(2^h\theta' + \eta) n^{|g \setminus F|}$ embeddings ϕ of $(H[F \cup g], H'[F \cup g])$ in (G, G') with $\phi|_F = \phi_i$ and $\phi(g) \in B^i \cup C^i$. Each of these extends to at most $n^{|V(H)\setminus(F\cup g)|}$ choices of ϕ^*. There are at most 2^h choices of g, so the claimed bound follows. Thus we choose ϕ^*_i randomly from at least $\delta n^{|V(H)\setminus F|}$ options.

Next we fix $g \in H_r \setminus H[F]$, $e \in G_r$ and estimate $\partial^{i'}_g \Phi^*(e)$. By assumption there is $f \in H_s[F]$ with $f \setminus g \neq \emptyset$. Let $j = |f \setminus g|$. Since $\partial^i_f \Phi^*$ is $\theta N n^{r-s}$-bounded for $i < i'$, the number of $i < i'$ with $|\phi_i(f) \setminus e| = j$ is at most $2^r n^{j-1} \cdot \theta N n^{r-s+1}$. For each such i, since $|e \setminus \phi_i(f)| = r - s + j$, there are at most $n^{|V(H)|-|F|-(r-s+j)}$ choices of ϕ^*_i with $\phi^*_i(g) = e$. Writing $\mathcal{F}$ for the algebra generated by the previous choices, we have $\mathbb{P}[\phi^*_i(g) = e \mid \mathcal{F}] < 2^r \delta^{-1} n^{-(r-s+j)}$. Thus $\mathbb{E}[\partial^{i'}_g \Phi^*(e) \mid \mathcal{F}] < 2^{2r} \delta^{-1} \theta N$. Fix any basic line

L in $\mathbb{F}_{p^a}^r$, let $J = \sum_{e \in G_r} \partial_g^{i'} \Phi^*(e) \partial M(e)$ and consider $X = \sum_{e \in L} v(J)(e)$. Applying the previous estimate to each $e' \in G_r$ with $v(e') \cap L \neq \emptyset$ and $e \in \partial M(e')$ we have $\mathbb{E}[X \mid \mathcal{F}] < 2^h \delta^{-1} \theta N n$. Then by Lemma 2.7 whp $X < \theta' N n$, so $\partial_g^i \Phi^*$ is M-linearly $\theta' N$-bounded for all $g \in H_r \setminus H[F]$ as required. □

Next we need a linear analogue of Lemma 4.9.

Lemma 5.22. *Suppose $J \subseteq G_r^*$ is null, M-simple and M-linearly θ-bounded. Then there is $\Phi \in \mathbb{Z}^{X_O(G^*)}$ such that $\partial \Phi = J$ and $\partial_\pm \Phi$ are M-simple and M-linearly θ'-bounded.*

Proof. By Proposition 5.18(i) J is θ-bounded, so by Lemma 4.9 with $N = 1$ there is a simple $\Phi' \in \mathbb{Z}^{X_O(G)}$ such that $\partial \Phi' = J$ and $\partial_\pm \Phi'$ are θ'-bounded. Now we apply the same procedure as in the proof of Lemma 4.16, using the version of the extension process from Definition 5.20. With the same notation as in that proof, we consider the $(\mathcal{E}, 1, r, J, M)$-processes, where $\mathcal{E} = (E_i : i \in [t])$ and $E_i = (\phi_i^*, F^*, H^*)$ is such that $\phi_i^*[F^*]$, $i \in [t]$ runs through all pairs of octahedra using an edge in $\partial_+ \Phi \cap \partial_- \Phi$ (counted with multiplicities), and $\mathcal{E}$ ranges through r isomorphism types according to the intersection of a pair. We obtain Φ from Φ' by replacing each pair (ϕ, ϕ') by Δ as defined in the proof of Lemma 4.16. Then for every unforced use of an edge by an octahedron ϕ in Φ' we have replaced ϕ. Thus Φ is M-simple by Definition 5.20, and each of $\partial_\pm \Phi$ is dominated by the sum of J and the edge boundaries of the processes, so whp M-linearly θ'-bounded by Lemma 5.21. □

Now we describe an operation for rearranging certain K_q^r-decompositions, which we call a 'shuffle'. We will use shuffles to modify the moves introduced in Definition 4.19 so that they become simple. They also motivate the cascades introduced in the next subsection. First we need some notation; we write

$$Q^I = (e^I M)^{-1} e^I \in \mathbb{F}_p^{r \times q}.$$

The following is immediate from the observation that $\text{rank}(MQ) \leq r$.

Proposition 5.23. *For any $Q \in \mathbb{F}_p^{r \times q}$ there are at most r choices of i with $e^i M Q = e^i$. Thus $(Q^I : I \in \binom{[q]}{r})$ are pairwise distinct.*

Next we define the underlying structure of the shuffle.

Definition 5.24. Let $X = (X_i : i \in [q])$ be a q-partite set with $X_i = \{x_i^v : v \in \mathbb{F}_p^q\}$ for $i \in [q]$. For $Q \in \mathbb{F}_p^{r \times q}$ we define q-partite q-sets f^Q and g^Q by $f^Q = \{x_i^{e^i M Q} : i \in [q]\}$ and $g^Q = \{x_i^{e^i M Q + e^i} : i \in [q]\}$. We let H^S be the q-partite q-complex on X generated by all sets f^Q and g^Q. We define $\Phi^S = \sum_Q (f^Q - g^Q) \in \mathbb{Z}^{X_q(H^S)}$, where we identify each $f \in H_q^S$ with an arbitrary fixed bijective map from $[q]$ to f.

The first part of the following lemma shows that $\{f^Q : Q \in \mathbb{F}_p^{r \times q}\}$ and $\{g^Q : Q \in \mathbb{F}_p^{r \times q}\}$ are both K_q^r-decompositions of the complete q-partite r-graph on X. We informally think of a 'shuffle' as the operation of replacing one by the other. It also shows that $\partial_{K_q^r} \Phi^S = 0$, as each q-partite r-set is counted once with each sign. The second part will be used in showing that the modified move is simple.

Lemma 5.25.

(i) Any q-partite r-set f in X is contained in a unique f^Q and a unique $g^{Q'}$.

(ii) Any q-partite $(r+1)$-set f in X is contained in at most one edge in H_q^S.

Proof. For (i), we can determine Q by $Q = (e^I M)^{-1} M^f$, where $f = (x_i^{v^i} : i \in I)$, and M^f has rows v^i, $i \in I$; then Q' is determined by $Q' = Q - Q^I$. For (ii), by (i) we only need to check that $|f^Q \cap g^{Q'}| \le r$ for all $Q, Q' \in \mathbb{F}_p^{r \times q}$. This holds by Proposition 5.23, as $f_i^Q = g_i^{Q'}$ if and only if $e^i M(Q - Q') = e^i$. □

Now we apply shuffles to construct the modified form of the move Φ_{qr} introduced in Definition 4.19. As before, we identify any q-set e with an arbitrary fixed bijective map from $[q]$ to e.

Definition 5.26. Let $G'^{q1} = G^{q1}$ and $\Phi'_{q1} = \Phi_{q1}$ be as in Definition 4.19. We define G'^{qr} and $\Phi'_{qr} \in \mathbb{Z}^{X_q(G'^{qr})}$ inductively as follows. Let $\Phi^*_{qr} \in \mathbb{Z}^{X_q(SG'^{(q-1)(r-1)})}$ be obtained by including ϕ^+, ϕ^- for each $\phi \in \Phi'_{(q-1)(r-1)}$, respectively with the same, opposite sign to ϕ. Let G'^{qr} be obtained from $SG'^{(q-1)(r-1)}$ by adding copies of H^S, such that for each $e \in \Phi^*_{qr}$ there is an embedding $\phi^e : H^S \to G^{qr}$, such that $\phi^e(f^0) = e$ and there are no other identifications of vertices between the copies of H^S. We define the *(q,r)-move* as

$$\Phi'_{qr} = \Phi^*_{qr} - \sum_{e \in SG'^{(q-1)(r-1)}_q} \Phi^*_{qr}(e) \phi^e(\Phi^S) \in \mathbb{Z}^{X_q(G'^{qr})}.$$

For any q and r we have the following key properties analogous to those in Propositions 4.20 and 4.21.

Proposition 5.27. *Φ'_{qr} is simple wrt K_q^r and $\partial_{K_q^r} \Phi'_{qr} = O(r)_r$.*

Proof. To show simplicity, by the same proof as in Proposition 4.20, it suffices to show that for every $e' \in G'^{qr}_{r+1}$ there is at most one $\phi \in \Phi'_{qr}$ with $e' \subseteq \phi([q])$. To see this, first note that every $e \in \Phi^*_{qr}$ does not appear in Φ'_{qr}, as it appears with opposite signs in Φ^*_{qr} and $\phi^e(\Phi^S)$. Thus by Lemma 5.25(ii), there is a unique $e \in \phi^*_{qr}$ and $f \in \phi^e(H_q^S)$ such that $e' \subseteq f$, as required.

Now we prove the second statement of the proposition by induction on r. It holds when $r = 1$ by Proposition 4.21 as $\Phi'_{q1} = \Phi_{q1}$. Next, since Φ^*_{qr} is constructed from $\Phi'_{(q-1)(r-1)}$ in the same way that Φ_{qr} was constructed from $\Phi_{(q-1)(r-1)}$, by the proof of Proposition 4.21 and induction we have $\partial \Phi^*_{qr} = S \partial \Phi'_{(q-1)(r-1)} = SO(r-1)_r = O(r)_r$. Then, as $\partial \Phi^S = 0$, we have $\partial \Phi'_{qr} = \partial \Phi^*_{qr} = O(r)_r$. □

We conclude this subsection with the linear analogue of Lemma 4.22.

Lemma 5.28. *Suppose $J \subseteq G^*_r$ is M-linearly θ-bounded, M-simple and K^r_q-divisible. Then there is $\Phi \in \mathbb{Z}^{X_q(G^*)}$ with $\partial\Phi = J$ such that $\partial_\pm\Phi$ are M-linearly θ'-bounded and M-simple and $\dim(Im(\phi)) = q$ for all $\phi \in \Phi$.*

Proof. Let $\theta \ll \theta_0 \ll \theta'_0 \ll \cdots \ll \theta_r \ll \theta'_r \ll \theta'$. We construct $\Phi = \sum_{i=0}^r \Phi^i$ such that $\partial_\pm\Phi^i$ are M-linearly θ_i-bounded and M-simple and $\dim(Im(\phi)) = q$ for all $\phi \in \Phi$. Also, defining $J^0 = J - \partial\Phi^0$ and $J^i = J^{i-1} - \partial\Phi^i$ for $i \in [r]$, each J^i will be K^r_q-divisible, M-linearly θ_i-bounded, M-simple and i-null. To construct Φ^0 we sum $\binom{q}{r}^{-1}|J|$ elements of $X_q(G)$, chosen sequentially uniformly at random from the full dimensional embeddings of $X_q(G)$, subject to the condition of remaining M-simple, i.e. whenever we choose some ϕ, in future choices for each $e \in \partial\phi$ we do not allow any ϕ' such that $\partial\phi' \cap \partial M(e) \neq \emptyset$. Then $J^0 = J - \partial\Phi^0$ is 0-null and whp $\partial_\pm\Phi^0$ are M-linearly θ_0-bounded (this follows from the proof of Lemma 5.21; we omit the details).

Now we construct Φ^i given J^{i-1} for $i \geq 1$. Note that $\binom{q-i}{r-i}$ divides $|J^{i-1}(e)|$ for all $e \in G_i$, and $J' := \binom{q-i}{r-i}^{-1}\partial_{K^i_r} J^{i-1} \in \mathbb{Z}^{G_i}$ is null and $\theta_{i-1}n^{r-i}$-bounded. If $i < r$ we apply Lemma 4.9 (with $N = n^{r-i}$) to obtain $\Psi \in \mathbb{Z}^{X_{O(i)}(G^*)}$ such that $\partial_{O(i)}\Psi = J'$ and $\partial_\pm\Psi$ are $\theta'_{i-1}n^{r-i}$-bounded. If $i = r$ we apply Lemma 5.22 (with $N = 1$) to obtain $\Psi \in \mathbb{Z}^{X_O(G^*)}$ such that $\partial\Psi = J' = J^{r-1}$ and $\partial_\pm\Psi$ are M-simple and M-linearly θ'_{r-1}-bounded.

Next we let $H = G'^{qi}$ be as in Definition 5.26, and $F \subseteq V(H)$ be such that $\partial_{K^i_q}\Phi'_{qi}$ is $O(i)_i$ on F (by Proposition 5.27). We choose $E_u = (\phi_u, F, H)$ and $s_u \in \{-1, 1\}$ for $u \in [t]$ such that $\Psi = \sum_u s_u\phi_u$. We also let $\mathcal{E} = (E_u : u \in [t])$ and $\mathcal{B} = (B^u : u \in [t])$, where B^u is the set of e such that $\partial_+\Phi^j(e)$ or $\partial_-\Phi^j(e)$ is non-zero for some $j < u$, and also all $e \in J' = J^{r-1}$ if $i = r$. We let $\Phi^* = (\phi^*_u : u \in [t])$ be the $(\mathcal{E}, 1, r, \mathcal{B}, M)$-process. Then we set $\Phi^i = \sum_{u\in[t]} s_u\phi^*_u(\Phi'_{qi}) \in \mathbb{Z}^{X_q(G)}$.

By Lemma 5.21 (with $N = 1$, $s = i$ and $\theta = \theta'_{i-1}$) whp $\partial_\pm\Phi^i$ are M-linearly θ_i-bounded (if $i = r$ this bound also uses the fact that $\partial_\pm\Psi$ are M-linearly θ'_{r-1}-bounded). Also $\dim(Im(\phi)) = q$ for all $\phi \in \Phi^i$ and $\sum_{j=0}^i \Phi^j$ is M-simple by Definition 5.20 and Proposition 5.27. Furthermore, for any $e \in G_i$ we have

$$|\partial_{K^r_q}\Phi^i(e)| = \tbinom{q-i}{r-i}\partial_{K^i_q}\Phi^i(e) = \partial_{K^i_r}J^{i-1}(e) = |J^{i-1}(e)|,$$

since $\partial_{K^i_q}\Phi^i = \partial_{O(i)}\Psi = J'$. Thus $J^i = J^{i-1} - \partial\Phi^i$ is i-null. It is also clear that J^i is K^r_q-divisible, M-linearly θ_i-bounded and M-simple.

Finally, setting $\Phi = \sum_{i=0}^r \Phi^i$, we have $\dim(Im(\phi)) = q$ for all $\phi \in \Phi$, $J^r = J - \partial\Phi$ is r-null, so identically zero, and $\partial_\pm\Phi$ are M-simple and M-linearly θ'-bounded. □

5.5. **Local modifications.** In the final subsection of the section we describe the procedure for local modifications to the template. We will require the following linear analogue of the extension process.

Definition 5.29. Suppose (G, G') and (H, H') are q-complex pairs, $V(G) = \mathbb{F}_{p^a}$, $F \subseteq V(H)$, $N \geq 1$, $\mathcal{B} = (B^i : i \in [t])$ with $B^i \subseteq G$, $E = (L, H, H')$ is a linear extension in (G, G'), and $\mathcal{E} = (E_i : i \in [t])$, where $E_i = (L^i, H, H')$

and L^i is a specialisation of L with base F such that L^i_F is an L-embedding of $(H[F], H'[F])$ in (G, G').

The $(\mathcal{E}, N, r, \mathcal{B})$*-process* is the random sequence $\Phi^* = (\phi^*_i : i \in [t])$, where ϕ^*_i is an L^i-embedding of (H, H') in (G, G') such that $(\phi^*_i)|_F = L^i_F$, and letting C^i be the set of $e \in G_r$ such that $e \in \phi^*_j(H_r \setminus H[F])$ for N choices of $j < r$, we choose ϕ^*_i uniformly at random subject to $\phi^*_i(f) \notin B^i \cup C^i$ for $f \in H \setminus H[F]$. If there is no such choice of ϕ^*_i then the process aborts.

Next we need the analogue of Lemma 4.11 for the linear extension process; this requires the following extra condition.

Definition 5.30. Let $E = (L, H, H^*)$ be a linear extension in (G, G^*) and $F \subseteq V(H)$. Given $f \in H_r[F]$ and $g \in H_r \setminus H[F]$, we say that (f, g) is F*-admissible* if $\dim(L_f) = \dim(L_g) = r$ and

$$R := Row(M(L_g)) \cap Row(M(L_f)) = Row(M(L_g)) \cap Row(M(L_F)).$$

We say that E is F*-admissible* if for every $g \in H_r \setminus H[F]$ with $\dim(L_g) = r$ there is $f \in H_r[F]$ such that (f, g) is F-admissible.

Lemma 5.31. *Let $E = (L, H, H^*)$ be an F-admissible M-closed linear extension in (G, G^*) with $|V(H)| \leq h$. Suppose $\mathcal{E} = (E^i : i \in [t])$, where $E_i = (L^i, H, H')$ and L^i is a (possibly random) basic specialisation of L with base F such that L^i_F is an r-generic L-embedding of $(H[F], H^*[F])$ in (G, G^*). Suppose $\mathcal{B} = (B^i : i \in [t])$ is a (possibly random) sequence with $B^i \subseteq G$. Let $N \geq 1$ and Φ^* be the $(\mathcal{E}, N, r, \mathcal{B})$-process. Let $\mathcal{B}^i$ be the 'bad' event that*

(i) $\partial^i_f \Phi^*$ is linearly θN-bounded for all $f \in H_r[F]$ with $\dim(L_f) = r$ and each $B^{i'}_j$, $i' \leq i$, $j \in [q]$ is linearly η-bounded, but

(ii) $\partial^i_g \Phi^*$ is not linearly $\theta' N$-bounded for some $g \in H_r \setminus H[F]$ with $\dim(L_g) = r$ or the process aborts before step i.

Then whp $\mathcal{B}^t$ does not hold.

Proof. We define a stopping time τ to be the first i for which $\mathcal{B}^i$ holds, or ∞ if there is no such i. We claim whp $\tau = \infty$. We estimate $\mathbb{P}(\tau = i')$ for some $i' < \infty$ as follows. For $i \leq i'$ we can assume that $\partial^i_f \Phi^*$ is linearly θN-bounded for all $f \in H_r[F]$ with $\dim(L_f) = r$ and each B^i_j, $j \in [q]$ is linearly η-bounded (otherwise $\mathcal{B}^{i'}$ cannot hold).

Fix $i < i'$ and consider the choice of ϕ^*_i. Since (G, G^*) is M-linearly $(2hc, h, r)$-typical we have

$$X_{E^i}(G, G^*) = (1 \pm 2hc) n^{\dim(L^i)} Z_{E^i} \pi_{U E^i}(G, G) > 2\delta n^{\dim(L^i)}.$$

We claim that at most $\delta n^{\dim(L^i)}$ choices of $\phi^* \in X_{E^i}(G^*)$ have $\phi^*_i(g) \in B^i \cup C^i$ for some $g \in H \setminus H[F]$. To see this, note that $\mathcal{B}^i$ does not hold, so $\partial^i_g \Phi^*$ is linearly $\theta' N$-bounded for all $g \in H_r \setminus H[F]$ (so C^i is linearly $2^h\theta'$-bounded). Now fix $g \in H \setminus H[F]$ and let $S^g = Span(F \cup g)$. Since L^i is basic, $Im(L^i_g)$ is basic, so by Proposition 5.18(ii) we have $|v(B^i \cup C^i) \cap Im(L^i_g)| < (2^h\theta' + \eta)|Im(L^i_g)|$. This bounds the number of choices of $\phi^*_i|_{S^g}$ such that $\phi^*_i(g) \in B^i \cup C^i$. For each such choice, by Lemma 5.7 there are at most $n^{\dim(L^i) - \dim(L^i_{S^g})}$ choices of $\phi^*_i \in X_{E^i}(G, G^*)$. Since $|Im(L^i_g)| = n^{\dim(L^i_{S^g})}$

the claimed bound follows. Thus we choose ϕ_i^* randomly from at least $\delta n^{\dim(L^i)}$ options.

Now consider any $g \in H_r \setminus H[F]$ and $f \in H_r[F]$ such that (f, g) is F-admissible. We fix $e \in Im(L_g)$ and estimate $\partial_g^{i'} \Phi^*(e)$. Let $S^g = Span(f \cup g)$. By Lemma 5.7 there is a basic specialisation L^e of L_{S^g} such that

$$Im(L^e) = \{v \in Im(L_{S^g}) : v[g] = e\},$$

and we have $\dim(L_f^e) = r - \dim(R)$, where

$$R = Row(M^{gz}) \cap Row(M^{fz}) = Row(M^{gz}) \cap Row(M^{Fz}),$$

as (f, g) is F-admissible. Since $g \not\subseteq F$ we have $\dim(R) < r$. Since $Im(L_f^e)$ is basic, by Proposition 5.18(ii) we have $\sum_{e' \in Im(L_f^e)} v(\partial_f^i \Phi^*)(e') \le \theta N n^{\dim(L_f^e)}$. For each i such that $L_f^i \in Im(L_f^e)$, as above we choose ϕ_i^* from at most $n^{\dim(L^i) - \dim(L_{S^g}^i)}$ options. Writing $\mathcal{F}$ for the algebra generated by the previous choices, we have $\mathbb{P}[\phi_i^*(g) = e \mid \mathcal{F}] < \delta^{-1} n^{-\dim(L_{S^g}^i)}$. By Lemma 5.7 we have $\dim(L_{S^g}^i) = r - \dim(R) = \dim(L_f^e)$, so

$$\mathbb{E}\partial_g^{i'} \Phi^*(e) < r!\theta N n^{\dim(L_f^e)} \cdot r!\delta^{-1} n^{-\dim(L_{S^g}^i)} = r!^2 \delta^{-1} \theta N,$$

where we include two factors of $r!$ for permutations of L_f and L_g. Now as in the proof of Lemma 5.21 whp $\partial_g^{i'} \Phi^*$ is linearly $\theta' N$-bounded, as required. □

Now we will describe a local modification procedure, which we call a 'cascade', which can rearrange the template decomposition so as to include any given q-set of which all r-subsets belong to the template. We cannot achieve this with a naive application of shuffles, as the linear constraints make the set of shuffles too sparse. Instead, we will construct it via a suitable combination of shuffles, illustrated in Figure 2. Our picture is for $q = 3$ and $r = 1$, i.e. matchings in 3-graphs, as it is difficult to draw a shuffle for $r > 1$, but for any r it is the correct picture from the viewpoint of the auxiliary hypergraph on r-sets.

The idea is that we want to modify the template decomposition (in this case it is a matching) so that it contains the dashed triple. To do so, we find a configuration as in the picture, where the bold solid lines are in the matching. We can flip these to the vertical lines while still covering the same vertices. Then we can flip the three vertical lines at the top of the picture to horizontal, obtaining a matching that contains the dashed triple. A key point of the construction is that, given the dashed triple, the set of choices for the six points of the other two horizontal triples has constant density in the set of all sextuples.

For a formal description, first we define the underlying linear extension for the cascade.

Definition 5.32.

(i) Let $X^\alpha = (X_i^\alpha : i \in [q])$, $\alpha \in [2]$ be q-partite sets with

$$X_i^1 = \{x^{1ib} : b \in \mathbb{F}_p^q\}$$

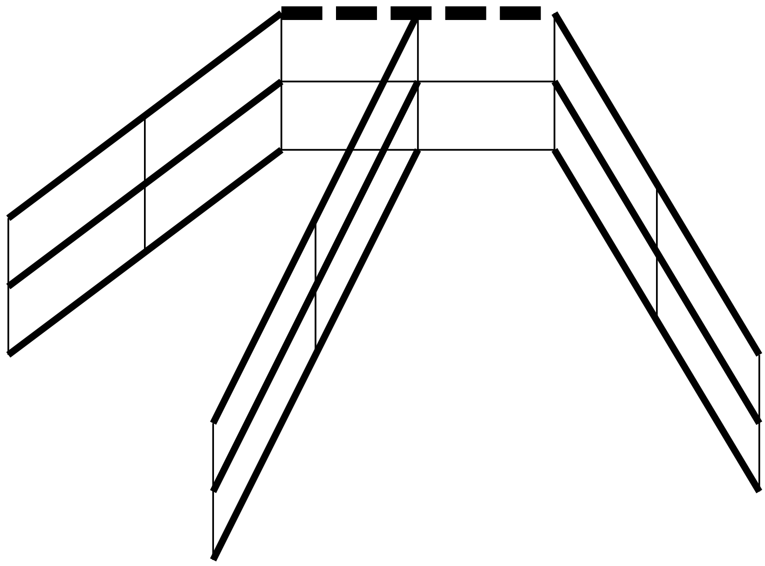

FIGURE 2. A cascade, viewed in the auxiliary hypergraph.

and

$$X_i^2 = \{x^{2iQb} : Q \in \mathbb{F}_p^{r\times q}, b \in \mathbb{F}_p^q\}$$

for $i \in [q]$. Let $X = (X_i : i \in [q])$ be obtained from the disjoint union of X^1 and X^2 by identifying x^{2iQe^i} with $x^{1i(e^i MQ)}$ for $i \in [q]$ and $Q \in \mathbb{F}_p^{r\times q}$.

(ii) For $Q, Q' \in \mathbb{F}_p^{r\times q}$ we define q-partite q-sets f^{1e}, f^{1Q}, $f^{2QQ'}$, g^{1Q} and $g^{2QQ'}$ by

$$f_i^{1e} = x^{1ie^i},\ f_i^{1Q} = x^{1i(e^i MQ)},\ f_i^{2QQ'} = x^{2iQ(e^i MQ')},$$

$$g_i^{1Q} = x^{1i(e^i MQ+e^i)}, \text{ and } g_i^{2QQ'} = x^{2iQ(e^i MQ'+e^i)}.$$

Let $F^C = \cup_{I\in\binom{[q]}{r}} f^{2Q^I Q^I}$ and let H^C be the q-partite q-complex on X generated by all sets f^{1e}, f^{1Q}, $f^{2QQ'}$, g^{1Q}, $g^{2QQ'}$.

(iii) Let $z = (z^1, t)$ be variables with

$$z^1 = (z_i^{1b} : i \in [q], b \in \mathbb{F}_p^q), \text{ and } t = (t_j^Q : j \in [r], Q \in \mathbb{F}_p^{r\times q}).$$

Let $E^C = (L^C, H^C)$ be the linear extension with variables z where for $i \in [q]$, $j \in [r]$, $b \in \mathbb{F}_p^q$, $Q \in \mathbb{F}_p^{r\times q}$, we have $L^C_{x^{1ib}} = z_i^{1b}$ and

$$L^C_{x^{2iQb}} = e^i M t^Q + b \cdot w^Q,$$

where $w^Q = (z_i^{1(e^i MQ)} - e^i M t^Q : i \in [q])$.

We record some basic properties of the cascade extension.

Proposition 5.33.

(i) E^C *is well-defined,*

(ii) $L^C_{f^{2QQ'}} = M(t^Q + Q'w^Q)$ *for* $Q, Q' \in \mathbb{F}_p^{r\times q}$,

(iii) $L^C_{f^{2Q^I Q^I}} = M(e^I M)^{-1} z^{1e^I}$, *where* $z^{1e^I} = (z_i^{1e^i} : i \in I)$,

(iv) $Span(f^{1e}) = F^C$.

Proof. For (i), we check that

$$L^C_{x^{2iQe^i}} = e^i M t^Q + e^i \cdot w^Q = z_i^{1(e^i MQ)} = L^C_{x^{1i(e^i MQ)}},$$

which is consistent with the identification of x^{2iQe^i} with $x^{1i(e^i MQ)}$. For (ii), we have $L^C_{f_i^{2QQ'}} = L^C_{x^{2iQ(e^i MQ')}} = e^i M(t^Q + Q'w^Q)$ for $i \in [q]$. In particular, when $Q = Q' = Q^I = (e^I M)^{-1} e^I$, since $e^i M Q^I = e^i$ for $i \in I$, we have

$$L^C_{f^{2Q^I Q^I}} = M(e^I M)^{-1}(e^I M t^{Q^I} + e^I w^{Q^I}) = M(e^I M)^{-1} z^{1e^I},$$

giving (iii). We also deduce (iv), as $L^C_{f_i^{1e}} = L^C_{x^{1ie^i}} = z_i^{1e^i}$ for $i \in [q]$. □

Now we define the cascade.

Definition 5.34. Suppose ϕ is an L^C-embedding of H^C in G^* with $\phi(v) = L^C_v(y)$ for some choice y of values in $\mathbb{F}_{p^a}$ for the variables z. The *y-cascade* is the following sequence of modifications to M^*:

(i) for each $Q, Q' \in \mathbb{F}_p^{r \times q}$ we replace $L^C_{f^{2QQ'}}(y)$ by $L^C_{g^{2QQ'}}(y)$,

(ii) for each $Q \in \mathbb{F}_p^{r \times q}$ we replace $L^C_{f^{1Q}}(y)$ by $L^C_{g^{1Q}}(y)$.

The purpose of the cascade is achieved by (ii), namely creating the set $L^C_{f^{1e}}(y) = L^C_{g^{10}}(y)$. It is well-defined, as $f^{1Q} = g^{2Q0}$ for each $Q \in \mathbb{F}_p^{r \times q}$, so after step (i) we have created the sets required to implement step (ii). We also have the following properties.

Lemma 5.35.

(i) any q-partite r-set f in X^1 *is contained in a unique* f^{1Q} *and a unique* g^{1Q^*},

(ii) any q-partite r-set f in X is contained in a unique $f^{2QQ'}$ *and a unique* $g^{2QQ'^*}$,

(iii) $\{f^{2QQ'} : Q, Q' \in \mathbb{F}_p^{r \times q}\}$ *forms a* K_q^r*-decomposition of* H_r^C,

(iv) $\{g^{2QQ'} : Q, Q' \in \mathbb{F}_p^{r \times q}, Q' \neq 0\} \cup \{g^{1Q} : Q \in \mathbb{F}_p^{r \times q}\}$ *forms a* K_q^r*-decomposition of* H_r^C,

(v) E^C *is M-closed, r-generic and* F^C*-admissible.*

Proof. The proofs of (i) and (ii) are similar to that of Lemma 5.25. For (i), Q is determined by $Q = (e^I M)^{-1} M^f$, where $f = (x^{1ib^i} : i \in I)$, and M^f has rows b^i, $i \in I$; then Q^* is determined by $Q^* = Q - Q^I$, so this proves (i). Note also that $x^{1ib^i} = x^{2iQe^i}$ for $i \in I$, so $f \subseteq f^{2QQ^I}$ and $f \subseteq g^{2Q0}$; again Q is determined by $Q = (e^I M)^{-1} M^f$, and similarly Q^I and 0 are uniquely determined. To finish the proof of (ii) it remains to consider the case $f \not\subseteq X^1$. By definition of H^C we have $f \subseteq f^{2QQ'}$ or $f \subseteq g^{2QQ'}$ for some $Q, Q' \in \mathbb{F}_p^{r \times q}$. Here Q is unique by choice of the identifications, and Q' is unique by the same argument as in (i), so this proves (ii). It follows immediately that $\{f^{2QQ'} : Q, Q' \in \mathbb{F}_p^{r \times q}\}$ forms a K_q^r-decomposition of H_r^C, so we have (iii). Similarly, $\{g^{2QQ'} : Q, Q' \in \mathbb{F}_p^{r \times q}\}$ forms a K_q^r-decomposition of H_r^C, and since $g^{2Q0} = f^{1Q}$, by (i) we deduce (iv).

For (v), we have $L^C_{f^{2QQ'}} = M(t^Q + Q'w^Q)$ by Proposition 5.33(ii), so M-closure holds by (iii) and $F^C = \cup_I f^{2Q^I Q^I}$, taking $\mathcal{M}$ to consist of all $f^{2QQ'}$ with (Q, Q') not of the form (Q^I, Q^I). Since M is generic, Proposition 5.33(ii) also implies that E^C is r-generic. Next, let $R \subseteq \mathbb{F}_{p^a}^z$ consist of all vectors that are supported on the variables z^{1e^i}, $i \in [q]$ (i.e. are zero outside of the corresponding columns). By Proposition 5.33(iii,iv) we have $Row(M(L^C_{F^C})) = R$. Now consider any $g \in H^C_r \setminus H^C[F^C]$ with $\dim(L_g) = r$. We have $g \subseteq f^{2QQ'}$ for some $Q, Q' \in \mathbb{F}_p^{r \times q}$. Then vectors in $Row(M(L^C_g))$ are supported on the variables t^Q_j, $j \in [r]$ and $z_i^{1(e^i MQ)}$, $i \in [q]$. By Proposition 5.23 there are at most r choices of i such that $e^i MQ = e^i$, so we can fix some $I \in \binom{[q]}{r}$ containing them all. Now any vector in $Row(M(L^C_g) \cap M(L^C_{F^C}))$ is supported on the variables $z_i^{1e^i}$, $i \in I$, so $(f_I^{2Q^I Q^I}, g)$ is F^C-admissible. □

6. Clique decompositions

In this section we prove our main result, which is the following generalisation of Theorem 1.4 to typical complexes. We say that a q-complex G is *K^r_q-divisible* if G_r is K^r_q-divisible.

Theorem 6.1. *Let $1/n \ll c \ll d, 1/h \ll 1/q \le 1/(r+1)$. Suppose G is a K^r_q-divisible (c,h)-typical q-complex on n vertices with $d_i(G) > d$ for $i \in [q]$. Then there is $\Phi \subseteq X_q(G)$ such that $\partial_{K^r_q} \Phi = G_r$.*

We prove Theorem 6.1 by induction on r. The base case is $r = 0$, when the statement is trivial. Indeed, for any q-complex G with $G_q \ne \emptyset$, taking $\Phi = \{\phi\}$ for any $\phi \in X_q(G)$ we obtain $\partial_{K^0_q} \Phi = \{\emptyset\} = G_0$.

Although the case $r = 1$ is covered by the general argument below, it is perhaps helpful to also describe a direct argument; we will sketch one here, but omit the details of the proof. We can assume $G_1 = V(G)$, and then we need to show that G_q has a perfect matching (note that $q \mid n$, since G is K^1_q-divisible). We choose a matching $B = \{b_1, \dots, b_{n/q}\} \in G_{q-1}$ by the random greedy algorithm, let $A = \{a_1, \dots, a_{n/q}\}$ denote the uncovered vertices, and let H be the bipartite graph on (A, B) where $a_i b_j \in H$ if $a_i \cup b_j \in G_q$. By typicality of G, writing $d = \prod_{i \in [q]} d_i(G)^{\binom{q}{i} - \binom{q-1}{i}}$, one can show that whp every vertex has degree $(1 \pm c')dn/q$ in H and every pair of vertices in the same part of H have $(1 \pm c')d^2 n/q$ common neighbours in H, where $c \ll c' \ll d$. It is well-known that this pseudorandomness condition implies that H has a perfect matching, so G_q has a perfect matching.

Next we record the following simple observation (the proof is obvious, so we omit it).

Proposition 6.2. *Suppose G is K^r_q-divisible, $j \in [r]$ and $e \in G_j$. Then $G(e)$ is K^{r-j}_{q-j}-divisible.*

Henceforth we assume $r \ge 1$ and that Theorem 6.1 holds for smaller values of r. This assumption is used in the following lemma.

Lemma 6.3. *Let $1/n \ll c \ll c' \ll c'' \ll d'' \ll d' \ll d, 1/h \ll 1/q \le 1/r$. Suppose G is a K^r_q-divisible (c,h)-typical q-complex on n vertices with*

$d_i(G) > d$ for $i \in [q]$. Suppose $G' = G[V']$, $V' \subseteq V(G)$ is such that (G, G') is (c', h)-typical with $d_i(G') > d'$ for $i \in [q]$. Then there is a simple $\Phi \subseteq X_q(G)$ such that $G''_r = G_r \setminus \partial_{K^r_q}\Phi \subseteq G'_r$, and $G'' = G'[G''_r]$ is a K^r_q-divisible (c'', h)-typical q-complex with $d_i(G'') > d''$ for $i \in [q]$.

Proof. We start by decomposing G into subgraphs according to intersection patterns with V'. We choose independent uniformly random injections $\pi_e : e \to [q]$, $e \in G_r$ and bijections $\pi_f : f \to [q]$, $f \in G_q$. For $e \in G_r$ we write $R_e = [q] \setminus \pi_e(e)$. We choose independent uniformly random functions $\tau_e : R_e \to \{0, 1\}$, $e \in G_r$ where

$$\mathbb{P}(\tau_e(i) = 1) = t := |V'|/|V(G)|$$

for all $i \in R_e$. For $v \in V(G)$ we let $i(v)$ be 1 if $v \in V'$ or 0 otherwise. For $x \in \{0, 1\}^q$ we let H^x be the q-complex generated by all $f \in G_q$ such that

(i) $i(v) = x_{\pi_f(v)}$ for all $v \in f$,
(ii) $\pi_e(v) = \pi_f(v)$ for all $v \in e \subseteq f$, $|e| = r$ and
(iii) $\tau_e(i) = x_i$ for all $e \subseteq f$, $|e| = r$, $i \in R_e$.

We let A^x be the auxiliary $\binom{q}{r}$-graph where $V(A^x)$ consists of all $e \in G_r$ with $i(v) = x_{\pi_e(v)}$ for all $v \in e$ and $\tau_e(i) = x_i$ for all $i \in R_e$, and A^x consists of all $\binom{q}{r}$-sets $\binom{f}{r}$ where $f \in H^x$. Note that for any $e \in G_r$, the choices of π_e and τ_e determine a unique x such that $e \in V(A^x)$, and we cannot have $e \subseteq f$ for any $f \in H^{x'}$, $x' \neq x$.

Next we show whp each A^x is approximately regular, so that we can apply the nibble. Fix $e \in G_r$, condition on π_e and τ_e being such that $e \in V(A^x)$, and write

$$x^1 = \sum_{i\in[q]} x_i,\ e^1 = \sum_{i\in\pi_e(e)} x_i,\ x^0 = q - x^1,\ e^0 = r - e^1.$$

Since (G, G') is (c', h)-typical, by Lemma 3.16 the number of choices for $f \in G_q$ with $e \subseteq f$ and $|f \cap V'| = x^1$ is

$$(1 \pm 4^{h!}c')d^* \tbinom{tn}{x^1-e^1}\tbinom{(1-t)n}{x^0-e^0}, \quad \text{where } d^* = \prod_{i\in[q]} d_i^{\binom{q}{i}-\binom{r}{i}}.$$

In the previous expression we modified the statement of Lemma 3.16 by using binomial coefficients rather than powers of n; otherwise we would overcount sets by a permutation factor of $(x^1 - e^1)!(x^0 - e^0)!$. This factor cancels, as for each such f, there are $(x^1 - e^1)!(x^0 - e^0)!$ choices of π_f that agree with π_e on e and map $(f \setminus e) \cap V'$ to $\{i \in R_e : x_i = 1\}$, and so map $(f \setminus e) \setminus V'$ to $\{i \in R_e : x_i = 0\}$. We choose any such π_f with probability $1/q!$, and for each r-set $e' \subseteq f$ with $e' \neq e$ we choose $\pi_{e'} = (\pi_f)|_{e'}$ with probability $z_{qr} = \prod_{i=1}^r (q + 1 - i)^{-1}$.

Finally, for each $v \in f$, the number of r-sets $e' \subseteq f$ with $v \notin e' \neq e$ is $\binom{q-1}{r}$ for $v \in e$ or $\binom{q-1}{r} - 1$ for $v \notin e$; writing $i = \pi_f(v)$, for each such e' we have $\mathbb{P}(\tau_{e'}(i) = x_i)$ equal to t if $x_i = 1$ or $1 - t$ if $x_i = 0$. We deduce that

$$\begin{aligned}\mathbb{E}|A^x(e)| =&(1 \pm 4^{h!}c')d^* \tbinom{tn}{x^1-e^1}\tbinom{(1-t)n}{x^0-e^0} \cdot (x^1 - e^1)!(x^0 - e^0)!/q! \\ &\cdot z_{qr}^{\binom{q}{r}-1} \cdot t^{x^1\binom{q-1}{r}}(1-t)^{x^0\binom{q-1}{r}} t^{-(x^1-e^1)}(1-t)^{-(x^0-e^0)}.\end{aligned}$$

The value of $|A^x(e)|$ is affected by at most 1 by the choice of any π_f with $e \subseteq f$, and at most n^{q-r-1} by the choice of any $\pi_{e'}$ or $\tau_{e'}$ with $e' \neq e$, so by Lemma 2.10 whp $|A^x(e)| = (1 \pm 5^{h!}c')\delta^x n^{q-r}$, where

$$\delta^x = d^* q!^{-1} z_{qr}^{\binom{q}{r}-1} (t^{x^1}(1-t)^{x^0})^{\binom{q-1}{r}}$$

is independent of e^0 and e^1, as required. This also gives $H_r^x = V(A^x)$.

All codegrees in each A^x are at most n^{q-r-1}, so we can apply the nibble. We do so in a carefully chosen order, and as we proceed we also separately cover the edges of G_r corresponding to vertices that are not covered by the nibbles. Fix $c' \ll c_0 \ll c_0' \ll \cdots \ll c_r \ll c_r' \ll c''$. We start by applying the nibble to all A^x with $x^0 \geq r$. By Lemma 2.17 for each such x we can choose a simple $\Phi^x \subseteq X_q(H^x)$ such that $H'^x_r = H_r^x \setminus \partial_{K_q^r} \Phi^x$ is c_0-bounded.

Next we cover all uncovered $e \in G_r$ with $e \subseteq V(G) \setminus V'$ by a greedy algorithm, where in each step we consider some e and choose some $\phi \in X_q(G)$ such that $\phi([r]) = e$, $\phi([q] \setminus [r]) \subseteq V'$ and $\partial_{K_q^r}\phi$ is edge-disjoint from all previous choices and from the edges covered by the nibbles. We say that an r-set is *full* if it has been covered during the algorithm, and for $i < r$ that an i-set is *full* if it belongs to at least $c_1 n$ full $(i+1)$-sets. Once a set is full we will not cover any more sets containing it. We claim that any set $f \subseteq V(G) \setminus V'$ with $|f| \leq r-1$ cannot be full. To see this, we estimate the number of r-sets e containing f covered during the algorithm. There are at most 2^q choices for x as above. Given x, each $\phi \in \Phi^x$ contains at most 2^q such e, and is uniquely determined by some $e' \in H'^x_r$ containing f. Each H'^x_r is c_0-bounded, so the number of r-sets containing f covered during the algorithm is at most $2^{2q} c_0 n^{r-|f|} < (c_1 n)^{r-|f|}$. Thus f cannot be full, as claimed.

To show that the algorithm can be completed, consider the step at which we consider some e. By Lemma 3.10(ii), $G^1 := G'[G(e)]$ is $(3^h c', h-r)$-typical with $d_i(G^1) > d^{2^q}$ for $i \in [q-r]$. Let G^2 be the set of $f \in G^1$ such that every r-subset of $e \cup f$ is uncovered. We can analyse G^2 by repeated applications of Lemma 3.19. Indeed, fix $g \subsetneq e$, and note that for any $e' \in G_r$ with $e' \cap e = g$, conditional on $\pi_{e'}$, under the random choice of $\tau_{e'}$, the probability of $e' \in H_r^x$ for some x with $x^0 < r$ is

$$p = \sum_{i < r - |g|} \tbinom{q-r}{i} t^{q-r-i} (1-t)^i,$$

independently of all other choices. Also, for any x with $x^0 \geq r$, the set of such e' in H_r^x is c_0'-bounded. Then we can apply Lemma 3.19, where J is p-binomial and J' is the set of $e' \setminus g$ where $e' \cap e = g$ and $e' \in H'^x_r$ for some x with $x^0 \geq r$ (note that J' is $2^q c_0'$-bounded). Repeating for all $g \subsetneq e$, whp G^2 is $(4^h c_1, h-r)$-typical with $d_i(G^2) > d^{3^q}$ for $i \in [q-r]$.

Now let G^3 be obtained from G^2 by deleting any set f such that $e \cup f$ contains a full set. Note that we do not delete $\emptyset$, as any subset of e of size at most $r-1$ is not full by the above claim, and e is not full as we have not yet covered it. We can analyse G^3 by repeated applications of Lemma 3.15. Indeed, we start by restricting to the set of vertices v such that $e \cup \{v\}$ does not contain any full set; for each subset e' of e there are at most $c_1 n$ vertices v such that $e' \cup \{v\}$ is full, so we delete at most $2^r c_1 n$ vertices.

Next, if $q \geq r+2$, we restrict to the set of pairs uv such that $e \cup \{u,v\}$ does not contain any full set; for each subset e' of e, since $e' \cup \{u\}$ is not full by the previous restriction, there are at most $c_1 n$ vertices v such that $e' \cup \{u,v\}$ is full, so we delete at most $2^r c_1 n$ pairs containing any remaining vertex u. Repeating this process, we deduce that G^3 is $(c'_1, h-r)$-typical with $d_i(G^1) > d^{3^q}/2$ for $i \in [q-r]$. Any edge of G^3_{q-r} gives a valid choice for ϕ, so the algorithm can be completed. At the end of the algorithm, any $(r-1)$-set e' belongs to at most $c_1 n + q$ covered r-sets, since there can be at most $c_1 n$ such r-sets before it becomes full, and we can overshoot this bound by at most q (to be precise, by at most $q-r$).

We continue the above process over several stages, which are similar, but also require the induction hypothesis of Theorem 6.1. Suppose at the start of stage $j \in [r-1]$ that we have covered all r-sets e with $|e \setminus V'| > r-j$ or $e \in V(A^x)$ with $x^0 > r-j$, and that $|A^x(e)| = (\delta^x \pm 2c_j)n^{q-r}$ for all $e \in V(A^x)$ with $x^0 \leq r-j$. Suppose also that every $(r-1)$-set belongs to at most $2c_j n$ r-sets that have been covered during a greedy algorithm (not counting those covered by a nibble). We apply the nibble to each A^x with $x^0 = r-j$, obtaining a simple $\Phi^x \subseteq X_q(H^x)$ such that $H'^x_r = H^x_r \setminus \partial_{K^r_q} \Phi^x$ is c'_j-bounded. Then we cover all uncovered $e \in G_r$ with $|e \setminus V'| = r-j$ by the following greedy algorithm.

We consider all $f \in G_{r-j}$ with $f \subseteq V \setminus V'$ sequentially. At the step when we consider f, we will show that there is a j-complex G^3 such that G^3_j is the set of $e \setminus f$ where e is an uncovered r-set containing f, to which we can apply the induction hypothesis of Theorem 6.1, obtaining $\Phi^f \subseteq X_{q-r+j}(G^3)$ such that $\partial_{K^j_{q-r+j}} \Phi^f = G^3_j$. Then for each $\phi \in \Phi^f$ we extend ϕ to $\phi^+ \in X_q(G)$ by including f in $Im(\phi^+)$, and add ϕ^+ to our decomposition.

We say that an r-set is *full* if it has been covered during the greedy algorithms, and for $i < r$ that an i-set is *full* if it belongs to at least $c_{j+1} n$ full $(i+1)$-sets. Once a set is full we will not cover any more sets containing it. We claim that any f' with $|f' \cap V'| \leq j$ cannot be full. To see this, note that each ϕ^+ containing f' gives rise to $\binom{q-|f'|}{r-|f'|}$ covered r-sets containing f', and ϕ^+ is uniquely determined by some $e' \in H'^x_r$ for some x with $x^0 = r-j$ with $|e' \setminus V'| = r-j$ and $f' \subseteq e' \subseteq \phi^+([q])$. Since each such H'^x_r is c'_j-bounded, there are at most $c'_j n^{r-|f'|}$ choices for e' given f' and x. Including at most $2c_j n^{r-|f'|}$ sets covered at previous stages, we see that the number of covered r-sets containing f' is at most $2^{2q} c'_j n^{r-|f'|} < (c_{j+1} n)^{r-|f'|}$, so f' cannot be full, as claimed.

To show that the algorithm can be completed, consider the step at which we consider some f. We define G^1, G^2, G^3 as in the 0th stage: let $G^1 = G'[G(f)]$; let G^2 be the set of $f' \in G^1$ such that every r-subset of $f \cup f'$ is uncovered; let G^3 be obtained from G^2 by deleting any set f' such that $f \cup f'$ contains a full set. Arguing as before, whp G^3 is $(c'_{j+1}, h-r)$-typical with $d_i(G^1) > d^{3^q}$ for $i \in [q-r+j]$. Also, G^3_j is the set of $e \setminus f$ where e is an uncovered r-set containing f, as any such e does not contain any full set by the claim. Then G^3 is K^j_{q-r+j}-divisible by Proposition 6.2. Now by the induction hypothesis of Theorem 6.1, we can choose $\Phi^f \subseteq X_{q-r+j}(G^3)$ such that $\partial_{K^j_{q-r+j}} \Phi^f = G^3_j$, as stated above.

At the end of the algorithm, we claim that any $(r-1)$-set e' belongs to at most $2c_{j+1}n$ r-sets e that have been covered during a greedy algorithm (recall that we need this at the start of stage $j+1$). To see this, first recall that we have at most $2c_j n$ such sets e from previous stages. Now consider stage j, and suppose first that $|e' \setminus V'| = r-j$. Then each such e determines a unique ϕ^+ with $e \subseteq Im(\phi^+)$, and each ϕ^+ with $e' \subseteq Im(\phi^+)$ gives $q-r+1$ such e. Since each H'^x_r with $x^0 = r-j$ is c'_j-bounded we obtain at most $2^q c'_j n$ sets e, so the claimed bound holds easily. On the other hand, if $|e' \setminus V'| < r-j$ then $|e' \cap V| \geq j$, so at any given step when we consider f, there is at most one $\phi \in \Phi^f$ with $e' \subseteq Im(\phi)$. We cover at most $c_{j+1}n$ r-sets containing e' before e' becomes full, and we can overshoot this bound by at most $q-r$, so again we deduce the claimed bound.

At the end of stage $r-1$ all edges of G_r not contained in V' have been covered. As above, $G^1 = G[V']$ is $(3^h c', h)$-typical with $d_i(G^1) > d^{2^q}$ for $i \in [q]$, and the restriction G'' of G^1 to the uncovered r-sets is (c'', h)-typical with $d_i(G'') > d''$ for $i \in [q]$. Also, G'' is K^r_q-divisible, as it is obtained from G by removing a partial K^r_q-decomposition, so this proves the lemma. □

Proof of Theorem 6.1. We argue by induction on r. The case $r = 0$ is obvious, so suppose $r \geq 1$. Next we reduce to the case $n = p^a$, for a prime p with $q \ll p \ll h$. Indeed, fix such p and define $a \in \mathbb{N}$ by $p^a \leq n < p^{a+1}$. Let $V' \subseteq V(G)$ be a random subset of size $n' = p^a$ and let $G' = G[V']$. By Lemma 3.18 whp (G, G') is $(c + n^{-1/4}, h)$-typical, with $d_i(G'') = (1 \pm n^{-1/4}) d_i(G')$ for $i > 1$. By Lemma 6.3 we can choose a simple $\Phi \subseteq X_q(G)$ such that $G''_r = G_r \setminus \partial_{K^r_q} \Phi \subseteq G'_r$, and $G'' = G'[G''_r]$ is a K^r_q-divisible (c'', h)-typical q-complex with $d_i(G'') > d''$ for $i \in [q]$, where $c \ll c'' \ll d'' \ll d$. Thus we can assume $n = p^a$.

Now we identify V with $\mathbb{F}_{p^a}$ and choose an M-template G^* of G satisfying Lemma 5.15, where without loss of generality σ is the identity permutation, and (G, G^*) is M-linearly $(2hc, h, r)$-typical and $(3^h c, h/q)$-typical with $d_i(G^*) = (1 \pm 9hqc) d_i(G)$ for $i \in [q] \setminus \{r\}$ and $d_r(G^*) = (1 \pm 9hqc) d^*_r$, where

$$d^*_r = d_r(G) \prod_{i=1}^{r} (q+1-i)^{1-\binom{q}{r}} d_i(G)^{\binom{q}{i}-\binom{r}{i}}.$$

Let Γ be the set of $e \in G_r$ with $\dim(e) = r$. Fix $c \ll c_1 \ll c_2 \ll c_3 \ll d, 1/h$. We start by choosing a simple $\Phi' \subseteq X_q(G)$ such that $G_r \setminus \Gamma \subseteq \partial\Phi' \subseteq G_r \setminus G^*_r$ and $\partial\Phi' \cap \Gamma$ is c_1-bounded wrt Γ.

To do this, we sequentially eliminate minimally dependent sets in order of increasing size. In stage $j \in [r]$, at each step we consider some $f \in G_j$ such that $\dim(f) < j$ and $\dim(f') = |f'|$ for all $f' \subsetneq f$. We will show that there is a $(r-j)$-complex G^2 such that G^2_{r-j} is the set of $e \setminus f$ where e is an uncovered r-set containing f, to which we can apply the induction hypothesis of Theorem 6.1, obtaining $\Phi^f \subseteq X_{q-j}(G^2)$ such that $\partial_{K^{r-j}_{q-j}} \Phi^f = G^2_{r-j}$. Then for each $\phi \in \Phi^f$ we extend ϕ to $\phi^+ \in X_q(G)$ by including f in $Im(\phi^+)$, and add ϕ^+ to Φ'.

Next we define several types of fullness as follows. Suppose $j \in [r]$, $x \in \{0,1\}^r$, $e \in G_r$ and $\pi : e \to [r]$ is bijective. We say that e is (j, x, π)-*full* if e has been covered during the algorithm when we considered some minimally dependent $f \in G_j$ with $e \cap f = \{v \in e : x_{\pi(v)} = 1\}$. Now suppose $i < r$, $e^1 \in G_i$ and $\pi^1 : e^1 \to [r]$ is injective. We say that e^1 is (j, x, π^1)-*full* if e^1 belongs to at least $c_1^2 n$ sets $e^2 \in G_{i+1}$ such that e^2 is (j, x, π^2)-full for some injection $\pi^2 : e^2 \to [r]$ with $\pi^2|_{e^1} = \pi^1$. Once any $e' \in G$ is (j, x, π')-full we will not allow any more r-sets e containing e' to become (j, x, π)-full for any bijection $\pi : e \to [r]$ with $\pi|_{e'} = \pi'$.

For any $e' \in G$, injection $\pi' : e' \to [r]$ and $x \in \{0,1\}^r$ such that $b := |e'| - \sum_{i \in \pi'(e')} x_i \le r - j$ we claim that e' cannot be (j, x, π')-full before the point in the algorithm where we consider some minimally dependent $f \subseteq e'$. To see this, we bound the number of r-sets e containing e' such that e was covered when we considered some minimally dependent $f \in G_j$ with $e \cap f = \{v \in e : x_{\pi(v)} = 1\}$, for some bijection $\pi : e \to [r]$ with $\pi|_{e'} = \pi'$. Note that

$$j - (|e'| - b) = j - \sum_{i \in \pi'(e')} x_i = |f \setminus e'| > 0$$

as we have not yet considered any minimally dependent subset of e'. As f is minimally dependent, there are $O(n^{j-(|e'|-b)-1})$ choices for f given e'. Since Φ^f is simple for K_{q-j}^{r-j} and $b \le r - j$ there are $O(n^{r-j-b})$ choices of $\phi \in \Phi^f$ with $e' \subseteq \phi^+([q])$. Each such ϕ gives rise to $O(1)$ choices of e and π. Thus the number of such (e, π) is $O(n^{r-|e'|-1}) < (c_1^2 n)^{r-|e'|}$, so e' cannot be (j, x, π')-full, as claimed.

To show that the algorithm can be completed, consider the step at which we consider some minimally dependent $f \in G_j$ as described above. Let $G^1 = G[G_r \setminus G_r^*](f)$. Let G^2 be obtained from G^1 by deleting any $e \in G^1$ such that there are $e^1 \subseteq e^2 \subseteq f \cup e$ with $|e^2| = r$ and a bijection $\pi : e^2 \to [r]$ such that e^1 is $(j, x(\pi, f), \pi|_{e^1})$-full, where

$$x(\pi, f)_i = 1_{\{v \in f\}} \text{ for } v \in e^2 \text{ and } i = \pi(v).$$

Note that for any $e \in G_{r-j}^2$ we have not covered $f \cup e$, as otherwise we would have deleted e in the definition of G^2 (with $e^1 = e^2 = f \cup e$). Conversely, if $e = e^2 \setminus f$ with $f \subseteq e^2 \in G_r^*$ and $e^1 \subseteq e^2$ and $\pi : e^2 \to [r]$ is bijective we claim that e^1 is not $(j, x(\pi, f), \pi|_{e^1})$-full. Indeed, since $\dim(e^2) = r$ there is no minimally dependent $f' \subseteq e^1$, and we have $|e^1| - \sum_{i \in \pi(e^1)} x(\pi, f)_i \le |e^2 \setminus f| = r - j$, so this holds by the above claim. Thus G_{r-j}^2 is the set of $e \setminus f$ where e is an uncovered r-set containing f. Then G^2 is K_{q-j}^{r-j}-divisible by Proposition 6.2.

By Lemmas 3.17 and 3.11, G^1 is $(5^{h!}c, h - j)$-typical with $d_i(G^1) > d^{2^q}$ for $i \in [q - j]$. (When applying Lemma 3.17 we take $G' = G^* \cup G_{<r}$, which only affects the density estimates by a factor of $1 + O(1/n)$.) As in the proof of Lemma 6.3, we analyse G^2 by repeated applications of Lemma 3.15. For $i \ge 0$, after the ith step we will have a complex $G^{1i} \subseteq G^1$ such that for every $e \in G^{1i}$ there is no $e^1 \subseteq e^2 \subseteq f \cup e$ with $|e^1| = i$ and $|e^2| = r$, bijection $\pi : e^2 \to [r]$ and $j \in [r]$ such that e^1 is $(j, x(\pi, f), \pi|_{e^1})$-full. By the above claim (applied with $e' = \emptyset$) this holds for $i = 0$ with $G^{i0} = G^1$.

Now suppose $i \geq 0$ and G^{1i} has the required property. Consider any $e \in G^{1i}$, $e^1 \subseteq e^2 \subseteq f \cup e$ with $|e^1| = i$ and $|e^2| = r$ and bijection $\pi : e^2 \to [r]$. By the ith step, e^1 is not $(j, x(\pi, f), \pi|_{e^1})$-full, so there are at most $c_1^2 n$ choices of $v \notin e^1 \cup f$ such that $e' = e^1 \cup \{v\}$ is $(j, x(\pi, f), \pi')$-full for some injection $\pi' : e' \to [r]$ with $\pi'|_{e^1} = \pi|_{e^1}$. Furthermore, for any e^1 there are at most $r!2^r$ choices for $\pi|_{e^1}$ and x_i for $i \in \pi|_{e^1}(e^1)$, so the number of deleted $(|e^1 \setminus f|+1)$-sets containing $e^1 \setminus f$ is at most $r!2^r c_1^2 n$. Thus we can construct $G^{1(i+1)}$ from G^{1i} by applying Lemma 3.15 repeatedly with $0 \leq s \leq i$ and $\theta = r!2^r c_1^2$. After step $q-j$ we deduce that G^2 is $(c_2, h-r)$-typical with $d_i(G^1) > d^{3^q}$ for $i \in [q-r]$. Now by the induction hypothesis of Theorem 6.1, we can choose $\Phi^f \subseteq X_{q-j}(G^2)$ such that $\partial_{K^{r-j}_{q-j}} \Phi^f = G^2_{r-j}$, as stated above.

At the end of the algorithm, by construction $\Phi' \subseteq X_q(G)$ is simple and $G_r \setminus \Gamma \subseteq \partial_{K^r_q} \Phi' \subseteq G_r \setminus G^*_r$. We claim that any $e' \in G^*_{r-1}$ belongs to at most $(2r)^r c_1^2 n$ sets of $\partial_{K^r_q} \Phi'$. To see this, note that since $\dim(e') = r-1$, whenever we considered a minimally dependent set f we had $|f \cup e'| \geq r$. Then there can be at most one $\phi \in \Phi^f$ such that $f \cup e' \subseteq Im(\phi^+)$, so we covered at most $q-r+1$ r-sets e containing e' at any step of the algorithm. For each $j \in [r]$, $x \in \{0,1\}^r$ and injection $\pi' : e' \to [r]$ we covered at most $c_1^2 n$ such e before e' became (j, x, π')-full, and we can overshoot by at most $q-r$, so we deduce the claimed bound. It follows that $\partial_{K^r_q} \Phi'$ is c_1-bounded.

Let G' be obtained from G by restricting to $G^*_{\leq r-1}$ and then deleting all sets that contain any set of $\partial_{K^r_q} \Phi'$. Then (G', G^*) is $(2c_1, h/q)$-typical by Lemmas 3.9 and 3.15. Next we can cover $G'_r \setminus G^*_r$ using the nibble and a greedy algorithm. Indeed, consider the auxiliary hypergraph A with $V(A) = G'_r \setminus G^*_r$ and $E(A)$ consisting of all $\binom{q}{r}$-sets $\partial_{K^r_q} \phi \subseteq V(A)$ with $\phi \in X_q(G')$. For any $e \in V(A)$, by Lemma 3.16 we have $|A(e)| = (1 \pm 3^{h!} c_1) d^\circ n^{q-r}$, where

$$d^\circ = (1 - d^*_r/d_r(G))^{\binom{q}{r}-1} \prod_{i=1}^{r} d_i(G)^{\binom{q}{i}-\binom{r}{i}}.$$

Since A has codegrees $O(n^{q-r-1})$, by Lemma 2.17 we can choose a matching Ψ in A covering all but $c_1 n^r$ vertices, such that $J = (G'_r \setminus G^*_r) \setminus \partial_{K^r_q} \Psi$ is c_1-bounded.

Now we cover J using copies of K^r_q in which every other edge is in G^*. Let $H = [q]^{\leq}$, $F = [r]$ and H' be the r-complex obtained from $\binom{[q]}{\leq r}$ by deleting the edge $[r]$. Let $\mathcal{E} = (E_i : i \in [|J|])$ be a sequence of rooted extensions $E_i = (\phi_i, F, H, H')$ in (G, G^*) such that $J = \{\phi_i(F) : i \in |J|\}$; this exists as $e' \in G^*$ for all $e' \subsetneq e \in J$. Let $B_0 = \partial \Phi' \cap G^*_r$ and $\Phi^* = (\phi^*_i : i \in [t])$ be the $(\mathcal{E}, 1, r, B_0, M)$-process. Let $\Phi'' = \Psi \cup \Phi^*$. Then $G'_r \setminus G^*_r \subseteq \partial \Phi''$. Let $J' = \partial \Phi'' \cap G^*_r$. Then J' is M-simple by Definition 5.20, and by Lemma 5.21 whp $J' := \partial \Phi'' \cap G^*_r$ is M-linearly c_2-bounded.

Note that J' is K^r_q-divisible, as we can write $J' = G^*_r - (G_r - \partial \Phi' - \partial \Phi'')$, where each summand is K^r_q-divisible. By Lemma 5.28 there is $\Psi \in \mathbb{Z}^{X_q(G^*)}$ with $\partial \Phi = J'$ such that $\partial_\pm \Psi$ are M-linearly c_3-bounded and M-simple and $\dim(Im(\psi)) = q$ for all $\psi \in \Psi$. Note that $\partial \Psi^- \subseteq \partial \Psi^+$. To complete the proof, we will modify M^* (the template q-sets) to form a new K_q^r-decomposition M° of G^* that includes Ψ^+. Then we can replace Ψ^+ by Ψ^- to complete the K_q^r-decomposition of G.

To do so, write $\Psi^+ = \{\psi_i : i \in T\}$ and let $\mathcal{E} = (E^i : i \in [T])$ (reusing notation) be a sequence of specialisations $E^i = (L^i, H^C)$ of the cascade extension $E^C = (L^C, H^C)$ with base F^C such that $L^i_{f^{1e}} = \psi_i$. To see that this exists, note that $\dim(L^C_{f^{1e}}) = q$ and $Span(f^{1e}) = F^C$ by Proposition 5.33, so we can apply Lemma 5.7 to set $L^i_{f^{1e}} = \psi_i$, and then $L^i_{F^C}$ is uniquely determined. Since L^C is basic, we can choose variables so that L^i is basic. Recall that E^C is M-closed and F^C-admissible by Lemma 5.35. Also, L^i is a generic specialisation of L^C, as $C(L^C) = 0$ and $\dim(Im(\psi_i)) = q$ by choice of Ψ, so L^i is simple and r-generic by Propositions 5.6(iv) and 5.9(ii). Let $B = \cup_{e \in \partial^+\Psi} \partial M(e)$; then B is linearly c_3-bounded, since $\partial_+\Psi$ is M-linearly c_3-bounded. We let $\Phi^* = (\phi^*_i : i \in [T])$ be the $(\mathcal{E}, 1, r, B)$-process; by Lemma 5.31 whp the process does not abort. Now we obtain M° from M^* by applying the y-cascade for each $i \in [T]$, where y is such that $\phi^*_i(v) = L^C_v(y)$ for all $v \in V(H^C)$. By definition of the process these cascades are edge-disjoint from each other and B, so they can all be completed. After applying the cascades, we have a K_q^r-decomposition M° that includes $\Psi^+ = \{\phi^*_i(f^{1e}) : i \in [T]\}$. Finally, $\Phi = \Phi' \cup \Phi'' \cup (M^\circ \setminus \Psi^+) \cup \Psi^-$ gives the required K_q^r-decomposition of G. □

As noted in the introduction, Theorem 6.1 implies the Existence Conjecture for Steiner systems. For designs, we need a generalisation to multicomplexes.

Definition 6.4. A *q-multicomplex* G is a q-complex where each $e \in G$ has a non-negative integer multiplicity $m_G(e)$. We say that G is a (c, m, r, q)-*multicomplex* if $m_G(e) = 1$ for all $e \in G \setminus G_r$, $m_G(e) \in [m]$ for all $e \in G_r$ and $\{e \in G_r : m_G(e) < m\}$ is c-bounded wrt G. We say that G is K_q^r-*divisible* if $\binom{q-i}{r-i}$ divides $|G_r(e)|$ (counted with multiplicities) for all $e \in G_i$, $0 \le i \le r$. We count extensions and define typicality by treating G as a complex, ignoring multiplicities.

Note that Definition 6.4 is rather ad hoc, but suffices for our purposes. One can make more natural general definitions for multicomplexes in which extensions are counted with multiplicity and typicality takes this into account. However, if we restrict to (c, m, r, q)-multicomplexes then these are essentially equivalent to the definitions ignoring multiplicity. We obtain the Existence Conjecture for designs from the following theorem applied with $G = \binom{[n]}{\le q}$ and $m_G(e) = \lambda$ for all $e \in G_r$. In the statement we identify G_r with $J \in \mathbb{N}^{G_r}$ where $J(e) = m_G(e)$ for $e \in G_r$.

Theorem 6.5. *Let $1/n \ll c \ll d, 1/h \ll 1/m \ll 1/q \le 1/(r+1)$. Suppose G is a K_q^r-divisible (c, h)-typical (c, m, r, q)-multicomplex on n vertices with $d_i(G) > d$ for $i \in [q]$. Then there is $\Phi \subseteq X_q(G)$ such that $\partial_{K_q^r}\Phi = G_r$.*

The proof of Theorem 6.5 is very similar to that of Theorem 6.1, so we will just briefly indicate the necessary modifications. The reduction to the prime power case is the same, where we use the induction hypothesis of

Theorem 6.5 to prove the multicomplex version of Lemma 6.3. The choice the template and algorithm to cover $G_r \setminus \Gamma$ is the same, again using the induction hypothesis of Theorem 6.5. Recall also that during the greedy algorithms we obtained the current complex by repeated applications of Lemma 3.15, so at every stage we have a (c', m, r, q)-multicomplex, where $c \ll c' \ll d, 1/h$. Next we apply the nibble and a greedy algorithm to reduce the multiplicity of every $e \in G'_r$ to be 1 or 0, such that the r-graph of 0-multiplicity edges is c'-bounded. To see that this can be achieved, we construct the auxiliary $\binom{q}{r}$-graph A similarly to before, using a vertex set that has $\max\{m_G(e) - 1, 0\}$ copies of each $e \in G'_r$; approximate regularity of A follows from typicality and the definition of (c, m, r, q)-multicomplexes. The remainder of the proof is the same as in Theorem 6.1.

Our results can be iterated to give many K^r_q-decompositions of G such that any q-set is used in at most one decomposition. In particular, by combining these decompositions we obtain designs with values of λ that grow with n (we cover a constant proportion of the possible values). We state two such theorems and outline the proof of the second. To obtain any λ^* in the full range of parameters for λ, we write N for the least common multiple of $\{\binom{q-i}{r-i} : 0 \le i \le r-1\}$ and modify Theorem 6.6 by taking all but one of the designs to have $\lambda = N$, and the last to have $\lambda = \lambda^* \bmod N$.

Theorem 6.6. *Let $1/n \ll \theta \ll 1/q \le 1/(r+1)$ and $\theta \ll 1/\lambda$. Suppose that $\binom{q-i}{r-i}$ divides $\lambda\binom{n-i}{r-i}$ for $0 \le i \le r-1$. Then there are θn^{q-r} pairwise disjoint designs with parameters (n, q, r, λ) on the same vertex set.*

Theorem 6.7. *Let $1/n \ll \theta \ll c \ll d, 1/h \ll 1/q \le 1/(r+1)$. Suppose G is a K^r_q-divisible (c, h)-typical q-complex on n vertices with $d_i(G) > d$ for $i \in [q]$. Then there are $\Phi_i \subseteq X_q(G)$ such that $\partial_{K^r_q}\Phi_i = G_r$ for $i \in [\theta n^{q-r}]$ and for each $e \in G_q$ there is at most one $i \in [\theta n^{q-r}]$, $\phi \in \Phi_i$ with $\phi([q]) = e$.*

To prove Theorem 6.7 we repeatedly apply Theorem 6.1, using a greedy algorithm, and taking account of full sets in a similar manner to before. Specifically, we say that a q-set is full if it has been used by some Φ_j, and for $r \le i < q$ that an i-set is full if it belongs to at least $\theta' n$ full $(i+1)$-sets, where $\theta \ll \theta' \ll c$. At each step we remove all sets that contain any full set from G. Clearly no r-set can be full, and Lemma 3.15 implies that at any step the current complex is $(2c, h)$-typical, so the algorithm can be completed.

Finally, we can obtain the following (crude) estimate for the number of K^r_q-decompositions of a typical complex. When G is complete, it shows that the number $S(n, q, r)$ of Steiner systems with parameters (n, q, r) (where n satisfies the divisiblity conditions) satisfies

$$\log S(n, q, r) = (1 + o(1))\binom{q}{r}^{-1}\binom{n}{r}(q - r)\log n.$$

Similar estimates hold for multicomplexes and designs.

Theorem 6.8. *Let $1/n \ll \theta \ll c \ll c' \ll d, 1/h \ll 1/q \le 1/(r+1)$. Suppose G is a K^r_q-divisible (c, h)-typical q-complex on n vertices with $d_i(G) > d$ for*

$i \in [q]$. *Then the number N of K_q^r-decompositions of G_r satisfies*

$$\log N = (1 \pm c')\binom{q}{r}^{-1}|G_r|(q-r)\log n.$$

To prove Theorem 6.8, first note that the upper bound holds by the trivial estimate on the number of subsets of G_q of size $\binom{q}{r}^{-1}|G_r|$. For the lower bound, we consider those K_q^r-decompositions of G_r obtained by first applying the nibble, and then applying Theorem 6.1 to the subcomplex G' induced by the r-sets not covered by the nibble. One can obtain whp G' is a K_q^r-divisible (c_1, h)-typical q-complex on n vertices with $d_i(G) > c_2$ for $i \in [q]$, where $c \ll c_1 \ll c_2 \ll c'$. (An alternative construction that requires less effort in analysing the nibble is to reserve a small random subset G_r^* of G_r, cover $G_r \setminus G_r^*$ by the nibble and a random greedy algorithm spilling into G_r^*, then apply Theorem 6.1 to the remainder of G_r^*.)

We estimate the number of such K_q^r-decompositions as follows. Let $t = O(1)$ be the number of bites in the nibble, and $(b_i : i \in [t])$, $(w_i : i \in [t])$ be such that whp the ith bite takes $(1 \pm c_0)b_i$ edges of the auxiliary hypergraph (which correspond to sets in G_q) and wastes at most w_i of them (those that intersect some other edge of the bite). We can take

$$\sum_{i\in[t]} b_i = (1 \pm c_0)\binom{q}{r}^{-1}|G_r| \quad \text{and} \quad \sum_{i\in[t]} w_i < c_0|G_r|,$$

where $c \ll c_0 \ll c_1$. The number B of choices for the bites in which the algorithm is successful satisfies $\log B = \sum_{i\in[t]}(1 \pm c_0)b_i \cdot (1+o(1))(q-r)\log n$. This overcounts the number of decompositions by a factor WP, where W is the number of choices for the wasted edges, and P is the number of ordered partitions of a decomposition into the t bites of the nibble and the final decomposition of G'. We can bound $\log W < c_0|G_r| \cdot 2(q-r)\log n$ and $\log P < 2|G_r|\log t$. The stated estimate on $\log N$ now follows.

7. Concluding remarks

We have shown that the divisibility conditions suffice for clique decomposition of (multi)complexes that satisfy a certain pseudorandomness condition, thus verifying the Existence Conjecture for designs. Our construction uses a randomised algorithm, which could in principle be implemented in a computerised search for explicit designs (perhaps it might work on a reasonable number of vertices, although our proof does not guarantee this).

We obtain estimates on the number of designs, although it would be desirable to make these more accurate. For example, Wilson [48] conjectured that the number of Steiner Triple Systems on n vertices (where n is 1 or 3 mod 6) satisfies $S(n,3,2) = ((1+o(1))n/e^2)^{n^2/6}$. Our analysis gives a similar expression with $1 + o(1)$ in the exponent, whereas Wilson obtained better upper and lower bounds having the form $(cn)^{n^2/6}$ with two values of c (assuming the van der Waerden permanent conjecture, which is now known to be true). Related questions for ‘high-dimensional permutations’ were recently posed by Linial and Luria [29].

Those readers familiar with Hypergraph Regularity Theory will note that our result applies to the so-called 'dense setting', whereas a result suitable for use with a hypergraph regularity decomposition would need to apply to the 'sparse setting'. One would also need a version adapted to 'multityped' typicality, corresponding to the structures that arise in a regularity decomposition. We believe that both issues can be addressed by our method of Randomised Algebraic Constructions, but the additional technical complications of the sparse setting are considerable, so we defer these to a future paper. One potential application is to a conjecture of Nash-Williams [31] on triangle decompositions in graphs of large minimum degree.

The more general setting will require a generalisation of Lemma 4.9: it is not hard to adapt the proof so as to allow coefficients in any finitely generated abelian group. Independently of its application to designs, the exactness property for generalised boundary operators in (pseudo)random simplicial complexes may merit further study from a topological point of view; it is reminiscent of recent results on the topology of random simplicial complexes (see e.g [2, 26]).

There are many potential directions of generalisation suggested by known results on graph decompositions. One could ask for H-decompositions for general r-graphs H, analogous to the results on integral H-decompositions obtained by Wilson [47]. One could also ask for a decomposition analogue of Baranyai's theorem, namely to partition all q-sets of a suitable q-complex into K^r_q-decompositions (we noted above that our results allow us to cover a constant fraction of the q-sets by decompositions).

Another viewpoint is to see the H-decomposition problem as a case of the hypergraph perfect matching problem, using the auxiliary hypergraphs we defined when applying the nibble. This is a powerful framework that encompasses many well-known problems in Design Theory (e.g. Ryser's Conjecture on transversals in Latin Squares is equivalent to the statement that certain auxiliary 3-graphs have perfect matchings). For dense hypergraphs, we have given structural characterisations of the perfect matching problem with Mycroft [23] and Knox and Mycroft [22] (see also the survey by Rödl and Ruciński [34] for many further references). However, the hypergraphs arising in design theory are typically very sparse, and it is NP-complete to determine whether a general (sparse) hypergraph has a perfect matching. Thus we expect that one must use their specific structure to some degree, but may also hope for some form of unifying general statement.

The perfect matching viewpoint also suggests some tantalising potential connections with Probability and Statistical Physics along the lines of results obtained by Kahn and Kayll [21] and Kahn [20] for matchings in graphs and Barvinok and Samorodnitsky [4] for matchings in hypergraphs. We formulate a couple of vague questions in this direction (and draw the reader's attention to the final section of [20] for many more questions): Can one give an asymptotic formula for the number of H-decompositions of a typical complex (with weights)? In a random decomposition, are the appearances of given edge-disjoint H's approximately uncorrelated?

Acknowledgements. I would like to thank Rick Wilson for encouraging me to work on the Existence Conjecture, and several participants of the recent Oberwolfach Combinatorics meeting for helpful comments.

References

[1] N. Alon, J. H. Kim and J. Spencer, Nearly perfect matchings in regular simple hypergraphs, *Israel J. Math.* 100:171–187 (1997).
[2] L. Aronshtam, N. Linial, T. Łuczak and R. Meshulam, Collapsibility and vanishing of top homology in random simplicial complexes, *Disc. Comp. Geometry*, to appear.
[3] K. Azuma, Weighted sums of certain dependent random variables, *Tôhoku Math. J.* 19:357–367 (1967).
[4] A. Barvinok and A. Samorodnitsky, Computing the partition function for perfect matchings in a hypergraph, *Combinatorics, Probability and Computing*, 20:815–825 (2011).
[5] C. J. Colbourn and J. H. Dinitz, *Handbook of Combinatorial Designs*, 2nd ed. Chapman & Hall / CRC, Boca Raton, 2006.
[6] P. Diaconis and S. Zabell, Closed form summation for classical distributions: variations on a theme of De Moivre, *Statistical Sci.* 61:284–302 (1991).
[7] P. Erdős and H. Hanani, On a limit theorem in combinatorial analysis, *Publicationes Mathematicae Debrecen* 10:10–13 (1963).
[8] A. Ferber, R. Hod, M. Krivelevich and B. Sudakov, A construction of almost Steiner systems, *J. Combinatorial Designs*, to appear.
[9] P. Frankl and V. Rödl, Near perfect coverings in graphs and hypergraphs, *European J. Combinatorics* 6:317–326 (1985).
[10] D. A. Freedman, On tail probabilities for martingales, *Ann. Probability* 3:100–118 (1975).
[11] W. T. Gowers, Hypergraph Regularity and the multidimensional Szemerédi Theorem, *Annals of Math.* 166:897–946 (2007).
[12] D. A. Grable, More-than-nearly perfect packings and partial designs, *Combinatorica* 19:221–239 (1999).
[13] R. L. Graham, S.-Y. R. Li, and W.-C. W. Li, On the structure of t-designs, *SIAM J. Alg. Disc. Meth.* 1:8–14 (1980).
[14] J. E. Graver and W. B. Jurkat, The module structure of integral designs, *J. Combinatorial Theory Ser. A* 15:75–90 (1973).
[15] T. Gustavsson, Decompositions of large graphs and digraphs with high minimum degree, Doctoral Dissertation, University of Stockholm, 1991.
[16] H. Hanani, The existence and construction of balanced incomplete block designs, *Annals Mathematical Statistics* 32:361–386 (1961).
[17] H. Hanani, A balanced incomplete block design, *Annals Mathematical Statistics* 36:711 (1965).
[18] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience, 2000.
[19] J. Kahn, Asymptotically good list-colorings, *J. Combinatorial Theory Ser. A* 73:1–59 (1996).
[20] J. Kahn, A normal law for matchings, *Combinatorica* 20:339-391 (2000).
[21] J. Kahn and M. Kayll, On the stochastic independence properties of hardcore distributions, *Combinatorica* 17:369-391 (1997).
[22] P. Keevash, F. Knox and R. Mycroft, Polynomial-time perfect matchings in dense hypergraphs, Proc. 45th ACM STOC (2013).
[23] P. Keevash and R. Mycroft, A geometric theory for hypergraph matchings, to appear in *Mem. Amer. Math. Soc.*
[24] J. H. Kim, Nearly optimal partial Steiner systems, *Electronic Notes in Discrete Mathematics* 7:74–77.
[25] A. Kostochka and V. Rödl, Partial Steiner systems and matchings in hypergraphs, *Random Structures and Algorithms* 13:335–347 (1997).
[26] D. Kozlov, The threshold function for vanishing of the top homology group of random d-complexes, *Proc. Amer. Math. Soc.* 138:4517–4527 (2010).

[27] G. Kuperberg, S. Lovett and R. Peled, Probabilistic existence of regular combinatorial objects, Proc. 44th ACM STOC (2012).

[28] N. Kuzjurin, On the difference between asymptotically good parkings and coverings, *European J. Combin.* 16:35–40 (1995).

[29] N. Linial and Z. Luria, An upper bound on the number of high-dimensional permutations, *Combinatorica*, to appear.

[30] C. McDiarmid, Concentration, in: Probabilistic Methods for Algorithmic Discrete Mathematics, *Alg. Combin.* 16:195–248 (1998).

[31] C. St. J. A. Nash-Williams, An unsolved problem concerning decomposition of graphs into triangles, *Combinatorial Theory and its Applications III*, North Holland (1970), 1179–1183.

[32] N. Pippenger and J. H. Spencer, Asymptotic behaviour of the chromatic index for hypergraphs, *J. Combinatorial Theory Ser. A* 51:24–42 (1989).

[33] V. Rödl, On a packing and covering problem, *European J. Combinatorics* 6:69–78 (1985).

[34] V. Rödl and A. Ruciński, Dirac-type questions for hypergraphs — a survey (or more problems for Endre to solve), *An Irregular Mind (Szemerédi is 70)* 21:1–30 (2010).

[35] V. Rödl, A. Ruciński and E. Szemerédi, Perfect matchings in large uniform hypergraphs with large minimum collective degree, *J. Combin. Theory Ser. A* 113:613–636 (2009).

[36] V. Rödl and M. Schacht, Regular partitions of hypergraphs: regularity lemma, *Combin. Probab. Comput.* 16:833–885 (2007).

[37] V. Rödl and J. Skokan, Regularity lemma for uniform hypergraphs, *Random Structures and Algorithms* 25:1–42 (2004).

[38] J. Spencer, Asymptotic packing via a branching process, *Random Structures and Algorithms* 7:167–172 (1995).

[39] L. Teirlinck, Non-trivial t-designs without repeated blocks exist for all t, *Discrete Math.* 65:301–311 (1987).

[40] V. A. Vatutin and V. G. Mikhailov, Limit theorems for the number of empty cells in an equiprobable scheme for group allocation of particles, *Theory Probab. Appl.* 27:734–743 (1982).

[41] V. Vu, New bounds on nearly perfect matchings in hypergraphs: higher codegrees do help, *Random Structures and Algorithms* 17:29–63 (2000).

[42] R. Wilson, The early history of block designs, *Rend. del Sem. Mat. di Messina* 9:267–276 (2003).

[43] R. M. Wilson, An existence theory for pairwise balanced designs I. Composition theorems and morphisms, *J. Combinatorial Theory Ser. A* 13:220–245 (1972).

[44] R. M. Wilson, An existence theory for pairwise balanced designs II. The structure of PBD-closed sets and the existence conjectures, *J. Combinatorial Theory Ser. A* 13:246–273 (1972).

[45] R. M. Wilson, An existence theory for pairwise balanced designs III. Proof of the existence conjectures, *J. Combinatorial Theory Ser. A* 18:71–79 (1975).

[46] R. M. Wilson, The necessary conditions for t-designs are sufficient for something, *Utilitas Math.* 4:207–215 (1973).

[47] R. M. Wilson, Signed hypergraph designs and diagonal forms for some incidence matrices, *Designs, Codes and Cryptography*, 17:289–297 (1999).

[48] R. M. Wilson, Nonisomorphic Steiner Triple Systems, *Math. Zeit.* 135:303–313 (1974).