Combinatorics, Probability and Computing (2020), **29**, pp. 868–885
doi:10.1017/S0963548320000280

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Pseudorandom hypergraph matchings†

Stefan Ehard[1], Stefan Glock[2*] and Felix Joos[3]

[1]Institut für Optimierung und Operations Research, Universität Ulm, Germany, [2]Institute for Theoretical Studies, ETH Zürich, Switzerland and [3]Institut für Informatik, Universität Heidelberg, Germany
*Corresponding author. Email: stefan.glock@eth-its.ethz.ch

(Received 23 July 2019; revised 19 March 2020; accepted 16 April 2020; first published online 22 July 2020)

Abstract
A celebrated theorem of Pippenger states that any almost regular hypergraph with small codegrees has an almost perfect matching. We show that one can find such an almost perfect matching which is 'pseudo-random', meaning that, for instance, the matching contains as many edges from a given set of edges as predicted by a heuristic argument.

2020 MSC Codes: Primary: 05C15, 05C65, 05C70; Secondary: 05D15, 05D40

1. Introduction

A hypergraph $\mathcal{H}$ consists of a vertex set $V(\mathcal{H})$ and an edge set $E(\mathcal{H}) \subseteq 2^{V(\mathcal{H})}$. If all edges have size r, then $\mathcal{H}$ is called *r-uniform*, or simply an *r-graph*. A *matching* in $\mathcal{H}$ is a collection of pairwise disjoint edges, and a *cover of* $\mathcal{H}$ is a set of edges whose union contains all vertices. A matching is *perfect* if it is also a cover. These concepts are widely applicable, as 'almost all combinatorial questions can be reformulated as either a matching or a covering problem of a hypergraph' [11], and their study is thus of great relevance in combinatorics and beyond.

Results such as Hall's theorem and Tutte's theorem that characterize when a graph has a perfect matching are central to graph theory. However, for each $r \geqslant 3$, it is NP-complete to decide whether a given r-uniform hypergraph has a perfect matching [19]. It is thus of great importance to find sufficient conditions that guarantee a perfect matching in an r-uniform hypergraph. This problem has received a lot of attention over the years. For instance, one line of research has focused on minimum degree conditions that guarantee a perfect matching (see *e.g.* [1, 14, 23, 37] and the survey [36]). Another important direction has been to study perfect matchings in random hypergraphs. The so-called Shamir's problem, to determine the threshold for which the (binomial) random k-graph has a perfect matching with high probability, was open for over 25 years resisting numerous efforts, until it was famously solved by Johansson, Kahn and Vu [15]. Moreover, Cooper, Frieze, Molloy and Reed [5] determined when regular hypergraphs have a perfect matching with high probability. It would be very interesting to obtain such results not only for random hypergraphs but to find pseudorandomness conditions that (deterministically) guarantee a perfect matching. Aside from some partial results (*e.g.* [10, 13, 28]), this seems wide open.

†The research leading to these results was partially supported by the EPSRC, grants EP/N019504/1 (S. Glock) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 339933727 (F. Joos).

Many of the aforementioned results are proved by first obtaining an *almost* perfect matching and then using some clever ideas to complete it. It turns out that almost perfect matchings often exist under weaker conditions. For example, in the minimum degree setting, the threshold for finding an almost perfect matching is often smaller than that of finding a perfect matching. Also, there is a well-known theorem that yields almost perfect matchings under astonishingly mild pseudorandomness conditions. Mostly referred to as Pippenger's theorem, any almost regular hypergraph with small codegrees has an almost perfect matching. Both the result itself and also its proof method, the so-called 'semi-random method' or 'Rödl nibble', have had a tremendous impact on combinatorics. We add to this body of research by showing the existence of 'pseudorandom' matchings in this setting. We note that our result does not improve previous bounds on the *size* of a matching that can be obtained. Rather, our focus is on the *structure* of such a matching within the hypergraph it is contained in.

In Section 1.1 we revisit Pippenger's theorem. In Section 1.2 we discuss a theorem of Alon and Yuster, which can be viewed as an intermediate step. In Section 1.3 we will motivate and state our main results.

1.1 Pippenger's theorem

Pippenger never published his theorem, and it was really the culmination of the efforts of various researchers in the 1980s. Most notably, in 1985, Rödl [35] proved a long-standing conjecture of Erdős and Hanani on approximate Steiner systems. A *(partial)* (n, k, t)*-Steiner system* is a set $\mathcal{S}$ of k-subsets of some n-set V such that every t-subset of V is contained in (at most) one k-set in $\mathcal{S}$. In 1853, Steiner asked for which parameters such systems exist, a question that has intrigued mathematicians for more than 150 years and was only answered recently by Keevash [20]. In 1963, Erdős and Hanani asked whether, for fixed k, t, one can always find an 'approximate Steiner system', that is, a partial (n, k, t)-Steiner system covering all but $o(n^t)$ of the t-sets, as $n \to \infty$. This was proved by Rödl using the celebrated 'nibble' method, with some ideas descending from [2, 26]. Frankl and Rödl [8] observed that in fact a much more general theorem holds, which applies to almost regular hypergraphs with small codegrees. Pippenger's version stated below is a slightly stronger and cleaner version. For a hypergraph $\mathcal{H}$, we let $v(\mathcal{H})$ and $e(\mathcal{H})$ denote the number of vertices and edges of $\mathcal{H}$, respectively, and for vertices $u, v \in V(\mathcal{H})$ we define the degree

$$\deg_{\mathcal{H}}(v) := |\{e \in E(\mathcal{H}) \colon v \in e\}|$$

and codegree

$$\deg_{\mathcal{H}}(uv) := |\{e \in E(\mathcal{H}) \colon \{u, v\} \subseteq e\}|.$$

Let

$$\Delta(\mathcal{H}) := \max_{v \in V(\mathcal{H})} \deg_{\mathcal{H}}(v), \quad \delta(\mathcal{H}) := \min_{v \in V(\mathcal{H})} \deg_{\mathcal{H}}(v), \quad \Delta^c(\mathcal{H}) := \max_{u \neq v \in V(\mathcal{H})} \deg_{\mathcal{H}}(uv)$$

denote the *maximum degree*, *minimum degree* and *maximum codegree* of $\mathcal{H}$, respectively.

Theorem 1.1 (Pippenger). *For $r \in \mathbb{N}$ and $\varepsilon > 0$, there exists $\mu > 0$ such that any r-uniform hypergraph $\mathcal{H}$ with $\delta(\mathcal{H}) \geqslant (1 - \mu)\Delta(\mathcal{H})$ and $\Delta^c(\mathcal{H}) \leqslant \mu\Delta(\mathcal{H})$ has a matching that covers all but at most an ε-fraction of the vertices.*

To see why this generalizes Rödl's result, fix n, k, t and construct a hypergraph $\mathcal{H}$ with vertex set $\binom{[n]}{t}$ where every k-set $X \subseteq [n]$ induces the edge $\binom{X}{t}$. Note that perfect matchings in $\mathcal{H}$ correspond to (n, k, t)-Steiner systems. Clearly $\mathcal{H}$ is $\binom{k}{t}$-uniform. Moreover, every vertex has degree

$$\binom{n-t}{k-t} = \Theta(n^{k-t}) \quad \text{and} \quad \Delta^c(\mathcal{H}) = \binom{n-t-1}{k-t-1} = o(n^{k-t}).$$

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

Thus, for sufficiently large n, Pippenger's theorem implies the existence of a matching $\mathcal{M}$ in $\mathcal{H}$ that covers all but $o(n^t)$ of the vertices, which corresponds to a partial (n, k, t)-Steiner system which covers all but $o(n^t)$ of the t-sets. Frankl and Rödl [8] also applied (their version of) this theorem to obtain similar results for other combinatorial problems, for instance the existence of Steiner systems in vector spaces. Keevash [22] raised the meta question of whether there exists a general theorem that provides sufficient conditions for a sparse 'design-like' hypergraph to admit a perfect matching (for a notion of 'design-like' that captures Steiner systems, for example, but hopefully many more structures). Since such hypergraphs will likely be (almost) regular and have small codegree, the existence of an almost perfect matching follows from Pippenger's theorem, and a natural approach would be to use the absorbing method to complete such a matching to a perfect one. This of course can be extremely challenging since the relevant auxiliary hypergraphs are generally very sparse.

1.2 The Alon–Yuster theorem

In the case of Steiner systems, the absorbing method has been used successfully to answer Steiner's question [12, 20]. Very roughly speaking, the idea of an absorbing approach is to set aside a 'magic' absorbing structure, then to obtain an approximate Steiner system, and finally to employ the magic absorbing structure to clean up. One (minor, but still relevant) challenge is that the leftover of the approximate Steiner system must be 'well-behaved'. More precisely, instead of the global condition that the number of uncovered t-sets is $o(n^t)$, one needs the stronger local condition that for every fixed $(t-1)$-set, the number of uncovered t-sets containing this $(t-1)$-set is $o(n)$. Fortunately, Alon and Yuster [3, Theorem 1.2], by building on a theorem of Pippenger and Spencer [34], provided a tool achieving this. They showed that any almost regular hypergraph with small codegrees contains a matching that is 'well-behaved' in the sense that it not only covers all but a tiny proportion of the entire vertex set, but also has this property with respect to a specified collection of not too many not too small vertex subsets. The precise statement is technical and allows for certain trade-offs between set sizes, number of sets, and degree conditions. To give a concrete example, if the r-uniform almost regular hypergraph $\mathcal{H}$ has N vertices, $\Delta^c(\mathcal{H}) \leqslant \Delta(H)/\log^{9r} N$ and we consider a family $\mathcal{F}$ of at most $N^{\log N}$ vertex subsets, each of size at least $N^{2/5}$, then there exists a matching in $\mathcal{H}$ which covers all but $o(|F|)$ vertices from F for each $F \in \mathcal{F}$.

In the above application to Steiner systems, for every $(t-1)$-set S, consider the set $U_S \subseteq V(\mathcal{H})$ of all t-sets containing S. A matching in $\mathcal{H}$ which covers almost all vertices of U_S then corresponds to a partial Steiner system which covers all but $o(n)$ of the t-sets containing S, as desired.

1.3 Pseudorandom matchings

The purpose of this paper is to provide a tool that is (qualitatively) a generalization of the Alon–Yuster theorem and gives much more control on the matching obtained. The need for such a tool arose in recent work of the authors on graph embeddings. In Section 4 we will discuss further applications of our result in more detail.

To motivate this, suppose for simplicity that we are given a D-regular hypergraph and want to find an (almost) perfect matching $\mathcal{M}$. Moreover, we wish $\mathcal{M}$ to be 'pseudorandom', that is, to have certain properties that we expect from an idealized random matching. In a perfect matching, at a fixed vertex, exactly one edge needs to be included in the matching, and assuming that each edge is equally likely to be chosen, we may heuristically expect that every edge of $\mathcal{H}$ is in a random perfect matching with probability $1/D$. Thus, given a (large) set $E \subseteq E(\mathcal{H})$ of edges, we expect $|E|/D$ matching edges in E. More generally, given a set X, a *weight function on* X is a function $\omega\colon X \to \mathbb{R}_{\geqslant 0}$. For a subset $X' \subseteq X$, we define $\omega(X') := \sum_{x \in X'} \omega(x)$. If ω is a weight function on $E(\mathcal{H})$, the above heuristic would imply that we expect from a 'pseudorandom' matching $\mathcal{M}$ that $\omega(\mathcal{M}) \approx \omega(E(\mathcal{H}))/D$. The following is a simplified version of our main theorem (Theorem 1.3), which asserts that a hypergraph with small codegrees has a matching that is pseudorandom in the above sense.

Theorem 1.2. *Suppose $\delta \in (0, 1)$ and $r \in \mathbb{N}$ with $r \geqslant 2$, and let $\varepsilon := \delta/50r^2$. Then there exists Δ_0 such that for all $\Delta \geqslant \Delta_0$, the following holds. Let $\mathcal{H}$ be an r-uniform hypergraph with $\Delta(\mathcal{H}) \leqslant \Delta$ and $\Delta^c(\mathcal{H}) \leqslant \Delta^{1-\delta}$ as well as $e(\mathcal{H}) \leqslant \exp(\Delta^{\varepsilon^2})$. Suppose that $\mathcal{W}$ is a set of at most $\exp(\Delta^{\varepsilon^2})$ weight functions on $E(\mathcal{H})$. Then there exists a matching $\mathcal{M}$ in $\mathcal{H}$ such that $\omega(\mathcal{M}) = (1 \pm \Delta^{-\varepsilon})\omega(E(\mathcal{H}))/\Delta$ for all $\omega \in \mathcal{W}$ with $\omega(E(\mathcal{H})) \geqslant \max_{e \in E(\mathcal{H})} \omega(e)\Delta^{1+\delta}$.*

We remark that a similar statement when $\mathcal{W}$ has bounded size and without polynomial error bounds is implied by a theorem of Kahn [18]. It has since been observed that the proof in [18] also gives the more general statement (see *e.g.* [21]). Here we prove a more general theorem which allows weight functions not only on edges but on tuples of edges. This allows us, for instance, to specify a set of pairs of edges, and control how many pairs will be contained in the matching (see Section 4 for applications). In particular, this provides a proof of Theorem 1.2, which we state here for completeness and convenient use in future research.

Let us discuss a few aspects of this theorem. First, note that we do not require $\mathcal{H}$ to be almost regular. The theorem can be applied with any (sufficiently large) Δ, and in Section 4 we will discuss in more detail the usefulness of this and the fact that $v(\mathcal{H})$ plays no role in the parametrization of the theorem. If $\mathcal{H}$ is almost regular, an almost perfect matching can be obtained by applying the theorem with $\Delta = \Delta(\mathcal{H})$ to the weight function $\omega \equiv 1$. This yields that

$$|\mathcal{M}| \geqslant (1 - o(1))\frac{e(\mathcal{H})}{\Delta(\mathcal{H})} \geqslant (1 - o(1))v(\mathcal{H})/r,$$

where the last inequality uses that

$$re(\mathcal{H}) = \sum_{x \in V(\mathcal{H})} \deg_{\mathcal{H}}(x) = (1 \pm o(1))v(\mathcal{H})\Delta(\mathcal{H}).$$

We remark that, while Pippenger's theorem only needs $\Delta^c(\mathcal{H}) = o(\Delta)$, we need a stronger condition to apply concentration inequalities. For the same reason, we also need that $\omega(E(\mathcal{H}))$ is not too small (relative to the maximum possible weight). As a result, our theorem also allows stronger conclusions in that the error term $\Delta^{-\varepsilon}$ decays polynomially with Δ.

Note that Theorem 1.2 is (qualitatively) more general than the Alon–Yuster theorem. Indeed, suppose $\mathcal{H}$ is an almost regular hypergraph and we are given a collection $\mathcal{V}$ of subsets $U \subseteq V(\mathcal{H})$ and want to ensure that $\mathcal{M}$ covers each $U \in \mathcal{V}$ almost completely. For each target subset $U \in \mathcal{V}$, we can define a weight function ω_U by setting $\omega_U(e) := |e \cap U|$. Note that

$$\omega_U(E(\mathcal{H})) = \sum_{x \in U} \deg_{\mathcal{H}}(x) = (1 \pm o(1))|U|\Delta(\mathcal{H}).$$

Thus, since $\omega_U(\mathcal{M}) = (1 \pm o(1))\omega_U(E(\mathcal{H}))/\Delta(\mathcal{H})$ by Theorem 1.2, we deduce that

$$|U \cap V(\mathcal{M})| = \omega_U(\mathcal{M}) = (1 \pm o(1))\omega_U(E(\mathcal{H}))/\Delta(\mathcal{H}) \geqslant (1 - o(1))|U|,$$

implying that almost all vertices of U are covered by $\mathcal{M}$. More generally (still assuming that $\mathcal{H}$ is almost regular), if we are given weight functions $p\colon V(\mathcal{H}) \to \mathbb{R}_{\geqslant 0}$ (*e.g.* $p_U(v) := \mathbb{1}_{v \in U}$), then, setting $\omega_p(e) := \sum_{v \in e} p(v)$, we obtain

$$\sum_{v \in V(\mathcal{M})} p(v) = (1 \pm o(1)) \sum_{v \in V(\mathcal{H})} p(v). \tag{1.1}$$

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

Note that the boundedness condition on the edge weight in Theorem 1.2 translates to the condition that

$$\max_{v\in V(\mathcal{H})} p(v) = o\left(\sum_{v\in V(\mathcal{H})} p(v)\right).$$

We now state our main result, for which we need to introduce a bit more notation. Given a set X and an integer $\ell \in \mathbb{N}$, an *ℓ-tuple weight function on X* is a function $\omega\colon \binom{X}{\ell} \to \mathbb{R}_{\geqslant 0}$, that is, a weight function on $\binom{X}{\ell}$. For a subset $X' \subseteq X$, we then define

$$\omega(X') := \sum_{S\in\binom{X'}{\ell}} \omega(S).$$

Moreover, if $\mathcal{X} \subseteq \binom{X}{\ell}$, we write $\omega(\mathcal{X})$ for $\sum_{S\in\mathcal{X}} \omega(S)$ as for usual weight functions. For $k \in [\ell]_0$ and a tuple $T \in \binom{X}{k}$, define

$$\omega(T) := \sum_{S\supseteq T} \omega(S), \quad \text{and let } \|\omega\|_k := \max_{T\in\binom{X}{k}} \omega(T). \tag{1.2}$$

Suppose $\mathcal{H}$ is an r-uniform hypergraph and ω is an ℓ-tuple weight function on $E(\mathcal{H})$. Clearly, if $\mathcal{M}$ is a matching, then a tuple of edges which do not form a matching will never contribute to $\omega(\mathcal{M})$. We thus say that ω is *clean* if $\omega(\mathcal{E}) = 0$ whenever $\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}$ is not a matching.

The following is our main result, which readily implies Theorem 1.2.

Theorem 1.3. *Suppose $\delta \in (0,1)$ and $r, L \in \mathbb{N}$ with $r \geqslant 2$, and let $\varepsilon \leqslant \delta/50L^2r^2$. Then there exists Δ_0 such that for all $\Delta \geqslant \Delta_0$ the following holds. Let $\mathcal{H}$ be an r-uniform hypergraph with $\Delta(\mathcal{H}) \leqslant \Delta$ and $\Delta^c(\mathcal{H}) \leqslant \Delta^{1-\delta}$ as well as $e(\mathcal{H}) \leqslant \exp(\Delta^{\varepsilon^2})$. Suppose that for each $\ell \in [L]$ we are given a set $\mathcal{W}_\ell$ of clean ℓ-tuple weight functions on $E(\mathcal{H})$ of size at most $\exp(\Delta^{\varepsilon^2})$, such that $\omega(E(\mathcal{H})) \geqslant \|\omega\|_k \Delta^{k+\delta}$ for all $\omega \in \mathcal{W}_\ell$ and $k \in [\ell]$.*

Then there exists a matching $\mathcal{M}$ in $\mathcal{H}$ such that $\omega(\mathcal{M}) = (1 \pm \Delta^{-\varepsilon})\omega(E(\mathcal{H}))/\Delta^\ell$ for all $\ell \in [L]$ and $\omega \in \mathcal{W}_\ell$.

We will prove Theorem 1.3 in Section 3, after stating some preliminary results in the next section. In Section 4 we will discuss applications of our main result.

2. Preliminaries

Our main tool is the next theorem of Molloy and Reed on the chromatic index of a hypergraph with small codegrees, improving on earlier work of Pippenger and Spencer as well as Kahn. Pippenger and Spencer [34] strengthened Theorem 1.1 by showing that under the same assumptions one can even obtain an almost optimal edge-colouring of $\mathcal{H}$, using $(1+o(1))\Delta$ colours. (The existence of an almost perfect matching then follows by averaging over the colour classes.) Kahn [17] generalized this to list colourings, and Molloy and Reed improved the $o(1)$-term. For simplicity, we only state their result for normal colourings.

Theorem 2.1 (Molloy and Reed [30, Theorem 2]). *Let $1/\Delta \ll \delta, 1/r$. Suppose $\mathcal{H}$ is an r-uniform hypergraph satisfying $\Delta^c(\mathcal{H}) \leqslant \Delta^\delta$ and $\Delta(\mathcal{H}) \leqslant \Delta$. Then the edge set $E(\mathcal{H})$ can be decomposed into $\Delta + \Delta^{1-(1-\delta)/r}\log^5\Delta$ edge-disjoint matchings.*

Note here that $\mathcal{H}$ is not required to be almost regular. In fact, this assumption can also be omitted from the Pippenger–Spencer theorem since any given r-uniform hypergraph $\mathcal{H}$ can be embedded into a $\Delta(\mathcal{H})$-regular hypergraph $\mathcal{H}'$ with $\Delta^c(\mathcal{H}') = \Delta^c(\mathcal{H})$, and any colouring of $\mathcal{H}'$ induces a colouring of $\mathcal{H}$ with the same number of colours.

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

We also make use of several probabilistic tools to establish concentration of a random variable X. If X is the sum of independent Bernoulli variables, we use the following well-known Chernoff-type bound.

Theorem 2.2 (Chernoff's bound). *Suppose $X_1, \dots, X_m$ are independent random variables taking values in $\{0, 1\}$. Let $X := \sum_{i=1}^m X_i$. Then, for all $\lambda > 0$,*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geqslant \lambda] \leqslant 2 \exp\left(-\frac{\lambda^2}{2(\mathbb{E}[X] + \lambda/3)}\right).$$

Similarly, if X is a function of several independent Bernoulli variables and does not depend too much on any of the variables, we use the following 'bounded differences inequality'.

Theorem 2.3 (McDiarmid's inequality: see [29]). *Suppose $X_1, \dots, X_m$ are independent Bernoulli random variables and suppose $b_1, \dots, b_m \in [0, B]$. Suppose X is a real-valued random variable determined by $X_1, \dots, X_m$ such that changing the outcome of X_i changes X by at most b_i for all $i \in [m]$. Then, for all $\lambda > 0$, we have*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geqslant \lambda] \leqslant 2 \exp\left(-\frac{2\lambda^2}{B \sum_{i=1}^m b_i}\right).$$

In one of our proofs we consider exposure martingales; that is, suppose we have a random variable X that is determined by independent random variables $Y_1, \dots, Y_n$ and we define $X_t := \mathbb{E}[X \mid Y_1, \dots, Y_t]$. Then it is well known that $(X_t)_{t \geqslant 0}$ is a martingale, the so-called *exposure martingale* for X. Note that $X_0 = \mathbb{E}[X]$ and $X_n = X$. Now Freedman's martingale concentration inequality can be used to obtain concentration of X around its mean.

Lemma 2.1 (Freedman's inequality [9]). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t \geqslant 0}$ be a filtration of $\mathcal{F}$. Let $(X_t)_{t \geqslant 0}$ be a martingale adapted to $(\mathcal{F}_t)_{t \geqslant 0}$. Suppose $\sum_{t \geqslant 0} \mathbb{E}[|X_{t+1} - X_t| \mid \mathcal{F}_t] \leqslant \sigma$ and that $|X_{t+1} - X_t| \leqslant C$ for all t. Then, for any $\lambda > 0$,*

$$\mathbb{P}[|X_t - X_0| \geqslant \lambda \textit{ for some } t] \leqslant 2 \exp\left(-\frac{\lambda^2}{2C(\lambda + \sigma)}\right).$$

For $a, b, c \in \mathbb{R}$, we write $a = b \pm c$ whenever $a \in [b - c, b + c]$. For $a, b, c \in (0, 1]$, we sometimes write $a \ll b \ll c$ in our statements, meaning that there are increasing functions $f, g : (0, 1] \to (0, 1]$ such that whenever $a \leqslant f(b)$ and $b \leqslant g(c)$, then the subsequent result holds. We assume that large numbers are integers if that does not affect the argument.

3. Proof

We first sketch our proof. For simplicity, we first consider only the setting of Theorem 1.2. We split $\mathcal{H}$ randomly into p vertex-disjoint induced subgraphs $\mathcal{H}_1, \dots, \mathcal{H}_p$ and let $\mathcal{H}'$ be the union of those. With high probability, $\Delta(\mathcal{H}_i) \approx \Delta(\mathcal{H}) p^{-(r-1)}$ for each i, and for a given weight function ω, we have $\omega(E(\mathcal{H}')) \approx \omega(E(\mathcal{H})) p^{-(r-1)}$. After fixing such a partition, we utilize the theorem of Molloy and Reed to find, for each $i \in [p]$, a partition of $E(\mathcal{H}_i)$ into $M \approx \Delta(\mathcal{H}) p^{-(r-1)}$ matchings. Finally, we select a matching from each partition uniformly at random, and let $\mathcal{M}$ be the union of these matchings. Clearly every edge in $\mathcal{H}'$ is contained in $\mathcal{M}$ with probability M^{-1}, so $\mathbb{E}[\omega(\mathcal{M})] = \omega(E(\mathcal{H}')) M^{-1} \approx \omega(E(\mathcal{H}))/\Delta(\mathcal{H})$. Moreover, the individual effect of the matching chosen in $\mathcal{H}_i$ is relatively small, so we could hope to use McDiarmid's inequality to establish concentration. So far, this approach is the same as that taken by Alon and Yuster. However, the individual effects of the matchings chosen in $\mathcal{H}_i$ are in fact still too large in our setting to apply McDiarmid's inequality. One important new ingredient in our proof is that we partition

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

each $\mathcal{H}_i$ further into edge-disjoint subgraphs $\mathcal{H}_{i,1}, \ldots, \mathcal{H}_{i,q}$ such that $\omega(E(\mathcal{H}_{i,j}))$ is of magnitude $\omega(E(\mathcal{H}_i))/q$, and then apply Theorem 2.1 to each $\mathcal{H}_{i,j}$. This gives, as above, a partition of $\mathcal{H}_i$ into matchings, from which we still choose one uniformly at random. However, the individual effect of each matching chosen has now been drastically reduced, which allows us to apply McDiarmid's inequality with the desired parameters.

In the setting of Theorem 1.2, the partition of each $\mathcal{H}_i$ into edge-disjoint subgraphs $\mathcal{H}_{i,1}, \ldots, \mathcal{H}_{i,q}$ could be done easily with a generalized Chernoff bound. However, in the setting of Theorem 1.3, we are not aware of a conventional concentration inequality that suits our needs for this step (in particular, since q is rather large). Thus we first prove a tool that will achieve this for us. Roughly speaking, what we require is the following. Let $\mathcal{H}$ be a 'directed' ℓ-graph on V, that is, a collection of ordered ℓ-subsets of V. Let $f\colon V \to [q]$ be obtained by choosing $f(v) \in [q]$ uniformly at random for each vertex v independently. For each directed edge $e = (v_1, \ldots, v_\ell)$, let $f(e) := (f(v_1), \ldots, f(v_\ell))$. For a fixed 'pattern' $\alpha \in [q]^\ell$, let X_α denote the number of $e \in E(\mathcal{H})$ with $f(e) = \alpha$. Clearly, for each edge e, we have $\mathbb{P}[f(e) = \alpha] = q^{-\ell}$, and thus $\mathbb{E}[X_\alpha] = q^{-\ell} e(\mathcal{H})$. We would like to know that X_α is concentrated around its mean, even when q is quite large.

For simplicity we will only consider the case when $\mathcal{H}$ is an ℓ-graph, the vertex set V is ordered, and each edge of $\mathcal{H}$ obtains its direction from the ordering of V. Thus our set-up is as follows. Let $(V, <)$ be an ordered set. Let $f\colon V \to [q]$ be obtained by choosing $f(v) \in [q]$ uniformly at random for each $v \in V$ independently. For each ℓ-set $e = \{v_1, \ldots, v_\ell\}$ with $v_1 < \cdots < v_\ell$, let $f(e) := (f(v_1), \ldots, f(v_\ell))$. For a fixed 'pattern' $\alpha \in [q]^\ell$, let $E_\alpha = E_\alpha(f)$ denote the (random) set of all $e \in \binom{V}{\ell}$ with $f(e) = \alpha$. Given an ℓ-tuple weight function ω on V, the following theorem shows that the random variable $\omega(E_\alpha)$ is concentrated around its mean.

Theorem 3.1. *Suppose* $(V, <)$, f, ℓ, α, ω *are as above. Suppose that* $g \geqslant 24\ell^3(\ell + 1 + \log|V|)$. *Define*

$$M := q^{-\ell} \max_{k \in [\ell]} \{\|\omega\|_k q^k g^{k-1}\}.$$

Then, for any $\lambda > 0$, *we have*

$$\mathbb{P}[|\omega(E_\alpha) - \mathbb{E}[\omega(E_\alpha)]| \geqslant \lambda] \leqslant 2^\ell \exp\left(-\frac{\lambda^2}{12\ell^2 M(\lambda + \mathbb{E}[\omega(E_\alpha)])}\right) + \exp\left(-\frac{g}{24\ell^2}\right).$$

Proof. Let $n := |V|$ and let $v_1 < \cdots < v_n$ be the ordered elements of V and write $\alpha = (\alpha_1, \ldots, \alpha_\ell)$. For $t \in [n]_0$, let

$$X_t := \mathbb{E}[\omega(E_\alpha) \mid f(v_1), \ldots, f(v_t)]$$

(and $X_t := X_n$ for $t \geqslant n$). Hence $X = (X_t)_{t \geqslant 0}$ is the so-called exposure martingale for $\omega(E_\alpha)$, where the labels $f(v_i)$ are revealed one by one. In particular, $X_0 = \mathbb{E}[\omega(E_\alpha)]$ and $X_n = \omega(E_\alpha)$.

For $k \in [\ell]$ and a k-tuple weight function ω' on V, let

$$M_k(\omega') := q^{-k} \max_{i \in [k]} \{\|\omega'\|_i q^i g^{i-1}\}.$$

Note that we have

$$M_k(\omega') q^k \leqslant M_\ell(\omega') q^\ell. \tag{3.1}$$

Let $M_k := M_k(\omega)$ and note that $M = M_\ell$.

We prove the theorem by induction on ℓ (with $(V, <)$ and g being fixed). Thus assume first that $\ell = 1$. (This case is also contained in the inductive step below with no inductive hypothesis being needed, but the short proof here may serve as a warm-up.) Observe that

$$X_t(f) - X_{t-1}(f) = \omega(\{v_t\})(\mathbb{1}_{f(v_t)=\alpha_1} - 1/q) \quad \text{for } t \in [n].$$

Hence we can directly apply Freedman's inequality to obtain (observe that $M_1 = \|\omega\|_1$)

$$\mathbb{P}[|\omega(E_\alpha) - \mathbb{E}[\omega(E_\alpha)]| \geqslant \lambda] \leqslant 2\exp\left(-\frac{\lambda^2}{2\|\omega\|_1(\lambda + \sum_{t\in[n]} 2\omega(\{v_t\})/q)}\right)$$
$$\leqslant 2\exp\left(-\frac{\lambda^2}{4M_1(\lambda + \mathbb{E}[\omega(E_\alpha)])}\right),$$

as desired.

Suppose now that $\ell \geqslant 2$. In order to apply induction, we need to introduce some more notation. For $t \in [n]$ and $k \in [\ell - 1]_0$, let $\omega^{t,k}\colon \binom{V}{k} \to [0, \infty)$ be defined as (where $j_1 < \ldots < j_k$)

$$\omega^{t,k}(\{v_{j_1}, \ldots, v_{j_k}\}) := \sum_{\substack{j_{k+1}<\ldots<j_\ell \\ j_k<j_{k+1}=t}} \omega(\{v_{j_1}, \ldots, v_{j_\ell}\}).$$

Moreover, let $\omega^{\leqslant t,k}\colon \binom{V}{k} \to [0, \infty)$ be defined by $\omega^{\leqslant t,k}(S) := \sum_{s\leqslant t} \omega^{s,k}(S)$ for all $S \in \binom{V}{k}$. Note that

$$\omega^{\leqslant n,k}(V) = \omega(V) \quad \text{and} \quad \|\omega^{\leqslant n,k}\|_i \leqslant \|\omega\|_i \quad \text{for all } i \in [k]. \tag{3.2}$$

For $k \in [\ell - 1]_0$, let $\alpha[k] := (\alpha_1, \ldots, \alpha_k)$, and define $E_{\alpha[k]} = E_{\alpha[k]}(f)$ as the random set of all k-sets $\{v_{j_1}, \ldots, v_{j_k}\}$ for which $f(v_{j_i}) = \alpha_i$ for all $i \in [k]$, where $j_1 < \cdots < j_k$. For clarity, we briefly discuss the case $k = 0$, when $\omega^{t,0}$ is the function that maps $\emptyset$ to $\sum_{t<j_2<\ldots<j_\ell} \omega(\{v_t, v_{j_2} \ldots, v_{j_\ell}\})$. In particular, we have for all $t \in [n]$ that

$$\omega^{t,0}(\emptyset) \leqslant \omega(\{v_t\}) \leqslant \|\omega\|_1 = M_1, \tag{3.3}$$
$$\omega^{\leqslant t,0}(\emptyset) \leqslant \omega(V). \tag{3.4}$$

Note also that $E_{\alpha[0]} = \{\emptyset\}$.

The purpose of these definitions lies in the following formula for the one-step change of the process X: for $t \in [n]$, we have

$$X_t(f) - X_{t-1}(f) = \sum_{k\in[\ell-1]_0} \omega^{t,k}(E_{\alpha[k]}(f)) \cdot (\mathbb{1}_{f(v_t)=\alpha_{k+1}} - 1/q) \cdot q^{-(\ell-(k+1))}.$$

Clearly $|\mathbb{1}_{f(v_t)=\alpha_{k+1}} - 1/q| \leqslant 1$ and $\mathbb{E}[|\mathbb{1}_{f(v_t)=\alpha_{k+1}} - 1/q|] = 2(1 - 1/q)/q \leqslant 2/q$. Hence, for the absolute change and expected absolute change of the process X in one step, we obtain the following bounds:

$$|X_t - X_{t-1}| \leqslant \sum_{k\in[\ell-1]_0} \omega^{t,k}(E_{\alpha[k]}) \cdot q^{k+1-\ell}, \tag{3.5}$$
$$\mathbb{E}[|X_t - X_{t-1}| \mid f(v_1), \ldots, f(v_{t-1})] \leqslant \sum_{k\in[\ell-1]_0} 2\omega^{t,k}(E_{\alpha[k]}) \cdot q^{k-\ell}. \tag{3.6}$$

Note that $\omega^{t,k}(E_{\alpha[k]})$ is itself a random variable when $k > 0$. Unfortunately, its deterministic upper bound is not good enough to apply Freedman's inequality directly to the martingale $(X_t)_{t\geqslant 0}$. We apply a common trick by defining a stopped process $Y = (Y_t)_{t\geqslant 0}$ which is equal to X as long as the random variables $\omega^{t,k}(E_{\alpha[k]})$ behave nicely, and then 'freezes'. We can then apply Freedman's inequality to Y. Finally, we need to show that the process is unlikely to freeze, implying that the concentration result for Y transfers to X. For this, we employ the statement inductively with $\omega^{t,k}, \omega^{\leqslant n,k}, \alpha[k]$.

We define two types of stopping times for X. For $k \in [\ell - 1]$, let

$$\tau'_k := \min_{t\in[n-1]} \left\{\omega^{\leqslant t+1,k}(E_{\alpha[k]}) \geqslant \omega(V)q^{-k} + \lambda q^{\ell-k}\right\} \wedge n. \tag{3.7}$$

Moreover, for $k \in [\ell-1]$ and $t \in [n-1]$, define

$$\tau_k^t := \begin{cases} t & \text{if } \omega^{t+1,k}(E_{\alpha[k]}) \geqslant 2M_{k+1}, \\ n & \text{otherwise.} \end{cases} \tag{3.8}$$

Let $\tau := \min_{t\in[n],k\in[\ell-1]}\{\tau_k', \tau_k^t\}$. Note that $\omega^{t+1,k}(E_{\alpha[k]})$ is fully determined by $f(v_1),\ldots,f(v_t)$, since $\omega^{t+1,k}(S)=0$ whenever S contains a vertex v_j with $j \geqslant t+1$. Thus τ is indeed a stopping time for X. We define $Y=(Y_t)_{t\geqslant 0}$ by $Y_t := X_{t\wedge\tau}$, and let $\Delta Y_t := Y_t - Y_{t-1}$. By the optional stopping theorem Y is also a martingale (see *e.g.* [25]), and thus we can apply Freedman's inequality. To this end, we next bound the absolute and expected one step change for Y.

We claim that $|\Delta Y_t| \leqslant 2\ell M_\ell$ for all t. Indeed, if $t \geqslant \tau+1$, then trivially $|\Delta Y_t|=0$, and whenever $t \leqslant \tau$, then

$$|\Delta Y_t| \overset{(3.5)}{\leqslant} \sum_{k\in[\ell-1]_0} \omega^{t,k}(E_{\alpha[k]})\cdot q^{k+1-\ell} \overset{(3.3),(3.8)}{\leqslant} \sum_{k\in[\ell-1]_0} 2M_{k+1}\cdot q^{k+1-\ell} \overset{(3.1)}{\leqslant} 2\ell M_\ell.$$

Similarly

$$\begin{aligned}\sum_{t\geqslant 1}\mathbb{E}[|\Delta Y_t|\,|\,f(v_1),\ldots,f(v_{t-1})] &\overset{(3.6)}{\leqslant} \sum_{t\in[\tau]}\sum_{k\in[\ell-1]_0} 2\omega^{t,k}(E_{\alpha[k]})\cdot q^{k-\ell} \\ &= \sum_{k\in[\ell-1]_0} 2\omega^{\leqslant\tau,k}(E_{\alpha[k]})\cdot q^{k-\ell} \\ &\overset{(3.4),(3.7)}{\leqslant} \sum_{k\in[\ell-1]_0} 2(\omega(V)q^{-k}+\lambda q^{\ell-k})\cdot q^{k-\ell} \\ &= 2\ell(\omega(V)q^{-\ell}+\lambda).\end{aligned}$$

Thus we can apply Freedman's inequality to obtain

$$\begin{aligned}\mathbb{P}[|Y_n-Y_0|\geqslant\lambda] &\leqslant 2\exp\left(-\frac{\lambda^2}{4\ell M_\ell(\lambda+2\ell(\omega(V)q^{-\ell}+\lambda))}\right) \\ &\leqslant 2\exp\left(-\frac{\lambda^2}{12\ell^2 M_\ell(\lambda+\mathbb{E}[\omega(E_\alpha)])}\right).\end{aligned}$$

It remains to show that $Y_n = X_n$ with high probability. We first consider the stopping times τ_k'. Fix $k\in[\ell-1]$ and note that

$$\mathbb{E}[\omega^{\leqslant n,k}(E_{\alpha[k]})] = \omega^{\leqslant n,k}(V)/q^k = \omega(V)/q^k$$

by (3.2). We apply the induction hypothesis to $\omega^{\leqslant n,k}$, with $\lambda q^{\ell-k}$ and k playing the roles of λ and ℓ, and obtain

$$\begin{aligned}\mathbb{P}[\tau_k'<n] &\leqslant \mathbb{P}[\omega^{\leqslant n,k}(E_{\alpha[k]})\geqslant \mathbb{E}[\omega^{\leqslant n,k}(E_{\alpha[k]})]+\lambda q^{\ell-k}] \\ &\leqslant 2^k\exp\left(-\frac{\lambda^2 q^{2(\ell-k)}}{12k^2 M_k(\omega^{\leqslant n,k})(\lambda q^{\ell-k}+\mathbb{E}[\omega^{\leqslant n,k}(E_{\alpha[k]})])}\right)+\exp\left(-\frac{g}{24k^2}\right) \\ &\leqslant 2^k\exp\left(-\frac{\lambda^2}{12k^2 M_\ell(\lambda+\mathbb{E}[\omega(E_\alpha)])}\right)+\exp\left(-\frac{g}{24k^2}\right),\end{aligned}$$

where we have used the fact that $\mathbb{E}[\omega^{\leqslant n,k}(E_{\alpha[k]})] = q^{\ell-k}\mathbb{E}[\omega(E_\alpha)]$ and $M_k(\omega^{\leqslant n,k}) \leqslant M_k(\omega) \leqslant q^{\ell-k}M_\ell$ by (3.2) and (3.1).

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

Next we consider the stopping times τ_k^t. Let $k \in [\ell-1]$ and $t \in [n-1]$. Observe that $\|\omega^{t,k}\|_i \leqslant \|\omega\|_{i+1}$ for all $i \in [k]$. Hence

$$\frac{M_{k+1}(\omega)}{M_k(\omega^{t,k})} = \frac{q^{-k-1}\max_{i\in[k+1]}\{\|\omega\|_i q^i g^{i-1}\}}{q^{-k}\max_{i\in[k]}\{\|\omega^{t,k}\|_i q^i g^{i-1}\}} \geqslant \frac{g\max_{i\in[k+1]}\{\|\omega\|_i q^i g^{i-1}\}}{\max_{i\in[k+1]\setminus\{1\}}\{\|\omega\|_i q^i g^{i-1}\}} \geqslant g.$$

Note that

$$\mathbb{E}[\omega^{t,k}(E_{\alpha[k]})] = q^{-k}\omega^{t,k}(V) \leqslant q^{-k}\|\omega\|_1 \leqslant M_{k+1}.$$

Thus, using induction for $\omega^{t,k}$ with M_{k+1} and k playing the roles of λ and ℓ, we deduce that

$$\begin{aligned}\mathbb{P}\big[\tau_k^t < n\big] &\leqslant \mathbb{P}\big[\omega^{t,k}(E_{\alpha[k]}) \geqslant 2M_{k+1}\big]\\ &\leqslant 2^k \exp\left(-\frac{M_{k+1}}{24k^2 M_k(\omega^{t,k})}\right) + \exp\left(-\frac{g}{24k^2}\right)\\ &\leqslant (2^k+1)\exp\left(-\frac{g}{24k^2}\right).\end{aligned}$$

A union bound now implies that

$$\begin{aligned}\mathbb{P}[\tau < n] &\leqslant \sum_{k=1}^{\ell-1}\left(2^k \exp\left(-\frac{\lambda^2}{12k^2 M_\ell(\lambda + \mathbb{E}[\omega(E_\alpha)])}\right) + (1 + n(2^k+1))\exp\left(-\frac{g}{24k^2}\right)\right)\\ &\leqslant (2^\ell - 2)\exp\left(-\frac{\lambda^2}{12\ell^2 M_\ell(\lambda + \mathbb{E}[\omega(E_\alpha)])}\right) + 2^{\ell+1} n \exp\left(-\frac{g}{24(\ell-1)^2}\right).\end{aligned}$$

Since $(\ell-1)^{-2} - \ell^{-2} \geqslant \ell^{-3}$ and $g/24\ell^3 \geqslant \log(2^{\ell+1}n)$ by assumption, we can finally conclude that

$$\begin{aligned}\mathbb{P}[|\omega(E_\alpha) - \mathbb{E}[\omega(E_\alpha)]| > \lambda] &\leqslant \mathbb{P}[|Y_n - Y_0| \geqslant \lambda] + \mathbb{P}[\tau < n]\\ &\leqslant 2^\ell \exp\left(-\frac{\lambda^2}{12\ell^2 M_\ell(\lambda + \mathbb{E}[\omega(E_\alpha)])}\right) + \exp\left(-\frac{g}{24\ell^2}\right).\end{aligned}$$

This completes the proof. □

We are now ready to prove Theorem 1.3. The proof proceeds in three steps as outlined at the beginning of this section.

Proof of Theorem 1.3. We can assume that $\varepsilon = \delta/50L^2r^2$.

Step 1: Random vertex partition.

Let $p := \Delta^{20Lr\varepsilon}$. We will first partition $V(\mathcal{H})$ into p subsets $V_1, \ldots, V_p$. For each $i \in [p]$, let $\mathcal{H}_i := \mathcal{H}[V_i]$. For an edge $e \in E(\mathcal{H})$, let $\tau(e) = i$ if $e \in E(\mathcal{H}_i)$, and let $\tau(e) = 0$ if no such i exists. For a tuple $\mathcal{E} = (e_1, \ldots, e_\ell) \in \binom{E(\mathcal{H})}{\ell}$, define the multiset $\tau(\mathcal{E}) := \{\tau(e_1), \ldots, \tau(e_\ell)\}$. Let $\mathcal{J}_\ell$ be the set of all multisets of size ℓ with elements in $[p]$. For $J \in \mathcal{J}_\ell$, let supp(J) be the underlying set. We further define $\pi(J)$ as the number of functions $f\colon [\ell] \to \text{supp}(J)$ with $\{f(1), \ldots, f(\ell)\} = J$. For all $\ell \in [L]$ and $J \in \mathcal{J}_\ell$, we define E_J as the set of all $\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}$ with $\tau(\mathcal{E}) = J$.

We claim that there exists a partition $V_1, \ldots, V_p$ of $V(\mathcal{H})$ such that the following hold:

(a) $\Delta(\mathcal{H}_i) \leqslant (1 + \Delta^{-2\varepsilon})\Delta/p^{r-1}$ for all $i \in [p]$,
(b) $\omega(E_J) = (1 \pm \Delta^{-2\varepsilon})\omega(E(\mathcal{H}))(\pi(J)/p^{r\ell})$ for all $\ell \in [L]$, $\omega \in \mathcal{W}_\ell$ and $J \in \mathcal{J}_\ell$.

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

This can be seen using a probabilistic argument. For every vertex $x \in V(\mathcal{H})$ independently, choose an index $i \in [p]$ uniformly at random and assign x to V_i. We now show that (a) and (b) hold with high probability, implying that such a partition exists.

For (a), consider a vertex $x \in V(\mathcal{H})$ and $i \in [p]$. Let X be the number of edges e containing x for which $e \setminus \{x\} \subseteq V_i$. For each edge e containing x, we have $\mathbb{P}[e \setminus \{x\} \subseteq V_i] = (1/p)^{r-1}$. Thus $\mathbb{E}[X] = \deg_{\mathcal{H}}(x)/p^{r-1} \leqslant \Delta/p^{r-1}$. Note that for any other vertex $y \neq x$, the random label that we choose for y affects X by at most $\deg_{\mathcal{H}}(xy) \leqslant \Delta^c(\mathcal{H})$. Note that

$$\sum_{y \in V(\mathcal{H})\setminus\{x\}} \deg_{\mathcal{H}}(xy) = \deg_{\mathcal{H}}(x)(r-1) \leqslant \Delta r.$$

Thus, using McDiarmid's inequality, we deduce that

$$\mathbb{P}[X - \mathbb{E}[X] \geqslant \Delta^{1-2\varepsilon}/p^{r-1}] \leqslant 2\exp\left(-\frac{2\Delta^{2-4\varepsilon}}{\Delta^c(\mathcal{H})\Delta r p^{2r-2}}\right) \leqslant 2\exp\left(-\Delta^{\delta-45Lr^2\varepsilon}\right) \leqslant \exp\left(-\Delta^{\varepsilon}\right).$$

With a union bound over all (non-isolated) vertices (there are at most $re(\mathcal{H}) \leqslant r\exp(\Delta^{\varepsilon^2})$ non-isolated vertices) and $i \in [p]$, we can infer that with high probability (a) holds.

For (b), consider $\ell \in [L]$, $\omega \in \mathcal{W}_\ell$ and $J \in \mathcal{J}_\ell$. For an edge $e \in E(\mathcal{H})$ and $i \in [p]$, we have $\mathbb{P}[e \in E(\mathcal{H}_i)] = p^{-r}$. Thus, for $\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}$, we have $\mathbb{P}[\tau(\mathcal{E}) = J] = \pi(J)p^{-r\ell}$ if the edges in $\mathcal{E}$ are pairwise disjoint, and $\omega(\mathcal{E}) = 0$ otherwise since ω is clean. Hence $\mathbb{E}[\omega(E_J)] = \omega(E(\mathcal{H}))(\pi(J)/p^{r\ell})$. We now establish concentration. For any vertex x, the random label chosen for x affects $\omega(E_J)$ by at most $\omega(E_x^\ell)$, where E_x^ℓ is the set of all $\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}$ for which x is contained in some edge of $\mathcal{E}$. Note that

$$\omega(E_x^\ell) \leqslant \Delta\|\omega\|_1 \text{ for all } x \in V(\mathcal{H}), \quad \text{and} \sum_{x \in V(\mathcal{H})} \omega\left(E_x^\ell\right) = r\ell\omega(E(\mathcal{H})).$$

Thus we can use McDiarmid's inequality to conclude that

$$\begin{aligned}
\mathbb{P}[\omega(E_J) \neq (1 \pm \Delta^{-2\varepsilon})\mathbb{E}[\omega(E_J)]] &\leqslant 2\exp\left(-\frac{2\mathbb{E}[\omega(E_J)]^2}{\Delta\|\omega\|_1 r\ell\omega(E(\mathcal{H}))\Delta^{4\varepsilon}}\right) \\
&\leqslant 2\exp\left(-\frac{\omega(E(\mathcal{H}))}{\|\omega\|_1\Delta^{1+45L^2r^2\varepsilon}}\right) \\
&\leqslant 2\exp\left(-\Delta^{\delta-45L^2r^2\varepsilon}\right) \leqslant \exp\left(-\Delta^{\varepsilon}\right),
\end{aligned}$$

which together with a union bound over all $\ell \in [L]$, $\omega \in \mathcal{W}_\ell$ and $J \in \mathcal{J}_\ell$ proves (b).

Step 2: Random edge partition.

Let $\mathcal{H}' := \bigcup_{i \in [p]} \mathcal{H}_i$. For each $i \in [p]$, we now partition $\mathcal{H}_i$ further into $q := \Delta^{1-20(r-1+1/4L)Lr\varepsilon}$ edge-disjoint subgraphs $\mathcal{H}_{i,1}, \ldots, \mathcal{H}_{i,q}$. Note that

$$p^{r-1}q = \Delta^{1-5r\varepsilon} \quad \text{and} \quad p^r q \geqslant \Delta^{1+15Lr\varepsilon}. \tag{3.9}$$

We do so (for all i at once) by choosing a function $f\colon E(\mathcal{H}') \to [q]$ and then let $\mathcal{H}_{i,j}$ consist of all edges $e \in E(\mathcal{H}_i)$ with $f(e) = j$, for all $i \in [p], j \in [q]$.

For $\ell \in [L]$, $J \in \mathcal{J}_\ell$ and a function $\sigma\colon \mathrm{supp}(J) \to [q]$, let $E_{J,\sigma}$ be the set of all $\mathcal{E} \in E_J$ for which $\sigma(\tau(e)) = f(e)$ for all $e \in \mathcal{E}$.

We claim that there exists a choice of f such that the following hold:

(A) $\Delta(\mathcal{H}_{i,j}) \leqslant (1 + 2\Delta^{-2\varepsilon})\Delta/qp^{r-1}$ for all $i \in [p], j \in [q]$,
(B) $\Delta^c(\mathcal{H}_{i,j}) \leqslant \Delta^{\varepsilon}$ for all $i \in [p], j \in [q]$,
(C) $\omega(E_{J,\sigma}) \leqslant 2\ell!\omega(E(\mathcal{H}))/q^\ell p^{r\ell}$ for all $\ell \in [L]$, $\omega \in \mathcal{W}_\ell$, $J \in \mathcal{J}_\ell$ and $\sigma\colon \mathrm{supp}(J) \to [q]$.

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

This again can be seen using a probabilistic argument. For each $e \in E(\mathcal{H}')$ independently, choose $f(e) \in [q]$ uniformly at random.

For (A), fix $i \in [p], j \in [q]$ and a vertex $x \in V(\mathcal{H}_i)$. Note that

$$\mathbb{E}[\deg_{\mathcal{H}_{i,j}}(x)] = \deg_{\mathcal{H}_i}(x)/q \leqslant (1+\Delta^{-2\varepsilon})\Delta/qp^{r-1}$$

by (a). Thus, by Chernoff's bound, we have

$$\mathbb{P}[\deg_{\mathcal{H}_{i,j}}(x) - \mathbb{E}[\deg_{\mathcal{H}_{i,j}}(x)] \geqslant \Delta^{1-2\varepsilon}/qp^{r-1}] \leqslant 2\exp\left(-\frac{\Delta^{1-4\varepsilon}}{3qp^{r-1}}\right) \overset{(3.9)}{\leqslant} \exp(-\Delta^{\varepsilon}).$$

Similarly, for (B), fix $i \in [p], j \in [q]$ and two distinct vertices $x, y \in V(\mathcal{H}_i)$. Note that $\mathbb{E}[\deg_{\mathcal{H}_{i,j}}(xy)] = \deg_{\mathcal{H}_i}(xy)/q \leqslant \Delta^c(\mathcal{H})/q \leqslant 1$. Thus, by Chernoff's bound, we have

$$\mathbb{P}[\deg_{\mathcal{H}_{i,j}}(xy) \geqslant \Delta^{\varepsilon}] \leqslant 2\exp(-\Delta^{\varepsilon}).$$

To prove (C), consider $\ell \in [L]$, $\omega \in \mathcal{W}_\ell$, $J \in \mathcal{J}_\ell$ and $\sigma : \mathrm{supp}(J) \to [q]$. First note that

$$\mathbb{E}[\omega(E_{J,\sigma})] = \omega(E_J)/q^\ell \leqslant \frac{3}{2}\ell!\omega(E(\mathcal{H}))/q^\ell p^{r\ell}$$

by (b). We now aim to employ Theorem 3.1 with $E(\mathcal{H}')$ playing the role of V. Let $<$ be an ordering of $E(\mathcal{H}')$ in which the edges of $\mathcal{H}_i$ precede those of $\mathcal{H}_{i'}$ whenever $i < i'$. Write $J = \{j_1, \dots, j_\ell\}$ such that $j_1 \leqslant \dots \leqslant j_\ell$ and define $\alpha := (\sigma(j_1), \dots, \sigma(j_\ell)) \in [q]^\ell$. Hence, for $\mathcal{E} \in E_J$, we have $\mathcal{E} \in E_{J,\sigma}$ if and only if $f(e_i) = \sigma(j_i)$ for all $i \in [\ell]$, where $\mathcal{E} = \{e_1, \dots, e_\ell\}$ with $e_1 < \dots < e_\ell$. Consequently, with notation as in Theorem 3.1, we have $E_{J,\sigma} = E_J \cap E_\alpha$. Thus $\omega(E_{J,\sigma}) = \omega_J(E_\alpha)$, where $\omega_J(\mathcal{E}) := \omega(\mathcal{E})\mathbb{1}_{\mathcal{E}\in E_J}$.

We now apply Theorem 3.1 with $E(\mathcal{H}')$, ℓ, ω_J, $\frac{1}{2}\ell!\omega(E(\mathcal{H}))/q^\ell p^{r\ell}$, $\Delta^{2\varepsilon}$ playing the roles of V, ℓ, ω, λ, g, respectively. For $k \in [\ell]$, we have (recall that $\omega(E(\mathcal{H})) \geqslant \|\omega\|_k \Delta^{k+\delta}$ by assumption)

$$\|\omega_J\|_k q^k g^{k-1} \leqslant \|\omega\|_k \Delta^k \leqslant \omega(E(\mathcal{H}))\Delta^{-\delta}.$$

Hence we infer that (note that $\mathbb{E}[\omega_J(E_\alpha)] + \lambda \leqslant 4\lambda$)

$$\begin{aligned}
\mathbb{P}[\omega_J(E_\alpha) \geqslant \mathbb{E}[\omega_J(E_\alpha)] + \lambda] &\leqslant 2^\ell \exp\left(-\frac{\lambda}{48\ell^2 q^{-\ell}\omega(E(\mathcal{H}))\Delta^{-\delta}}\right) + \exp\left(-\frac{\Delta^{2\varepsilon}}{24\ell^2}\right) \\
&\leqslant 2^\ell \exp\left(-\frac{\Delta^\delta}{96\ell p^{r\ell}}\right) + \exp\left(-\frac{\Delta^{2\varepsilon}}{24\ell^2}\right) \\
&\leqslant \exp(-\Delta^\varepsilon).
\end{aligned}$$

A union bound implies that the random choice of f satisfies (A), (B) and (C) simultaneously with positive probability. From now on, fix such a function f.

Step 3: Random matchings.

Let $\widetilde{\Delta} := (1+2\Delta^{-2\varepsilon})\Delta/qp^{r-1} \geqslant \Delta^{5r\varepsilon}$ by (3.9) and $M := (1+\Delta^{-2\varepsilon})\widetilde{\Delta}$. Note that

$$p^{r-1}qM = (1 \pm 4\Delta^{-2\varepsilon})\Delta. \tag{3.10}$$

By (A), we have $\Delta(\mathcal{H}_{i,j}) \leqslant \widetilde{\Delta}$. Moreover, by (B), $\Delta^c(\mathcal{H}_{i,j}) \leqslant \Delta^\varepsilon \leqslant \widetilde{\Delta}^{1/5r}$. Thus, for all $i \in [p], j \in [q]$, we can apply Theorem 2.1 (with $\delta = 1/2$, say) to obtain a partition of $E(\mathcal{H}_{i,j})$ into M matchings. This yields a partition of each $E(\mathcal{H}_i)$ into $q \cdot M$ matchings $\mathcal{M}_{i,1}, \dots, \mathcal{M}_{i,qM}$.

Now, for each $i \in [p]$ independently, pick an index $s_i \in [qM]$ uniformly at random, and define

$$\mathcal{M} := \bigcup_{i\in[p]} \mathcal{M}_{i,s_i}.$$

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

Clearly $\mathcal{M}$ is a matching in $\mathcal{H}' \subseteq \mathcal{H}$. Moreover, every edge of $\mathcal{H}'$ belongs to $\mathcal{M}$ with probability $1/qM$.

Now consider $\ell \in [L]$ and $\omega \in \mathcal{W}_\ell$. We first determine the expected value of $\omega(\mathcal{M})$. By linearity,

$$\mathbb{E}[\omega(\mathcal{M})] = \sum_{\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}} \omega(\mathcal{E})\mathbb{P}[\mathcal{E} \subseteq \mathcal{M}].$$

We analyse this sum according to the different types of $\mathcal{E}$. For $k \in [\ell]$, let $\mathcal{J}_{\ell,k}$ be the set of all $J \in \mathcal{J}_\ell$ with $|\mathrm{supp}(J)| = k$. Consider $\mathcal{E} \in \binom{E(\mathcal{H})}{\ell}$ and let $J := \tau(\mathcal{E})$. Note that if $0 \in J$, then some edge in $\mathcal{E}$ does not belong to $\mathcal{H}'$ and hence $\mathbb{P}[\mathcal{E} \subseteq \mathcal{M}] = 0$. Hence we can assume that $J \in \mathcal{J}_\ell$. If $J \in \mathcal{J}_{\ell,\ell}$, then the edges in $\mathcal{E}$ belong to $\mathcal{M}$ independently with probability $1/qM$, and hence $\mathbb{P}[\mathcal{E} \subseteq \mathcal{M}] = (qM)^{-\ell}$. Now suppose $J \in \mathcal{J}_{\ell,k}$ for some $k \in [\ell - 1]$. By the definition of $\mathcal{M}$, if $e, e' \in \mathcal{E}$ with $\tau(e) = \tau(e')$, then $\mathbb{P}[\mathcal{E} \subseteq \mathcal{M}] = 0$ if $e \in E(\mathcal{H}_{\tau(e),j})$ and $e' \in E(\mathcal{H}_{\tau(e),j'})$ for distinct j, j'. Hence we can further assume that $\mathcal{E} \in E_{J,\sigma}$ for some $\sigma : \mathrm{supp}(J) \to [q]$. We then have $\mathbb{P}[\mathcal{E} \subseteq \mathcal{M}] \in \{0, (qM)^{-k}\}$. Altogether, we deduce that

$$\mathbb{E}[\omega(\mathcal{M})] = \sum_{J \in \mathcal{J}_{\ell,\ell}} \omega(E_J)(qM)^{-\ell} \pm \sum_{k=1}^{\ell-1} \sum_{J \in \mathcal{J}_{\ell,k}, \sigma : \mathrm{supp}(J) \to [q]} \omega(E_{J,\sigma})(qM)^{-k}.$$

We will show that the first sum is the dominant term. Clearly $|\mathcal{J}_{\ell,\ell}| = \binom{p}{\ell}$. Thus, using (b), we infer that

$$\sum_{J \in \mathcal{J}_{\ell,\ell}} \omega(E_J)(qM)^{-\ell} = \binom{p}{\ell} \cdot (1 \pm \Delta^{-2\varepsilon}) \frac{\ell! \omega(E(\mathcal{H}))}{p^{r\ell}} \cdot \frac{1}{(qM)^\ell} \overset{(3.10)}{=} (1 \pm \Delta^{-3\varepsilon/2})\omega(E(\mathcal{H}))/\Delta^\ell.$$

For $k \in [\ell - 1]$, employing (C) and

$$|\mathcal{J}_{\ell,k}| = \binom{p}{k}\binom{\ell-1}{k-1},$$

we deduce that

$$\sum_{J \in \mathcal{J}_{\ell,k}, \sigma : \mathrm{supp}(J) \to [q]} \omega(E_{J,\sigma})(qM)^{-k} \leqslant p^k 2^\ell q^k \frac{2\ell! \omega(E(\mathcal{H}))}{q^\ell p^{r\ell}} \cdot \frac{1}{(qM)^k} \leqslant \frac{\omega(E(\mathcal{H}))}{\Delta^{\ell+14\varepsilon}},$$

where in the last inequality we used

$$\frac{p^k q^k}{q^\ell p^{r\ell} (qM)^k} = \frac{1}{(p^r q)^{\ell-k} (p^{r-1} qM)^k}$$

together with $(p^r q)^{\ell-k} \geqslant \Delta^{\ell-k+15\varepsilon}$ by (3.9) and $(p^{r-1}qM)^k \geqslant \frac{1}{2}\Delta^k$ by (3.10). Putting everything together, we obtain

$$\mathbb{E}[\omega(\mathcal{M})] = (1 \pm 2\Delta^{-3\varepsilon/2})\omega(E(\mathcal{H}))/\Delta^\ell.$$

Finally, we need to bound the effect of each random variable s_i. Note that each outcome of the variables $s_1, \ldots, s_p$ induces a function $\sigma : [p] \to [q]$, where $\sigma(i)$ is the unique $j \in [q]$ for which $\mathcal{M}_{i,s_i}$ was one of the matchings coming from $E(\mathcal{H}_{i,j})$, and each tuple $\mathcal{E} \subseteq \mathcal{M}$ satisfies $\mathcal{E} \in E_{J,\sigma|_{\mathrm{supp}(J)}}$, where $J = \tau(\mathcal{E}) \in \mathcal{J}_\ell$. Since changing the value of s_i only affects those $\mathcal{E}$ with $i \in \tau(\mathcal{E})$, we have that the effect of s_i on $\omega(\mathcal{M})$ is at most

$$\max_{\sigma : [p] \to [q]} \sum_{J \in \mathcal{J}_\ell : i \in J} \omega(E_{J,\sigma|_{\mathrm{supp}(J)}}) \overset{(C)}{\leqslant} p^{\ell-1} \frac{2\ell! \omega(E(\mathcal{H}))}{q^\ell p^{r\ell}} \overset{(3.9)}{=} \frac{2\ell! \omega(E(\mathcal{H}))}{p\Delta^{(1-5r\varepsilon)\ell}} \leqslant \frac{\omega(E(\mathcal{H}))}{\Delta^{\ell+14Lr\varepsilon}}.$$

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

Thus, using McDiarmid's inequality, we deduce that

$$\mathbb{P}[\omega(\mathcal{M}) \neq (1 \pm \Delta^{-2\varepsilon})\mathbb{E}[\omega(\mathcal{M})]] \leqslant 2\exp\left(-\frac{2\Delta^{-4\varepsilon}\mathbb{E}[\omega(\mathcal{M})]^2}{p \cdot (\omega(E(\mathcal{H}))/\Delta^{\ell+14Lr\varepsilon})^2}\right)$$
$$\leqslant 2\exp\left(-\frac{\Delta^{28Lr\varepsilon-4\varepsilon}}{p}\right)$$
$$\leqslant \exp(-\Delta^{\varepsilon}).$$

A union bound over all $\ell \in [L]$ and $\omega \in \mathcal{W}_\ell$ completes the proof. □

4. Applications

In this section we provide a small exposition of applications of Theorem 1.3. In Section 4.1 we deduce the existence of approximate Steiner systems that behave 'randomly', for example with respect to subgraph statistics. Then we briefly explain how we apply Theorem 1.3 in two forthcoming papers [6, 7] on rainbow embeddings and approximate decompositions.

4.1 Pseudorandom Steiner systems

Recall that an (n, k, t)-Steiner system is a set $\mathcal{S}$ of k-subsets of some n-set V such that every t-subset of V is contained in exactly one k-set in $\mathcal{S}$. We now view such an $\mathcal{S}$ as a k-graph. Note that any subgraph of $\mathcal{S}$ has the following property: any two of its edges intersect in less than t vertices; we will simply say that such graphs are *t-avoiding*. For $t = 2$, such hypergraphs are often called 'linear' or 'simple'. Now, for a fixed t-avoiding k-graph F, we may ask how many (labelled) copies of F exist in $\mathcal{S}$. Since $|\mathcal{S}| = \binom{n}{t}/\binom{k}{t}$, the edge density of $\mathcal{S}$ is (for large n) approximately $p := (k-t)!n^{-k+t}$. In a random k-graph with this density, we would expect $p^{e(F)}n^{v(F)}$ labelled copies of F. Of course this only makes sense when $(-k+t)e(F) + v(F) > 0$, or equivalently, when the average degree of F is less than $k/(k-t)$. Moreover, in order to be able to obtain precise counts for F, one needs this condition for all non-empty subgraphs of F. We thus define the *maximum average degree of F*, denoted $\mathrm{mad}(F)$, as the maximum of $ke(F')/v(F')$ over all non-empty subgraphs F' of F. For two k-graphs F, G, let $\mathrm{inj}(F, G)$ be the number of labelled copies of F in G, that is, the number of injections $f\colon V(F) \to V(G)$ for which $f(e) \in E(G)$ for all $e \in E(F)$.

As one application of Theorem 1.3, we show that there exist approximate Steiner systems whose subgraph statistics resemble the random model.

Theorem 4.1. *Suppose $1/n \ll \varepsilon \ll 1/k, 1/v$ and $t \in \{2, \dots, k-1\}$. Let $\mathcal{F}$ be the family of all t-avoiding k-graphs F with $v(F) \leqslant v$ and $\mathrm{mad}(F) < k/(k-t)$, and let $p := (k-t)!n^{-k+t}$. There exists a partial (n, k, t)-Steiner system $\mathcal{S}$ with $|\mathcal{S}| \geqslant (1 - n^{-\varepsilon})\binom{n}{t}/\binom{k}{t}$ such that*

$$\mathrm{inj}(F, \mathcal{S}) = (1 \pm n^{-\varepsilon})p^{e(F)}n^{v(F)} \quad \textit{for all } F \in \mathcal{F}.$$

Proof. Choose a new constant $\delta > 0$ such that $1/n \ll \varepsilon \ll \delta \ll 1/k, 1/v$.

For $e \in \binom{[n]}{k}$, let $\pi(e) := \binom{e}{t}$. Define $\mathcal{H}$ as the $\binom{k}{t}$-uniform hypergraph from Section 1.1, that is,

$$V(\mathcal{H}) := \binom{[n]}{t} \quad \text{and} \quad E(\mathcal{H}) := \left\{\pi(e) \colon e \in \binom{[n]}{k}\right\}.$$

Clearly π is a bijection between $E(K_n^k) = \binom{[n]}{k}$ and $E(\mathcal{H})$, which also naturally induces a bijection between the subsets of $E(K_n^k)$ and the subsets of $E(\mathcal{H})$. Crucially, observe that matchings in $\mathcal{H}$ correspond to t-avoiding subgraphs of K_n^k, and thus to partial (n, k, t)-Steiner systems. Hence finding an almost perfect matching in $\mathcal{H}$ produces an approximate (n, k, t)-Steiner system. This is the

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

well-known fact discussed in Section 1.1. Our aim is now to show that a 'pseudorandom' almost perfect matching, obtained by applying our Theorem 1.3 to $\mathcal{H}$, produces an approximate Steiner system $\mathcal{S}$ with the desired subgraph statistics.

Recall that $\mathcal{H}$ is $\binom{n-t}{k-t}$-regular. Moreover, $\Delta^c(\mathcal{H}) \leqslant n^{k-t-1} \leqslant \Delta(\mathcal{H})^{1-\delta}$.

Now fix $F \in \mathcal{F}$ and let $\ell := e(F)$. Define the ℓ-tuple weight function ω_F on $E(\mathcal{H})$ as follows: for an ℓ-set $\mathcal{E} = \{\pi(e_1), \ldots, \pi(e_\ell)\}$ of edges of $\mathcal{H}$, let $\omega_F(\mathcal{E})$ be the number of injections $f\colon V(F) \to [n]$ for which $\{f(e)\colon e \in E(F)\} = \{e_1, \ldots, e_\ell\}$. Hence, for $G \subseteq K_n^k$ with $V(G) = [n]$, we have $\mathrm{inj}(F, G) = \omega_F(\pi(E(G)))$. In particular,

$$\omega_F(E(\mathcal{H})) = \mathrm{inj}\big(F, K_n^k\big) = (1 \pm n^{-0.9})n^{v(F)}.$$

Note also that ω_F is clean since F is t-avoiding.

Fix $\ell' \in [\ell]$ and a set of ℓ' edges $\pi(e_1), \ldots, \pi(e_{\ell'})$ in $\mathcal{H}$. Let $v' := |e_1 \cup \cdots \cup e_{\ell'}|$. The number of injections $f\colon V(F) \to [n]$ for which $\{e_1, \ldots, e_{\ell'}\} \subseteq \{f(e)\colon e \in E(F)\}$ is at most $v(F)!n^{v(F)-v'}$. Since $\mathrm{mad}(F) < k/(k-t)$, we have $v' \geqslant \ell'(k-t) + 1$, and thus (assuming $\delta \leqslant 1/2k$, say)

$$\|\omega_F\|_{\ell'} \Delta^{\ell'+\delta} \leqslant v!n^{v(F)-v'} \cdot n^{(k-t)(\ell'+\delta)} \leqslant v!n^{v(F)-1/2} \leqslant \omega_F(E(\mathcal{H})).$$

Thus we can apply Theorem 1.3 with $\Delta := 1/p = n^{k-t}/(k-t)! \geqslant \Delta(\mathcal{H})$ and 2ε in place of ε to obtain a matching $\mathcal{M}$ such that $\omega_F(\mathcal{M}) = (1 \pm \Delta^{-2\varepsilon})\omega_F(E(\mathcal{H}))/\Delta^{e(F)}$ for all $F \in \mathcal{F}$. Let $\mathcal{S} := \pi^{-1}(\mathcal{M})$. Note that $\mathcal{S}$ is a partial (n, k, t)-Steiner system, which we now view as a k-graph on $[n]$. Moreover, for any $F \in \mathcal{F}$, we have

$$\mathrm{inj}(F, \mathcal{S}) = \omega_F(\mathcal{M}) = (1 \pm \Delta^{-2\varepsilon})(1 \pm n^{-0.9})n^{v(F)}p^{e(F)} = (1 \pm n^{-\varepsilon})p^{e(F)}n^{v(F)},$$

as desired.

Finally, note that the k-graph F_0 consisting of only one edge is trivially t-avoiding and $\mathrm{mad}(F_0) = 1$. Thus, by the above, $\mathrm{inj}(F_0, \mathcal{S}) \geqslant (1 - n^{-\varepsilon})pn^k$. We conclude that

$$|\mathcal{S}| = \mathrm{inj}(F_0, \mathcal{S})/k! \geqslant (1 - n^{-\varepsilon})pn^k/k! \geqslant (1 - n^{-\varepsilon})\binom{n}{t}\Big/\binom{k}{t},$$

completing the proof. □

One could also ensure that the residual t-graph of uncovered t-sets is quasirandom.[1] The results from [12, 20] can then be used to complete $\mathcal{S}$ to a Steiner system. The lower bound on the number of F-copies would then still hold. However, such a completion step, even if only applied to $o(n^t)$ t-sets, could drastically increase the number of F-copies. For simplicity, we thus omitted such a completion entirely. Needless to say, variations of this theorem can be obtained in the same way, for instance asking for the number of 'rooted' copies.

4.2 Rainbow problems

In [6] we consider subgraph embeddings in edge-coloured graphs with the additional requirement that the embedded subgraph is 'rainbow', meaning that any two edges in the subgraph have distinct colours. Such rainbow embeddings have applications to various other problems. For instance, Montgomery, Pokrovskiy and Sudakov [32, 33] recently used rainbow embeddings to solve Ringel's conjecture from 1963 (which states that any tree with n edges decomposes K_{2n+1}). We consider the classic setting of the blow-up lemma due to Komlós, Sárközy and Szemerédi [27]. Given a multipartite graph G, where the bipartite graphs between two parts are 'quasirandom', and

[1] It is possible to define weight functions which achieve this directly, but perhaps the most convenient way is to first take out a sparse random set of t-sets, then apply the above proof to the remaining t-sets, and finally combine the leftover with the random reservoir.

a bounded degree graph H with a fitting vertex partition, H can be embedded as a spanning subgraph of G. We show that this is still true when G is edge-coloured and we want to find a rainbow copy of H, assuming certain boundedness conditions on the edge-colouring which can be seen to be almost optimal.

To achieve this, we employ Theorem 1.2 as a crucial tool. In the following, we briefly explain how we apply Theorem 1.2 and exploit the weight functions in our proof. To this end, we consider the following toy example. Suppose G is the complete bipartite graph with bipartition (U, V) and $|U| = |V| = n$. Suppose further that $c\colon E(G) \to C$ is a proper edge-colouring of G. Our aim is to find an almost perfect rainbow matching. When the colouring is optimal, that is, it uses only n colours, then finding such a matching of size $n-1$ is equivalent to the famous Ryser–Brualdi–Stein conjecture on almost transversals in Latin squares (see [31] and references therein).

In order to apply our theorem, we formulate the problem as a hypergraph matching problem. Let $\mathcal{H}$ be the hypergraph with vertex set $U \cup V \cup C$ and edge set $\{\{u, v, c(uv)\}\colon uv \in E(G)\}$. The key property of $\mathcal{H}$ is the following bijection between the set of all rainbow matchings in G and the set of all matchings in $\mathcal{H}$ – we simply assign a rainbow matching M in G to the matching $\mathcal{M} := \{\{u, v, c(uv)\}\colon uv \in M\}$ in $\mathcal{H}$. Clearly $\Delta(\mathcal{H}) = n$ and $\Delta^c(\mathcal{H}) = 1$. The existence of an almost perfect rainbow matching in G now follows from known results. For instance, Theorem 2.1 yields a decomposition of $E(\mathcal{H})$ into $(1 + o(1))n$ hypergraph matchings, and as $e(\mathcal{H}) = n^2$, there must be a hypergraph matching $\mathcal{M}$ of size $(1 - o(1))n$ in $\mathcal{H}$ and in turn a rainbow matching M in G of this size.

In the proof of our rainbow blow-up lemma, we also seek almost perfect matchings in bipartite graphs. However, we need much more control over these matchings, which we achieve using our new Theorem 1.2. Our proof proceeds in several rounds where, in each round, we embed essentially all vertices that need to be embedded into a particular cluster of our multipartite graph. Each such embedding step is modelled as finding a rainbow matching M in an auxiliary bipartite 'candidacy graph'. Although these candidacy graphs are more complicated and have more complex colour constraints than our toy example above, they can still be handled using hypergraph matchings in a similar way. However, in order to perform the embedding rounds repeatedly, we need to ensure that certain quasirandomness properties are preserved throughout the procedure, which depend on the previous embeddings. In our toy example this would mean, for instance, that for some specified sets $U' \subseteq U$, $V' \subseteq V$, we need $|E(G[U', V']) \cap M| \approx |U'||V'|/n$, and more generally that for sets $E' \subseteq E(G)$, we need $|E' \cap M| \approx |E'|/n$. This can be ensured by utilizing weight functions as in Theorem 1.2 by defining $\omega_{E'}(e \cup \{c(e)\}) = \mathbb{1}_{e \in E'}$ for all $e \in E(G)$. It is very important here that Theorem 1.2 applies to hypergraphs which are not necessarily almost regular, since the colouring can be arbitrary. For the same reason, it is useful that $v(\mathcal{H})$ plays no role in the parametrization of the theorem.

4.3 Decompositions

Kim, Kühn, Osthus and Tyomkyn [24] proved that in the setting of the original blow-up lemma as described above, the quasirandom multipartite graph does not only contain any single graph of bounded degree with the same multipartite structure, but can even be almost decomposed into any collection of such bounded degree graphs. This result has already found fruitful applications [4, 16]. The first and third author give an alternative and in particular much shorter proof for this decomposition result in [7].

The overall strategy is to use Theorem 1.3 in a similar way as for the rainbow embeddings described in Section 4.2. To this end, the decomposition problem is transformed into a rainbow embedding problem as follows. Suppose G is a graph and $\mathcal{H}$ is a collection of graphs on at most $|V(G)|$ vertices. Define a new graph $\mathcal{G}$ by taking $|\mathcal{H}|$ disjoint copies $(G_H)_{H \in \mathcal{H}}$ of G and colour

https://doi.org/10.1017/S0963548320000280 Published online by Cambridge University Press

every copy of a particular edge of G with a unique colour. Hence a collection of edge-disjoint copies of the graphs in $\mathcal{H}$ in G is equivalent to a rainbow embedding of the disjoint union of the graphs in $\mathcal{H}$ into $\mathcal{G}$ where each $H \in \mathcal{H}$ is embedded into G_H.

Acknowledgement

We thank Candida Bowtell and Katherine Staden for drawing our attention to Kahn's paper [18], and the referee for careful reading and helpful comments.

References

[1] Aharoni, R., Georgakopoulos, A. and Sprüssel, P. (2009) Perfect matchings in r-partite r-graphs. *European J. Combin.* **30** 39–42.
[2] Ajtai, M., Komlós, J. and Szemerédi, E. (1981) A dense infinite Sidon sequence. *European J. Combin.* **2** 1–11.
[3] Alon, N. and Yuster, R. (2005) On a hypergraph matching problem. *Graphs Combin.* **21** 377–384.
[4] Condon, P., Kim, J., Kühn, D. and Osthus, D. (2019) A bandwidth theorem for approximate decompositions. *Proc. Lond. Math. Soc.* **118** 1393–1449.
[5] Cooper, C., Frieze, A., Molloy, M. and Reed, B. (1996) Perfect matchings in random r-regular, s-uniform hypergraphs. *Combin. Probab. Comput.* **5** 1–14.
[6] Ehard, S., Glock, S. and Joos, F. (2019) A rainbow blow-up lemma for almost optimally bounded edge-colourings. arXiv:1907.09950
[7] Ehard, S. and Joos, F. (2020) A short proof of the blow-up lemma for approximate decompositions. arXiv:2001.03506
[8] Frankl, P. and Rödl, V. (1985) Near perfect coverings in graphs and hypergraphs. *European J. Combin.* **6** 317–326.
[9] Freedman, D. A. (1975) Packing Hamilton cycles in random and pseudo-random hypergraphs. *Random Struct. Algorithms* **41** 1–22.
[10] Frieze, A. and Krivelevich, K. (2012) Packing Hamilton cycles in random and pseudo-random hypergraphs. *Random Struct. Algorithms* **41** 1–22.
[11] Füredi, Z. (1988) Matchings and covers in hypergraphs. *Graphs Combin.* **4** 115–206.
[12] Glock, S., Kühn, D., Lo, A. and Osthus, D. The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *Mem. Amer. Math. Soc.*, to appear.
[13] Hàn, H., Han, J. and Morris, P. (2020) Factors and loose Hamilton cycles in sparse pseudo-random hypergraphs. In *14th Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 702–717, SIAM.
[14] Hàn, H., Person, Y. and Schacht, M. (2009) On perfect matchings in uniform hypergraphs with large minimum vertex degree. *SIAM J. Discrete Math.* **23** 732–748.
[15] Johansson, A., Kahn, J. and Vu, V. (2008) Factors in random graphs. *Random Struct. Algorithms* **33** 1–28.
[16] Joos, F., Kim, J., Kühn, D. and Osthus, D. (2019) Optimal packings of bounded degree trees. *J. Eur. Math. Soc.* **21** 3573–3647.
[17] Kahn, J. (1996) Asymptotically good list-colorings. *J. Combin. Theory Ser. A* **73** 1–59.
[18] Kahn, J. (1996) A linear programming perspective on the Frankl–Rödl–Pippenger theorem. *Random Struct. Algorithms* **8** 149–157.
[19] Karp, R. (1972) Reducibility among combinatorial problems. In *Complexity of Computer Computations* (R. E. Miller *et al.*, eds), pp. 85–103, Springer.
[20] Keevash, P. (2014) The existence of designs. arXiv:1401.3665
[21] Keevash, P. (2018) The existence of designs II. arXiv:1802.05900
[22] Keevash, P. (2018) Hypergraph matchings and designs. *Proc. Int. Cong. Math.* **3** 3099–3122.
[23] Keevash, P. and Mycroft, R. (2015) *A Geometric Theory for Hypergraph Matching*, Vol. 233, no. 1098 of Memoirs of the American Mathematical Society, AMS.
[24] Kim, J., Kühn, D., Osthus, D. and Tyomkyn, M. (2019) A blow-up lemma for approximate decompositions. *Trans. Amer. Math. Soc.* **371** 4655–4742.
[25] Klenke, A. (2014) *Probability Theory*, second edition, Springer.
[26] Komlós, J., Pintz, J. and Szemerédi, E. (1982) A lower bound for Heilbronn's problem. *J. London Math. Soc.* (2) **25** 13–24.
[27] Komlós, J., Sárközy, G. N. and Szemerédi, E. (1997) Blow-up lemma. *Combinatorica* **17** 109–123.
[28] Lenz, J. and Mubayi, D. (2016) Perfect packings in quasirandom hypergraphs I. *J. Combin. Theory Ser. B* **119** 155–177.
[29] McDiarmid, C. (1989) On the method of bounded differences. In *Surveys in Combinatorics, 1989 (Norwich, 1989)*, Vol. 141 of London Mathematical Society Lecture Note Series, pp. 148–188, Cambridge University Press.
[30] Molloy, M. and Reed, B. (2000) Near-optimal list colorings. *Random Struct. Algorithms* **17** 376–402.

[31] Montgomery, R., Pokrovskiy, A. and Sudakov, B. (2019) Decompositions into spanning rainbow structures. *Proc. Lond. Math. Soc.* **119** 899–959.

[32] Montgomery, R., Pokrovskiy, A. and Sudakov, B. (2020) A proof of Ringel's conjecture. arXiv:2001.02665

[33] Montgomery, R., Pokrovskiy, A. and Sudakov, B. Embedding rainbow trees with applications to graph labelling and decomposition. *J. Eur. Math. Soc.*, to appear.

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

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

[36] Rödl, V. and Ruciński, A. (2010) Dirac-type questions for hypergraphs: a survey (or more problems for Endre to solve). In *An Irregular Mind* (I. Bárány and J. Solymosi, eds), Vol. 21 of Bolyai Society Mathematical Studies, pp. 561–590, János Bolyai Mathematical Society.

[37] Rödl, V., Ruciński, A. and Szemerédi, E. (2009) Perfect matchings in large uniform hypergraphs with large minimum collective degree. *J. Combin. Theory Ser. A* **116** 613–636.

Cite this article: Ehard S, Glock S and Joos F (2020). Pseudorandom hypergraph matchings. *Combinatorics, Probability and Computing* **29**, 868–885. https://doi.org/10.1017/S0963548320000280