Optimal thresholds for Latin squares, Steiner Triple Systems, and edge colorings

Vishesh Jain*   Huy Tuan Pham†

Abstract

Given a graph G, a random (k, n)-list assignment L for edges of G is an assignment of an independent, uniformly random set $L(e) \in \binom{[n]}{k}$ of colors to each edge e and a proper L-list coloring of G is a proper edge-coloring where the color of an edge e belongs to $L(e)$. We show that for a random $(O(\log n), n)$-list assignment L for edges of the complete bipartite graph $K_{n,n}$, there is a an L-list coloring of $K_{n,n}$ with high probability. We also prove analogous results for the thresholds of Steiner triple systems and Latin squares in random (binomial) hypergraphs. All of our results are optimal up to absolute constants, and resolve several related conjectures of Johansson, Luria-Simkin, Casselgren-Häggkvist, Simkin, and Kang-Kelly-Kühn-Methuku-Osthus.

A key contribution of our work is to show that in natural settings, the Lovász Local Lemma – a central tool in probabilistic combinatorics to establish the existence of objects with desired properties – can also be used to design optimally "spread" distributions on such objects. This is made possible by carefully exploiting the local uniformity property of the so-called Lovász Local Lemma distribution, an important observation that has recently been utilized in finding efficient algorithms for sampling approximately uniformly random solutions to constraint satisfaction problems. In conjunction with the recently proved Kahn-Kalai conjecture, this opens the door to obtaining optimal threshold results for the appearance of many interesting objects.

1   Introduction

Given a finite set X and $p \in (0, 1)$, let $X_p \subseteq X$ be a random subset where each element of X is sampled independently with probability p. For a non-trivial monotone property $\mathcal{P} \subseteq 2^X$ over subsets of X, the threshold $p_c(\mathcal{P})$ of $\mathcal{P}$ is the value p^* at which $\mathbb{P}(X_{p^*} \in \mathcal{P}) = 1/2$. When $X = \binom{[n]}{k}$, X_p is precisely the Erdős-Rényi random k-uniform hypergraph $G^{(k)}(n, p)$, for which estimating thresholds of interesting combinatorial properties has been a major direction in combinatorics, going back to a seminal result of Erdős and Rényi that the threshold for the appearance of perfect matchings in $G^{(2)}(n, p)$ is $\log n/n$ [6]. The determination of the threshold for the appearance of perfect matchings in k-uniform hypergraphs for $k \geq 3$ is a notorious problem of Shamir, for which the threshold $\Theta((\log n)/n^{k-1})$ was established in the celebrated work of Johansson, Kahn and Vu [14].

A closely related object to perfect matchings in 3-uniform hypergraphs, which constitutes another way of generalizing the definition of perfect matchings in graphs to 3-uniform hypergraphs, is a Steiner triple system. Whereas perfect matchings require that every vertex appears in exactly one hyperedge, a Steiner triple system requires that every *pair* of vertices appears in exactly one hyperedge. The tripartite analogue of a Steiner triple system is a Latin square, namely, a Latin square may equivalently be defined as a tripartite 3-uniform hypergraph with vertex parts of size n, where each pair of vertices in different parts is contained in exactly one hyperedge. The question of determining the threshold for which $G^3((n, n, n), p)$ – the random tripartite 3-uniform hypergraph on parts of size n each – contains a Latin square was raised by Johansson in 2006 [13] and has been popularized by Simkin in the past few years as his favorite open problem. Luria and Simkin [19] conjectured that this threshold is $\Theta(\log n/n)$. It is easily seen using a coupon-collector argument (see, e.g., [21]) that the threshold is $\Omega(\log n/n)$, so that the challenge is in proving the upper bound.

We mention three related conjectures:

- Simkin [22] conjectured that the threshold for containment of a Steiner triple system in the random 3-uniform hypergraph $G^3(n, p)$, with $n \equiv 1, 3 \mod 6$, is $\Theta((\log n)/n)$.

- Given a graph G, a random (k, n)-list assignment L for edges of G is an assignment of an independent, uniformly random set $L(e) \in \binom{[n]}{k}$ of colors to each edge e and a proper L-list coloring of G is a proper

*Department of Mathematics, Statistics, and Computer Science, University of Illinois Chicago.

†Department of Mathematics, Stanford University.

edge-coloring where the color of an edge e belongs to $L(e)$. Casselgren and Häggkvist [5] conjectured that for a random $(O(\log n), n)$-list assignment L for edges of $K_{n,n}$, there is an L-list coloring of $K_{n,n}$ with probability at least $1/2$. We remark that in the setting of *vertex* coloring graphs with a number of colors slightly larger than the maximum degree, a similar problem has been studied by the algorithms community under the name of 'palette sparsification', with applications to streaming/sublinear/distributed algorithms for graph vertex coloring (see, e.g. [1, 4, 10]).

- A non-partite version of the above, namely that a random $O((\log n), 2n-1)$-list assignment L for edges of K_{2n} admits an L-list coloring of K_{2n} with probability at least $1/2$, was conjectured by Kang, Kelly, Kühn, Methuku and Osthus [15].

We refer the reader to [21, 15] for further discussion of the history, as well as for more precise versions of the aforementioned conjectures.

In this paper, we resolve all of these conjectures.

THEOREM 1. *There exists an absolute constant $C > 0$ such that for all $n \geq C$, each of the following statements hold with probability at least $1/2$.*

1. *The random 3-uniform hypergraph $G^{(3)}(n, C(\log n)/n)$ contains a Steiner triple system, provided $n \equiv 1, 3$ mod 6.*

2. *The random tripartite 3-uniform $G^{(3)}((n, n, n), C(\log n)/n)$ contains a Latin square.*

3. *A random $(C \log n, n)$-list assignment L for edges of $K_{n,n}$ admits a proper L-list coloring.*

4. *A random $(C \log n, 2n-1)$-list assignment L for edges of K_{2n} admits a proper L-list coloring.*

1.1 Background and related work Recently, Park and Pham [20] proved the Kahn–Kalai conjecture; a weaker fractional version conjectured by Talagrand [23] was obtained earlier in work of Frankston, Kahn, Narayanan, and Park [7] building on the sunflower breakthrough of Alweiss, Lovett, Wu and Zhang [3]. The (fractional) Kahn-Kalai conjecture allows one (among other things) to get a much simpler proof of the resolution of Shamir's problem by Johansson-Kahn-Vu [14]. To state the result of [20], we say that $\mathcal{H} \subseteq 2^X$ is p-small if there exists $\mathcal{G} \subseteq 2^X$ such that any set in $\mathcal{H}$ contains a set from $\mathcal{G}$ and $\sum_{G \in \mathcal{G}} p^{|G|} < \frac{1}{2}$. Moreover, a distribution μ supported on $\mathcal{H}$ is said to be p-spread if for all $S \subseteq X$, $\mu(\{W \in \mathcal{H} : S \subseteq W\}) \leq 2p^{|S|}$. Using linear programming duality, Talagrand [23] observed that if $\mathcal{H}$ supports a p-spread distribution μ, then $\mathcal{H}$ is not p-small.

THEOREM 2. ([20]) *There exists $C > 0$ such that the following holds. Let $\mathcal{H} \subseteq 2^X$ and $p \in (0, 1)$ be such that $\mathcal{H}$ is not p-small. Then, $X_{Cp \log |X|}$ contains a set in $\mathcal{H}$ with high probability (i.e. probability going to 0 as $|X| \to \infty$). In particular, the same conclusion holds if there exists a distribution μ supported on $\mathcal{H}$ which is p-spread.*

Given Theorem 2, a natural approach to all of these conjectures is to show that the corresponding property is not $O(1/n)$-small, potentially by establishing the stronger statement that the relevant collection of subsets supports an $O(1/n)$-spread distribution. For Shamir's problem, simple counting shows that the uniform distribution on perfect matchings of a k-uniform hypergraph on n vertices is $O_k(1/n)$-spread. On the other hand, for complicated structures such as Steiner triple systems and Latin squares, known enumeration results are unfortunately not precise enough to imply anything non-trivial about the spread of the uniform distribution (this is related to the fact that, even following the breakthrough work of Keevash [16] and the many developments since then, we have asymptotically correct estimates only for the *logarithm* of the number of Steiner triple systems and other designs, see the discussion in [15, 21, 7]). In a recent breakthrough work, Sah, Sawhney and Simkin [21] demonstrated the existence of an $O(n^{o(1)}/n)$-spread distribution on Steiner triple systems in $K_n^{(3)}$ ($n \equiv 1, 3$ mod 6) and Latin squares in $K_{n,n,n}^{(3)}$, thereby establishing the corresponding conjectures about the threshold within a subpolynomial factor $n^{o(1)}$. Their proof constructs such a spread distribution using a clever "spread boosting" argument utilizing the iterative absorption framework of Kühn, Osthus, and collaborators [8]. In a beautiful work using a different and simpler iterative absorption scheme, Kang, Kelly, Kühn, Methuku and Osthus [15] improved the spread parameter to $O(\log n/n)$, which is a single logarithmic factor off from the conjectured optimal bound.

1.2 Our contribution One of the main tools in probabilistic combinatorics to establish the *existence* of objects with desired properties is the Lovász Local Lemma (LLL). However, since the probability of success in the Lovász Local Lemma is typically exponentially small in the number of events, the change of measure factor between the so-called LLL distribution (i.e. the distribution obtained by conditioning the realisation from the product distribution to satisfy the desired properties) and the product distribution is exponentially large. This is related to the challenge in making the Lovász Local Lemma algorithmic – an area that has seen tremendous success (see, e.g., [2, Chapter 5]). A key idea in our work is to show that in natural settings, the LLL distribution is essentially as spread as the underlying product distribution. This opens up the possibility of constructing optimal or near-optimal spread distributions on various objects, and correspondingly, obtaining optimal or near-optimal thresholds for the appearance of such objects. In order to do so, we leverage the so-called local uniformity property of the LLL distribution (Theorem 6), first used in works on the algorithmic Lovász Local Lemma ([9]) and more recently, a crucial ingredient of algorithms for approximately sampling from the LLL distribution. In particular, we remark that the distributions constructed in our proof of Theorem 3 below can be approximately sampled from in polynomial time, due to recent progress in algorithms for sampling from the LLL distribution, in particular [11, 12].

Concretely, in order to prove Theorem 1, we show how to construct a $O(1/n)$-spread distribution on ordered one factorizations of $K_{n,n}$ (i.e. a decomposition of the edges of $K_{n,n}$ into a tuple $(M_1, \ldots, M_n)$ of disjoint perfect matchings). Here, and later, for a d-regular graph G and $d'|d$, we say that a probability distribution on decompositions $(H_1, \ldots, H_{d/d'})$ of G into d'-regular subgraphs is q-spread if for all $S_1, \ldots, S_{d/d'} \subseteq E(G)$,

$$\mathbb{P}[S_1 \subseteq E(H_1), \ldots, S_{d/d'} \subseteq E(H_d)] \leq q^{|S_1| + \cdots + |S_{d/d'}|}.$$

In fact, we prove the following more general result for dense quasirandom bipartite graphs:

THEOREM 3. *For any $\varepsilon \in (0, 1/2)$, there exist sufficiently large constants $N_0 \in \mathbb{N}$ and $S \in 2^{\mathbb{N}} - 1$ for which the following holds. Let $\mathbb{G} = (A, B, E)$ be a bipartite graph with $|A| = |B| = n$ which is D_0-regular for $D_0 \geq \varepsilon n \geq N_0$ and 0-nice (Theorem 8). For $r \in \mathbb{N}$, let $D_r := D_0/S^r$. Then, for all $r \in \mathbb{Z}_{\geq 0}$ such that $D_r \geq N_0$, there exists a probability distribution $\boldsymbol{\mathcal{P}}_r$ on decompositions $\mathcal{P}_r$ of $E(\mathbb{G})$ into regular subgraphs with degree $(1 + o_n(1))D_r$, with spread $2\varepsilon^{-1} \cdot D_r/n$.*

The above theorem provides a quantitative strengthening of the key [15, Theorem 1.10], which applies to a subset of the bipartite graphs considered here and crucially, loses a multiplicative factor of $\log n$ in the spreadness. By following exactly the same reductions in [15] (see [15, Theorems 1.5,1.6,1.7]), we can deduce Theorem 1 from Theorem 3.

1.3 Techniques In [15], the authors devised an iterative edge-absorption scheme to construct an $O(\log n/n)$-spread distribution on decompositions of a regular, nearly-complete, bipartite graph into regular subgraphs of degree $O(\log n)$. Briefly, using standard concentration techniques, they first (essentially uniformly) decompose the graph into nearly-regular subgraphs of average degree $\Theta(\log n)$ satisfying some additional expansion properties, and then iteratively correct these subgraphs to be regular by combining a novel and elegant iterative edge-absorption procedure with a classical result about the existence of (large) regular subgraphs of a given graph (Theorem 4). The initial decomposition is clearly spread, and the key in [15] is to show that the edge-absorption procedure approximately preserves the spreadness. At a very high level, our approach is to recursively use a simpler version of the framework of [15], which does not use iterative absorption, until the regular subgraphs in the decomposition have degrees $O(1)$. This requires a number of new ideas, which we now discuss.

The LLL distribution The reason why the argument of [15] stops at subgraphs of degree $\Theta(\log n)$ is that for sub-logarithmic degrees, one cannot use the union bound to guarantee near-regularity of the initial decomposition. However, if one could replace the argument of [15] by an iterative procedure, where the degree of the new subgraphs is only a constant factor (say) smaller than the degree of the original graph, one might hope that replacing the union bound by the Lovász Local Lemma (LLL) would immediately do the trick. Unfortunately, this is not the case – in our setting, for graphs of sublogarithmic degree, the LLL only guarantees that the initial decomposition is nearly-regular with probability exponentially small in n (up to logarithmic factors), so that conditioning on near-regularity (the resulting distribution is called the LLL distribution in this case) could change the measure by an exponentially large amount, thereby completely destroying the spreadness. To overcome this, we use the

insight that we require the LLL distribution only for some specific events. For these events, we carefully use a refined comparison between the initial distribution and the LLL distribution (Theorem 6) to show that spreadness does not degrade much. We believe that this technique of using event-specific comparison bounds for the LLL distribution will be generally useful in the study of thresholds (via constructing spread measures).

Using spread to perform the union bound In order to find a sufficiently large regular subgraph inside the initial nearly-regular graph, we need to show that the nearly-regular graph has the property of not having 'too many' edges between small sets of vertices (Theorem 4). For logarithmic degree, this can be accomplished by a simple union bound over all small sets; however, this fails for sublogarithmic degree, and unfortunately the dependencies are too numerous to employ the LLL. We overcome this issue using the spreadness of the iterates. Roughly, in Theorem 15, we show using a direct union bound argument that if a regular graph of degree D is sampled from an $O(D/n)$-spread distribution, then with high probability, a random decomposition into nearly-regular subgraphs has the requisite expansion property.

Lossless edge-absorption In [15], the iterative edge-absorption procedure blows up the spread of the initial decomposition into nearly-regular subgraphs by a factor of approximately 2. While this is certainly sufficient for their result, it is not amenable to recursion, since to go to subgraphs of degree $O(1)$, we must recurse $\Omega(\log^* n)$ times. To circumvent this issue, we devise a simple, yet refined, edge-absorption procedure (Theorem 11), which only uses edges from a randomly chosen, polynomially small part of the graph to make corrections. This ensures that the spread of the corrected distribution is approximately the same as the spread of the initial decomposition into nearly-regular subgraphs, thereby permitting recursion.

1.4 Concurrent and independent work Keevash has independently and concurrently proved Theorem 1 and essentialy the same result as Theorem 3 (cf. [17, Section 2.4] and Theorem 8) in [17]. His proof is substantially different from ours, relying instead on a very careful analysis of a randomized greedy process.

2 Preliminaries

We will use the following lemma, which follows easily from the max-flow min-cut theorem (see [18]), and was also employed in [15].

LEMMA 4. *Consider a bipartite graph $G = (A, B, E)$. Given $f : A \to \mathbb{Z}_{\geq 0}$ and $g : B \to \mathbb{Z}_{\geq 0}$ with $\sum_{a \in A} f(a) = \sum_{b \in B} g(b)$, there exists a spanning subgraph H of G with $d_H(a) = f(a)$ for all $a \in A$ and $d_H(b) = g(b)$ for all $b \in B$ if and only if the following holds: for every subset $A' \subseteq A$ and $B' \subseteq B$, $|E_G(A', B')| \geq \sum_{a \in A'} f(a) - \sum_{b \notin B'} g(b)$.*

A key ingredient in our work is the following comparison between the so-called "Lovász Local Lemma distribution" and the product distribution on a collection of random variables. (see, e.g., Theorem 2.1 of [9]). The proof of this result follows directly from the inductive proof of the Lovász Local Lemma (see, e.g. [2]).

DEFINITION 5. *Given independent random variables $\{X_i\}_{i \in I}$ and events $\mathcal{E}_j$ ($j \in J$), where each $\mathcal{E}_j$ depends on a subset $S_j \subseteq I$ of variables, we say that a graph Γ on vertex set J is a dependency graph if it holds that $S_j \cap S_{j'} = \emptyset$ whenever j and j' are not adjacent in Γ.*

PROPOSITION 6. *With notation as in Theorem 5, denote by $\mathbf{P}$ the usual product measure on the random variables $\{X_i\}_{i \in I}$ and by $\mathbb{P}$, the conditional measure $\mathbf{P}(\cdot \mid \cap_j \mathcal{E}_j^c)$. Assume that $\mathbf{P}(\mathcal{E}_j) \leq p$ for all $j \in J$ and that the maximum degree of Γ, denoted by Δ, satisfies $4p\Delta \leq 1$. Given an event $\mathcal{E}$ depending on a subset of variables $S \subseteq I$, letting N be the number of events $\mathcal{E}_j$ ($j \in J$) with $S_j \cap S \neq \emptyset$, we have*

$$\mathbb{P}[\mathcal{E}] \leq \mathbf{P}[\mathcal{E}] \exp(6pN).$$

Finally, we will use the following standard version of the Chernoff bound.

LEMMA 7. *Let $X_1, \ldots, X_m$ be independent $\mathrm{Bernoulli}(p)$ random variables. Then, for $\delta > 0$,*

$$\mathbb{P}[X_1 + \cdots + X_m > (1+\delta)pm] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}} \right)^{pm} \leq \exp\left(-\frac{\delta^2 pm}{2+\delta} \right)$$

$$\mathbb{P}[X_1 + \cdots + X_m < (1-\delta)pm] \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}} \right)^{pm} \leq \exp\left(-\frac{\delta^2 pm}{2} \right).$$

3   Proof of Theorem 3

3.1   Recursion Throughout this section, we use the notation appearing in the statement of Theorem 3. In the following definition, S denotes a positive integer which, in our application, will be chosen to be sufficiently large depending on ε. We define

$$q_r := D_r^{-1/8}, \quad \delta_r := 2^{10} \cdot S \cdot \sum_{r' \leq r} q_r.$$

DEFINITION 8. (r-NICENESS) *For $r \in \mathbb{Z}_{\geq 0}$, we say that a bipartite graph $G = (A, B, E)$ with $|A| = |B| = n$ is r-nice if the following properties hold:*

(N1) G is D_G-regular, with

$$D_G = \exp(\pm\delta_{r-1}) \cdot D_r \text{ i.e. } D_G \in [\exp(-\delta_{r-1}) \cdot D_r, \exp(\delta_{r-1}) \cdot D_r]$$

(N2) For all $A' \subseteq A, B' \subseteq B$ with $|B'| \geq |A'| > n/S$ and $n - |B'| = 1.01|A'|$,

$$|E_G(A', B \setminus B')| \leq \exp(1 + \delta_{r-1}) \cdot D_G|A'| \cdot \max\left(\frac{1}{3}, \frac{n - |B'|}{n}\right),$$

and similarly with the role of A' and B' interchanged.

For $r \in \mathbb{Z}_{\geq 0}$ and a regular bipartite graph $\mathbb{G} = (A, B, E)$ with $|A| = |B| = n$ and degree D_0, a decomposition $\mathcal{P}$ of $E(\mathbb{G})$ is said to be r-nice if $|\mathcal{P}| = D_0/D_r(= S^r)$ and every $G \in \mathcal{P}$ is r-nice.

Theorem 3 follows immediately from the following proposition, whose proof is the content of the remainder of this paper.

THEOREM 9. *For any $\varepsilon \in (0, 1/2)$, there exist sufficiently large constants $N_0 \in \mathbb{N}, S \in 2^{\mathbb{N}} - 1$ for which the following holds. Let $\mathbb{G} = (A, B, E)$ be a bipartite graph with $|A| = |B| = n$ which is D_0-regular for $D_0 \geq \varepsilon n \geq N_0$ and 0-nice (Theorem 8). For $r \in \mathbb{N}$, let $D_r := D_0/S^r$. Then, for all $r \in \mathbb{Z}_{\geq 0}$ such that $D_r \geq N_0$, there exists a probability distribution $\boldsymbol{\mathcal{P}}_r$ on decompositions $\mathcal{P}_r$ of $E(\mathbb{G})$, supported on r-nice decompositions, with spread $2\varepsilon^{-1} \cdot D_r/n$.*

For $r = 0$, note that the distribution supported on the trivial decomposition consisting of only one part satisfies the conclusion of Theorem 9. The distributions for $r \geq 1$ will be constructed recursively in Theorem 15, starting from this choice of $\boldsymbol{\mathcal{P}}_0$.

3.2   Admissibility The recursive construction used to prove Theorem 9 is based on a procedure to partition a regular bipartite graph G into regular subgraphs. Throughout this subsection, we consider $r \in \mathbb{Z}_{\geq 0}$ and a D_G-regular bipartite graph $G = (A, B, E)$ with $|A| = |B| = n$ and $D_G \geq N_0$, which is r-nice (Theorem 8) with respect to the sequence D_r in the statement of Theorem 9. Recall that $S \in \mathbb{N}$ will be chosen to be sufficiently large.

Consider the following collection of random variables:

- For each edge $e \in G$, $\pi(e)$ is supported in $[S](= [s_0 - 1])$.

- For each edge $e \in G$, $\xi(e)$ is supported in $\{0, 1\}$.

To this collection of random variables, we associate the following collection of subgraphs of G:

- For $i \in [S]$, H_i denotes the subgraph of G consisting of all edges e for which $\pi(e) = i$. Succinctly, $H_i = \pi^{-1}(i)$.

- For $i \in [S]$, $H_i^+ = H_i \cap \{e : \xi(e) = 1\} = H_i \cap \xi^{-1}(1)$.

- Let $H^+ = H_1^+ \cup \cdots \cup H_S^+$.

The next definition collects the properties we will require of these random variables.

DEFINITION 10. (r-ADMISSIBILITY) *With notation as above, a realisation of the collection of random variables* $\{\pi(e), \xi(e)\}_{e \in G}$ *is said to be r-admissible if the following properties hold:*

(R1) for all $v \in V(G)$, for all $i \in [S]$,

$$|d_{H_i}(v) - D_G/S| \leq 9\sqrt{\log D_G \cdot D_G/S};$$

(R2) for all $v \in V(G)$, for all $i \in [S]$,

$$|d_{H_i^+}(v) - q_r D_G/S| \leq 9\sqrt{\log D_G \cdot D_G/S};$$

(E1) for all $i \in [S]$ and all $A' \subseteq A, B' \subseteq B$ with either $|B'| \geq |A'| \geq 4n/5$ or $4n/5 \geq |A'| \geq n/S$ and $n - |B'| < 1.01|A'|$

$$|E_{H_i^+}(A', B')| \geq (9/10) \cdot |E_G(A', B')| q_r/S,$$

(E2) for all $i \in [S]$ and all $A' \subseteq A, B' \subseteq B$ with $|A'| \leq n/S$ and $|B'| = 1.01|A'|$

$$|E_{H_i^+}(A', B')| \leq (D_G q_r/2S) \cdot |A'|,$$

(E3) for all $i \in [S]$ and all $A' \subseteq A, B' \subseteq B$ with $|B'| \geq |A'| \geq n/S$ and $n - |B'| = 1.01|A'|$,

$$|E_{H_i}(A', B \setminus B')| \leq \exp(1 + q_r + \delta_{r-1}) D_G |A'| \beta'/S,$$

where $\beta' = \max(\kappa, 1 - |B'|/n)$,

(E4) for all $i \in [S]$ and all $A' \subseteq A, B' \subseteq B$ with $|B'| \geq |A'| \geq n/S$ and $n - |B'| = 1.01|A'|$,

$$\left| E_{H_i^+}(A', B \setminus B') \right| \leq 4 D_G |A'| \beta' q_r/S,$$

where $\beta' = \max(\kappa, 1 - |B'|/n)$,

and similarly the properties $(E1), (E2), (E3), (E4)$ with the roles of A, A' and B, B' interchanged.

The motivation for the definition of r-admissibility comes from the next proposition which shows that given an r-admissible realisation, one can decompose an r-nice graph G into $(r + 1)$-nice regular subgraphs of degree approximately D_G/S in a manner that will turn out to be sufficiently 'spread'.

PROPOSITION 11. *Given any r-admissible realisation of $\{\pi(e), \xi(e)\}_{e \in G}$, there exists a disjoint collection of regular subgraphs $\{R_i\}_{i \in [S]}$ of G satisfying the following properties:*

(P1) $G = R_1 \cup \cdots \cup R_S$;

(P2) for all $i \in [S - 1]$, $H_i \setminus H_i^+ \subseteq R_i \subseteq H_i$;

(P3) $H_S \subseteq R_S \subseteq H_S \cup H^+$;

(P4) for all $i \in [S]$, the degree of each R_i is in $(1 \pm 2Sq_r)D_G/S$;

(P5) for all $i \in [S]$, R_i is $(r + 1)$-nice i.e. it satisfies properties $(N1)$ and $(N2)$ in Theorem 8 with $D_G := D_{R_i}$ and D_{r+1}.

Proof. Let $K_i := H_i \setminus H_i^+$. Suppose that for all $i \in [S - 1]$, there exists a regular subgraph satisfying

$$(3.1) \qquad\qquad K_i \subseteq R_i \subseteq H_i.$$

Then, we set $R_S := G \setminus (R_1 \cup \cdots \cup R_{S-1})$. Note that since G is regular and $R_1, \ldots, R_{S-1}$ are disjoint regular subgraphs of G, R_S is also regular.

We claim that for any r-admissible realisation, we can always find regular subgraphs satisfying eq. (3.1). Before proving this claim, let us verify that the construction thus obtained satisfies the desired properties. $(P1), (P2), (P3)$ are satisfied by construction. It remains to verify that $(P4)$ and $(P5)$ hold. For $(P4)$, we have for all $i \in [S]$ that

$$\delta_{H_i} - \Delta_{H_i^+} \le d_{R_i} \le \Delta_{H_i} + \Delta_{H^+},$$

where $d_{G'}, \Delta_{G'}, \delta_{G'}$ denote respectively the average, maximum and minimum degree of a subgraph G' of G. Using $(R1), (R2)$, we get that

$$\underbrace{\frac{D_G}{S} - 18\sqrt{\log D_G \cdot \frac{D_G}{S}} - q_r \frac{D_G}{S}}_{\ge \frac{D_G}{S} - 4q_r \frac{D_G}{S}} \le d_{R_i} \le \frac{D_G}{S} + 9S\sqrt{\log D_G \cdot \frac{D_G}{S}} + q_r D_G$$

$$\le \frac{D_G}{S} + 2Sq_r \frac{D_G}{S},$$

where we have used that $q_r = D_r^{-1/8}$ and $D_r \ge N_0$, a sufficiently large absolute constant. In particular, using that $D_G = \exp(\pm\delta_{r-1})D_r$, we get that $D_{R_i} = \exp(\pm\delta_r)D_{r+1}$, thereby verifying $(N1)$ in $(P5)$. For $(N2)$ in $(P5)$, note that by $(E3), (E4)$, for all $A' \subseteq A, B' \subseteq B$ with $|B'| \ge |A'| > n/S$, $n - |B'| = 1.01|A'|$, and $\beta' = \max(1/3, 1 - |B'|/n)$, we have that

$$|E_{R_i}(A', B \setminus B')| \le |E_{H_i}(A', B \setminus B')| + \sum_{i \in [S]} |E_{H_i^+}(A', B \setminus B')|$$

$$\le (\exp(1 + q_r + \delta_{r-1}) + 4Sq_r)\frac{D_G}{S}|A'|\beta'$$

$$\le \exp(1 + q_r + \delta_{r-1})(1 + 4Sq_r)^2 D_{R_i}|A'|\beta'$$

$$\le \exp(1 + \delta_r)D_{R_i}|A'|\beta',$$

where we have used that $q_r = D_r^{-1/8}$ and $D_r \ge N_0$, a sufficiently large absolute constant.

Finally, we establish the existence of regular subgraphs satisfying eq. (3.1) by applying Theorem 4 on the graph $H_i^+ = (A, B, E(H_i^+))$ with $f(a) = d - d_{K_i}(a)$ and $g(b) = d - d_{K_i}(b)$ for $d := d_{K_i} + 10^5\sqrt{(D_G/S)\log D_G}$, where for notational convenience, we assume that second summand is an integer. Observe that this choice of f and g is valid. Indeed, since $|A| = |B|$, we have that $\sum_{a \in A} f(a) = \sum_{b \in B} g(b)$ and moreover, by $(R1), (R2)$,

$$d - d_{K_i}(v) \ge d_{K_i} + 10^5\sqrt{(D_G/S)\log D_G} - \Delta_{K_i} > (10^5 - 72)\sqrt{(D_G/S)\log D_G},$$

so that $f, g \ge 0$. For later use, note also that by a similar computation,

$$d - d_{K_i}(v) \le d_{K_i} + 10^5\sqrt{(D_G/S)\log D_G} - \delta_{K_i} < (10^5 + 72)\sqrt{(D_G/S)\log D_G}.$$

It remains to show that for all $A' \subseteq A$ and $B' \subseteq B$,

$$(3.2) \qquad \left|E_{H_i^+}(A', B')\right| \ge \Delta(A', B') := \sum_{a \in A'}(d - d_{K_i}(a)) - \sum_{b \notin B'}(d - d_{K_i}(b)).$$

Since

$$\sum_{a \in A'}(d - d_{K_i}(a)) - \sum_{b \notin B'}(d - d_{K_i}(b)) = \sum_{b \in B'}(d - d_{K_i}(b)) - \sum_{a \notin A'}(d - d_{K_i}(a)),$$

it suffices by symmetry to assume that $|B'| \ge |A'|$.

We consider a few different cases. In each case, we will use that $D_r \ge N_0$, a sufficiently large absolute constant.

Case 1: $|A'| \ge 4n/5$. In this case, $|B'| \ge |A'| \ge 4n/5$, so that by $(E1)$,

$$\left|E_{H_i^+}(A', B')\right| \ge (9/10) \cdot |E_G(A', B')| q_r/S \ge (9/10) \cdot (D_G|A'| - D_G(n - |B'|))q_r/S \ge 2D_G q_r n/(5S),$$

and hence

$$\Delta(A', B') \leq (10^5 + 72)|A'|\sqrt{(D_G/S)\log D_G} - (10^5 - 72)(n - |B'|)\sqrt{(D_G/S)\log D_G}$$
$$\leq (10^5 + 72)n\sqrt{(D_G/S)\log D_G} \leq 2D_G q_r n/(5S) \leq \left| E_{H_i^+}(A', B') \right|.$$

Case 2: $n/S \leq |A'| < 4n/5$. We may assume that $n - |B'| < 1.01|A'|$, since otherwise,

$$\Delta(A', B') \leq \left((10^5 + 72)|A'| - (10^5 - 72)(n - |B'|)\right) \cdot \sqrt{(D_G/S)\log D_G} < 0$$

and eq. (3.2) trivially holds. Assuming that $n - |B'| < 1.01|A'|$, we have by $(E1)$ and the r-niceness of G that

$$\left| E_{H_i^+}(A', B') \right| \geq (9/10) \cdot |E_G(A', B')| \, q_r/S \geq (9/10) \cdot q_r/S \cdot (D_G|A'| - |E_G(A', B \setminus B')|)$$
$$\geq D_G|A'|q_r/(100S) \geq (10^5 + 72)|A'|\sqrt{(D_G/S)\log D_G} \geq \Delta(A', B').$$

Case 3: $|A'| < n/S$. As before, we may assume that $n - |B'| < 1.01|A'|$. By (R2) and (E2), we have

$$\left| E_{H_i^+}(A', B') \right| \geq \delta_{H_i^+}|A'| - |E_{H_i}^+(A', B \setminus B')| \geq \left(\frac{D_G q_r}{S} - 9\sqrt{(D_G/S)\log D_G} - \frac{D_G q_r}{2S} \right)|A'|$$
$$\geq \frac{D_G q_r}{4S}|A'| \geq (10^5 + 72)|A'|\sqrt{(D_G/S)\log D_G} \geq \Delta(A', B').$$

$\square$

REMARK 12. *Our calculation in Case 1 and Case 2 in the above proof shows that for all $A' \subseteq A$ and $B' \subseteq B$ satisfying either $|B'| \geq |A'| \geq 4n/5$ or $4n/5 \geq |A'| \geq n/S$ and $n - |B'| < 1.01|A'|$, we have*

$$|E_G(A', B')| \geq D_G|A'|/100,$$

and similarly with the roles of A' and B' interchanged.

3.3 The building block Having established Theorem 11, our goal now is to construct a distribution on $\{\pi(e), \xi(e)\}_{e \in G}$ such that for r-nice graphs G, the random variables are r-admissible with sufficiently high probability and moreover, if G is drawn from an $O(D_r/n)$-spread distribution, then the resulting distribution on $(r+1)$-nice graphs is $O(D_{r+1}/n)$-spread. As in the previous subsection, we consider $r \in \mathbb{Z}_{\geq 0}$ and a D_G-regular bipartite graph $G = (A, B, E)$ with $|A| = |B| = n$ and $D_G \geq N_0$, which is r-nice (Theorem 8) with respect to the sequence D_r in the statement of Theorem 9. For $v \in A \cup B$, consider the following events:

- $\mathcal{R}_1(v)$ denotes the event that for some $i \in [S]$, v does not satisfy $(R1)$ in Theorem 10.

- $\mathcal{R}_2(v)$ denotes the event that for some $i \in [S]$, v does not satisfy $(R2)$ in Theorem 10.

Let $\boldsymbol{P}_G$ denote the product distribution on $\{\pi(e), \xi(e)\}_{e \in G}$, where each $\pi(e)$ is distributed uniformly in $[S]$ and each $\xi(e) \sim \text{Bernoulli}(q_r)$. We define the probability distribution $\mathbb{P}_G$ on $\{\pi(e), \xi(e)\}_{e \in G}$ to be the conditional distribution

$$\mathbb{P}_G := \boldsymbol{P}_G \left[\cdot \mid (\cap_{v \in V(G)} \mathcal{R}_1(v)^c) \bigcap (\cap_{v \in V(G)} \mathcal{R}_2(v)^c) \right].$$

Using the Chernoff bound (Theorem 7) and Theorem 6, we show that with very high probability, a sample from $\mathbb{P}_G$ satisfies all properties of admissibility except possibly $(E2)$.

LEMMA 13. *Let r and G be as above. With probability at least $1 - \exp(-n)$, a random realisation of $\{\pi(e), \xi(e)\}_{e \in G}$ drawn from the distribution $\mathbb{P}_G$ satisfies all properties in Theorem 10, except possibly $(E2)$.*

Proof. Note that $(R1)$ and $(R2)$ are always satisfied by construction. In order to control the probability that at least one of $(E1), (E3), (E4)$ fails, we will use Theorem 6. To this end, we begin by observing that a direct application of the Chernoff bound (Theorem 7) and union bound shows that for all $v \in V(G)$ and $j \in [\gamma]$,

(3.3) $$\max\left(\boldsymbol{P}_G[\mathcal{R}_1(v)], \boldsymbol{P}_G[\mathcal{R}_2(v)]\right) \leq D_r^{-20};$$

For the collection of events $\{\mathcal{R}_1(v), \mathcal{R}_2(v)\}_{v \in V(G)}$ and the product measure $\boldsymbol{P}_G$ on $\{\pi(e), \xi(e)\}_{e \in G}$, it is easily verified that a dependency graph Γ (in the sense of Theorem 5) is given by G itself, so that the condition in Theorem 6 is satisfied. By the Chernoff bound (Theorem 7) and Theorem 12, for any $A' \subseteq A, B' \subseteq B$ satisfying either $|B'| \geq |A'| \geq 4n/5$ or $4n/5 \geq |A'| \geq n/S$ and $n - |B'| < 1.01|A'|$, we have for any $i \in [S]$ that

$$\boldsymbol{P}_G \left[\underbrace{\left| E_{H_i^+}(A', B') \right| < (9/10) \cdot E_G(A', B') q_r / S}_{\mathcal{E}_1(A', B', i)} \right] \leq \exp\left(-\frac{E_G(A', B') q_r}{200 S} \right) \leq \exp\left(-\frac{D_G |A'| q_r}{20000 S} \right).$$

Since there are $4n$ events of the form $\mathcal{R}_1(v)$ or $\mathcal{R}_2(v)$, it follows from Theorem 6 that

$$\mathbb{P}'_G[\mathcal{E}_1(A', B', i)] \leq \exp\left(-\frac{D_G |A'| q_r}{20000 S} \right) \exp\left(24n D_r^{-20} \right) \leq \exp\left(-\frac{D_G |A'| q_r}{40000 S} \right) \leq \exp\left(-D_r^{3/4} |A'| \right),$$

so that a union bound over the relevant choices of A', B' and i (together with the case where the roles of A' and B' are interchanged) shows that

$$\mathbb{P}_G[(E1) \text{ fails}] \leq \exp(-n D_r^{1/2}).$$

Since G is r-nice, we have by $(N2)$ that for all $A' \subseteq A, B' \subseteq B$ with $|B'| \geq |A'| > n/S$ and $n - |B'| = 1.01|A'|$,

$$|E_G(A', B \setminus B')| \leq \exp(1 + \delta_{r-1}) D_G |A'| \underbrace{\max\left(\frac{1}{3}, \frac{n - |B'|}{n} \right)}_{\beta'}.$$

Therefore, by the Chernoff bound (Theorem 7),

$$\boldsymbol{P}_G \left[|E_{H_i}(A', B \setminus B')| > \exp(1 + 10 q_r + \delta_{r-1}) D_G |A'| \beta' / S \right] \leq \exp\left(-\frac{q_r^2 D_G |A'| \beta'}{S} \right) \leq \exp\left(-D_r^{5/8} |A'| \right),$$

so once again, using Theorem 6 and the union bound over relevant choices of A', B', i (together with the case where the roles of A' and B' are interchanged) shows that

$$\mathbb{P}_G[(E3) \text{ fails}] \leq \exp(-n D_r^{1/2}).$$

A similar computation shows that

$$\mathbb{P}_G[(E4) \text{ fails}] \leq \exp(-n D_r^{1/2}).$$

☐

It remains to show that the probability that $(E2)$ is violated by a sample from $\mathbb{P}_G$ is sufficiently small. While this is not necessarily true for all r-nice graphs G, we show in the key Theorem 15 that if G is drawn from an $O(D_r/n)$-spread distribution over r-nice graphs, then with very high probability over the choice of G, $\mathbb{P}_G$ does have this property.

DEFINITION 14. (R-EXCELLENCE) *For $r \in \mathbb{Z}_{\geq 0}$, we say that a bipartite graph $G = (A, B, E)$ with $|A| = |B| = n$ is r-excellent if G is r-nice and moreover, the probability that a random realisation $\{\pi(e), \xi(e)\}_{e \in G}$ sampled from $\mathbb{P}_G$ is r-admissible is at least $1 - n^{-50}$.*

For $r \in \mathbb{Z}_{\geq 0}$ and a regular bipartite graph $\mathbb{G} = (A, B, E)$ with $|A| = |B| = n$ and degree D_0, we say that a decomposition $\mathcal{P}$ of $E(\mathbb{G})$ is r-excellent if $|\mathcal{P}| = D_0/D_r(= S^r)$ and every $G \in \mathcal{P}$ is r-excellent.

PROPOSITION 15. *For any $C > 0$, there exist N_0 and S such that the following holds. With notation as above, let $\boldsymbol{G}_r$ be a probability distribution supported on r-nice bipartite graphs with vertex sets (A, B) of size $|A| = |B| = n$. If $\boldsymbol{G}_r$ is $C \cdot D_r/n$-spread, where $D_r \geq N_0$, then the probability that $G \sim \boldsymbol{G}_r$ is r-excellent is at least $1 - n^{-50}$.*

Proof. By Theorem 13, it suffices to show that with probability at least $1 - n^{-200}$, the collection $\{\pi(e), \xi(e)\}_{e \in G}$ sampled from the measure $\boldsymbol{G}_r \times \mathbb{P}_G$ (i.e. first sample G from $\boldsymbol{G}_r$ and then sample $\{\pi(e), \xi(e)\}_{e \in G}$ from $\mathbb{P}_G$) satisfies (E2). If (E2) fails, then there exists $i \in [S]$ and $A' \subseteq A, B' \subseteq B$ with $|A'| \leq n/S$ and $|B'| = 1.01|A'|$ (or A', B' with their roles switched) such that

$$\left| E_{H_i^+}(A', B') \right| > \frac{D_G q_r}{2S}|A'| > \frac{D_r q_r}{4S}|A'|.$$

Let $\mathcal{T}$ denote the collection of all subgraphs $T = (A', B', E(T))$ of $K_{n,n}$ with $|A'| \leq n/S$, $|B'| = 1.01|A'|$, and $|E(T)| = D_r q_r |A'|/4S$. For fixed $i \in [S]$ and any $T \in \mathcal{T}$, note that given an r-nice $G \supseteq T$, the event $T \subseteq H_i^+$ shares variables with at most $4|E(T)|$ events of the form $\mathcal{R}_1(v), \mathcal{R}_2(v)$. Therefore, using the spreadness of $\boldsymbol{G}_r$, eq. (3.3), and Theorem 6, we have that

$$\begin{aligned}
(\boldsymbol{G}_r \times \mathbb{P}_G)\left[T \subseteq H_i^+\right] &= \boldsymbol{G}_r[T \subseteq G] \cdot \mathbb{P}_G[T \subseteq H_i^+ \mid T \subseteq G] \\
&\leq \left(\frac{CD_r}{n}\right)^{|E(T)|} \cdot \boldsymbol{P}_G[T \subseteq H_i^+ \mid T \subseteq G] \cdot e^{4|E(T)|D_r^{-20}} \\
&\leq \left(\frac{e^{4D_r^{-20}} CD_r}{n}\right)^{|E(T)|} \cdot \left(\frac{q_r}{S}\right)^{|E(T)|} \leq \left(\frac{2Cq_r D_r}{nS}\right)^{|E(T)|}.
\end{aligned}$$

Therefore, by the union bound, and assuming that $D_r \geq N_0 \geq 2^{20}S^2$ and S is sufficiently large compared to C, it follows that

$$\begin{aligned}
(\boldsymbol{G}_r \times \mathbb{P}_G)\left[\bigcup_{T \in \mathcal{T}}\{T \subseteq H_i^+\}\right] &\leq \sum_{T \in \mathcal{T}} \left(\frac{2Cq_r D_r}{nS}\right)^{|E(T)|} \leq \sum_{k=1}^{n/S} \binom{n}{k}\binom{n}{1.01k}\binom{1.01k^2}{D_r q_r k/4S} \cdot \left(\frac{2Cq_r D_r}{nS}\right)^{D_r q_r k/4S} \\
&\leq \sum_{k=1}^{n/S} \left(\frac{4n}{k}\right)^{2k} \cdot \left(\frac{100Ck}{n}\right)^{D_r q_r k/4S} \leq \sum_{k=1}^{n/S} \left(\frac{100Ck}{n}\right)^{D_r q_r k/8S} \leq n^{-400}.
\end{aligned}$$

$\square$

3.4 Putting everything together Finally, we combine everything to show how to construct an $O(D_{r+1}/n)$-spread distribution on $(r+1)$-nice decompositions, starting with an $O(D_r/n)$-spread distribution on r-nice decompositions. As discussed earlier, this completes the proof of Theorem 9. Inductively, the following proposition establishes a probability distribution $\boldsymbol{\mathcal{P}}_r$ on decompositions $\mathcal{P}_r$ of $E(\mathbb{G})$ which is $\varepsilon^{-1} \exp(\sum_{r' < r} 5Sq_{r'})D_r/n$-spread. Note that, for N_0 sufficiently large, $\sum_{r' < r} q_r < 1/(100S)$ whenever $D_r \geq N_0 \geq 2^{20}S^{10}$ and thus $\boldsymbol{\mathcal{P}}_r$ is $2\varepsilon^{-1}D_r/n$-spread for any such r.

PROPOSITION 16. *With notation as in the statement of Theorem 9, let $\boldsymbol{\mathcal{P}}_r$ denote a probability distribution supported on r-nice decompositions of $E(\mathbb{G})$, and let $\boldsymbol{\mathcal{P}}_{r+1}$ be the distribution on $(r+1)$-nice decompositions of $E(\mathbb{G})$ defined as follows: first sample $\mathcal{P}$ from the conditional distribution $\boldsymbol{\mathcal{P}}_r \mid r$-excellent, then sample $\{\pi(e), \xi(e)\}_{e \in G}$ from the conditional distribution $\mathbb{P}_G \mid r$-admissible for each $G \in \mathcal{P}$, and finally use the procedure in Theorem 11. For any $C \leq 2\varepsilon^{-1}$, if $\boldsymbol{\mathcal{P}}_r$ is $C \cdot D_r/n$-spread, then $\boldsymbol{\mathcal{P}}_{r+1}$ is $C \cdot e^{5Sq_r} \cdot D_{r+1}/n$-spread.*

Proof. Let $\boldsymbol{\mathcal{P}}_r = (\boldsymbol{G}_1, \ldots, \boldsymbol{G}_{S^r})$. $D_{\boldsymbol{G}_i} = \exp(\pm \delta_{r-1}) \cdot D_r$. Since $\boldsymbol{\mathcal{P}}_r$ is $C \cdot D_r/n$-spread, the same holds by marginalization for each of the distributions $\boldsymbol{G}_1, \ldots, \boldsymbol{G}_{S^r}$. Therefore, by Theorem 15 and the union bound, the probability that $\mathcal{P}_r$ drawn from $\boldsymbol{\mathcal{P}}_r$ is r-excellent is at least $1 - n^{-49}$; in particular, the conditional distribution $\widetilde{\boldsymbol{\mathcal{P}}}_r := \boldsymbol{\mathcal{P}}_r \mid r$-excellent is spread with parameter at most $(1 - n^{-49})^{-1} \cdot C \cdot D_r/n \leq e^{q_r} CD_r/n$.

Recall that $\boldsymbol{\mathcal{P}}_{r+1} = (\boldsymbol{G}_{i,j})_{i \in [S^r], j \in [S]}$ is obtained by first sampling $(\boldsymbol{G}_1, \ldots, \boldsymbol{G}_{S^r})$ from $\widetilde{\boldsymbol{\mathcal{P}}}_r$ and then decomposing $\boldsymbol{G}_i$ into $\boldsymbol{G}_{i,1} \cup \cdots \cup \boldsymbol{G}_{i,S}$, independently for each $i \in [S^r]$, by sampling $\{\pi(e), \xi(e)\}_{e \in \boldsymbol{G}_i}$ from $\widetilde{\mathbb{P}}_{\boldsymbol{G}_i} := \mathbb{P}_{\boldsymbol{G}_i} \mid r$-admissible and using the procedure in Theorem 11. Let $\{T_{i,j}\}_{i \in [S^r], j \in [S]}$ be disjoint subsets of $E(\mathbb{G})$. For $i \in [S^r]$, let $T_i = \bigcup_{j \in [S]} T_{i,j}$. Let $I \subseteq [S^r]$ denote the subset of indices i for which $T_i \neq \emptyset$. Then,

$$(3.4) \qquad \Pr\left[\bigcap_{i \in [S^r], j \in [S]} \{T_{i,j} \subseteq \boldsymbol{G}_{i,j}\}\right] = \Pr\left[\bigcap_{i \in I, j \in [S]} \{T_{i,j} \subseteq \boldsymbol{G}_{i,j}\} \mid \bigcap_{i \in I}\{T_i \subseteq \boldsymbol{G}_i\}\right] \cdot \Pr\left[\bigcap_{i \in I}\{T_i \subseteq \boldsymbol{G}_i\}\right].$$

For the second term in the product, using the spreadness of $\widetilde{\boldsymbol{\mathcal{P}}}_r$, we have

$$(3.5) \qquad \Pr\left[\bigcap_{i\in I}\{T_i \subseteq \boldsymbol{G}_i\}\right] \leq \left(\frac{e^{q_r}CD_r}{n}\right)^{\sum_i |T_i|}.$$

For the first term in the product, we have the upper bound

$$(3.6) \qquad \prod_{i\in I}\widetilde{\mathbb{P}}_{\boldsymbol{G}_i}[\cap_{j\in[S]}T_{i,j} \subseteq \boldsymbol{G}_{i,j} \mid T_i \subseteq \boldsymbol{G}_i] \leq (1-n^{-50})^{-|I|} \cdot \prod_{i\in I}\mathbb{P}_{\boldsymbol{G}_i}[\cap_{j\in[S]}T_{i,j} \subseteq \boldsymbol{G}_{i,j} \mid T_i \subseteq \boldsymbol{G}_i],$$

where we have used that each $\boldsymbol{G}_i$ is supported on r-excellent graphs. Recalling that for $j \in [S]$, $\boldsymbol{G}_{i,j}$ is always contained in $H_j \cup H^+$, we can upper bound the ith term in the product by

$$\mathbb{P}_{\boldsymbol{G}_i}[\cap_{j\in[S]} \cap_{e\in T_{i,j}} \{(\pi(e) = j) \cup (\xi(e) = 1)\} \mid T_i \subseteq \boldsymbol{G}_i]$$
$$(3.7) \qquad\qquad \leq e^{20|T_i|D_r^{-20}}\left(\frac{1}{S} + q_r\right)^{|T_i|} \leq \left(\frac{e^{3Sq_r}}{S}\right)^{|T_i|},$$

where in the first inequality, we use Theorem 6 and eq. (3.3). Combining eqs. (3.4) to (3.7), we have that

$$\Pr\left[\bigcap_{i\in[S^r],j\in[S]}\{T_{i,j} \subseteq \boldsymbol{G}_{i,j}\}\right] \leq \left(\frac{e^{q_r}CD_r}{n}\right)^{\sum_i |T_i|} \cdot \left(\frac{e^{3Sq_r}}{S}\right)^{\sum_i |T_i|} \cdot (1-n^{-50})^{-|I|}$$
$$\leq \left(\frac{e^{5Sq_r}CD_r}{nS}\right)^{\sum_i |T_i|} \leq \left(\frac{e^{5Sq_r}CD_{r+1}}{n}\right)^{\sum_{i,j} |T_{i,j}|},$$

as desired. $\quad\square$

Acknowledgements.

References

[1] N. Alon and S. Assadi, Palette sparsification beyond $(\Delta + 1)$ vertex coloring, Approximation,randomization, and combinatorial optimization. Algorithms and techniques, LIPIcs. Leibniz Int. Proc. Inform., vol.176, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2020,Art. No. 6, 22.

[2] N. Alon and J. Spencer, The probabilistic method, fourth ed., Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2016.

[3] R. Alweiss, S. Lovett, K. Wu, and J. Zhang, Improved bounds for the sunflower lemma, *Ann. of Math.* **194**(3) (2021), 795–815.

[4] S. Assadi, Y. Chen, and S. Khanna, Sublinear algorithms for $(\Delta + 1)$ vertex coloring, *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SIAM, Philadelphia, PA*, (2019), 767—786.

[5] C. J. Casselgren and R. Häggkvist, Coloring complete and complete bipartite graphs from random lists, *Graphs Combin.* **32** (2016), 533–542.

[6] P. Erdős and A. Rényi, On the existence of a factor of degree one of a connected random graph, *Acta Math. Acad. Sci. Hungar.* **17** (1966), 359–728.

[7] K. Frankston, J. Kahn, B. Narayanan, and J. Park, Thresholds versus fractional expectation-thresholds, *Ann. of Math.* **194**(2) (2021), 475–495.

[8] S. Glock, D. Kühn, A. Lo, and D. Osthus, The existence of designs via iterative absorption: Hypergraph F-designs for arbitrary F, *Mem. Amer. Math. Soc.* (to appear), 95pp.

[9] B. Haeupler, B. Saha, and A. Srinivasan, New constructive aspects of the Lovász local lemma, *Journal of the ACM (JACM)*, vol. 58, no. 6, pp. 1–28, 2011.

[10] M. Halldórsson, F. Kuhn, A. Nolin, and T. Tonoyan, Near-optimal distributed degree+ 1 coloring, *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, (2022), 450–463.

[11] K. He, C. Wang, and Y. Yin, Sampling Lovász Local Lemma For General Constraint Satisfaction Solutions In Near-Linear Time, arXiv:2204.01520 (2022).

[12] V. Jain, H.T. Pham, and T.D. Vuong, Towards the sampling Lovász Local Lemma, *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 173–183, 2022.

[13] A. Johansson, Triangle factors in random graphs (2006), manuscript.

[14] A. Johansson, J. Kahn, and V. H. Vu, Factors in random graphs, *Random Struct. Algorithms* **33**(1) (2008), 1–28.

[15] D. Y. Kang, T. Kelly, D. Kühn, A. Methuku, D. Osthus, Thresholds for Latin squares and Steiner triple systems: Bounds within a logarithmic factor, arXiv:2206.14472 (2022).

[16] P. Keevash, The existence of designs, arXiv:1401.7265 (2014).

[17] P. Keevash, The optimal edge-colouring threshold, arXiv:2212.04397 (2022).

[18] L. Lovász, Combinatorial problems and exercises, second ed., AMS Chelsea Publishing, Providence, RI, 2007.

[19] Z. Luria and M. Simkin, On the threshold problem for Latin boxes, *Random Struct. Algorithms* **55** (2019), 926–949.

[20] J. Park and H. T. Pham, A proof of the Kahn–Kalai conjecture, arXiv:2203.17207 (2022).

[21] A. Sah, M. Sawhney, and M. Simkin, Threshold for Steiner triple systems, arXiv:2204.03964 (2022).

[22] M. Simkin, $(n, k, k - 1)$-Steiner systems in random hypergraphs, arXiv:1711.01975 (2017).

[23] M. Talagrand, Are many small sets explicitly small?, Proceedings of the 2010 ACM International Symposium on Theory of Computing (2010), 13–35.