NOTE ON RANDOM LATIN SQUARES AND THE TRIANGLE REMOVAL PROCESS

MATTHEW KWAN, ASHWIN SAH, AND MEHTAAB SAWHNEY

Abstract. This is a companion note to the paper "Almost all Steiner triple systems have perfect matchings" (arXiv:1611.02246). That paper contains several general lemmas about random Steiner triple systems; in this note we record analogues of these lemmas for random Latin squares, which in particular are necessary ingredients for our recent paper "Large deviations in random Latin squares" (arXiv:2106.11932). Most important is a relationship between uniformly random order-n Latin squares and the triangle removal process on the complete tripartite graph $K_{n,n,n}$.

1. Introduction

An order-n *Latin square* is usually defined as an $n \times n$ array of the numbers between 1 and n (we call these *symbols*), such that each row and column contains each symbol exactly once. In [5], Kwan introduced some general probabilistic techniques for studying so-called *Steiner triple systems*, and described how these techniques can be extended to Latin squares. The purpose of this note is to record complete proofs of various lemmas about random Latin squares, which are analogues of the lemmas in [5]. In particular, these lemmas are ingredients in our recent paper on large deviations on random Latin squares [6].

We emphasise that the proofs in this note are almost exactly the same as the proofs of corresponding lemmas in [5]; the goal of this note is completeness, not new ideas. Also, we refer the reader to [5, 6] for further references, motivation and background on this topic.

First, it will be more convenient for us to make a slightly different (equivalent) definition of a Latin square, in terms of 3-uniform hypergraphs.

Definition 1.1 (Latin squares)**.** Define

$$R = R_n = \{1, \ldots, n\}, \qquad C = C_n = \{n+1, \ldots, 2n\}, \qquad S = S_n = \{2n+1, \ldots, 3n\}.$$

We call the elements of R, C, S *rows*, *columns* and *symbols* respectively. Then, a *partial Latin square* (of order n) is a 3-partite 3-uniform hypergraph with 3-partition $V := R \cup C \cup S$, such that no pair of vertices is involved in more than one edge. Let $\mathcal{L}_m$ be the set of partial Latin squares with m hyperedges. A *Latin square* is a partial Latin square with exactly $N := n^2$ hyperedges (this is the maximum possible, and implies that every pair of vertices in different parts is contained in exactly one edge). Let $\mathcal{L}$ be the set of Latin squares.

Definition 1.2 (Ordered Latin squares)**.** Let $\mathcal{O}$ be the set of ordered Latin squares (i.e., Latin squares with an ordering on their set of hyperedges), and let $\mathcal{O}_m$ be the set of ordered partial Latin squares with m hyperedges. For $L \in \mathcal{O}_m$ and $i \leq m$, let L_i be the ordered partial Latin square consisting of just the first i hyperedges of L.

Definition 1.3 (Triangle removal process)**.** The (3-partite) *triangle removal process* is defined as follows. Start with the complete 3-partite graph $K_{n,n,n}$ on the vertex set $R \cup C \cup S$. At each step,

Part of this research was done while Kwan was working at Stanford University and ETH Zurich, and was supported in part by NSF grant DMS-1953990 and SNSF project 178493. Sah was supported by The Paul & Daisy Soros Fellowship. Sah and Sawhney were supported by NSF Graduate Research Fellowship Program DGE-1745302.

consider the set of all triangles in the current graph, select one uniformly at random, and remove it. Note that after m steps of this process, the removed triangles can be interpreted as an ordered partial Latin square $L \in \mathcal{O}_m$ (unless we run out of triangles before the mth step). Let $\mathbb{R}(n,m)$ be the distribution on $\mathcal{O}_m \cup \{*\}$ obtained from m steps of the triangle removal process (where "$*$" corresponds to the event that we run out of triangles). Note that it also makes sense to run the triangle removal process starting from some $G \subseteq K_{n,n,n}$ instead of starting from $K_{n,n,n}$ itself.

Definition 1.4 (Quasirandomness)**.** For this definition (and occasionally henceforth) we write V^1, V^2, V^3 instead of R, C, S for the three parts of $K_{n,n,n}$, and we write $V = V^1 \cup V^2 \cup V^3$ for the vertex set of $K_{n,n,n}$. The *density* of a subgraph $G \subseteq K_{n,n,n}$ is defined to be $d(G) = e(G)/(3N)$. A subgraph $G \subseteq K_{n,n,n}$ is (ε, h)*-quasirandom* if for each $q \in \{1,2,3\}$, every set $A \subseteq V \setminus V^q$ with $|A| \le h$ has $(1 \pm \varepsilon)d(G)^{|A|}n$ common neighbours in V^q. For a (possibly ordered) partial Latin square L, let $G(L)$ be the graph consisting of those edges of $K_{n,n,n}$ which are not included in any hyperedge of L (so if $m = N$ $(= n^2)$ then $G(L)$ is always the empty graph, and if $m = 0$ then always $G(L) = K_{n,n,n}$). Let $\mathcal{L}_m^{\varepsilon,h}$ be the set of partial Latin squares $P \in \mathcal{L}_m$ such that $G(P)$ is (ε, h)-quasirandom, and let $\mathcal{O}_m^{\varepsilon,h} \subseteq \mathcal{O}_m$ be the set of ordered partial Latin squares $L \in \mathcal{O}_m$ such that $L_i \in \mathcal{L}_i^{\varepsilon,h}$ for each $i \le m$.

Definition 1.5 (Binomial random hypergraph)**.** Let $\mathbb{G}^3(n,p)$ be the probability distribution on 3-partite 3-uniform hypergraphs with vertex set $R \cup C \cup S$, where every possible hyperedge respecting the 3-partition is included with probability p (so, the expected number of edges is pn^3).

Now, our lemmas are as follows. Recall that $N = n^2$. The first lemma states that quasirandom partial Latin squares have similar amounts of completions, up to multiplicative factors of $\exp(O(n^{2-\Omega(1)}))$. It is proved in Section 5.

Lemma 1.6. *For an ordered partial Latin square $L \in \mathcal{O}_m$, let $\mathcal{O}^*(L) \subseteq \mathcal{O}$ be the set of ordered Latin squares L^* such that $L^*_m = L$. For sufficiently large $h \in \mathbb{N}$ and any $a > 0$, there is $b = b(a,h) > 0$ such that the following holds. For any fixed $\alpha \in (0,1)$, if $\varepsilon = n^{-a}$ then any $L, L' \in \mathcal{O}_{\alpha N}^{\varepsilon,h}$ satisfy*

$$\frac{|\mathcal{O}^*(L)|}{|\mathcal{O}^*(L')|} \le \exp\Big(O\Big(n^{2-b}\Big)\Big).$$

The second lemma states that quasirandom partial Latin squares are output by the triangle removal process with comparable probabilities. It is proved in Section 3.

Lemma 1.7. *The following holds for any fixed $a \in (0,2)$ and $\alpha \in (0,1)$. Let $\varepsilon = n^{-a}$, let $L, L' \in \mathcal{O}_{\alpha N}^{\varepsilon,2}$ and let $\boldsymbol{L} \sim \mathbb{R}(n, \alpha N)$. Then*

$$\frac{\Pr(\boldsymbol{L} = L)}{\Pr(\boldsymbol{L} = L')} \le \exp\big(O\big(n^{2-a}\big)\big).$$

The third lemma essentially states that for any Latin square, most random subsets of its edges look quasirandom. It is proved in Section 2.

Lemma 1.8. *The following holds for any fixed $h \in \mathbb{N}$, $\alpha \in (0,1)$ and $a \in (0,1/2)$. Let $\varepsilon = n^{-a}$, consider any Latin square L, and uniformly at random order its hyperedges to obtain an ordered Latin square $\boldsymbol{L} \in \mathcal{O}$. Then* $\Pr\Big(\boldsymbol{L}_{\alpha N} \notin \mathcal{O}_{\alpha N}^{\varepsilon,h}\Big) = \exp\big(-\Omega\big(n^{1-2a}\big)\big)$.

The next lemma shows how to compare the triangle removal process to a nicer independent model (with deletions). It is proved in Section 4.

Lemma 1.9. *Let $\mathcal{P}$ be a property of unordered partial Latin squares that is monotone increasing in the sense that $L \in \mathcal{P}$ and $L' \supseteq L$ implies $L' \in \mathcal{P}$. Fix $\alpha \in (0,1)$, let $\boldsymbol{L} \sim \mathbb{R}(n, \alpha N)$, let $\boldsymbol{G} \sim \mathbb{G}^3(n,p)$*

for $p = \alpha/n$ and let $\boldsymbol{L}^*$ be the partial Latin square obtained from $\boldsymbol{G}$ by deleting (all at once) every hyperedge which intersects another hyperedge in more than one vertex. Then

$$\Pr(\boldsymbol{L} \notin \mathcal{P} \text{ and } \boldsymbol{L} \neq *) = O(\Pr(\boldsymbol{L}^* \notin \mathcal{P})).$$

As a final lemma, we verify that the triangle removal process succeeds, i.e., produces a partial Latin square instead of $*$, with probability $1 - o(1)$ (and in fact produces a quasirandom output). It is proved in Section 6, based on Theorem 6.1, which is a general analysis of the triangle removal process also needed in Section 5.

Lemma 1.10. *The following holds for any $h \in \mathbb{N}$. There is a constant $a = a(h) > 0$ such that if $\alpha \in (0, 1)$ and $\varepsilon = n^{-a}$, then for $\boldsymbol{L} \sim \mathbb{R}(n, \alpha N)$ we have*

$$\Pr(\boldsymbol{L} \notin \mathcal{O}^{\varepsilon,h}_{\alpha N} \text{ or } \boldsymbol{L} = *) = o(1).$$

1.1. **Notation.** We use standard asymptotic notation throughout. Here and for the rest of the paper, asymptotics are as $n \to \infty$. For functions $f = f(n)$ and $g = g(n)$:

- $f = O(g)$ means there is a constant C such that $|f| \leq C|g|$,
- $f = \Omega(g)$ means there is a constant $c > 0$ such that $f \geq c|g|$,
- $f = \Theta(g)$ means that $f = O(g)$ and $f = \Omega(g)$,
- $f = o(g)$ means that $f/g \to 0$.
- By "asymptotically almost surely", or "a.a.s.", we mean that the probability of an event is $1 - o(1)$. In particular, to say that a.a.s. $f = o(g)$ means that for any $\varepsilon > 0$, a.a.s. $f/g < \varepsilon$.

Also, following [4], the notation $f = 1 \pm \varepsilon$ means $1 - \varepsilon \leq f \leq 1 + \varepsilon$.

We also use standard graph theory notation: $V(G)$ and $E(G)$ are the sets of vertices and (hyper)edges of a (hyper)graph G, and $v(G)$ and $e(G)$ are the cardinalities of these sets. The subgraph of G induced by a vertex subset U is denoted $G[U]$, the degree of a vertex v is denoted $\deg_G(v)$, and the subgraph obtained by deleting v is denoted $G - v$.

For a positive integer n, we write $[n]$ for the set $\{1, 2, \ldots, n\}$. For a real number x, the floor and ceiling functions are denoted $\lfloor x \rfloor = \max\{i \in \mathbb{Z} : i \leq x\}$ and $\lceil x \rceil = \min\{i \in \mathbb{Z} : i \geq x\}$. We will however mostly omit floor and ceiling signs and assume large numbers are integers, wherever divisibility considerations are not important. All logarithms are in base e.

Finally, we remark that throughout the paper we adopt the convention that random variables (and random objects more generally) are printed in bold.

2. Randomly ordered Latin squares

In this section we prove Lemma 1.8.

Proof of Lemma 1.8. Recall that $N = n^2$ and consider $m \leq \alpha N$. Note that $\boldsymbol{L}_m$ (as an unordered partial Latin square) is a uniformly random subset of m hyperedges of L. Also note that

$$d(G(\boldsymbol{L}_m)) = \frac{3N - 3m}{3N} = 1 - \frac{m}{N}.$$

We can obtain a random partial Latin square almost equivalent to $\boldsymbol{L}_m$ by including each hyperedge of L with independent probability m/N. Let $\boldsymbol{L}'$ denote the partial Latin square so obtained, and let $\boldsymbol{G}' = G(\boldsymbol{L}')$. Now, fix $q \in \{1, 2, 3\}$ and fix a set A of at most h vertices not in V^q. It suffices to prove

$$\left| \bigcap_{w \in A} N_q(w) \right| = \big(1 \pm n^{-a}\big)\Big(1 - \frac{m}{N}\Big)^{|A|} n, \tag{2.1}$$

with probability $1 - \exp\big(-\Omega\big(n^{1-2a}\big)\big)$, where $N_q(w)$ is the neighbourhood of w in V_q, in the graph $\boldsymbol{G}'$. Indeed, the so-called Pittel inequality (see [3, p. 17]) would imply that the same estimate holds with essentially the same probability if we replace $\boldsymbol{L}'$ with $\boldsymbol{L}_m$ (thereby replacing $\boldsymbol{G}'$ with $G(\boldsymbol{L}_m)$). We would then be able to finish the proof by applying the union bound over all $m \le \alpha N$ and all choices of A.

Note that there are at most $\binom{|A|}{2} = O(1)$ hyperedges of L that include more than one vertex in A (by the defining property of a Latin square). Let U be the set of vertices involved in these atypical hyperedges, plus the vertices in A, so that $|U| = O(1)$. Let $\boldsymbol{N} = \left|\left(\bigcap_{w\in A} N_q(w)\right)\setminus U\right|$. For every $v \in V^q \setminus U$ and $w \in A$ there is exactly one hyperedge e_v^w in L containing v and w, whose presence in $\boldsymbol{L}'$ would prevent v from contributing to $\boldsymbol{N}$. For each fixed $v \in V^q \setminus U$ the hyperedges e_v^w, for $w \in A$, are distinct by definition of U, so

$$\Pr\left(v \in \bigcap_{w\in A} N_q(w)\right) = \left(1 - \frac{m}{N}\right)^{|A|},$$

and thus by linearity of expectation $\mathbb{E}\boldsymbol{N} = (1 - m/N)^{|A|}(n - |U|)$. Now, $\boldsymbol{N}$ is determined by the presence of at most $(n - |U|)|A| = O(n)$ hyperedges in $\boldsymbol{L}'$, and changing the presence of each affects $\boldsymbol{N}$ by at most $2 = O(1)$. So, by the Azuma–Hoeffding inequality (see [3, Section 2.4]),

$$\begin{aligned}\Pr\left(\left|\boldsymbol{N} - \left(1 - \frac{m}{N}\right)^{|A|} n\right| > n^{-a}\left(1 - \frac{m}{N}\right)^{|A|} n - |U|\right) &\le \exp\left(-\Omega\left(\frac{\left(n^{-a}(1-\alpha)^h n\right)^2}{n}\right)\right)\\ &= \exp\left(-\Omega\left(n^{1-2a}\right)\right).\end{aligned}$$

Finally, we recall that $\left|\left(\bigcap_{w\in A} N_q(w)\right)\right| = \boldsymbol{N} \pm |U|$, which completes the proof of (2.1). □

3. Approximate uniformity of the triangle removal process

In this section we prove Lemma 1.7. We first make the simple observation that the number of triangles in a quasirandom graph G can be easily estimated in terms of the density of G.

Proposition 3.1. *Let $G \subseteq K_{n,n,n}$ be an $(\varepsilon, 2)$-quasirandom graph on n vertices with $\varepsilon \in (0, 1]$. Then the number of triangles in G is $(1 \pm O(\varepsilon))n^3 d(G)^3$.*

Proof. For every vertex $v \in R$, its degree in C is $(1 \pm \varepsilon)nd(G)$. So, there are $(1 \pm \varepsilon)n^2 d(G)$ edges between R and C. Then, for each such edge, the number of ways to add a vertex $u \in S$ to create a triangle is $(1 \pm \varepsilon)nd(G)^2$. The desired result follows. □

Now we are ready to prove Lemma 1.7.

Proof of Lemma 1.7. Each $G(L_i)$ has

$$\left(1 \pm O\left(n^{-a}\right)\right)\left(1 - \frac{i}{N}\right)^3 n^3$$

triangles, by $(n^{-a}, 2)$-quasirandomness and Proposition 3.1. We therefore have

$$\Pr(\boldsymbol{L} = L) = \prod_{i=0}^{\alpha N - 1} \frac{1}{(1 \pm O(n^{-a}))(1 - i/N)^3 n^3},$$

and a similar expression holds for $\Pr(\boldsymbol{L} = L')$. Taking quotients term-by-term gives

$$\begin{aligned}\frac{\Pr(\boldsymbol{L} = L)}{\Pr(\boldsymbol{L} = L')} &\le \left(1 + O\left(n^{-a}\right)\right)^{\alpha N}\\ &\le \exp\left(O\left(n^{2-a}\right)\right)\end{aligned}$$

as desired. □

4. A coupling lemma

In this section we prove Lemma 1.9.

Proof of Lemma 1.9. Observe that $\boldsymbol{L}^*$ can be coupled with $\boldsymbol{L}$ in such a way that, if $e(\boldsymbol{G}) \le \alpha N$, then either $\boldsymbol{L}^* \subseteq \boldsymbol{L}$ or $\boldsymbol{L} = *$. Indeed, an equivalent way to define the triangle removal process (and thus the distribution of $\boldsymbol{L}$) is to take a uniformly random ordering of the triangles in $K_{n,n,n}$, go through the triangles in order, and accept each triangle if it is edge-disjoint from previously accepted triangles. Note that a random ordering of the hyperedges of $\boldsymbol{G}$ can be viewed as the first $e(\boldsymbol{G}) \sim \mathrm{Bin}(n^3, \alpha/n)$ elements of a random ordering of the set of triangles of $K_{n,n,n}$, and the triangle removal process with this ordering produces a superset of $\boldsymbol{L}^*$ whenever $e(\boldsymbol{G}) \le \alpha N$ and $\boldsymbol{L} \ne *$ (since every triangle in $\boldsymbol{L}^*$ by definition does not share an edge with the prior triangles).

It follows from this and the monotonicity of $\mathcal{P}$ that

$$\Pr(\boldsymbol{L} \notin \mathcal{P} \text{ and } \boldsymbol{L} \ne *) \le \Pr(\boldsymbol{L}^* \notin \mathcal{P} \,|\, e(\boldsymbol{G}) \le \alpha N).$$

Next, since $e(\boldsymbol{G})$ has a binomial distribution with mean $\mathbb{E}e(\boldsymbol{G}) = n^3\alpha/n = \alpha N$, it is easy to see that $\Pr(e(\boldsymbol{G}) \le \alpha N) = \Omega(1)$. It follows that

$$\Pr(\boldsymbol{L} \notin \mathcal{P} \text{ and } \boldsymbol{L} \ne *) \le \Pr(\boldsymbol{L}^* \notin \mathcal{P}) / \Pr(e(\boldsymbol{G}) \le \alpha N) = O(1) \cdot \Pr(\boldsymbol{L}^* \notin \mathcal{P}). \qquad \square$$

5. Counting completions of partial Latin squares

In this section we prove Lemma 1.6. As always, recall that $N = n^2$.

For a partial Latin square $L \in \mathcal{L}_{\alpha N}$, let $\mathcal{L}^*(L)$ be the number of full Latin squares that include L. We want to determine $|\mathcal{O}^*(L)| = (N - \alpha N)!|\mathcal{L}^*(L)|$ up to a factor of $e^{n^{2-b}}$ (for some $b > 0$).

First, we can get an upper bound via the *entropy method*. Before we begin the proof, we briefly remind the reader of the basics of the notion of entropy. For random elements $\boldsymbol{X}, \boldsymbol{Y}$ with supports $\operatorname{supp} \boldsymbol{X}$, $\operatorname{supp} \boldsymbol{Y}$, we define the (base-$e$) *entropy*

$$H(\boldsymbol{X}) = -\sum_{x \in \operatorname{supp} \boldsymbol{X}} \Pr(\boldsymbol{X} = x) \log(\Pr(\boldsymbol{X} = x))$$

and the *conditional entropy*

$$H(\boldsymbol{X} \,|\, \boldsymbol{Y}) = \sum_{y \in \operatorname{supp} \boldsymbol{Y}} \Pr(\boldsymbol{Y} = y) H(\boldsymbol{X} \,|\, \boldsymbol{Y} = y).$$

We will use two basic properties of entropy. First, we always have $H(\boldsymbol{X}) \le \log |\operatorname{supp} \boldsymbol{X}|$, with equality only when $\boldsymbol{X}$ has the uniform distribution on its support. Second, for any sequence of random elements $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$, we have

$$H(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) = \sum_{i=1}^{n} H(\boldsymbol{X}_i \,|\, \boldsymbol{X}_1, \dots, \boldsymbol{X}_{i-1}).$$

See for example [1] for an introduction to the notion of entropy and proofs of the above two facts.

Theorem 5.1. *For any $a \in (0, 2)$, any $\alpha \in [0, 1]$, and any $L \in \mathcal{L}_{\alpha N}^{n^{-a}, 2}$,*

$$|\mathcal{L}^*(L)| \le \left(\left(1 + O\left(n^{-a} + n^{-1/2}\right)\right)\left(\frac{1-\alpha}{e}\right)^2 n\right)^{N(1-\alpha)}.$$

Proof. Let $\boldsymbol{L}^* \in \mathcal{L}^*(L)$ be a uniformly random completion of L. We will estimate the entropy $H(\boldsymbol{L}^*) = \log |\mathcal{L}^*(L)|$ of $\boldsymbol{L}^*$.

Let $G = G(L)$. For each $e = \{x, y\} \in G[R \cup C]$, let $\{x, y, \boldsymbol{z}_e\}$ be the hyperedge that includes e in $\boldsymbol{L}^*$ (i.e., in the $n \times n$ array formulation of a Latin square, $\boldsymbol{z}_e$ is the symbol in the cell corresponding to

$\{x,y\}$). So, the sequence $(\boldsymbol{z}_e)_{e\in G[R\cup C]}$ determines $\boldsymbol{L}^*$. For any ordering $\prec$ on the edges of $G[R\cup C]$, we have

$$H(\boldsymbol{L}^*) = \sum_{e\in G[R\cup C]} H\big(\boldsymbol{z}_e \,\big|\, (\boldsymbol{z}_{e'} : e' \prec e)\big). \tag{5.1}$$

Now, a sequence $\lambda \in [0,1]^{R\times C}$ with all λ_e distinct induces an ordering $\prec_\lambda$ on the edges of $G[R\cup C]$, with $e' \prec_\lambda e$ when $\lambda_{e'} > \lambda_e$. Let $\boldsymbol{R}_e(\lambda)$ be an upper bound on $|\mathrm{supp}(\boldsymbol{z}_e \,|\, \{\boldsymbol{z}_{e'} : \lambda_{e'} > \lambda_e\})|$ defined as follows for $e = \{x,y\}$. $\boldsymbol{R}_e(\lambda)$ is 1 plus the number of vertices $v \in S\backslash\{\boldsymbol{z}_e\}$ such that $\{x,v\},\{y,v\} \in G$, and $\lambda_{e'} < \lambda_e$ for both the $e' \in G[R\cup C]$ included in the hyperedges that include $\{x,v\}$ and $\{y,v\}$ in $\boldsymbol{L}^*$. (In the $n\times n$ array formulation of a Latin square, this is just the number of symbols whose position has not yet been revealed in the row and column specified by e). Note that $\boldsymbol{R}_e(\lambda)$ is random depending on all of $(\boldsymbol{z}_e)_{e\in G[R\cup C]}$ (hence $\boldsymbol{L}^*$), even though the support we are bounding is only a function of $(\boldsymbol{z}_{e'})_{e'\prec_\lambda e}$.

Since $\boldsymbol{R}_e(\lambda)$ is an upper bound on $|\mathrm{supp}(\boldsymbol{z}_e \,|\, \boldsymbol{z}_{e'} : \lambda_{e'} > \lambda_e)|$, we have

$$H(\boldsymbol{z}_e \,|\, \{\boldsymbol{z}_{e'} : \lambda_{e'} > \lambda_e\}) \le \mathbb{E}[\log \boldsymbol{R}_e(\lambda)]. \tag{5.2}$$

It follows from (5.1) applied to $\prec_\lambda$ and (5.2) that

$$H(\boldsymbol{L}^*) \le \sum_{e\in G[R\cup C]} \mathbb{E}[\log \boldsymbol{R}_e(\lambda)].$$

This is true for any fixed λ, so it is also true if λ is chosen randomly, as follows. Let $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_e)_{e\in G}$ be a sequence of independent random variables, where each $\boldsymbol{\lambda}_e$ has the uniform distribution in $[0,1]$. (With probability 1 each $\boldsymbol{\lambda}_v$ is distinct from the others). Then

$$H(\boldsymbol{L}^*) \le \sum_{e\in G[R\cup C]} \mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda})].$$

Next, for any $L^* \in \mathcal{L}^*(L)$ and $\lambda_e \in [0,1]$, let

$$R_e^{L^*,\lambda_e} = \mathbb{E}[\boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{L}^* = L^*,\, \boldsymbol{\lambda}_e = \lambda_e].$$

(Note that $\boldsymbol{\lambda}_e = \lambda_e$ occurs with probability zero, so formally we should condition on $\boldsymbol{\lambda}_e = \lambda_e \pm \mathrm{d}\lambda_e$ and take limits in what follows, but there are no continuity issues so we will ignore this detail). Now, in G, by $(n^{-a},2)$-quasirandomness, x and y have $(1+O(n^{-a}))(1-\alpha)^2 n$ common neighbours (in S) other than $\boldsymbol{z}_e$. By the definition of $\boldsymbol{R}_e(\boldsymbol{\lambda})$ and linearity of expectation, we have

$$R_e^{L^*,\lambda_e} = 1 + \big(1+O\big(n^{-a}\big)\big)(1-\alpha)^2\lambda_e^2 n.$$

By Jensen's inequality,

$$\mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{L}^* = L^*,\, \boldsymbol{\lambda}_e = \lambda_e] \le \log R_e^{L^*,\lambda_e}.$$

We then have

$$\begin{aligned}\mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{L}^* = L^*] &\le \mathbb{E}\Big[\log R_e^{L^*,\boldsymbol{\lambda}_e}\Big] \\ &= \int_0^1 \log\Big(1 + \big(1+O\big(n^{-a}\big)\big)(1-\alpha)^2\lambda_e^2 n\Big)\,\mathrm{d}\lambda_e.\end{aligned}$$

For $C > 0$ we can compute

$$\int_0^1 \log\big(1+Ct^2\big)\,\mathrm{d}t = \log(1+C) - 2 + \frac{2\arctan\sqrt{C}}{\sqrt{C}}, \tag{5.3}$$

so (taking $C = (1+O(n^{-a}))(1-\alpha)^2 n$) we deduce

$$\mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda}) \,|\, \boldsymbol{L}^* = L^*] \le \log\Big((1-\alpha)^2 n\Big) - 2 + O\Big(n^{-a} + n^{-1/2}\Big).$$

We conclude that

$$\begin{aligned}\log|\mathcal{L}^*(L)| &= H(\boldsymbol{L}^*)\\ &\le \sum_{e\in G[R\cup C]} \mathbb{E}[\log \boldsymbol{R}_e(\boldsymbol{\lambda})]\\ &\le (N-\alpha N)\Big(\log\Big((1-\alpha)^2 n\Big) - 2 + O\Big(n^{-a}+n^{-1/2}\Big)\Big),\end{aligned}$$

which is equivalent to the theorem statement. □

For the lower bound, we will count ordered Latin squares.

Theorem 5.2. *Fixing sufficiently large $h\in\mathbb{N}$ and any $a>0$, there is $b=b(a,h)>0$ such that the following holds. For any fixed $\alpha\in(0,1)$ and any $L\in\mathcal{O}_{\alpha N}^{n^{-a},h}$,*

$$|\mathcal{O}^*(L)| \ge \left(\Big(1-O\Big(n^{-b}\Big)\Big)\left(\frac{1-\alpha}{e}\right)^2 n\right)^{N(1-\alpha)}(N-\alpha N)!.$$

To prove Theorem 5.2 we will need an analysis of the triangle removal process (which we provide in Section 6) and the following immediate consequence of [4, Theorem 1.5], which counts completions of Latin squares.

Theorem 5.3. *There are $h\in\mathbb{N}$, $\varepsilon_0, a\in(0,1)$ and $n_0,\ell\in\mathbb{N}$ such that if $L\in\mathcal{L}_m^{\varepsilon,h}$ is a partial Latin square with $n\ge n_0$, $d(G(L)) = 1-m/N \ge n^{-a}$ and $\varepsilon\le\varepsilon_0 d(G)^\ell$, then L can be completed to a Latin square.*

Proof of Theorem 5.2. Let $h\ge 2$, ℓ, ε_0 be as in Theorem 5.3. Let $c>0$ be smaller than $a\cdot b(a,h)$ in the notation of Theorem 6.1, and smaller than the "a" in Theorem 5.3. Let $\varepsilon = n^{-c/\ell}/\varepsilon_0$ and $\varepsilon' = n^{-c}$ and $M=(1-\varepsilon)N$. Let $\boldsymbol{L}^*\in\mathcal{O}_M\cup\{*\}$ be the result of running the triangle removal process on $G(L)$ to build a partial Latin square extending L, until there are M hyperedges. Let $\mathcal{O}^*$ be the set of M-hyperedge (ε',h)-quasirandom ordered partial Latin squares $L^*\in\mathcal{O}_M^{\varepsilon',h}$ extending L. The choice of c ensures that by Theorem 6.1 we a.a.s. have $\boldsymbol{L}^*\in\mathcal{O}^*$, and then by Theorem 5.3 each $L^*\in\mathcal{O}^*$ can be completed to an ordered Latin square.

Now, by Proposition 3.1 and quasirandomness coming from the output of Theorem 6.1, for each $L^*\in\mathcal{O}^*$ the number of triangles in each $G(L_i^*)$ is

$$\big(1\pm O\big(n^{-c}\big)\big)(1-i/N)^3 n^3,$$

so

$$\Pr(\boldsymbol{L}^*=L^*) \le \prod_{i=\alpha N}^{M-1}\frac{1}{(1-O(n^{-c}))(1-i/N)^3 n^3}.$$

As discussed, using Theorem 6.1 we have

$$\sum_{L^*\in\mathcal{O}^*}\Pr(\boldsymbol{L}^*=L^*) = 1-o(1),$$

so

$$\begin{aligned}|\mathcal{O}^*| &\ge (1-o(1))\prod_{i=\alpha N}^{M-1}\big(1-O\big(n^{-c}\big)\big)\left(1-\frac{i}{N}\right)^3 n^3\\ &= \big(\big(1-O\big(n^{-c}\big)\big)n^3\big)^{M-\alpha N}\exp\left(3\sum_{i=\alpha N}^{M-1}\log\left(1-\frac{i}{N}\right)\right).\end{aligned}$$

Now, note that

$$\sum_{i=\alpha N}^{M-1} \frac{1}{N} \log\left(1 - \frac{i+1}{N}\right) \le \int_{\alpha}^{(1-\varepsilon)} \log(1-t)\,\mathrm{d}t \le \sum_{i=\alpha N}^{M-1} \frac{1}{N} \log\left(1 - \frac{i}{N}\right).$$

We compute

$$\begin{aligned}\sum_{i=\alpha N}^{M} \left(\log\left(1-\frac{i}{N}\right) - \log\left(1-\frac{i+1}{N}\right)\right) &= \sum_{i=\alpha N}^{M} \log\left(1 + \frac{1}{N-(i+1)}\right)\\ &\le \sum_{i=\alpha N}^{M} \frac{1}{N-(i+1)}\\ &= O(\log n),\end{aligned}$$

so, noting that $\int \log s\,\mathrm{d}s = s(\log s - 1)$,

$$\begin{aligned}3\sum_{i=\alpha N}^{M} \log\left(1-\frac{i}{N}\right) &= 3N\int_{\alpha}^{(1-\varepsilon)} \log(1-t)\,\mathrm{d}t + O(\log n)\\ &= 3N\int_{\varepsilon}^{(1-\alpha)} \log s\,\mathrm{d}s + O(\log n)\\ &= 3N((1-\alpha)(\log(1-\alpha)-1) - \varepsilon(\log\varepsilon - 1)) + O(\log n),\end{aligned}$$

$$\exp\left(3\sum_{i=\alpha N}^{M} \log\left(1-\frac{i}{N}\right)\right) = \left(\left(1+O\left(n^{-c/\ell}\log n\right)\right)\frac{1-\alpha}{e}\right)^{3N(1-\alpha)}.$$

For $b < c/\ell$, it follows from this and $M = (1-\varepsilon)N$ that

$$\begin{aligned}|\mathcal{O}^*| &\ge \left(\left(1-O\left(n^{-b}\right)\right)\frac{n^3(1-\alpha)^3}{e^3}\right)^{(1-\alpha)N}\\ &= \left(\left(1-O\left(n^{-b}\right)\right)\left(\frac{1-\alpha}{e}\right)^2 n\right)^{(1-\alpha)N}(N-\alpha N)!.\end{aligned}$$

Recalling that each $L^* \in \mathcal{O}^*$ can be completed to a full Latin square, the desired result follows. □

Now, it is extremely straightforward to prove Lemma 1.6.

Proof. Let $b \le \min\{a, 1/2\}$ and $h \ge 2$ satisfy Theorem 5.2. By Theorem 5.1 we have

$$|\mathcal{O}^*(L)| = |\mathcal{L}^*(L)|(N-\alpha N)! \le \left(\left(1+O\left(n^{-b}\right)\right)\left(\frac{1-\alpha}{e}\right)^2 n\right)^{N(1-\alpha)}(N-\alpha N)!,$$

and by Theorem 5.2 we have

$$|\mathcal{O}^*(L')| \ge \left(\left(1-O\left(n^{-b}\right)\right)\left(\frac{1-\alpha}{e}\right)^2 n\right)^{N(1-\alpha)}(N-\alpha N)!.$$

Dividing these bounds gives

$$\frac{|\mathcal{O}^*(L)|}{|\mathcal{O}^*(L')|} \le \left(1+O\left(n^{-b}\right)\right)^{N(1-\alpha)} \le \exp\left(O\left(n^{2-b}\right)\right).$$ □

6. An analysis of the triangle removal process

In this section we prove Theorem 6.1, which was used in Section 5. We also deduce Lemma 1.10 from it, which will complete the proofs of all of our claims regarding random Latin squares.

The triangle removal process is defined as follows. We start with a graph $G \subseteq K_{n,n,n}$ with say $3N - 3m$ edges, then iteratively delete (the edges of) a triangle chosen uniformly at random from all triangles in the remaining graph. Let

$$G = \boldsymbol{G}(m), \boldsymbol{G}(m+1), \ldots$$

be the sequence of random graphs generated by this process. This process cannot continue forever, but we "freeze" the process instead of aborting it: if $\boldsymbol{G}(\boldsymbol{M})$ is the first graph in the sequence with no triangles, then let $\boldsymbol{G}(i) = \boldsymbol{G}(\boldsymbol{M})$ for $i \geq \boldsymbol{M}$.

Our objective in this section is to show that if G is quasirandom then the triangle removal process is likely to maintain quasirandomness and unlikely to freeze until nearly all edges are gone.

Theorem 6.1. *For all $h \geq 2$ and $a > 0$ there are $b = b(a,h) > 0$ and $\varepsilon' = \varepsilon'(a,h) > 0$ such that the following holds. Let $n^{-a} \leq \varepsilon < \varepsilon'$ and suppose $G \subseteq K_{n,n,n}$ is an (ε, h)-quasirandom graph with $N - 3m$ edges. Then a.a.s. $\boldsymbol{M} \geq \left(1 - \varepsilon^b\right)N$ and moreover for each $m \leq i \leq \left(1 - \varepsilon^b\right)N$, the graph $\boldsymbol{G}(i)$ is $\left(\varepsilon^b, h\right)$-quasirandom.*

Note that $K_{n,n,n}$ is $(O(1/n), h)$-quasirandom for any fixed h, so in particular when we start the triangle removal process from $G = K_{n,n,n}$ it typically runs almost to completion. This is encapsulated by Lemma 1.10, which we deduce before moving on to the proof of Theorem 6.1.

Proof of Lemma 1.10. Apply Theorem 6.1 to $G = K_{n,n,n}$, which is $(O(1/n), h)$-quasirandom, with its "a" set to 1. Letting $a = b(1,h)/2$, the result immediately follows since α is a constant. □

To prove Theorem 6.1, it will be convenient to use Freedman's inequality [2, Theorem 1.6], as follows. (This was originally stated for martingales, but it also holds for supermartingales with the same proof). Here and in what follows, we write $\Delta X(i)$ for the one-step change $X(i+1) - X(i)$ in a variable X.

Lemma 6.2. *Let $\boldsymbol{X}(0), \boldsymbol{X}(1), \ldots$ be a supermartingale with respect to a filtration $(\mathcal{F}_i)$. Suppose that $|\Delta \boldsymbol{X}(i)| \leq K$ for all i, and let $V(i) = \sum_{j=0}^{i-1} \mathbb{E}\left[(\Delta \boldsymbol{X}(j))^2 \,\middle|\, \mathcal{F}_j\right]$. Then for any $t, v > 0$,*

$$\Pr(\boldsymbol{X}(i) \geq \boldsymbol{X}(0) + t \text{ and } V(i) \leq v \text{ for some } i) \leq \exp\left(-\frac{t^2}{2(v + Kt)}\right).$$

Proof of Theorem 6.1. For $q \in \{1,2,3\}$ and a set $A \subseteq V \setminus V^q$ of at most h vertices, let $\boldsymbol{Y}_A(i) = \left|\bigcap_{w \in A} N_q^{(i)}(w)\right|$, where $N_q^{(i)}(w)$ is the number of neighbours of w in V^q, in the graph $\boldsymbol{G}(i)$. Let $p(i) = (1 - i/N)$ and let $p^k(i) = (1 - i/N)^k$, so that $p^{|A|}(i)n$ is the predicted trajectory of each $\boldsymbol{Y}_A(i)$.

Fix some large C and small c to be determined. We will choose $b < c/(C+1)$ so that $e(i) := p(i)^{-C}\varepsilon^c \leq \varepsilon^b$ for $i \leq N\left(1 - \varepsilon^b\right)$. This means that if the conditions

$$\boldsymbol{Y}_A(i) \leq p^{|A|}(i)n(1 + e(i)),$$
$$\boldsymbol{Y}_A(i) \geq p^{|A|}(i)n(1 - e(i)),$$

are satisfied for all A, then $\boldsymbol{G}(i)$ is $(e(i), h)$-quasirandom (therefore $\left(\varepsilon^b, h\right)$-quasirandom).

Let $\boldsymbol{T}'$ be the smallest index $i \geq m$ such that for some A, the above equations are violated (let $\boldsymbol{T}' = \infty$ if this never happens). Let $\boldsymbol{T} = \boldsymbol{T}' \wedge N\left(1 - \varepsilon^b\right)$, where $\wedge$ denotes the minimum. Define the stopped processes

$$\begin{aligned}
\boldsymbol{Y}_A^+(i) &= \boldsymbol{Y}_A(i \wedge \boldsymbol{T}) - p^{|A|}(i \wedge \boldsymbol{T})n(1 + e(i \wedge \boldsymbol{T})),\\
\boldsymbol{Y}_A^-(i) &= -\boldsymbol{Y}_A(i \wedge \boldsymbol{T}) + p^{|A|}(i \wedge \boldsymbol{T})n(1 - e(i \wedge \boldsymbol{T})).
\end{aligned}$$

We want to show that for each A and each $s \in \{+, -\}$, the process $\boldsymbol{Y}_A^s = (\boldsymbol{Y}_A^s(i), \boldsymbol{Y}_A^s(i+1), \dots)$ is a supermartingale, and then we want to use Lemma 6.2 and the union bound to show that a.a.s. each $\boldsymbol{Y}_A^s$ only takes negative values.

To see that this suffices to prove Theorem 6.1, note that if $i < \boldsymbol{T}$ then by Proposition 3.1 the number of triangles in $\boldsymbol{G}(i)$ is

$$\boldsymbol{Q}(i) = (1 \pm O(e(i)))p^3(i)n^3,$$

which is positive as $e(i) \le \varepsilon^b$ and as long as we choose $\varepsilon'(a, h)$ small enough.

This means $\boldsymbol{T} \le \boldsymbol{M}$, so the event that each $\boldsymbol{Y}_A^s$ only takes negative values contains the event that each $\boldsymbol{G}(i)$ is non-frozen and sufficiently quasirandom for $i \le N\big(1 - \varepsilon^b\big)$.

Let $\boldsymbol{R}_A(i) = \bigcap_{w \in A} N_q^{(i)}(w)$, so that $\boldsymbol{Y}_A(i) = |\boldsymbol{R}_A(i)|$. Fix A, and consider $x \in \boldsymbol{R}_A(i)$, for $i < \boldsymbol{T}$. The only way we can have $x \notin \boldsymbol{R}_A(i+1)$ is if we remove a triangle containing an edge $\{x, w\}$ for some $w \in A$. Now, for each $w \in A$, the number of triangles in $\boldsymbol{G}(i)$ containing the edge $\{x, w\}$ is $(1 \pm O(e(i)))p^2(i)n$ by $(e(i), 2)$-quasirandomness. The number of triangles containing x and more than one vertex of A is $O(1)$. So, if b is small enough then for realizations of $\boldsymbol{G}(i)$ with $i < \boldsymbol{T}$ we have

$$\begin{aligned}
\Pr(x \notin \boldsymbol{R}_A(i+1) \,|\, \boldsymbol{G}(i)) &= \frac{1}{\boldsymbol{Q}(i)}\left(\sum_{w \in A}(1 \pm O(e(i)))p^2(i)n - O(1)\right)\\
&= (1 \pm O(e(i)))\frac{|A|}{p(i)N}.
\end{aligned}$$

For $i < \boldsymbol{T}$ we have $|\boldsymbol{R}_A(i)| = (1 \pm e(i))p^{|A|}(i)n$, so by linearity of expectation

$$\begin{aligned}
\mathbb{E}[\Delta \boldsymbol{Y}_A(i) \,|\, \boldsymbol{G}(i)] &= -(1 \pm O(e(i)))\frac{|A|p^{|A|-1}(i)n}{N}\\
&= -\frac{|A|p^{|A|-1}(i)n}{N} + O\left(\frac{e(i)p^{|A|-1}(i)}{n}\right).
\end{aligned}$$

Note also that we have the bound $|\Delta \boldsymbol{Y}_A(i)| \le 2 = O(1)$ (with probability 1). Also, for fixed k, we have

$$\begin{aligned}
\Delta p^k(i) &= \left(1 - \frac{i+1}{N}\right)^k - \left(1 - \frac{i}{N}\right)^k\\
&= \left(1 - \frac{i}{N}\right)^k\left(\left(\frac{N-i-1}{N-i}\right)^k - 1\right)\\
&= p^k(i)\left(\left(1 - \frac{1}{N-i}\right)^k - 1\right)\\
&= p^k(i)\left(-\frac{k}{N-i} + O\left(\frac{1}{(N-i)^2}\right)\right)\\
&= -\frac{kp^{k-1}(i)}{N}\left(1 + O\left(\frac{p(i)}{n^2}\right)\right)\\
&= -\frac{kp^{k-1}(i)}{N} + o\left(\frac{e(i)p^{k-1}(i)}{n^2}\right),
\end{aligned}$$

and with ep^k denoting the pointwise product $i \mapsto e(i)p^k(i)$, we then have

$$\begin{aligned}\Delta\Big(ep^k\Big)(i) &= \varepsilon^c \Delta p^{k-C}(i)\\ &= \varepsilon^c \Theta\bigg(\frac{(C-k)p^{k-C-1}(i)}{N}\bigg)\\ &= \Theta\bigg(\frac{(C-k)e(i)p^{k-1}(i)}{n^2}\bigg).\end{aligned}$$

Also, $\Delta(ep^k)(i) > 0$ for $C > k$. For large C it thus follows that

$$\mathbb{E}\big[\Delta \boldsymbol{Y}_A^+(i) \,\big|\, \boldsymbol{G}(i)\big] = \mathbb{E}[\Delta \boldsymbol{Y}_A(i) \,|\, \boldsymbol{G}(i)] - \Delta p^{|A|}(i)n - \Delta\Big(ep^{|A|}\Big)(i)n \le 0,$$

and similarly

$$\mathbb{E}\big[\Delta \boldsymbol{Y}_A^-(i) \,\big|\, \boldsymbol{G}(i)\big] \le 0$$

for $i < \boldsymbol{T}$. (For $i \ge \boldsymbol{T}$ we trivially have $\Delta \boldsymbol{Y}_A^s(i) = 0$) Since each $\boldsymbol{Y}_A^s$ is a Markov process, it follows that each is a supermartingale. Now, we need to bound $\Delta \boldsymbol{Y}_A^s(i)$ and $\mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big]$, which is easy given the preceding calculations. First, recalling that $\Delta \boldsymbol{Y}_A(i) = O(1)$ and noting that $\Delta p^k(i), \Delta\big(ep^k\big)(i) = O(1/N)$ we immediately have $|\Delta \boldsymbol{Y}_A^s(i)| = O(1)$. Noting in addition that $\mathbb{E}[\Delta \boldsymbol{Y}_A(i) \,|\, \boldsymbol{G}(i)] = O(1/n)$, we have

$$\mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big] = O(\mathbb{E}[\Delta \boldsymbol{Y}_A^s(i) \,|\, \boldsymbol{G}(i)]) = O\bigg(\frac{1}{n}\bigg).$$

Since $\boldsymbol{T} \le N$, we also have

$$\sum_{i=m}^{\infty} \mathbb{E}\Big[(\Delta \boldsymbol{Y}_A^s(i))^2 \,\Big|\, \boldsymbol{G}(i)\Big] = O\bigg(\frac{N}{n}\bigg) = O(n).$$

Provided $c < 1$ and C is large enough (and recalling that $\varepsilon < \varepsilon'(a,h)$), applying Lemma 6.2 with $t = e(m)p^{|A|}(m)n - \varepsilon p^{|A|}(m)n = \Omega(n\varepsilon^c)$ and appropriate $v = O(n)$ then gives

$$\Pr(\boldsymbol{Y}_A^s(i) > 0 \text{ for some } i) \le \exp\big(-O\big(n\varepsilon^{2c}\big)\big).$$

Indeed, recall that we start at step $i = m$, and the initial quasirandomness conditions show that $\boldsymbol{Y}_A^s(m) \le \varepsilon p^{|A|}(m)n - e(m)p^{|A|}(m)n$.

So, if $2c < 1/a \le \log_{1/\varepsilon} n$, the union bound over all A, s finishes the proof. □

References

[1] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley & Sons, 2012.
[2] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probability **3** (1975), 100–118.
[3] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Cambridge University Press, 2000.
[4] P. Keevash, *The existence of designs II*, arXiv:1802.05900.
[5] M. Kwan, *Almost all Steiner triple systems have perfect matchings*, Proc. Lond. Math. Soc. (3) **121** (2020), no. 6, 1468–1495.
[6] M. Kwan, A. Sah, and M. Sawhney, *Large deviations in random Latin squares*, arXiv:2106.11932.

Institute of Science and Technology (IST) Austria
Email address: matthew.kwan@ist.ac.at

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA
Email address: {asah,msawhney}@mit.edu