

LARGE DEVIATIONS IN RANDOM LATIN SQUARES

MATTHEW KWAN, ASHWIN SAH, AND MEHTAAB SAWHNEY

ABSTRACT. In this note, we study large deviations of the number \mathbf{N} of *intercalates* (2×2 combinatorial subsquares which are themselves Latin squares) in a random $n \times n$ Latin square. In particular, for constant $\delta > 0$ we prove that $\Pr(\mathbf{N} \leq (1 - \delta)n^2/4) \leq \exp(-\Omega(n^2))$ and $\Pr(\mathbf{N} \geq (1 + \delta)n^2/4) \leq \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$, both of which are sharp up to logarithmic factors in their exponents. As a consequence, we deduce that a typical order- n Latin square has $(1 + o(1))n^2/4$ intercalates, matching a lower bound due to Kwan and Sudakov and resolving an old conjecture of McKay and Wanless.

1. INTRODUCTION

A *Latin square* (of order n) is an $n \times n$ array filled with the numbers 1 through n (we call these *symbols*), such that every symbol appears exactly once in each row and column. Latin squares are a fundamental type of combinatorial design, and in their various guises they play an important role in many contexts (ranging, for example, from group theory, to experimental design, to the theory of error-correcting codes). A classical introduction to the subject of Latin squares can be found in [33]. More recently, Latin squares have also played a role in the “high-dimensional combinatorics” program spearheaded by Linial, where they can be viewed as the first nontrivial case of a “high-dimensional permutation”¹ (see for example [42–44]).

There are still a number of surprisingly basic questions about Latin squares that remain unanswered, especially with regard to statistical aspects. For example, there is still a big gap between the best known upper and lower bounds on the number of order- n Latin squares (see for example [52, Chapter 17]), and there is no known algorithm that (provably) efficiently generates a random order- n Latin square². Perhaps the main difficulty is that Latin squares are extremely “rigid” objects: in general there is very little freedom to make local perturbations to change one Latin square into another.

Despite this difficulty, there are a number of theorems that have been rigorously proved about random Latin squares (and a larger number of conjectures and speculations); see for example [11–13, 27, 40, 41, 43–46, 53, 54]. A large portion of this work has focused on existence and enumeration of various types of substructures. As perhaps the simplest nontrivial example, an *intercalate* in a Latin square L is a order-2 Latin (combinatorial) subsquare. That is, it is a pair of rows $i < j$ and a pair of columns $x < y$ such that $L_{i,x} = L_{j,y}$ and $L_{i,y} = L_{j,x}$ (see Figure 1). It is a classical fact that (for all orders except 2 and 4) there exist Latin squares with no intercalates [38, 39, 47]. However, in 1999 McKay and Wanless [45] proved that with probability $1 - \exp(n^{2-o(1)})$ a random order- n Latin square has at least one intercalate, and that with probability $1 - o(1)$ there are at least $n^{3/2-o(1)}$ intercalates. In the same paper, they conjectured that the typical number of intercalates is $(1 + o(1))n^2/4$. More recently, Kwan and Sudakov [41] proved the lower bound in this conjecture—that random Latin squares typically have *at least* this many intercalates (see also [13] for previous progress on this conjecture). In the present paper we finally resolve McKay and Wanless’ conjecture in full.

Kwan was supported by NSF grant DMS-1953990. Sah and Sawhney were supported by NSF Graduate Research Fellowship Program DGE-1745302.

¹To see the analogy to permutation matrices, note that a Latin square can equivalently, and more symmetrically, be viewed as an $n \times n \times n$ zero-one array such that every axis-aligned line sums to exactly 1.

²Jacobson and Matthews [30] and Pittenger [48] designed Markov chains that converge to the uniform distribution, but it is not known whether these Markov chains mix rapidly.

4	1	5	3	2
5	3	2	1	4
2	4	1	5	3
3	5	4	2	1
1	2	3	4	5

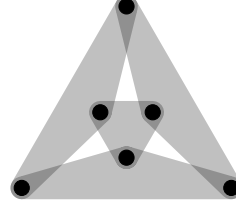


Figure 1. On the left is an example of a Latin square of order 5 with an intercalate in bold (in the first and second rows, and the second and fourth columns). On the right is depicted the 3-uniform hypergraph representation of an intercalate.

Theorem 1.1. *Let \mathbf{L} be a uniformly random order- n Latin square. Then, with probability $1 - o(1)$, the number of intercalates in \mathbf{L} is $(1 + o(1))n^2/4$.*

It is natural to draw an analogy to small subgraph counts in random graphs and hypergraphs. For example, in an Erdős–Rényi random graph $\mathbb{G}(n, p)$, the number of triangles is typically close to its expected value of $\binom{n}{3}p^3$ (as may be proved with a routine application of Chebyshev’s inequality). There is no obvious way to compute almost any kind of expected value in random Latin squares, but this point of view at least gives a heuristic explanation for why one should expect Theorem 1.1 to hold, as follows. An order- n Latin square can be equivalently viewed as a 3-partite 3-uniform hypergraph with parts of size n (corresponding to rows, columns, and symbols), satisfying the property that every pair of vertices in different parts is included in exactly one (hyper)edge. In this setting an intercalate is a subgraph isomorphic to a particular 4-edge hypergraph; see Figure 1. Now, every Latin square has exactly n^2 edges, so by symmetry, in a random Latin square each of the n^3 possible edges is present with probability $1/n$. If we imagine that each of these edges were present with probability $1/n$ *independently*, then the expected number of intercalates would be $2\binom{n}{2}^3(1/n)^6 = (1 + o(1))n^2/4$.

There are a huge number of questions about subgraph counts in random graphs and hypergraphs that have natural analogues for random Latin squares. One particularly influential direction is the study of *large deviations*. For example, what is the probability that a random graph $\mathbb{G}(n, p)$ has more than twice as many triangles as expected? What is the probability it has fewer than half as many as expected? These types of questions have been intensely studied and are intimately related to the development of many important techniques in graph theory and probability theory; see for example the monograph of Chatterjee [15] and the more recent works [1, 4, 5, 16, 28]. Beyond Theorem 1.1, we are able to prove the following near-optimal bounds on large deviation probabilities for intercalates in random Latin squares.

Theorem 1.2. *Fix a constant $\delta > 0$. Let \mathbf{N} be the number of intercalates in a uniformly random order- n Latin square \mathbf{L} . Then*

- (a) $\Pr(\mathbf{N} \leq (1 - \delta)n^2/4) \leq \exp(-\Omega(n^2))$,
- (b) $\Pr(\mathbf{N} \geq (1 + \delta)n^2/4) \leq \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$.

Moreover, these bounds are best-possible up to logarithmic factors in the exponent:

- (c) $\Pr(\mathbf{N} \leq (1 - \delta)n^2/4) \geq \exp(-O(n^2 \log n))$ for $\delta \leq 1$,
- (d) $\Pr(\mathbf{N} \geq (1 + \delta)n^2/4) \geq \exp(-O(n^{4/3} \log n))$.

Note that Theorem 1.1 is a direct corollary of Theorem 1.2(a–b). Another direct corollary of Theorem 1.2(a) is that $\Pr(\mathbf{N} = 0) \leq \exp(-\Omega(n^2))$, improving McKay and Wanless’ aforementioned bound of $\exp(-n^{2-o(1)})$. We remark that the lower tail bound proved by Kwan and Sudakov [41] was of the form $\Pr(\mathbf{N} \leq (1 - \delta)n^2/4) \leq \exp(-\Omega(\sqrt{n}/\log n))$.

From the form of the upper and lower tail probabilities in Theorem 1.2, one can already begin to get an idea for why the upper bound in Theorem 1.1 is more difficult than the lower bound. In general, for subgraph counts in random graphs and hypergraphs, lower tails tend to behave in a relatively simple “Gaussian-like” way, while upper tails tend to be quite different due to “clustering” behaviour (for example, in some regimes the “most likely way” for a random graph

to have a large number of triangles is for it to contain a large clique that has many triangles on its own). This phenomenon is often referred to as the “infamous upper tail” (see [31] for a survey). In the setting of [Theorem 1.2](#), it seems that the “most likely way” for a random Latin square to have a large number of intercalates is for it to contain a configuration similar to the multiplication table of an abelian 2-group $(\mathbb{Z}/2\mathbb{Z})^q$ (which may be interpreted as a Latin square of order 2^q), for suitably chosen q .

We remark that as a naïve approach to try to prove [Theorem 1.2](#), we might try to study the independent random hypergraph model mentioned earlier (in which each edge is present with probability $1/n$ independently), and to condition on the (hopefully not too unlikely) event that our random hypergraph is in fact a Latin square. For example, it is possible to study large deviations in random regular graphs with a related approach [6, 26] (although the details are highly nontrivial). However, the property of being a Latin square is extremely restrictive, and there does not seem to be any simple independent model that produces a Latin square property with probability greater than about $(1/\sqrt{n})^{n^2}$ (which is vanishingly small compared to the large deviation probabilities in [Theorem 1.2\(a–b\)](#)). Therefore, we employ some techniques not commonly seen in large deviations theory.

The upper and lower tails in [Theorem 1.2](#) are handled quite differently. For the lower tail, we employ the powerful machinery of Keevash (see [34–37]) originally developed for his celebrated proof of the *existence of designs* conjecture. Using Keevash’s machinery, Kwan [40] developed a general method for comparing random Latin squares with a stochastic graph process called the *triangle removal process*. It has been observed by Simkin [51] that this method is suitable for bounding lower tail probabilities, but to prove the strong bound in [Theorem 1.2\(a\)](#), we need to refine Kwan’s method (introducing an additional averaging technique).

For the upper tail bound in [Theorem 1.2\(b\)](#), instead of working directly with random Latin squares we work with random *Latin rectangles* (a Latin rectangle is a $k \times n$ array, for some $k \leq n$, filled with the symbols 1 through n , such that every number appears at most once in each row and column). As observed by McKay and Wanless, we can use estimates on the permanent (Bregman’s theorem [9] and the Egorychev–Falikman theorem [18, 21]) to compare random Latin rectangles with random Latin squares. To study random Latin rectangles we use the method of *switchings* (in which we study the typical effect of random perturbations to a Latin rectangle), in connection with a general enumeration theorem of Godsil and McKay [25] and the so-called *deletion method* of Rödl and Ruciński (see [32, 49]), adapted to this highly non-independent situation.

1.1. Further Directions. There are a few natural questions left open by our work. Let \mathbf{N} be the number of intercalates in a random order- n Latin square.

- Can we improve our understanding of the large deviation probabilities for \mathbf{N} , and sharpen the logarithmic factors³ in [Theorem 1.2](#)? In particular, it seems that [Theorem 1.2\(c\)](#) is improvable, but the difficulty lies in finding a general way to complete partial Latin squares to Latin squares without introducing too many intercalates. It seems that Keevash’s machinery may not be suitable for this, but the more recent approach of “iterative absorption” due to Glock, Kühn, Lo and Osthus [24] (see also [3]) may be helpful here.
- It would be nice to obtain a more accurate understanding of the expected value $\mathbb{E}\mathbf{N}$, and to say more about the distribution of \mathbf{N} (in particular, it is not even obvious how to estimate the variance of \mathbf{N}). With the ideas in this paper it is possible to find an explicit interval of length $n^{2-\Omega(1)}$ in which \mathbf{N} typically lies, but we suspect the true behaviour is that \mathbf{N} has an asymptotic Gaussian distribution with standard deviation $\Theta(n)$.
- We have studied 2×2 Latin subsquares; of course it is natural to consider subsquares of higher order. McKay and Wanless [45] conjectured that the expected number of

³We note that in the case of triangles in random graphs, it was an longstanding open problem to find the correct logarithmic factor in the exponent of the upper tail probability. This was famously solved by Chatterjee [14] and DeMarco and Kahn [17].

3×3 Latin subsquares is $1/18 + o(1)$ (we would further conjecture that the distribution is asymptotically Poisson with this mean), and they suggested that Latin subsquares of higher order should typically not appear at all. They also proved that $n/2 \times n/2$ subsquares are vanishingly unlikely in a random order- n Latin square (this is the largest a proper Latin subsquare could possibly be). We suspect it may not be too hard to show that a typical order- n Latin square does not contain a proper Latin subsquare of order greater than $n^{1/2+\varepsilon}$ (for any constant $\varepsilon > 0$), and it would be interesting to go beyond this. Of course, it is also possible to study more general subgraph statistics: for any fixed partial Latin square H , we can ask about the number of copies of H in a random Latin square.

- A *Steiner triple system* of order n is a 3-uniform hypergraph on a vertex set of size n , such that every pair of vertices is included in exactly one edge. These objects are natural “non-partite” analogues of Latin squares, and are even more difficult to study (to our knowledge, the only nontrivial results about random Steiner triple systems can be found in [2, 22, 40, 51]). In the setting of Steiner triple systems, the 4-edge hypergraph we have been calling an intercalate is usually called a *Pasch configuration*. Pasch configurations represent the smallest nontrivial “girth” obstruction for Steiner triple systems (Erdős conjectured that there exist Steiner triple systems with arbitrarily high girth; see [19]), and they provide one of very few ways to “switch” between different Steiner triple systems. Simkin [51] adapted some ideas of Kwan [40] to prove that a random Steiner triple system typically has at least $(1 - o(1))n^2/24$ Pasch configurations (and the ideas in this paper are suitable for proving near-optimal bounds on the lower-tail probabilities), but due to the “infamous upper tail” it will require new ideas to prove a corresponding bound for the upper tail.
- We would also like to draw attention to a few other interesting open problems in the area of random Latin squares that are a bit less directly related to the results in this paper. Linial and Luria [43] conjectured that random Latin squares typically satisfy an expansion property closely resembling the expander mixing lemma (see [41] for progress on this conjecture) and Cavenagh, Greenhill and Wanless [13] conjectured that a fixed pair of rows in a random Latin square can be very closely approximated (in some precise sense) by a uniformly random derangement (see also [11] for further discussion).

1.2. Notation. We use standard asymptotic notation throughout, as follows. For functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ to mean that there is a constant C such that $|f| \leq C|g|$, $f = \Omega(g)$ to mean that there is a constant $c > 0$ such that $f(n) \geq c|g(n)|$ for sufficiently large n , $f = \Theta(g)$ to mean that that $f = O(g)$ and $f = \Omega(g)$, and $f = o(g)$ to mean that $f/g \rightarrow 0$ as $n \rightarrow \infty$. Also, following [34], the notation $f = 1 \pm \varepsilon$ means $1 - \varepsilon \leq f \leq 1 + \varepsilon$.

We will use the convention that random objects (for example, random variables or random graphs) are printed in bold.

Acknowledgements. We thank Zachary Hunter for pointing out some important typographical errors.

2. APPROXIMATION FOR RANDOM LATIN SQUARES

In this section we state and prove a refined version of a theorem due to Kwan [40] ([Theorem 2.4](#)), using machinery due to Keevash [35] to approximate a random Latin square with the so-called triangle removal process. This will be the main technical ingredient for the proof of [Theorem 1.2\(a\)](#).

To say a bit more about our contribution: Kwan’s original approximation theorem ([40, Theorem 2.4]) is not capable of proving that any events hold with probability less than e^{-n} , so is not sufficient for proving the extremely strong lower tail bound in [Theorem 1.2\(a\)](#). Our improvement comes from an averaging/double counting technique (for the reader familiar with the proof of [40, Theorem 2.4], instead of conditioning on an outcome of a random subset of

a random Latin square, we average over many subsets). This averaging technique is closely related to the “distance to hyperplane” lemma in work of Rudelson and Vershynin [50], which is ubiquitous in random matrix theory.

First, we need some definitions, including the (equivalent) hypergraph formulation of a Latin square.

Definition 2.1. Define

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

We call the elements of R , C and S *rows*, *columns*, and *symbols* respectively. A *partial Latin square* (of order n) is a 3-partite 3-uniform hypergraph with 3-partition $R \cup C \cup S$, 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 edges. A *Latin square* is a partial Latin square with exactly n^2 edges (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 2.2. 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, 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 set of removed triangles can be interpreted as a partial Latin square $L \in \mathcal{L}_m$ (unless we run out of triangles before the m th step). Let $\mathbb{L}(n, m)$ be the distribution on $\mathcal{L}_m \cup \{*\}$ obtained from m steps of the triangle removal process (where “*” corresponds to the event that we run out of triangles).

Definition 2.3. Let $\mathcal{T}_m \subseteq \mathcal{L}_m$ be a property of m -edge partial Latin squares and let $\mathcal{T} \subseteq \mathcal{L}$ be a property of Latin squares. Say \mathcal{T}_m is ρ -*inherited* from \mathcal{T} if for any $L \in \mathcal{T}$, taking $\mathbf{L}_m \subseteq L$ as a uniformly random subset of m edges of L , we have $\mathbf{L}_m \in \mathcal{T}_m$ with probability at least ρ .

Now, our approximation theorem is as follows.

Theorem 2.4. *There is an absolute constant $\gamma > 0$ such that the following holds. Consider $\mathcal{T}_m \subseteq \mathcal{L}_m$ and $\mathcal{T} \subseteq \mathcal{L}$ such that \mathcal{T}_m is $1/2$ -inherited from \mathcal{T} . Let $\mathbf{P} \sim \mathbb{L}(n, m)$ be a partial Latin square obtained by m steps of the triangle removal process, and let $\mathbf{L} \in \mathcal{L}$ be a uniformly random order- n Latin square. Then*

$$\Pr(\mathbf{L} \in \mathcal{T}) \leq \exp(n^{2-\gamma}) \Pr(\mathbf{P} \in \mathcal{T}_m).$$

In our proof of [Theorem 2.4](#), we will cite some lemmas from [\[40\]](#) that are stated in the setting of Steiner triple systems, but the generalisations to Latin squares are straightforward (see [\[40, Section 7\]](#) for a discussion of the minor changes that are necessary).

2.1. Counting completions of partial Latin squares. First, we need the fact that all partial Latin squares satisfying a certain quasirandomness property extend to a Latin square in a comparable number of ways. This is proved with the entropy method, the triangle removal process, and Keevash’s machinery.

Definition 2.5. For this definition we write V^1, V^2, V^3 instead of R, C, S for the three parts of $K_{n,n,n}$. A subgraph $G \subseteq K_{n,n,n}$ with density $p(G) = e(G)/n^3$ is (ε, h) -*quasirandom* if for each $i \in \{1, 2, 3\}$, every set $A \subseteq V \setminus V^i$ with $|A| \leq h$ has $(1 \pm \varepsilon)(e(G)/(3n^2))^{|A|}n$ common neighbours in V^i . For a partial Latin square $P \in \mathcal{L}_m$, let $G(P)$ be the graph consisting of those edges of $K_{n,n,n}$ which are not included in some edge of P (so if $m = n^2$ then $G(P)$ is always the empty graph, and if $m = 0$ then always $G(P) = K_{n,n,n}$). 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.

Lemma 2.6 ([\[40, Lemma 2.5\]](#)). *Fixing a sufficiently large constant $h \in \mathbb{N}$ and fixing a constant $a > 0$, there is $b = b(a, h) > 0$ such that the following holds. Fix a constant $\alpha \in (0, 1)$, let*

$\varepsilon = n^{-a}$ and $m \leq \alpha n^2$, let $P, P' \in \mathcal{L}_m^{\varepsilon, h}$, let $\mathbf{L} \in \mathcal{L}$ be a random Latin square, and let \mathbf{L}_m be a uniformly random set of m edges of \mathbf{L} . Then

$$\frac{\Pr(\mathbf{L}_m = P)}{\Pr(\mathbf{L}_m = P')} \leq \exp(n^{2-b}).$$

2.2. The triangle removal process. Next, we need the fact that the triangle removal process produces every quasirandom partial Latin square with a comparable probability. This follows from the fact that quasirandom graphs have a predictable number of triangles.

Definition 2.7. An *ordered* partial Latin square is a partial Latin square $P \in \mathcal{P}_m$ together with an ordering on its edge set. Since the triangle removal process removes triangles sequentially, we can actually interpret $\mathbb{L}(n, m)$ as a distribution on ordered partial Latin squares with m edges. Let $\mathcal{O}_m^{\varepsilon, h}$ be the set of ordered partial Latin squares $P \in \mathcal{L}_m$ such that, for each $i \leq m$, writing P_i for the partial Latin square consisting of the first i edges of P , the graph $G(P_i)$ is (ε, h) -quasirandom.

Lemma 2.8 ([40, Lemma 2.7]). *The following holds for any fixed constant $a \in (0, 2)$. Let $\varepsilon = n^{-a}$, let $P, P' \in \mathcal{O}_m^{\varepsilon, 2}$ and let $\mathbf{P} \sim \mathbb{L}(n, m)$. Then*

$$\frac{\Pr(\mathbf{P} = P)}{\Pr(\mathbf{P} = P')} \leq \exp(O(n^{2-a})).$$

We also need the fact that the triangle removal process is likely to produce quasirandom partial Latin squares. This follows from a very simple and crude analysis (as in [40, Theorem 4.1]). We note that with modern techniques it is possible to prove a much stronger theorem (see [7]), but this will not be necessary for our application.

Lemma 2.9. *For any constant $h \in \mathbb{N}$ there is a constant $a = a(h) > 0$ such that the following holds. Fix $\alpha \in (0, 1)$, let $m \leq \alpha n^2$, let $\varepsilon = n^{-a}$ and let $\mathbf{P} \sim \mathbb{L}(n, m)$. Then $\Pr(\mathbf{P} \notin \mathcal{O}_m^{\varepsilon, h}) = o(1)$.*

2.3. Randomly ordered Latin squares. Finally, we need to know that a random ordering of any Latin square is likely to satisfy our quasirandomness property. This follows from a simple Chernoff bound calculation.

Lemma 2.10 ([40, Lemma 2.6]). *The following holds for any fixed constants $h \in \mathbb{N}$, $\alpha \in (0, 1)$ and $a \in (0, 1/2)$. Let $m \leq \alpha n^2$ and $\varepsilon = n^{-a}$, consider any Latin square L , and let \mathbf{L}_m be a random ordering of a random set of m edges of L . Then $\Pr(\mathbf{L}_m \notin \mathcal{O}_m^{\varepsilon, h}) = o(1)$.*

2.4. Putting everything together. Finally, we prove [Theorem 2.4](#).

Proof of Theorem 2.4. Let h be as in [Lemma 2.6](#), let $a = a(h)$ be as in [Lemma 2.9](#), let $\varepsilon = n^{-a}$, and let $b = b(a, h)$ be as in [Lemma 2.6](#).

Let N_{pair} be the number of pairs (L, L_m) where $L \in \mathcal{T}$ is a Latin square satisfying property \mathcal{T} , and L_m is an ordered partial Latin square consisting of m edges of L , which satisfies $\mathcal{T}_m \cap \mathcal{O}_m^{\varepsilon, h}$. Then

$$N_{\text{pair}} \geq (1/2 - o(1))n^2(n^2 - 1) \dots (n^2 - m + 1)|\mathcal{T}|,$$

by [Lemma 2.10](#) and the definition of being $1/2$ -inherited.

Let $N_{\text{ext}} = n^2(n^2 - 1) \dots (n^2 - m + 1)|\mathcal{L}|/|\mathcal{O}_m^{\varepsilon, h}|$ be an upper bound on the average number of ways to extend a partial Latin square $P \in \mathcal{O}_m^{\varepsilon, h}$ to a Latin square. By [Lemma 2.6](#), we have

$$N_{\text{pair}} \leq \exp(n^{2-b})|\mathcal{T}_m \cap \mathcal{O}_m^{\varepsilon, h}|N_{\text{ext}}.$$

It follows that

$$\Pr(\mathbf{L} \in \mathcal{T}) = \frac{|\mathcal{T}|}{|\mathcal{L}|} \leq (2 + o(1)) \exp(n^{2-b}) \frac{|\mathcal{T}_m \cap \mathcal{O}_m^{\varepsilon, h}|}{|\mathcal{O}_m^{\varepsilon, h}|}.$$

Using [Lemma 2.8](#), we have

$$\frac{|\mathcal{T}_m \cap \mathcal{O}_m^{\varepsilon, h}|}{|\mathcal{O}_m^{\varepsilon, h}|} \leq \exp(O(n^{2-a})) \Pr(\mathbf{P} \in \mathcal{T}_m \mid \mathbf{P} \in \mathcal{O}_m^{\varepsilon, h}),$$

and using [Lemma 2.9](#), we have

$$\Pr(\mathbf{P} \in \mathcal{T}_m \mid \mathbf{P} \in \mathcal{O}_m^{\varepsilon, h}) \leq \frac{\Pr(\mathbf{P} \in \mathcal{T}_m)}{\Pr(\mathbf{P} \in \mathcal{O}_m^{\varepsilon, h})} = (1 + o(1)) \Pr(\mathbf{P} \in \mathcal{T}_m).$$

The desired result follows (taking $\gamma < \min\{b, a\}$). \square

3. LATIN RECTANGLES

In this section we recall the notion of a Latin rectangle and some useful facts about them. The results in this section will be used in the proofs of the upper tail bounds [Theorem 1.2](#)(b,d).

Definition 3.1. A *Latin rectangle* (of order n , with k rows) is a $k \times n$ array containing the symbols $1, \dots, n$, such that every symbol appears at most once in each row and column. (So, if $k = n$ this is the same as a Latin square). A *partial* Latin rectangle is a $k \times n$ array satisfying the same property, but where some of the cells are allowed to be empty.

There is also an equivalent hypergraph formulation of a Latin rectangle.

Definition 3.2. Recall the sets $R = R_n, C = C_n, S = S_n$ from [Definition 2.1](#) and, for $k \leq n$, in addition define

$$R^{(k)} = \{1, \dots, k\} \subseteq R.$$

A *partial Latin rectangle* (of order n , with k rows) is a 3-partite 3-graph with tripartition $R^{(k)} \cup C \cup S$ such that no pair of vertices is involved in more than one edge. A *Latin rectangle* is a partial Latin rectangle with exactly kn edges (which is the maximum possible). Let $\mathcal{Q}^{(k)}$ denote the set of all such Latin rectangles (we omit the superscript when k is clear from context). We note that one can similarly define $C^{(k)} = \{n+1, \dots, n+k\}$ and $S^{(k)} = \{2n+1, \dots, 2n+k\}$, and symmetrically define a notion of a Latin rectangle with k columns or with k symbols.

We will switch back and forth between the two equivalent definitions in [Definitions 3.1](#) and [3.2](#), depending on which is more convenient at the time (this will be clear from context).

3.1. Counting completions of Latin rectangles. The primary reason Latin rectangles will be important for us is that every Latin rectangle can be completed in roughly the same number of ways to a Latin square. The following lemma is from [\[45, Proposition 4\]](#). It is proved by iteratively applying Bregman's theorem [\[9\]](#) and the Egorychev–Falikman theorem [\[18, 21\]](#) to give upper and lower bounds on the number of ways to add an extra row to a given Latin rectangle.

Theorem 3.3. Let Q, Q' be two Latin rectangles with order n and the same number k of rows. Let \mathbf{L} be a random Latin square and let \mathbf{L}_k be the Latin rectangle consisting of its first k rows. Then

$$\frac{\Pr(\mathbf{L}_k = Q)}{\Pr(\mathbf{L}_k = Q')} = e^{O(n(\log n)^2)}.$$

3.2. Subset probabilities in random Latin rectangles. The following theorem provides estimates on the probability that a given set of entries is present in a random Latin rectangle. It is a direct consequence of a theorem of Godsil and McKay ([\[25, Theorem 4.7\]](#)), and is proved using the switching method.

Theorem 3.4. Let P be a partial Latin rectangle, let $d_i(P)$ denote the number of entries of P in row i , and let $\Delta = \max_{1 \leq i \leq k} d_i(P)$. Let \mathbf{L} be a uniformly random Latin rectangle. Then

$$\Pr(P \subseteq \mathbf{L}) = \left(\frac{1 + O(k/(n - 2k - \Delta))}{n} \right)^{|P|}.$$

4. LOWER BOUNDS

In this section we prove [Theorem 1.2\(c-d\)](#), lower-bounding the large deviation probabilities for the number of intercalates \mathbf{N} in a random Latin square.

Proof of [Theorem 1.2\(c\)](#). As noted in the introduction, for all orders except 2 and 4 there is a Latin square with no intercalates (this combines results of Kotzig, Lindner and Rosa [38], McLeish [47], and Kotzig and Turgeon [39]). On the other hand, the total number of order- n Latin squares is clearly at most $n^{n^2} = \exp(O(n^2 \log n))$. The desired result follows. \square

Proof of [Theorem 1.2\(d\)](#). Let $k = 2^q$ be the smallest power of two such that $k \binom{k}{2}/2 \geq (1 + \delta)n^2/4$. Let L be the Latin square corresponding to the multiplication table of $(\mathbb{Z}/2\mathbb{Z})^q$ (where we fix some correspondence between elements of this group and the integers $1, \dots, k$). Then, L has order $k = \Theta(n^{2/3})$, and it is easy to see that it has $k \binom{k}{2}/2 \geq (1 + \delta)n^2/4$ intercalates (see for example [10]). Let $\mathbf{Q} \in \mathcal{Q}$ be a uniformly random order- n Latin rectangle with k rows. By [Theorem 3.4](#), with probability at least $((1 - o(1))/n)^{k^2} = \exp(-O(n^{4/3} \log n))$, our special Latin square L appears in the first k columns of \mathbf{Q} .

Let \mathbf{L}_r be the Latin rectangle consisting of the first k rows of our random Latin square \mathbf{L} . By [Theorem 3.3](#), the probabilities of different outcomes of \mathbf{L}_r differ by a factor of only $e^{O(n(\log n)^2)}$, so with probability at least $\exp(-O(n^{4/3} \log n + n(\log n)^2)) = \exp(-O(n^{4/3} \log n))$, our special Latin square L appears in the first k rows and columns of \mathbf{L} , in which case $\mathbf{N} \geq (1 + \delta)n^2/4$. \square

5. UPPER-BOUNDING THE LOWER TAIL

In this section we prove [Theorem 1.2\(a\)](#). We will apply [Theorem 2.4](#) with \mathcal{T} being the property of having “too few” intercalates. First, we establish that this property is likely to be inherited by random subsets.

Lemma 5.1. *Fix $\alpha, \delta \in [0, 1]$, let $\mathcal{T}^\delta \subseteq \mathcal{L}$ be the property that a Latin square $L \in \mathcal{L}$ has at most $(1 - \delta)n^2/4$ intercalates, and for $m = \alpha n^2$ let $\mathcal{T}_m^\delta \subseteq \mathcal{L}_m$ be the property that a partial Latin square $P \in \mathcal{L}_m$ has at most $\alpha^4(1 - \delta/2)n^2/4$ intercalates. Then \mathcal{T}_m^δ is $1/2$ -inherited from \mathcal{T}^δ .*

Proof. Let $L \in \mathcal{T}^\delta$ and let \mathbf{L}_m consist of m random edges of L . Let \mathcal{I} be the set of intercalates in L , for $I \in \mathcal{I}$ let $\mathbf{1}_I$ be the indicator random variable for the event that $I \subseteq L_m$, and let $\mathbf{X} = \sum_{I \in \mathcal{I}} \mathbf{1}_I$ be the number of intercalates in \mathbf{L}_m . For each $I \in \mathcal{I}$ we have $\mathbb{E}\mathbf{1}_I = \alpha^4 + O(1/n)$, so $\mathbb{E}\mathbf{X} \leq \alpha^4(1 - \delta + o(1))n^2/4$. Also, for each pair of disjoint $I, J \in \mathcal{I}$ we have $\text{Cov}(\mathbf{1}_I, \mathbf{1}_J) = O(1/n)$. In every Latin square, every intercalate intersects at most $4n$ other intercalates, so there are $O(n^3)$ intersecting pairs of intercalates in \mathcal{I} , meaning $\text{Var } \mathbf{X} = O(n^3)$. By Chebyshev’s inequality, we conclude that

$$\Pr(\mathbf{L}_m \in \mathcal{T}) = \Pr(\mathbf{X} < \alpha^4(1 - \delta/2)n^2/4) = 1 - o(1) > 1/2,$$

meaning that \mathcal{T}_m^δ is $1/2$ -inherited from \mathcal{T}^δ . \square

Before we continue with the proof we record some auxiliary lemmas.

5.1. A coupling lemma. It is not very easy to study the triangle removal process directly, so the following coupling lemma is useful in combination with [Theorem 2.4](#). It has the same proof as [40, Lemma 2.10]. Let $\mathbb{G}^{(3)}(n, p)$ be the random 3-partite 3-graph on the vertex set $R \cup C \cup S$ obtained by including all possible edges with probability p independently.

Lemma 5.2. *Let \mathcal{T} be a property of unordered partial Latin squares that is monotone increasing in the sense that $P \in \mathcal{T}$ and $P' \supseteq P$ implies $P' \in \mathcal{T}$. Fix $\alpha \in (0, 1)$, let $\mathbf{P} \sim \mathbb{L}(n, \alpha n^2)$, let $\mathbf{G} \sim \mathbb{G}^{(3)}(n, \alpha/n)$ and let \mathbf{G}^* be the partial Latin square obtained from \mathbf{G} by deleting every edge which intersects another edge in more than one vertex (all at once). Then*

$$\Pr(\mathbf{P} \in \mathcal{T}) \leq O(1) \Pr(\mathbf{G}^* \in \mathcal{T}).$$

We remark that one can prove a similar coupling lemma for monotone decreasing properties (see [22, Lemma 2.6]), though this will not be necessary for us.

5.2. A concentration inequality. The following concentration inequality may be deduced from an inequality of Freedman [23]. It appears as [40, Theorem 2.11].

Theorem 5.3. *Let $\omega = (\omega_1, \dots, \omega_N)$ be a sequence of independent, identically distributed random variables with $\Pr(\omega_i = 1) = p$ and $\Pr(\omega_i = 0) = 1 - p$. Let $f : \{0, 1\}^N \rightarrow \mathbb{R}$ satisfy the Lipschitz condition $|f(\omega) - f(\omega')| \leq K$ for all pairs $\omega, \omega' \in \{0, 1\}^N$ differing in exactly one coordinate. Then*

$$\Pr(|f(\omega) - \mathbb{E}f(\omega)| > t) \leq \exp\left(-\frac{t^2}{4K^2Np + 2Kt}\right).$$

5.3. Putting everything together. We are now ready to prove Theorem 1.2(a).

Proof of Theorem 1.2(a). Let $\alpha > 0$ be some constant that is sufficiently small with respect to δ , let $m = \alpha n^2$, let \mathbf{G}, \mathbf{G}^* be as in Lemma 5.2, and let \mathcal{T}_m^δ be the property that a partial Latin square (not necessarily with exactly m edges) has at most $(1 - \delta/2)\alpha^4 n^2/4$ intercalates. We bound $\Pr(\mathbf{G}^* \in \mathcal{T}_m^\delta)$ using a “maximum disjoint family” technique essentially due to Bollobás [8]. Let \mathbf{N} be the number of intercalates in \mathbf{G}^* , let \mathbf{N}' be the maximum size of a collection of disjoint intercalates in \mathbf{G}^* , and let \mathbf{N}_2 be the number of pairs of distinct intercalates in \mathbf{G} which share an edge. Observe that $\mathbf{N}' \geq \mathbf{N} - \mathbf{N}_2$.

Now, we estimate $\mathbb{E}\mathbf{N}'$. There are $2\binom{n}{2}^3$ ways to specify an intercalate, and each is present in \mathbf{G}^* with probability $(\alpha/n)^4(1 - \alpha/n)^{12(n-1)-8}$. So,

$$\mathbb{E}\mathbf{N} = (e^{-12\alpha} + o(1))\alpha^4 n^2/4.$$

For $\mathbb{E}\mathbf{N}_2$, we observe that there are $O(n^7)$ ways to specify a pair of intercalates that share two edges, and each such pair is present in \mathbf{G} with probability $(\alpha/n)^6$. There are $2\binom{n}{2}^3 \cdot 4(n-2)^3$ ways to specify an ordered pair of intercalates that share one edge, and each such pair is present in \mathbf{G} with probability $(\alpha/n)^7$. So,

$$\mathbb{E}\mathbf{N}_2 = (\alpha^3 + o(1))\alpha^4 n^2.$$

If α is sufficiently small (in terms of δ) then $\mathbb{E}\mathbf{N}' \geq \mathbb{E}\mathbf{N} - \mathbb{E}\mathbf{N}_2 \geq \alpha^4(1 - \delta/4)n^2/4$.

We next claim that \mathbf{N}' is a 3-Lipschitz function of the edges of the random hypergraph \mathbf{G} . Indeed, adding an edge to \mathbf{G} can increase \mathbf{N}' by at most one, and removing an edge from \mathbf{G} can increase \mathbf{N}' by at most three (by adding up to three edges to \mathbf{G}^*). So, by Theorem 5.3 we have

$$\Pr(\mathbf{G}^* \in \mathcal{T}_m^\delta) \leq \Pr(\mathbf{N}' \leq \mathbb{E}\mathbf{N}' - \alpha^4 \delta n^2/16) \leq \exp(-\Omega(n^2)).$$

It follows from Lemma 5.2 that if $\mathbf{P} \sim \mathbb{L}(n, \alpha n^2)$ then $\Pr(\mathbf{P} \in \mathcal{T}_m^\delta) \leq \exp(-\Omega(n^2))$. The desired result follows from Lemma 5.1 and Theorem 2.4. \square

6. UPPER-BOUNDING THE UPPER TAIL

In this section we prove Theorem 1.2(b). We will work mostly with random Latin rectangles, and use Theorem 3.3 to transfer our results to random Latin squares. Recall that we defined two equivalent notions of a Latin rectangle (Definitions 3.1 and 3.2); we will use both perspectives in this section.

6.1. Deletion. The first step in the proof of Theorem 1.2(b) is to adapt the deletion method of Rödl and Ruciński (see [32, 49]), using Theorem 3.4, to reduce to the case where one has a small subset of edges which contributes a large number of the intercalates. To effectively apply Theorem 3.4, for now we restrict our attention to a small number of rows, columns and symbols.

Lemma 6.1. *Fix a sufficiently small constant $\delta > 0$, and let $k = \lfloor \delta^2 n \rfloor$. Let \mathbf{L} be a uniformly random order- n Latin square, and let $\mathbf{L}^{(k)}$ be the subhypergraph induced by $R^{(k)} \cup C^{(k)} \cup S^{(k)}$ (that is to say, $\mathbf{L}^{(k)}$ consists of the entries in the first k columns and the first k rows, involving the first k symbols). Then with probability $1 - \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$ there is a set $\mathbf{E}_0 \subseteq \mathbf{L}^{(k)}$ of $n^{4/3}(\log n)^{2/3}$ edges such that $\mathbf{L}^{(k)} \setminus \mathbf{E}_0$ contains at most $(1 + \delta/2)k^6/(4n^4)$ intercalates.*

Proof. We first note that this event is purely a function of the first k rows of \mathbf{L} , and therefore by [Theorem 3.3](#) it suffices to prove the same event for a uniformly random order- n Latin rectangle $\mathbf{Q} \in \mathcal{Q}$ with k rows (as the relative change of measure $\exp(O(n(\log n)^2))$ is simply swallowed into the error term). Let $\mathbf{Q}^{(k)}$ be the subhypergraph of \mathbf{Q} induced by $R^{(k)} \cup C^{(k)} \cup S^{(k)}$.

Now let \mathcal{A} be the event that the desired property fails (i.e., for each set $E \subseteq \mathbf{Q}^{(k)}$ of size $r = n^{4/3}(\log n)^{2/3}$, the partial Latin rectangle $\mathbf{Q}^{(k)} \setminus E$ contains at least $N = (1 + \delta/2)k^6/(4n^4)$ intercalates). Let Z be the number of $\kappa = \lfloor r/4 \rfloor$ -element sequences of disjoint intercalates in $\mathbf{Q}^{(k)}$. If \mathcal{A} holds, then $Z \geq N^\kappa$, since we may choose κ intercalates sequentially with at least N choices each time. On the other hand, there are $2^{\binom{k}{2}^3} = (1 + o(1))k^6/4$ potential intercalates that can appear in $\mathbf{Q}^{(k)}$, and by [Theorem 3.4](#) each κ -element sequence of disjoint intercalates appears in \mathbf{Q} with probability at most $((1 + O(\delta^2))/n^4)^\kappa \leq ((1 + \delta/4)/n^4)^\kappa$ (for small δ). So, $\mathbb{E}Z \leq ((1 + \delta/3)k^6/(4n^4))^\kappa$ by linearity of expectation, and by Markov's inequality it follows that $\Pr(\mathcal{A}) \leq \mathbb{E}Z/N^\kappa \leq \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$. \square

6.2. A combinatorial decomposition. Given [Lemma 6.1](#), we now wish to understand the probability that there is a small set of edges participating in many intercalates. To make this analysis tractable, we need a lemma decomposing any set of edges into well-behaved subsets.

Definition 6.2. A *star* is a hypergraph all of whose hyperedges contain a common vertex. A *matching* is a hypergraph all of whose hyperedges are disjoint.

Lemma 6.3. *For any $r \in \mathbb{N}$, every 3-uniform hypergraph with m hyperedges can be partitioned into a combination of at most m/r stars and at most $3r + m/r$ matchings, each of which have at most r edges.*

Proof. As long as there is a vertex incident to r edges, take r of these edges as a star (we obtain at most m/r stars in this way). After no more deletions are possible, we now have a 3-uniform hypergraph with all degrees less than r . We can greedily find a proper edge-colouring of this hypergraph with at most $3(r - 1) + 1 \leq 3r$ colours. Each of the colour classes is a matching. Finally, arbitrarily decompose the matchings into sub-matchings each with at most r edges, which introduces at most m/r new matchings. \square

6.3. Switching for stars and matchings. By applying [Lemma 6.3](#) to the set \mathbf{E}_0 provided by [Lemma 6.1](#), it now suffices to bound the probability that there is a small star or matching which participates in many intercalates. We will handle both cases separately, with similar switching-based proofs (in random Latin rectangles; afterwards we will use [Theorem 3.3](#) to deduce a result for random Latin squares). Our application of the switching method will be rather simple and completely elementary, but we remark that Fack and McKay [20] and Hasheminezhad and McKay [29] have proved very general theorems with which one can analyse more complicated switching operations.

First, the following lemma will be used to handle stars (note that in the context of Latin rectangles, a star is a set of entries corresponding to a single row, column or symbol).

Lemma 6.4. *Let $k \leq n/10$, and let $\mathbf{Q} \in \mathcal{Q}$ be a uniformly random order- n Latin rectangle with k rows. Let $N(\mathbf{Q})$ be the number of intercalates in \mathbf{Q} which involve the first row. Then $\Pr(N(\mathbf{Q}) \geq t) \leq (k/t)^{\Omega(t)}$ for $t \geq 20k$.*

Proof. Let $\mathcal{Q}(\ell) \subseteq \mathcal{Q}$ be the set of Latin rectangles $Q \in \mathcal{Q}$ for which there are exactly $N(Q) = \ell$ intercalates involving the first row.

Consider the following switching operation: select a row $i \in R^{(k)} \setminus \{1\}$ (i.e., not the first row) and a pair of columns $x, y \in C$, and swap the contents of columns x and y in row i . Note that it is possible that the resulting $k \times n$ array is no longer a Latin rectangle (columns x and y may now contain a repeated symbol). We next compute some upper and lower bounds on the number of ways to switch from a Latin rectangle $Q \in \mathcal{Q}(\ell)$ to a Latin rectangle $Q' \in \mathcal{Q}(\ell')$, for $\ell \neq \ell'$.

In the hypergraph formulation of a Latin rectangle, our switching introduces two new edges and removes two edges. In a Latin rectangle, every edge outside the first row participates in at most one intercalate involving the first row, so it is only possible to switch from $\mathcal{Q}(\ell)$ to $\mathcal{Q}(\ell')$ if $|\ell - \ell'| \leq 2$.

Next, we observe that for any Latin rectangle $Q \in \mathcal{Q}$, there are at most $(k-1)n \leq kn$ switchings which create an intercalate involving the first row. Indeed, first note that swapping entries in columns x and y of a given row can never create an intercalate involving the first row and both x and y . Then, for every column z , we consider the number of switchings that create an intercalate involving z without actually swapping an entry in column z . Such an intercalate must involve one of the $k-1$ rows other than the first, and for each such row i , there is at most one switching that actually creates the desired intercalate (in row i , we must swap the column x satisfying $Q_{i,x} = Q_{1,z}$ with the column y satisfying $Q_{1,y} = Q_{i,z}$).

Now, given a Latin rectangle $Q \in \mathcal{Q}(\ell)$, there are ℓ intercalates involving the first row. Note that each such intercalate is destroyed by at least $2(n-2k)$ switchings which maintain the Latin rectangle property. Indeed, consider one of the two edges of the intercalate not in the first row (in row i , column x and symbol s , say). There are at most k columns which already include s , and at most k columns whose symbol in row i already appears in column x . For any of the (at least $n-2k$) other rows y , we can swap columns x and y in row i to destroy the desired intercalate. Now, the intercalates involving the first row are edge-disjoint outside of this first row, but a given switching could remove two different intercalates at once (or destroy an intercalate by interchanging its entries outside the first row). So, there are a total of at least $\ell(n-2k)$ distinct switchings which maintain the Latin rectangle property and remove an intercalate. We have just observed that at most kn of these switchings also introduce an intercalate, so for any $\ell \geq 2$ we deduce

$$(\ell(n-2k) - kn) |\mathcal{Q}(\ell)| \leq kn (|\mathcal{Q}(\ell-1)| + |\mathcal{Q}(\ell-2)|).$$

This implies $|\mathcal{Q}(\ell)| \leq (4k/\ell) \max\{|\mathcal{Q}(\ell-1)|, |\mathcal{Q}(\ell-2)|\}$ for $\ell \geq 10k$. Iterating this, we see that for $\ell \geq 20k$ we have

$$|\mathcal{Q}(\ell)| \leq \frac{4k}{\ell} \cdot \frac{4k}{\ell-2} \cdots \frac{4k}{\ell-2\lceil(\ell-10k-2)/2\rceil} \cdot \max\{|\mathcal{Q}(10k)|, |\mathcal{Q}(10k-1)|\} \leq \left(\frac{k}{\ell}\right)^{\Omega(\ell)} |\mathcal{Q}|.$$

(To justify the second inequality, note that at least $\ell/4$ terms in the product are at most $8k/\ell$). This implies that $\Pr(N(\mathbf{Q}) = \ell) \leq (k/\ell)^{\Omega(\ell)}$, and the desired result follows by summing over $\ell \geq t$. \square

To handle matchings, we first use a similar switching argument to handle intercalates which are “mostly disjoint” from the vertices of the matching, other than the necessary included edge. Note that in the context of Latin rectangles, a matching is a set of entries such that no pair shares a row, column or symbol. Such a set is also called a *partial transversal*.

Lemma 6.5. *Let $k \leq n/10$ and let $\mathbf{Q} \in \mathcal{Q}$ be a uniformly random order- n Latin rectangle with k rows. Fix a set M of $r = k/6$ disjoint triples in $R^{(k)} \times C \times S$ (which may or may not appear as edges in \mathbf{Q}). Say an intercalate is good if it includes one of the triples in M as an edge, and its other three vertices are completely disjoint from the vertices in M . Let $N(\mathbf{Q})$ be the number of good intercalates in \mathbf{Q} . Then $\Pr(N(\mathbf{Q}) \geq t) \leq (k/t)^{\Omega(t)}$ for $t \geq 20k$.*

Proof. Similarly to the proof of Lemma 6.4, we partition the set of all $k \times n$ Latin rectangles \mathcal{Q} into subsets $\mathcal{Q}(\ell)$ depending on the number ℓ of good intercalates they contain. Without loss of generality we may assume that M involves the first r rows, the first r columns and the first r symbols. We consider the same switching operation as before, but we only consider swaps in rows $i > r$ (i.e., we never switch in the rows where the entries of M live). As in the proof of Lemma 6.4, we need to prove estimates on the number of ways to switch between different $\mathcal{Q}(\ell)$. The arguments will be very similar, so we will be brief with the details.

This time, it is only possible to switch between $\mathcal{Q}(\ell)$ and $\mathcal{Q}(\ell')$ if $|\ell - \ell'| \leq 6$. This is because any given entry outside the first r rows can be involved in at most 3 good intercalates (it must share a row, column or symbol with an edge in M).

The same considerations as before show that for any $Q \in \mathcal{Q}$, there are at most rk switchings which create a good intercalate. Also, if we consider any $Q \in \mathcal{Q}(\ell)$, there are at least $\ell(n - 2k)/3$ switchings which destroy an intercalate in Q (the reason we divide by three is that a single entry can participate in at most three good intercalates). We deduce that, for $\ell \geq 6$,

$$(\ell(n - 2k)/3 - rk) |\mathcal{Q}(\ell)| \leq rk \sum_{r=1}^6 |\mathcal{Q}(\ell - r)|,$$

and we can then iterate this bound to conclude the proof in essentially the same way as [Lemma 6.4](#). \square

Now, a simple combinatorial argument allows us to infer a bound not requiring disjointness.

Lemma 6.6. *There is a constant $C_{6.6} > 0$ such that the following holds. Let $k \leq n/10$ and let $\mathbf{Q} \in \mathcal{Q}$ be a uniformly random order- n Latin rectangle with k rows. Fix a set F of $r \leq k/6$ disjoint triples in $R^{(k)} \times C \times S$ (which may or may not appear as edges in \mathbf{Q}). Let $N(\mathbf{Q})$ be the number of intercalates in \mathbf{Q} which include an edge in F . Then $\Pr(N(\mathbf{Q}) \geq s) \leq (k/s)^{\Omega(s)}$ for $s \geq C_{6.6}k$.*

Proof. We claim that if any Latin rectangle $Q \in \mathcal{Q}$ has at least $s \geq C_{6.6}k$ intercalates involving edges in F , then there is a subset $M \subseteq Q$ such that there are at least $s/16$ intercalates in Q which are good with respect to M (i.e., they involve an edge in M , and the three vertices outside this edge are completely disjoint from M). This suffices to prove the lemma: if H is large enough then [Lemma 6.5](#) and the union bound show that with probability at least $1 - 2^r(16k/s)^{\Omega(s/16)} = 1 - (k/s)^{\Omega(s)}$, our random Latin rectangle \mathbf{R} has the property that there is no subset $M \subseteq Q$ for which there are at least $s/16$ intercalates in \mathbf{R} which are good with respect to M .

To prove the claim, we use the probabilistic method. Consider any Latin rectangle $Q \in \mathcal{Q}$, and let M be a random subset of F obtained by including each element of F independently with probability $1/2$. For each intercalate I involving an edge in $e \in F$, note that I is good with respect to M with probability at least $1/16$. Indeed, note that there are at most 3 edges in $F \setminus \{e\}$ which intersect I . The probability that $e \in M$ and the other intersecting edges are not in M is at least $(1/2) \cdot (1/2)^3 = 1/16$.

By linearity of expectation, the expected number of intercalates which are good with respect to M is at least $s/16$, so there is an outcome of M such that there are at least $s/16$ good intercalates. This completes the proof of the claim. \square

We conclude this subsection by using [Theorem 3.3](#) to deduce from [Lemmas 6.4](#) and [6.6](#) a corresponding result for random Latin squares.

Lemma 6.7. *Let $k \leq n/10$ and $r \leq k/6$. Let $\mathbf{L} \in \mathcal{L}$ be a uniformly random order- n Latin square, and let $\mathbf{L}^{(k)}$ be the subhypergraph induced by $R^{(k)} \cup C^{(k)} \cup S^{(k)}$ (i.e., the first k rows, columns and symbols). Let $K_{k,k,k}^{(3)}$ be the complete 3-uniform 3-partite hypergraph with parts $R^{(k)}, C^{(k)}, S^{(k)}$, and fix a star or matching $F \subseteq K_{k,k,k}^{(3)}$ with r edges. Let \mathbf{N}_F be the number of intercalates in $\mathbf{L}^{(k)}$ which include an edge in F . Then $\Pr(\mathbf{N}_F \geq s) \leq \exp(O(n(\log n)^2))(k/s)^{\Omega(s)}$ for $s \geq C_{6.6}k$, where $C_{6.6}$ is the constant in [Lemma 6.6](#).*

Proof. If F is a matching, we may assume without loss of generality that it involves the first r rows. The desired result then follows from [Lemma 6.6](#) and [Theorem 3.3](#) (recall that by [Theorem 3.3](#), we lose a factor of at most $\exp(O(n(\log n)^2))$ when changing measure from a random $k \times n$ Latin rectangle to the first k rows of a random Latin square).

If F is a star, without loss of generality we may assume that all of its edges are in the first row (recall that there is a symmetry between the rows, columns and symbols of a Latin square). We then apply [Lemma 6.4](#) and [Theorem 3.3](#) in the same way. \square

6.4. Completing the proof. We are now ready to bound the upper tail deviation probability.

Proof of [Theorem 1.2\(b\)](#). We may assume δ is sufficiently small (the desired bound only becomes stronger as we make δ smaller). Let $k = \lfloor \delta^2 n \rfloor$, and let $\mathbf{L}^{(k)}$ be the subhypergraph induced by $R^{(k)} \cup C^{(k)} \cup S^{(k)}$ (i.e., the first k rows, columns and symbols). Let \mathbf{N}_k be the number of intercalates in $\mathbf{L}^{(k)}$; we will prove that $\Pr(\mathbf{N}_k \geq (1 + 2\delta/3)k^6/(4n^4)) \leq \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$. To see that this suffices, note that by symmetry and the union bound, it would follow that $\binom{n}{k}^3 \exp(-\Omega(n^{4/3}(\log n)^{2/3})) = \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$ is an upper bound on the probability that there is *any* choice of k rows, columns and symbols which contains more than $(1 + 2\delta/3)k^6/(4n^4)$ intercalates. But if an order- n Latin square L contains at least $(1 + \delta)n^2/4$ intercalates, then by averaging there is some subset of k rows, k columns and k symbols inducing at least $(1 + 2\delta/3)k^6/(4n^4)$ intercalates.

So, we study intercalates in $\mathbf{L}^{(k)}$. Let $K_{k,k,k}^{(3)}$ be the complete 3-uniform 3-partite hypergraph with parts $R^{(k)}, C^{(k)}, S^{(k)}$, and for a set of edges $E \subseteq K_{k,k,k}^{(3)}$, let \mathbf{N}_E be the number of intercalates in $\mathbf{L}^{(k)}$ involving an edge of E . For every possible outcome of $\mathbf{L}^{(k)}$, let $\mathbf{E}_0 \subseteq \mathbf{L}^{(k)}$ be a subset of $m = n^{4/3}(\log n)^2/3$ edges of $\mathbf{L}^{(k)}$ such that $\mathbf{N}_{\mathbf{E}_0}$ is maximised. By [Lemma 6.1](#), it suffices to show that $\Pr(\mathbf{N}_{\mathbf{E}_0} \geq (\delta/6)k^6/(4n^4)) \leq \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$.

Let $r = \sqrt{m} = n^{1/3}(\log n)^{1/3}$. By [Lemma 6.3](#), we can always partition \mathbf{E}_0 into at most $3r + m/r = 4r$ stars and matchings each with at most r edges. In order to have $\mathbf{N}_{\mathbf{E}_0} \geq (\delta/6)k^6/(4n^4)$, there must be some r -edge star or matching F with

$$\mathbf{N}_F \geq \frac{(\delta/6)k^6/(4n^4)}{4r} = \Omega(n^{4/3}(\log n)^{-1/3}).$$

But this occurs with probability at most $\binom{n}{k}^3 \exp(-\Omega(n^{4/3}(\log n)^{2/3})) = \exp(-\Omega(n^{4/3}(\log n)^{2/3}))$ by [Lemma 6.7](#) and the union bound. \square

REFERENCES

- [1] F. Augeri, *Nonlinear large deviation bounds with applications to Wigner matrices and sparse Erdős-Rényi graphs*, Ann. Probab. **48** (2020), no. 5, 2404–2448.
- [2] L. Babai, *Almost all Steiner triple systems are asymmetric*, Ann. Discrete Math. **7** (1980), 37–39, Topics on Steiner systems.
- [3] B. Barber, S. Glock, D. Kühn, A. Lo, R. Montgomery, and D. Osthus, *Minimalist designs*, Random Structures Algorithms **57** (2020), no. 1, 47–63.
- [4] A. Basak and R. Basu, *Upper tail large deviations of regular subgraph counts in Erdős-Rényi graphs in the full localized regime*, arXiv:1912.11410.
- [5] B. B. Bhattacharya, S. Ganguly, E. Lubetzky, and Y. Zhao, *Upper tails and independence polynomials in random graphs*, Adv. Math. **319** (2017), 313–347.
- [6] S. Bhattacharya and A. Dembo, *Upper tail for homomorphism counts in constrained sparse random graphs*, Random Structures Algorithms (to appear).
- [7] T. Bohman, A. Frieze, and E. Lubetzky, *Random triangle removal*, Adv. Math. **280** (2015), 379–438.
- [8] B. Bollobás, *The chromatic number of random graphs*, Combinatorica **8** (1988), no. 1, 49–55.
- [9] L. M. Brègman, *Certain properties of nonnegative matrices and their permanents*, Dokl. Akad. Nauk SSSR **211** (1973), 27–30.
- [10] J. M. Browning, P. J. Cameron, and I. M. Wanless, *Bounds on the number of small Latin subsquares*, J. Combin. Theory Ser. A **124** (2014), 41–56.
- [11] P. Cameron, *A niggling problem*, Cameron Counts, <https://cameroncounts.wordpress.com/2015/01/24/a-niggling-problem/>, 2015.
- [12] P. J. Cameron, *Almost all quasigroups have rank 2*, Discrete Math. **106/107** (1992), 111–115.
- [13] N. J. Cavenagh, C. Greenhill, and I. M. Wanless, *The cycle structure of two rows in a random Latin square*, Random Structures Algorithms **33** (2008), no. 3, 286–309.
- [14] S. Chatterjee, *The missing log in large deviations for triangle counts*, Random Structures Algorithms **40** (2012), no. 4, 437–451.

- [15] S. Chatterjee, *Large deviations for random graphs*, Lecture Notes in Mathematics, vol. 2197, Springer, Cham, 2017, Lecture notes from the 45th Probability Summer School held in Saint-Flour, June 2015, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School].
- [16] N. Cook and A. Dembo, *Large deviations of subgraph counts for sparse Erdős-Rényi graphs*, Adv. Math. **373** (2020), 107289, 53.
- [17] B. DeMarco and J. Kahn, *Upper tails for triangles*, Random Structures Algorithms **40** (2012), no. 4, 452–459.
- [18] G. P. Egorychev, *The solution of van der Waerden's problem for permanents*, Adv. in Math. **42** (1981), no. 3, 299–305.
- [19] P. Erdős, *Problems and results in combinatorial analysis*, Colloquio Internazionale sulle Teorie Combinatorie (Rome, 1973), Tomo II, 1976, pp. 3–17. Atti dei Convegni Lincei, No. 17.
- [20] V. Fack and B. D. McKay, *A generalized switching method for combinatorial estimation*, Australas. J. Combin. **39** (2007), 141–154.
- [21] D. I. Falikman, *Proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix*, Mat. Zametki **29** (1981), no. 6, 931–938, 957.
- [22] A. Ferber and M. Kwan, *Almost all Steiner triple systems are almost resolvable*, Forum Math. Sigma **8** (2020), Paper No. e39, 24.
- [23] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probability **3** (1975), 100–118.
- [24] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption*, Mem. Amer. Math. Soc., to appear.
- [25] C. D. Godsil and B. D. McKay, *Asymptotic enumeration of Latin rectangles*, J. Combin. Theory Ser. B **48** (1990), no. 1, 19–44.
- [26] B. Gunby, *Upper tails of subgraph counts in sparse regular graphs*, arXiv:2010.00658.
- [27] R. Häggkvist and J. C. M. Janssen, *All-even Latin squares*, Proceedings of the 6th Conference on Formal Power Series and Algebraic Combinatorics (New Brunswick, NJ, 1994), vol. 157, 1996, pp. 199–206.
- [28] M. Harel, F. Mousset, and W. Samotij, *Upper tails via high moments and entropic stability*, arXiv:1904.08212.
- [29] M. Hasheminezhad and B. D. McKay, *Combinatorial estimates by the switching method*, Combinatorics and graphs, Contemp. Math., vol. 531, Amer. Math. Soc., Providence, RI, 2010, pp. 209–221.
- [30] M. T. Jacobson and P. Matthews, *Generating uniformly distributed random Latin squares*, J. Combin. Des. **4** (1996), no. 6, 405–437.
- [31] S. Janson and A. Ruciński, *The infamous upper tail*, Random Structures Algorithms **20** (2002), no. 3, 317–342, Probabilistic methods in combinatorial optimization.
- [32] S. Janson and A. Ruciński, *The deletion method for upper tail estimates*, Combinatorica **24** (2004), no. 4, 615–640.
- [33] A. D. Keedwell and J. Dénes, *Latin squares and their applications*, second ed., Elsevier/North-Holland, Amsterdam, 2015, With a foreword to the previous edition by Paul Erdős.
- [34] P. Keevash, *The existence of designs*, arXiv:1401.3665.
- [35] P. Keevash, *The existence of designs II*, arXiv:1802.05900.
- [36] P. Keevash, *Counting designs*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 4, 903–927.
- [37] P. Keevash, *Hypergraph matchings and designs*, Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 3113–3135.
- [38] A. Kotzig, C. C. Lindner, and A. Rosa, *Latin squares with no subsquares of order two and disjoint Steiner triple systems*, Utilitas Math. **7** (1975), 287–294.
- [39] A. Kotzig and J. Turgeon, *On certain constructions for Latin squares with no Latin subsquares of order two*, Discrete Math. **16** (1976), no. 3, 263–270.
- [40] M. Kwan, *Almost all Steiner triple systems have perfect matchings*, Proc. Lond. Math. Soc. (3) **121** (2020), no. 6, 1468–1495.
- [41] M. Kwan and B. Sudakov, *Intercalates and discrepancy in random Latin squares*, Random Structures Algorithms **52** (2018), no. 2, 181–196.
- [42] N. Linial and Z. Luria, *An upper bound on the number of high-dimensional permutations*, Combinatorica **34** (2014), no. 4, 471–486.
- [43] N. Linial and Z. Luria, *Discrepancy of high-dimensional permutations*, Discrete Anal. (2016), Paper No. 11, 8.
- [44] N. Linial and M. Simkin, *Monotone subsequences in high-dimensional permutations*, Combin. Probab. Comput. **27** (2018), no. 1, 69–83.
- [45] B. D. McKay and I. M. Wanless, *Most Latin squares have many subsquares*, J. Combin. Theory Ser. A **86** (1999), no. 2, 322–347.
- [46] B. D. McKay and I. M. Wanless, *On the number of Latin squares*, Ann. Comb. **9** (2005), no. 3, 335–344.
- [47] M. McLeish, *On the existence of Latin squares with no subsquares of order two*, Utilitas Math. **8** (1975), 41–53.
- [48] A. O. Pittenger, *Mappings of Latin squares*, Linear Algebra Appl. **261** (1997), 251–268.

- [49] V. Rödl and A. Ruciński, *Threshold functions for Ramsey properties*, J. Amer. Math. Soc. **8** (1995), no. 4, 917–942.
- [50] M. Rudelson and R. Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math. **218** (2008), no. 2, 600–633.
- [51] M. Simkin, *Methods for analyzing random designs*, IAS special day on combinatorial design theory, http://math.huji.ac.il/~michaels/files/papers/pasch_in_sts.pdf, 2018.
- [52] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, second ed., Cambridge University Press, Cambridge, 2001.
- [53] G. H. J. van Rees, *Subsquares and transversals in Latin squares*, Ars Combin. **29** (1990), no. B, 193–204, Twelfth British Combinatorial Conference (Norwich, 1989).
- [54] I. M. Wanless, *Transversals in Latin squares: a survey*, Surveys in combinatorics 2011, London Math. Soc. Lecture Note Ser., vol. 392, Cambridge Univ. Press, Cambridge, 2011, pp. 403–437.

DEPARTMENT OF MATHEMATICS, STANFORD UNIVERSITY, STANFORD, CA
Email address: `mattkwan@stanford.edu`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA
Email address: `{asah,msawhney}@mit.edu`