PARITIES IN RANDOM LATIN SQUARES

MATTHEW KWAN, KALINA PETROVA, AND MEHTAAB SAWHNEY

ABSTRACT. In a Latin square, every row can be interpreted as a permutation, and therefore has a parity (even or odd). We prove that in a uniformly random $n \times n$ Latin square, the n row parities are very well approximated by a sequence of n independent unbiased coin flips: for example, the total variation error of this approximation tends to zero as $n \to \infty$. This resolves a conjecture of Cameron. In fact, we prove a generalisation of Cameron's conjecture for the *joint* distribution of the row parities, column parities and symbol parities (the latter are defined by the symmetry between rows, columns and symbols of a Latin square).

Along the way, we introduce several general techniques for the study of random Latin squares, including a new re-randomisation technique via "stable intercalate switchings", and a new approximation theorem comparing random Latin squares with a certain independent model.

1. INTRODUCTION

An $n \times n$ *Latin square* is an $n \times n$ array filled with n different "symbols" (usually taken to be the integers $1, \ldots, n$), with the property that each symbol appears exactly once in each row and each column. For example, the multiplication table of a group is always a Latin square; in general, Latin squares can be interpreted as multiplication tables of a class of algebraic structures called *quasigroups*. See for example [35] for an introduction to this vast subject.

Each row or column of an $n \times n$ Latin square L can be interpreted as a permutation of order n, which can be either even or odd. Let $N_{\text{row}}(L)$ be the number of odd row permutations, and let $N_{\text{col}}(L)$ be the number of odd column permutations. If L is the multiplication table of a group, then either $N_{\text{row}}(L) = N_{\text{col}}(L) = 0$ or $N_{\text{row}}(L) = N_{\text{col}}(L) = n/2$. However, for general Latin squares, the row and column parities can have much richer behaviour, and much is still unknown. For example, one of the most important conjectures in this direction is the *Alon–Tarsi conjecture*, which (in probabilistic language) says that if n is even, and $\mathbf{L}$ is a uniformly random $n \times n$ Latin square, then

$$\mathbb{P}\big[N_{\text{row}}(\mathbf{L}) \text{ is even}\big] = \mathbb{P}\big[N_{\text{row}}(\mathbf{L}) + N_{\text{col}}(\mathbf{L}) \text{ is even}\big] \neq \frac{1}{2}.$$

(The first equality is not part of the original conjecture; it was observed by Huang and Rota [30], in a paper where they also observed that the Alon–Tarsi conjecture has a number of surprising consequences in seemingly unrelated areas of mathematics; see [26] for a modern survey). It was first proved by Alpöge [3] that

$$\mathbb{P}\big[N_{\text{row}}(\mathbf{L}) \text{ is even}\big] = \frac{1}{2} + o(1). \tag{1.1}$$

as $n \to \infty$. In other words, if the Alon–Tarsi conjecture is true, then it is true "just barely".

Going far beyond (1.1), it has been suggested by Peter Cameron (in a variety of different sources; see for example [7, 10, 11, 12, 13, 14, 15]) that the row parities of a Latin square might be statistically completely unconstrained, in the sense that one can model the n row parities of a random Latin square by simply making n independent coin flips.

Conjecture 1.1. *Let $\mathbf{L}$ be a uniformly random $n \times n$ Latin square. Then the distribution of $N_{\text{row}}(\mathbf{L})$ is approximately the binomial distribution $\text{Bin}(n, 1/2)$, as $n \to \infty$.*

(Note that exchanging rows does not affect the distribution of $\mathbf{L}$, so the sequence of row parities $\vec{\xi}_{\text{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ has a permutation-invariant distribution. This means that if we condition on $N_{\text{row}}(\mathbf{L})$, then $\vec{\xi}_{\text{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ is a uniformly random sequence in $(\mathbb{Z}/2\mathbb{Z})^n$, constrained to have exactly $N_{\text{row}}(\mathbf{L})$ "1"s. That is to say, in order to understand the distribution of $\vec{\xi}_{\text{row}}(\mathbf{L})$, it is enough to understand the distribution of $N_{\text{row}}(\mathbf{L})$.)

To elaborate on the attribution/history of Conjecture 1.1: the starting point seems to have been the problem session at the *British Combinatorial Conference* in 1993, where Cameron asked a related question that was extended by Jeannette Janssen. However, at that time they did not seem to be very confident that the statement of Conjecture 1.1 was actually true (in the BCC problem list [10] they phrase the question as "is it true that..."). Cameron later posed the problem more assertively in a 2002 survey on permutations and permutation groups [7], and it seems he first referred to it as a "conjecture" in a 2003 lecture on random Latin squares [11]. However, as far as we can tell, he never stated the problem in a fully precise form (always using language like "approximately").

There are many different ways to compare distributions, to make rigorous sense of the word "approximately"; in this paper we confirm Conjecture 1.1 in many different senses. For example, one strong way to compare distributions is in terms of *total variation distance*: for two probability distributions μ, ν on the same space, write $\mathrm{d}_{\mathrm{TV}}(\mu, \nu) = \sup_A |\mu(A) - \nu(A)|$ (so for any event A, the probabilities of A with respect to μ and ν differ by at most $\mathrm{d}_{\mathrm{TV}}(\mu, \nu)$). The following theorem is a consequence of our main (technical) result.

Theorem 1.2. *Let* $\mathbf{L}$ *be a uniformly random* $n \times n$ *Latin square. Then*

$$\lim_{n \to \infty} \mathrm{d}_{\mathrm{TV}}\big(N_{\mathrm{row}}(\mathbf{L}), \mathrm{Bin}(n, 1/2)\big) = 0.$$

Equivalently, writing $\vec{\xi}_{\mathrm{row}}(\mathbf{L}) \in (\mathbb{Z}/2\mathbb{Z})^n$ *for the sequence of parities of rows of* $\mathbf{L}$*, and writing* $\mathrm{Unif}((\mathbb{Z}/2\mathbb{Z})^n)$ *for the uniform distribution on* $(\mathbb{Z}/2\mathbb{Z})^n$*, we have*

$$\lim_{n \to \infty} \mathrm{d}_{\mathrm{TV}}\big(\vec{\xi}_{\mathrm{row}}(\mathbf{L}), \mathrm{Unif}((\mathbb{Z}/2\mathbb{Z})^n)\big) = 0.$$

We can actually prove much more than Theorem 1.2; we will discuss this momentarily, but first it is worthwhile to briefly discuss why this conjecture took so long to be resolved (and, in our opinion, why it is so interesting). Generally speaking, it is quite easy to make plausible predictions about uniformly random Latin squares, by making various kinds of approximate independence assumptions (for example, Cameron's conjecture can be justified from the point of view that there is "no obvious reason" for the parities of different rows to be correlated). However, it is surprisingly difficult to rigorously prove anything nontrivial about uniformly random Latin squares, or even to study them empirically.

The main issues are that Latin squares do not enjoy any neat recursive structure, and they are very "rigid" objects, in the sense that there are only very limited ways to make a "local perturbation" to change a Latin square into another one. To highlight the difficulties here, we remark that (despite some very ambitious conjectures; see [1, Section 4.1]) we still have a rather poor understanding of the total number of $n \times n$ Latin squares (the best known upper and lower bounds differ by exponential factors; see [56, Section 17]), and there is no rigorously justified way to efficiently sample a uniformly random Latin square (there are certain ergodic Markov chains on the space of $n \times n$ Latin squares [31, 51], but these are not known to be rapidly mixing). By now, there are quite a few known results about random Latin squares; the proofs of many of these results have required fundamental new additions to a very limited toolbox of techniques, and Theorem 1.2 is no exception (we discuss existing and new techniques in Section 2).

1.1. **Row, column and symbol permutations.** Our proof techniques actually allow us to go beyond Cameron's conjecture, to approximate not just the distribution of $N_{\mathrm{row}}(\mathbf{L})$ but also the *joint* distribution between $N_{\mathrm{row}}(\mathbf{L})$, $N_{\mathrm{col}}(\mathbf{L})$ and a third parameter $N_{\mathrm{sym}}(\mathbf{L})$, which counts the number of odd *symbol permutations*. Perhaps the most natural way to define a symbol permutation is to reinterpret a Latin square in a way that emphasises the natural symmetry between rows, columns and symbols: indeed, observe that an $n \times n$ Latin square (with symbols $1, \ldots, n$) can be interpreted as an $n \times n \times n$ array in which every entry is "0" or "1", such that every two-dimensional "slice" of this array is a permutation matrix[1]. So, the row and column permutations correspond to slices in two of the three possible directions, and the symbol permutations correspond to slices in the third direction.

The row, column and symbol parities cannot vary completely freely: it was proved by Janssen [34] and Zappa [58] that for any Latin square L of order n, we have

$$N_{\mathrm{row}}(L) + N_{\mathrm{col}}(L) + N_{\mathrm{sym}}(L) = f(n) \pmod 2,$$

---

[1] The correspondence is that if the Latin square has symbol s in the (i, j)-entry, then we put a "1" in the (i, j, s)-entry of the corresponding $n \times n \times n$ array.

where

$$f(n) = \begin{cases} 0 & \text{if } n = 0 \text{ or } 1 \pmod 4 \\ 1 & \text{if } n = 2 \text{ or } 3 \pmod 4. \end{cases}$$

Subject to this constraint, we are able to prove the natural extension of Cameron's conjecture to row, column and symbol parities. Specifically, let μ^* be the conditional distribution of three independent $\text{Bin}(n, 1/2)$ random variables, given that their sum is $f(n)$ mod 2. We are able to prove that $(N_{\text{row}}(\mathbf{L}), N_{\text{col}}(\mathbf{L}), N_{\text{sym}}(\mathbf{L}))$ approximately has the distribution μ^* as $n \to \infty$, in many different senses.

Theorem 1.3. *Let $\mathbf{L}$ be a uniformly random $n \times n$ Latin square, and let*

$$\vec{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) = (N_{\text{row}}(\mathbf{L}), N_{\text{col}}(\mathbf{L}), N_{\text{sym}}(\mathbf{L})).$$

Then, the following hold as $n \to \infty$.

(1) *(Law of large numbers) We have the convergence in probability*

$$\frac{1}{n}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) \xrightarrow{p} (1/2, 1/2, 1/2),$$

i.e., $\vec{\mathbf{X}}$ satisfies the same law of large numbers as μ^*.

(2) *(Central limit theorem) We have the convergence in distribution*

$$\frac{1}{\sqrt{n/4}}(\mathbf{X}_1 - n/2, \mathbf{X}_2 - n/2, \mathbf{X}_3 - n/2) \xrightarrow{d} \mathcal{N}(0, I_3),$$

i.e., $\vec{\mathbf{X}}$ satisfies the same trivariate central limit theorem as μ^*.

(3) *(Local central limit theorem) For all $\vec{x} = (x_1, x_2, x_3)$ satisfying $x_1 + x_2 + x_3 = f(n)$ (mod 2), we have*

$$\mathbb{P}[\vec{\mathbf{X}} = \vec{x}] = 2 \cdot \frac{1}{(2\pi(n/4))^{3/2}} \exp\left(-\frac{(x_1 - n/2)^2 + (x_2 - n/2)^2 + (x_3 - n/2)^2}{2(n/4)}\right) + o(n^{-3/2}),$$

i.e., $\vec{\mathbf{X}}$ satisfies the same local *central limit theorem as μ^*.*

(4) *(Total variation convergence) We have*

$$d_{\text{TV}}(\vec{\mathbf{X}}, \mu^*) \to 0.$$

(5) *(Large deviation principle) Let $H_2 : \alpha \mapsto -\alpha \log_2 \alpha - (1 - \alpha) \log_2 (1 - \alpha)$ be the base-2 binary entropy function, and let $I(x_1, x_2, x_3) = 3 - H_2(x_1) - H_2(x_2) - H_2(x_3)$. We have*

$$-\inf_{\vec{x} \in E^\circ} I(\vec{x}) \leq \liminf n^{-1} \log_2 \mathbb{P}[n^{-1}\vec{\mathbf{X}} \in E] \leq \limsup n^{-1} \log_2 \mathbb{P}[n^{-1}\vec{\mathbf{X}} \in E] \leq -\inf_{\vec{x} \in \overline{E}} I(\vec{x})$$

for all Borel $E \subseteq \mathbb{R}^3$, where $\overline{E}$ and E° denote the closure and interior of E. That is to say, $\vec{\mathbf{X}}$ satisfies the same large deviation principle as μ^*.

Actually, all parts of Theorem 1.3 are really consequences of a single master theorem (Theorem 6.4), but we feel the statement of that master theorem is a little too technical for this introduction.

1.2. **Previous work.** We end this introduction with a brief discussion of previous work on Cameron's conjecture. The first theorem in this area was due to Häggkvist and Janssen [29], who proved that

$$\mathbb{P}[N_{\text{row}}(\mathbf{L}) = 0] \leq (\sqrt{3/4} + o(1))^n \approx 0.87^n \tag{1.2}$$

(note that Theorem 1.3(5) implies that[2] $\mathbb{P}[N_{\text{row}}(\mathbf{L}) = 0] \leq (1/2 + o(1))^n)$. Second, generalising (1.1), Cavenagh and Wanless [18] proved that $(N_{\text{row}}(\mathbf{L}), N_{\text{col}}(\mathbf{L}))$ is asymptotically equidistributed modulo 2 (this is a special case of Theorem 1.3(4)). Third, Cavenagh, Greenhill and Wanless [17] proved that the probability that the first two rows of $\mathbf{L}$ have the same parity lies between $1/4 - o(1)$ and $3/4 + o(1)$ (note that Theorem 1.3(1) implies that this probability is $1/2 + o(1)$, using the invariance of the distribution of $\mathbf{L}$ under permutations of the rows). Until now, these were the only rigorous pieces of evidence towards Cameron's conjecture.

Surprisingly, empirical evidence for small Latin squares has run *counter* to Cameron's conjecture (and therefore Theorem 1.3): it seems that the approximations in Theorem 1.3 only start to "kick in" for reasonably large n (though, there is a limit to what can be done empirically, because randomly

---

[2]Curiously, the methods in this paper do not directly imply the corresponding *lower* bound $\mathbb{P}[N_{\text{row}}(\mathbf{L}) = 0] \geq (1/2 + o(1))^n$. However, this does seem to be provable with some additional tricks; we intend to explore this in future work with Catherine Greenhill and Lenka Kopfová.

sampling Latin squares is a notoriously difficult problem). Cameron himself [15] observed that "evidence supports this conjecture fairly well, but the tails of the distribution seem a little heavier than it would predict", suggesting that there might be good reason that the Häggkvist–Janssen bound mentioned above is significantly larger than 2^{-n}. Alimohammadi, Diaconis, Roghani and Saberi [1] attempted to test Conjecture 1.1 using *sequential importance sampling* (in the same paper, they gave rigorous grounding to this method). They were not able to find any evidence of the total variation convergence[3] $d_{\mathrm{TV}}(N_{\mathrm{row}}(\mathbf{L}), \mathrm{Bin}(n, 1/2)) \to 0$, with random Latin squares of size $n \leq 15$.

1.3. **Notation.** We use standard asymptotic notation throughout: for functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ or $g = \Omega(f)$ to mean that there is a constant C such that $|f(n)| \leq C|g(n)|$ for sufficiently large n, we write $f = \Theta(g)$ to mean that $f = O(g)$ and $f = \Omega(g)$, and we write $f = o(g)$ or $g = \omega(f)$ to mean that $f/g \to 0$ as $n \to \infty$. We will often want to treat certain quantities as constants, for the purpose of asymptotic notation; we will always make this clear with phrasing like "fix a constant $\alpha > 0$".

Somewhat less standardly, for $\varepsilon > 0$ we write $f \pm \varepsilon$ to denote a quantity that differs from f by at most ε.

Regarding basic mathematical notation: 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 generally omit floor and ceiling symbols and assume large numbers are integers, wherever divisibility considerations are not important. We write $[n] = \{1, \ldots, n\}$, and all logarithms are base e unless explicitly stated otherwise.

Finally, since this paper features so many objects of different types, we adopt some typographical conventions: random objects are printed in bold (e.g., a random variable $\mathbf{X}$ or a random Latin square $\mathbf{L}$), and "ordered" objects are decorated with an arrow (e.g., vectors $\vec{x}$, and later in the paper we will consider ordered partial Latin squares $\vec{P}$).

1.4. **Acknowledgements.** The authors would like to thank Lenka Kopfová for helpful feedback on an earlier draft of this paper.

2. Discussion of proof techniques

In this section we give a high-level overview of the techniques which have previously been brought to bear on random Latin squares, and with this context we describe the new ideas that go into the proof of Theorem 6.4. (We end the section with an outline of the structure of the rest of the paper.)

As far as we know, all the rigorous work on random Latin squares uses one of three general classes of techniques.

(1) **Permutation-invariance.** The distribution of a uniformly random $n \times n$ Latin square is invariant under random permutation of the rows, columns, and symbols. So, one can attempt to study random Latin squares purely by considering the effect of such random permutations. This type of reasoning is rather limited, but does have some applications (see e.g. [9, 47]).

(2) **Enumeration.** The most obvious way to study the probability that a random Latin square satisfies a property $\mathcal{P}$ is to simply count the number of Latin squares satisfying $\mathcal{P}$, and divide by the total number of Latin squares. Unfortunately, we only have very crude methods to enumerate $n \times n$ Latin squares (the best known upper and lower bounds differ by a factor of about $\exp(n \log^2 n)$), so generally speaking this can only be used to study properties that are extremely unlikely (i.e., that occur with probability less than about $\exp(-n \log^2 n)$). Even for such properties, actually proving the necessary estimates is often a highly nontrivial matter (and often involves consideration of auxiliary random models, as we discuss in Section 2.1). For some examples of properties of random Latin squares proved by enumeration, see e.g. [5, 8, 28, 40, 42, 43, 44, 46, 50].

(3) **Cycle switching.** A common technique in probabilistic combinatorics is to study a random object by studying the effect of local "switching" operations. Unfortunately, it is very difficult to make a controlled local change to a Latin square. To give some idea of the difficulty, it is an open question to understand the minimum number $f(n)$ such that every $n \times n$ Latin square can be transformed into some different Latin square by changing $f(n)$ entries (all we know is that $f(n)$ has order of magnitude between $\log n$ and $\sqrt{n}$; see [16]). There is only really one type of switching operation that has successfully been used to study random Latin squares, namely *cycle*

---

[3]In fact, they did not even see any evidence of convergence with respect to the so-called *Wasserstein metric* (Wasserstein convergence is significantly weaker than total variation convergence).

switchings (and minor variants thereof). We will define cycle switchings properly in Section 2.3, but for now we just emphasise that they are very unwieldy; in general, if we want to make a particular local change using a cycle switching, we may be forced to make far-reaching changes to the rest of the Latin square. Nonetheless, since cycle switchings are one of very few available tools, they have played a crucial role in many of the known results about random Latin squares (see e.g. [2, 17, 18, 29, 44]).

Permutation-invariance alone is certainly not enough to prove Cameron's conjecture, since any permutation of the rows, columns and symbols affects all the row parities in the same way (though, permutation-invariance can be used to easily show that $\mathbb{P}[N_{\text{row}}(\mathbf{L})$ is even$] = 1/2$ when $n \geq 3$ is odd). At first sight, enumeration-based approaches also seem to be fundamentally unsuitable to prove Cameron's conjecture, since we are concerned with events that occur with non-negligible probability, and a multiplicative error of $\exp(n \log^2 n)$ would completely overwhelm the main term. So, cycle-switching techniques seem to be the most promising avenue towards Cameron's conjecture; indeed, such techniques underpin the results of Häggkvist–Janssen, Cavenagh–Wanless and Cavenagh–Greenhill–Wanless mentioned in the introduction. However, the usual type of switching analysis is far too crude to have any hope of proving anything like Theorem 1.2 (we discuss this further in Sections 2.3 and 2.4), and some fundamental new ideas are required.

In this paper, we use a *combination* of cycle-switching and enumeration-based techniques. To very briefly summarise our approach: we use enumeration-based techniques to prove a new approximation theorem, comparing a random Latin square to a certain Erdős–Rényi-type random hypergraph (in a stronger sense than previously available). This allows us to show that random Latin squares are extremely likely to have a very rich constellation of 2×2 subsquares with certain strong disjointness and canonicity/stability properties. The special properties of these 2×2 subsquares allow us to *independently* perform many cycle switchings, to "re-randomise" a random Latin square without biasing its distribution. We are then able to prove the desired results using the randomness of our independent switchings, via some linear algebra over the finite field $\mathbb{F}_2$.

In the following subsections, we describe the above ideas in more detail.

2.1. **Enumeration and approximation.** Let $|\mathcal{L}_n|$ be the number of $n \times n$ Latin squares. This number is known to be
$$|\mathcal{L}_n| = \left(\frac{n}{e^2} + o(n) \right)^{n^2},$$
and there are two known ways to prove this.

(1) We can build up an $n \times n$ Latin square in a *row-by-row* fashion: at each step we choose an option for the next row that does not conflict with previous rows. One can use celebrated *permanent estimates* of of Egorychev–Falikman [21, 22] and Bregman [6] to show that the number of choices for each row does not depend too strongly on previous choices; multiplying these bounds over all steps yields
$$\exp(-O(n \log n)) \leq \frac{|\mathcal{L}_n|}{(n/e^2)^{n^2}} \leq \exp(O(n \log^2 n))$$
(see [56, Theorem 17.3] for the details).

(2) We can instead build up an $n \times n$ Latin square in an *entry-by-entry* fashion: at each step we choose a row/column/symbol triple that does not conflict with the partial Latin square constructed so far, and add it. Unfortunately, the number of choices at each step can depend quite dramatically on previous choices; in particular it is possible to "get stuck", and find oneself in a position where there are no legal choices for the next row/column/symbol triple.

However, it turns out that "on average" (i.e., if one makes random choices at each step), this entry-by-entry process is quite well-behaved. Combining a number of very powerful tools, one can use this process to prove
$$\exp(-O(n^{2-\varepsilon})) \leq \frac{|\mathcal{L}_n|}{(n/e^2)^{n^2}} \leq \exp(O(n^{3/2})).$$
for some tiny[4] constant $\varepsilon > 0$. The details appear in [41], but are really an adaptation of a general enumeration approach systematised by Keevash [39]. Specifically, the upper bound is proved using the so-called *entropy method* of Radhakrishnan [52], in a form pioneered by

---

[4]It is unclear what is the optimal ε with this approach, but certainly ε cannot be greater than $1/2$ (this is a general barrier for techniques based on the *triangle removal process*, which we intend to explore further in upcoming work).

Linial and Luria [45]. For the lower bound, one combines an analysis of an instance of the so-called *triangle removal process* (first studied by Spencer [55] and Rödl and Thoma [53]), with a *completion theorem* of Keevash [37] (here one needs a "second generation" completion theorem, building on Keevash's earlier ideas in the setting of the *existence of designs conjecture* [36]; see also [19, 27, 38]).

The approaches in (1) and (2) suggest two different ways to study random Latin squares, by comparing them to two different auxiliary random models. First, a $k \times n$ *Latin rectangle* is a $k \times n$ array filled with n different symbols, such that each symbol appears exactly once in each row and at most once in each column. The ideas in (1) show that all $k \times n$ Latin rectangles can be completed to Latin squares in "about the same number of ways" (up to a multiplicative factor[5] of about $\exp(n \log^2 n)$), so if we could show that some property is overwhelmingly likely to hold in a random $k \times n$ Latin rectangle, then we could deduce that that property is also very likely to hold in a random $n \times n$ Latin square.

Similarly, consider the random process in (2) (which can be interpreted as an instance of the so-called *triangle removal process*), and suppose we run this process for just a few steps (until the partial Latin square is half-full, say). The ideas in (2) allow us to show that most outcomes of this random process can be completed to Latin squares in "about the same number of ways" (up to a multiplicative factor of about $\exp(n^{2-\varepsilon})$), so if we had some way to show that the first few steps of this random process are overwhelmingly likely to satisfy some property, then we could deduce that that property is also very likely to hold in a random $n \times n$ Latin square.

Both of these two methods have quite fundamental quantitative limitations (they can only be used to study properties that hold with overwhelming probability), but they have seen many applications. To compare the two methods: the first method has much better quantitative aspects, and it is "more elementary" (not depending on Keevash's sophisticated completion machinery). On the other hand, the second method is generally much easier to apply: the entry-by-entry random process can be straightforwardly coupled with an Erdős–Rényi random hypergraph, which allows one to take advantage of the huge body of techniques that have been developed to study random graphs[6]. The only known way to directly study random Latin rectangles is via delicate switching arguments (taking advantage of the fact that it is much easier to make local changes to a Latin rectangle than a Latin square).

2.2. **A new approximation lemma.** One of our key contributions in this paper is a new approximation lemma (Lemma 4.7) that combines the advantages of both the above methods (and is proved by a delicate combination of both methods). Roughly speaking, this lemma says that if one can show that the entry-by-entry random process described in the last subsection (i.e., the *triangle removal process*) satisfies a property with probability $1 - \varepsilon$, then that property holds in a random subset of a random Latin square with probability at least $1 - \varepsilon \exp(O(n \log^2 n))$. The reader may also be interested in an easy-to-apply corollary of our main approximation lemma (Corollary 5.5) comparing random Latin squares to Erdős–Rényi random hypergraphs.

For example, these tools could have been used to remove the switching analysis (or remove the dependence on general-purpose switching estimates) from [5,28,42,44,50], and remove the dependence on Keevash's completion machinery from theorems in [40,42]. For the present paper, our new approximation lemma is rather crucial: quantitative aspects are very important (see Remark 2.1), and the properties we need to consider are so complicated that it would have been a herculean task to study them via switching on Latin rectangles. We hope that our approximation lemma will also be broadly useful in future work on random Latin squares.

2.3. **Cycle switchings.** A cycle switching is specified by a pair of rows, a pair of columns or a pair of symbols, together with an entry of the Latin square "belonging to that pair". For example, consider two rows r_1, r_2 and an entry (r_1, c, s) (meaning that symbol s appears in row r_1 of column c). The first step of the cycle switching is to "move our entry into the other row": we put the symbol s into cell (r_2, c). However, there was already another symbol s' in that cell, which we need to move somewhere else; we move it to the only available cell, which is (r_1, c). But there is already an instance of s' in row r_1 (in some column c'), so we move that instance of s' to the cell (r_2, c'). But there was already some symbol s'' in that cell, which we move to (r_1, c'). At this point, if we are very lucky, we will have $s = s''$, in which case no further moves are necessary, and we have successfully switched to a new Latin square. If

---

[5]Of course, this multiplicative factor depends on k, but it turns out that most of the contribution comes from the last few rows, so one cannot gain much by taking k very close to n.

[6]The second method is also much more robust, and can be applied in much more general settings than random Latin squares, though this is not relevant for the present paper.

not, we can keep going, repeatedly moving entries between r_1 and r_2 as long as it takes to resolve all conflicts and reach a Latin square. (In the worst case, we could end up exchanging the entirety of rows r_1 and r_2).

The smallest possible cycle switching is called an *intercalate switching*. In the above example, it corresponds to the case where $s = s''$, so the entire switching affects just four entries of the Latin square. Alternatively, an intercalate switching can be viewed as "flipping a 2×2 Latin subsquare". In general, the switchings described above ("row switchings", where we are switching between two rows r_1 and r_2) can be viewed as "flipping a $2 \times k$ Latin subrectangle", for some k.

Cycle switchings have an easy-to-describe impact on the parities of rows, columns and symbols. For example, a row switching changes the parity of every column and symbol involved in the switching, and it may or may not change the parity of the two involved rows, depending on whether the length of the cycle switching is even or odd. More subtly, cycle switchings also have an impact on the structure of other cycle switchings (e.g., switching in a pair of rows can affect the switchings that are possible in some pair of columns), so one can consider combinations of cycle switchings with more complicated effects. It is even possible to combine multiple "partial" cycle switchings of different types (cf. the "cross-switch" operations in [17,29]).

One can use these types of switchings to estimate the *relative likelihoods* of various events. Indeed, given two sets of Latin squares $\mathcal{A}, \mathcal{B}$, we can design a switching operation to transform a Latin square in $\mathcal{A}$ into a Latin square in $\mathcal{B}$. Then, for $L_1 \in \mathcal{A}$ we can try to estimate the number of Latin squares in $\mathcal{B}$ that can be reached by such a switching, and for $L_2 \in \mathcal{B}$ we can try to estimate the number of Latin squares in $\mathcal{A}$ that can reach L_2 by such a switching. Dividing these two estimates gives us an estimate on $|\mathcal{B}|/|\mathcal{A}|$.

For example, for a sequence $\vec{x} \in \{0,1\}^k$, let $\mathcal{L}_{\vec{x}}$ be the set of $n \times n$ Latin squares for which the parities of the first k rows are described by $\vec{x}$. Using the above types of ideas, Häggkvist and Janssen [29] managed to show that if k is significantly smaller than $n/2$, we have $|\mathcal{L}_{\vec{x}}| \le 3|\mathcal{L}_{\vec{y}}|$ whenever $\vec{x}$ and $\vec{y}$ differ in a single coordinate, and they iterated this estimate to deduce (1.2). Obviously, this comes very far short of proving Cameron's conjecture; the condition on k and the extraneous factor of 3 are basically due to the fact that one has very little control over the structure of cycle switchings in an arbitrary Latin square. To make a desired change to the row parities it is necessary to make case distinctions with different combinations of cycle switchings that interact in different ways, and it is necessary to reserve a large subset of "junk rows" which may be affected by the cycle switchings in unpredictable ways.

2.4. **Individual intercalate switchings.** An important point is that we do not actually need to consider cycle switchings in *arbitrary* Latin squares: we can use the enumeration/approximation methods described in Sections 2.1 and 2.2 to prove that Latin squares are very likely to satisfy certain properties, and then take advantage of these properties in cycle switching arguments. This sounds like a natural approach, but to our knowledge it has never been employed before, because until recently the properties that could be proved using enumeration/approximation were extremely limited.

In particular, a very relevant direction is the study of intercalates (2×2 Latin subsquares) in random Latin squares. If an intercalate involves two even (respectively, odd) rows, then switching that intercalate reduces (respectively, increases) the number of even rows by exactly two. So, if we had very tight control of the number of intercalates among the even rows, and the number of intercalates among the odd rows, we could hope to estimate the ratios

$$\frac{\mathbb{P}[N_{\mathrm{row}}(\mathbf{L}) = x]}{\mathbb{P}[N_{\mathrm{row}}(\mathbf{L}) = x + 2]}, \tag{2.1}$$

for all $x \in \mathbb{N}$, by considering all the ways to switch a single intercalate. Note that these ratios fully determine the distribution of $N_{\mathrm{row}}(\mathbf{L})$ (except for a possible bias mod 2, but this is handled by (1.1)).

Recently, resolving a conjecture of McKay and Wanless [50], Kwan, Sah and Sawhney [42] showed how to use enumeration/approximation techniques (together with techniques from large deviations theory) to prove that $\mathbf{L}$ has $n^2/4 + o(n^2)$ intercalates with probability $1 - o(1)$. These techniques can be adapted to show that $\mathbf{L}$ typically has $k^2/4 + o(n^2)$ intercalates in any subset of k rows, which allows one to estimate the ratio in (2.1) up to a multiplicative factor of $1 + o(n/(x(n-x)))$. This is already enough to make new progress towards Cameron's conjecture: since binomial tails decay very rapidly, it is straightforward to deduce that $N_{\mathrm{row}}(\mathbf{L}) = n/2 + o(n)$ with probability $1 - o(1)$; that is, $N_{\mathrm{row}}(\mathbf{L})$ satisfies the same law of large numbers as $\mathrm{Bin}(n, 1/2)$. This implies Theorem 1.3(1), and with a bit more work it is also possible to prove the large deviation principle in Theorem 1.3(5) using this type of idea.

That is to say, by studying the effect of switching an intercalate, one can estimate the ratios (2.1) tightly enough to study the tails of $N_{\mathrm{row}}(\mathbf{L})$ (and $N_{\mathrm{col}}(\mathbf{L})$ and $N_{\mathrm{sym}}(\mathbf{L})$). Unfortunately, it is not feasible to study the bulk of the distribution of $N_{\mathrm{row}}(\mathbf{L})$ this way: since the fluctuations of $\mathrm{Bin}(n, 1/2)$ have order of magnitude $\sqrt{n}$, one would need to control the ratios (2.1) up to a factor of $1 + o(1/\sqrt{n})$; roughly speaking, this is comparable to showing that the number of intercalates $\mathbf{Y}$ in $\mathbf{L}$ satisfies $\mathbb{P}[|\mathbf{Y} - \mathbb{E}\mathbf{Y}| > t] = o(1)$ for some $t = o(n^{3/2})$. While we believe this to be true, it is beyond the reach of enumeration-based techniques: due to clustering phenomena in the upper tail of $\mathbf{Y}$, the relevant deviation probabilities are simply not small enough[7] to tolerate the super-exponential error terms described in Section 2.1. (This is an instance of Janson's "infamous upper tail" [33], and is discussed further in [42].)

2.5. **Multiple intercalate switchings.**

So, instead of studying the effect of a single intercalate switching, our approach is to make *many disjoint intercalate switches at the same time*.

A naïve approach for this is as follows. Using our new approximation theorem, together with techniques from [40,42], one can show that a random $n \times n$ Latin square $\mathbf{L}$ typically has a collection of disjoint intercalates $\mathcal{I}$ which "robustly span almost all the rows", in the sense that for every set R_0 of about[8] $\log^2 n$ rows, there is an intercalate in $\mathcal{I}$ which involves a row in R_0 and a row outside R_0. Some simple linear algebra over $\mathbb{F}_2$ then shows that if we were to randomly and independently switch all the intercalates in $\mathcal{I}$ (thereby obtaining a new random Latin square $\mathbf{L}'$), the resulting distribution of the n row parities would be uniform on some affine-linear subspace of $\mathbb{F}_2^n$ with codimension at most about $\log^2 n$, and it is easy to deduce a central limit theorem for $N_{\mathrm{row}}(\mathbf{L}')$. (With a more sophisticated argument, one can handle the joint distribution of $(N_{\mathrm{row}}(\mathbf{L}'), N_{\mathrm{col}}(\mathbf{L}'), N_{\mathrm{sym}}(\mathbf{L}'))$ and upgrade the central limit theorem to a *local* central limit theorem.)

Remark 2.1. We are really proving a central limit theorem for $N_{\mathrm{row}}(\mathbf{L}')$ conditioned on an outcome of $\mathbf{L}$. So, it is important that for different outcomes of $\mathbf{L}$, the corresponding conditional distributions of $N_{\mathrm{row}}(\mathbf{L}')$ are statistically indistinguishable. This would not be true if the codimension above were greater than $\sqrt{n}$, since the fluctuations of $\mathrm{Bin}(n, 1/2)$ have order of magnitude $\sqrt{n}$. So, the quantitative aspects of our new approximation theorem are very important here. (E.g., the approximation theorem in [40] would not have sufficed.)

Unfortunately, this only characterises the distribution of $N_{\mathrm{row}}(\mathbf{L}')$, which could in principle be very different to the distribution of $N_{\mathrm{row}}(\mathbf{L})$. The problem is that switching some intercalates can destroy other intercalates, or create new ones, so the probability of switching from L to L' could be different from the probability of switching from L' to L (i.e., the switching could "push probability mass" towards a subset of Latin squares, introducing bias to the distribution of $\mathbf{L}'$). Trying to quantify this bias takes us back to the issues described in Section 2.4.

The only way we were able to resolve this issue was to execute the entire argument above with a *stable/canonical* collection of disjoint intercalates. Specifically, we associate a collection of disjoint intercalates $\mathcal{I}(L)$ to each Latin square L, in such a way that if we switch any subset of these intercalates to obtain a new Latin square L', then $\mathcal{I}(L)$ and $\mathcal{I}(L')$ are exactly the same (except that obviously some of these intercalates are switched). So, the probability of switching from L to L', and the probability of switching from L' to L, are both exactly $2^{-|\mathcal{I}(L)|} = 2^{-|\mathcal{I}(L')|}$, and $\mathbf{L}'$ is a uniformly random Latin square.

In a typical Latin square, the intercalates heavily intersect each other (and therefore cannot be independently switched). Given a pair of intersecting intercalates, how do we decide which is "the canonical one" that we're allowed to switch? Our approach is to *randomly sparsify* the situation, to reduce the intersections between intercalates. Namely, we consider a random "template" $\mathbf{T}$, which describes a sparse subset of "switchable" row/column pairs. Then, given a Latin square L, we define a collection of intercalates $\mathcal{I}_{\mathbf{T}}(L)$ by starting with the collection of intercalates in L which use only row/column pairs in $\mathbf{T}$, then deleting the intercalates which are "non-canonical" (the precise definition is a little complicated, but e.g. we delete all intercalates in intersecting pairs, and we delete all intercalates that could introduce a new intercalate when switched). The idea is that this deletion step should not have a very severe effect, since $\mathbf{T}$ is sparse. We need to show that a typical outcome of $\mathbf{T}$ yields collections of intercalates $\mathcal{I}_{\mathbf{T}}(L)$ which "robustly span almost all the rows" for almost all L, with which we can implement the argument described above.

---

[7]For this to work, one would need to be able to estimate the number of $n \times n$ Latin squares $|\mathcal{L}_n|$ up to a subexponential $\exp(o(n))$ multiplicative error.

[8]We cannot go smaller than about $\log^2 n$, due to the multiplicative error of about $\exp(n \log^2 n)$ in our enumeration estimates.

Actually proving the necessary properties of $\mathcal{I}_\mathbf{T}(\mathbf{L})$ is a very delicate matter (much more so than any previous work on random Latin squares), largely due to the aforementioned "infamous upper tail" issue. Specifically, because enumerative estimates have super-exponential error terms, most steps of the proof need to hold with overwhelming probability, so we constantly need to worry about large deviation behaviour. Whenever we need an upper bound on some quantity, we need to be extremely careful to avoid situations where clustering phenomena cause upper tails to be too heavy.

2.6. **Outline of the paper.** The structure of the proof of Theorem 1.3, as distributed over the rest of the paper, is as follows. First (after Section 3, in which we record some standard concentration inequalities that will be used throughout the paper), in Section 4 we prove our new approximation theorem (Lemma 4.7) relating random Latin squares to the triangle removal process. In Section 5 we prove some simple lemmas relating the triangle removal process to Erdős–Rényi random hypergraphs, which will be useful in combination with Lemma 4.7.

In Section 6, we reduce all parts of Theorem 1.3 to a technical "master theorem" (Theorem 6.4) giving a precise description of the distribution of $(N_{\mathrm{row}}(\mathbf{L}), N_{\mathrm{col}}(\mathbf{L}), N_{\mathrm{sym}}(\mathbf{L}))$. Then, in Section 7, we introduce the notion of a "stable intercalate", which is an intercalate which can be safely switched without affecting any other intercalates. In Section 8 we show how to use linear-algebraic arguments with stable intercalate switchings to prove the master theorem described above, given a key lemma (Lemma 8.3) on stable intercalates in random Latin squares. Roughly speaking, the key lemma says that there is a "template" T (obtained via a sparse random set of row/column pairs), such that for a uniformly random Latin square, it is very likely that for any large-enough sets of rows, columns and symbols, we can find a stable intercalate (involving only entries in T) "inside" those sets.

The rest of the paper is devoted to the proof of Lemma 8.3. In Section 9 we break down this proof into two lemmas. First, Lemma 9.11 says that (inside any large-enough sets of rows, columns and symbols) there are likely to be many disjoint intercalates (saying nothing about whether they are stable). Second, Lemma 9.13 says that there are unlikely to be many entries which lie in certain "bad" arrangements of intercalates which could lead to our intercalates being non-stable. (Here, due to "infamous upper tail" issues, we cannot hope to consider the number of bad arrangements themselves, only the number of entries in them!)

The first of these lemmas (Lemma 9.11) is proved in Section 10, via a careful 2-step application of Freedman's martingale concentration inequality (using a "maximum disjoint family" technique of Bollobás [4]).

The second lemma (Lemma 9.13) is proved in Section 11, and is much more involved. Here, our job is to understand "bad" arrangements of intercalates; there is an enormous range of possibilities for the structure of such an arrangement, so the first step is a switching argument (unrelated to the main switching argument in Section 8), which shows that every bad arrangement of intercalates can be switched to obtain one of four "basic types", and therefore it suffices to restrict our attention to these types. We then carefully study how the four basic types of bad intercalate arrangements can emerge in the triangle removal process.

3. Concentration inequalities

Throughout this paper we will frequently use a general-purpose concentration inequality for functions of independent Bernoulli random variables. The following statement appears as [57, Corollary 6][9]; it follows from the martingale approach of Freedman [25].

Theorem 3.1. *Let $\vec{\mathbf{g}} = (\mathbf{g}_1, \ldots, \mathbf{g}_N)$ be a sequence of independent Bernoulli random variables with $\mathbb{P}[\mathbf{g}_i = 1] = p_i$. Let $f : \{0,1\}^N \to \mathbb{R}$ satisfy $|f(\vec{g}) - f(\vec{g}')| \le K_i$ for all pairs $\vec{g}, \vec{g}' \in \{0,1\}^N$ differing only in the i-th coordinate, and suppose $K_i \le K$ for all $i \in [N]$. Then for all $t \ge 0$, we have*

$$\mathbb{P}\Big[\big|f(\vec{\mathbf{g}}) - \mathbb{E}[f(\vec{\mathbf{g}})]\big| > t\Big] \le 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^N p_i K_i^2 + 2Kt/3}\right).$$

Note that, in order to obtain a strong bound from Theorem 3.1, we need to know that $f(\vec{g})$ is not very sensitive to changes in any individual g_i. In this respect Theorem 3.1 is similar to the much more well-known *bounded differences inequality* of McDiarmid [49] (which is a consequence of the Azuma–Hoeffding martingale concentration inequality). However, we need the stronger statement of Theorem 3.1 for its essentially-optimal dependence on $p_1, \ldots, p_N$.

---

[9]In [57], Theorem 3.1 is stated only for the upper tail of $f(\vec{\mathbf{g}})$; applying it to $-f(\vec{\mathbf{g}})$ yields the lower tail.

We also record a corollary of [Theorem 3.1](#) in the special case of weighted sums of Bernoulli random variables (this is basically the standard Chernoff bound).

Corollary 3.2. *Let $\mathbf{g}_1, \ldots, \mathbf{g}_N$ be independent Bernoulli random variables, consider weights $w_1, \ldots, w_N \in [0, \Delta]$, and let $\mathbf{X} = w_1\mathbf{g}_1 + \cdots + w_N\mathbf{g}_N$. Then, for any $\delta > 0$ we have*

$$\mathbb{P}\big[|\mathbf{X} - \mathbb{E}\mathbf{X}| \geq \delta\mathbb{E}\mathbf{X}\big] \leq 2\exp\left(-\frac{\delta^2\mathbb{E}\mathbf{X}}{(2 + 2\delta/3)\Delta}\right).$$

4. Approximating random Latin squares with the triangle removal process

In this section, we prove an approximation lemma that relates random Latin squares to the *triangle removal process* (which is a tractable probability distribution on *partial* Latin squares). To state this lemma, we need some preparations. First, it will be convenient for us to take an alternative perspective on Latin squares.

Fact 4.1. An $n \times n$ Latin square can be interpreted as a tripartite 3-uniform hypergraph (specifically, a subgraph of the complete tripartite 3-uniform hypergraph $K_{n,n,n}^{(3)}$), in which every pair of vertices in different parts features in exactly one hyperedge.

To see the correspondence, think of the three parts as the set of rows, the set of columns and the set of symbols; a hyperedge (r, c, s) means that symbol s appears in the c-th column of the r-th row.

From now on, we interchangeably take the hypergraph point of view and the $n \times n$ array point of view, depending on whichever is most convenient at any given moment (generally speaking, the hypergraph point of view is more expressive, and highlights the symmetry between rows, columns and symbols, but it is easier to draw pictures with the $n \times n$ array point of view).

Definition 4.2. Let $\mathcal{L}_n$ be the set of all $n \times n$ Latin squares.

A *partial* Latin square is a tripartite 3-uniform hypergraph in which every pair of vertices in different parts features in *at most* one hyperedge (alternatively, one can also interpret this as an $n \times n$ array in which only some of the entries are filled with symbols). Let $\mathcal{P}_{n,m}$ be the set of all $n \times n$ partial Latin squares with m hyperedges, and let $\mathcal{P}_n = \bigcup_{m \leq n^2} \mathcal{P}_{n,m} \supseteq \mathcal{L}_n$ be the set of all $n \times n$ partial Latin squares (with any number of hyperedges).

An *ordered* partial Latin square is a partial Latin square together with an ordering on its hyperedges. Let $\vec{\mathcal{P}}_{n,m}$ be the set of all $n \times n$ ordered partial Latin squares with m hyperedges, and let $\vec{\mathcal{P}}_n = \bigcup_{m \leq n^2} \vec{\mathcal{P}}_{n,m}$.

Now, our approximation lemma will compare uniformly random Latin squares to a distribution on (ordered) partial Latin squares; for this to make sense we need a notion of inheritance from a property of Latin squares to a property of (ordered) partial Latin squares.

Definition 4.3. Let $\vec{\mathcal{U}} \subseteq \vec{\mathcal{P}}_n$ be a property of ordered partial Latin squares and let $\mathcal{T} \subseteq \mathcal{L}_n$ be a property of Latin squares. For $m \leq n^2$, we say $\vec{\mathcal{U}}$ is (ρ, m)-*inherited* from $\mathcal{T}$ if for any $L \in \mathcal{T}$, taking $\vec{\mathbf{P}}_m(L)$ to be a uniformly random subset of m hyperedges of L, equipped with a uniformly random order, we have $\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}$ with probability at least ρ.

Remark 4.4. It will often be quite obvious how to specify the property $\vec{\mathcal{U}}$, for a desired event $\mathcal{T}$ (and often, the ordering will play no role). For example, if $\mathcal{T}$ is the property of having fewer than $(1 - \varepsilon)n^2/4$ intercalates, it would make sense to take $\vec{\mathcal{U}}$ to be the property of having fewer than $(1 - \varepsilon/2)(m/n^2)^4 n^2/4$ intercalates. However, as we will see later in the paper, there is sometimes reason to make more exotic choices for $\vec{\mathcal{U}}$.

Remark 4.5. Throughout the paper we will need to consider many different properties of (ordered) (partial) Latin squares, and we have attempted to adopt some notational conventions to help the reader keep track of the "type" of each property. Specifically, for properties of partial Latin squares, we will usually use the letter $\mathcal{U}$ (think "unfinished") and for properties of complete Latin squares we will usually use the letter $\mathcal{T}$ (think "total"). An arrow on top of a property indicates that it is a property of ordered (rather than unordered) partial Latin squares.

Now, the main lemma of this section concerns the *triangle removal process*, which we now define.

Definition 4.6. Let $\text{TRP}(n, m)$ be the distribution over $\vec{\mathcal{P}}_{n,m} \cup \{\perp\}$ obtained as follows. Start with the complete tripartite graph $K_{n,n,n}$. At each step, choose a uniformly random triangle (among all the triangles in the current graph), and remove it. After m steps of this process, the sequence of removed triangles can be interpreted as an ordered partial Latin square $\vec{R} \in \vec{\mathcal{P}}_{n,m}$ (unless we ran out of triangles at some point in the process, in which case we take $\vec{R} = \perp$).

We are now ready to state the main result of this section, comparing a uniformly random Latin square to the triangle removal process.

Lemma 4.7. *Fix constants $\alpha \in (0, 1)$ and $\rho \in (0, 1]$, and let $m = \alpha n^2$. Consider properties $\vec{\mathcal{U}} \subseteq \vec{\mathcal{P}}_{n,m}$ and $\mathcal{T} \subseteq \mathcal{L}_n$ such that $\vec{\mathcal{U}}$ is (ρ, m)-inherited from $\mathcal{T}$. Let $\vec{R} \sim \text{TRP}(n, m)$ be an ordered partial Latin square obtained by m steps of the triangle removal process, and let $L \sim \text{Unif}(\mathcal{L}_n)$ be a uniformly random $n \times n$ Latin square. Then*

$$\mathbb{P}[L \in \mathcal{T}] \leq \exp\left(\frac{n \log^2 n}{1 - \alpha - o(1)}\right) \mathbb{P}[\vec{R} \in \vec{\mathcal{U}}].$$

Lemma 4.7 is a quantitative improvement to [42, Theorem 2.4] (which is itself an adaptation of [40, Theorem 2.4]). Our proof of Lemma 4.7 bears some similarities to the proofs in [40, 42], but incorporates various additional ideas (related to our use of enumeration results that are stronger but much less robust).

Remark 4.8. It may seem that Lemma 4.7 is not suitable for studying "local" properties (e.g., properties concerning a particular vertex v). Such properties are important for the so-called *absorption method*, which (is not relevant for the main results in this paper but) has played a central role in previous work on random Latin squares. Indeed, it is not hard to see that with probability at least $\exp(-O(n))$, the random hypergraph $\vec{R}$ has no edges at all containing v, and this probability seems too large to have any hope of applying Lemma 4.7 to any interesting local property about v. However, one can generally overcome this issue with a judicious choice of $\vec{\mathcal{U}}$, cf. Remark 4.4. For example (very informally speaking), if $\mathcal{T}$ is the property that v does not participate in a desired local structure, we can take $\vec{\mathcal{U}}$ to be the property that v has high degree but still does not participate in the desired local structure.

The rest of this section is devoted to the proof of Lemma 4.7. We start with some further definitions.

Definition 4.9. For an (ordered) partial $n \times n$ Latin square P, let $G(P) \subseteq K_{n,n,n}$ be the graph which contains an edge uv (of $K_{n,n,n}$) if and only if there is no hyperedge of P containing u and v. That is to say, $G(P)$ is the graph of pairs that would need to be covered to complete P to a full Latin square.

For a graph $G \subseteq K_{n,n,n}$, let $\text{dens}(G) = e(G)/(3n^2)$.

Definition 4.10. We say that a graph $G \subseteq K_{n,n,n}$, is ε-*triangle-typical* if its number of triangles is $(1 \pm \varepsilon)n^3 \text{dens}(G)^3$. Let $\mathcal{P}_{n,m}^{\triangle:\varepsilon} \subseteq \mathcal{P}_{n,m}$ be the set of partial Latin squares $P \in \mathcal{P}_{n,m}$ such that $G(P)$ is ε-triangle-typical. Let $\vec{\mathcal{P}}_{n,m}^{\triangle:\varepsilon} \subseteq \vec{\mathcal{P}}_{n,m}$ be the set of ordered partial Latin squares $\vec{P} \in \vec{\mathcal{P}}_{n,m}$ such that for each $i \leq m$, if we consider the partial Latin square P_i consisting of the first i hyperedges of $\vec{P}$, then $P_i \in \mathcal{P}_{n,m}^{\triangle:\varepsilon}$.

Definition 4.11. We say that a graph $G \subseteq K_{n,n,n}$, is γ-*quasirandom* if for all pairs of vertices u, v in different parts, the number of common neighbours of u and v is $(1 \pm \gamma)n \text{dens}(G)^2$. Let $\mathcal{P}_{n,m}^{\gamma}$ be the set of partial Latin squares $P \in \mathcal{P}_{n,m}$ such that $G(P)$ is γ-quasirandom.

Remark 4.12. It is not hard to see that if a graph is γ-quasirandom, then it is necessarily $O(\gamma)$-triangle-typical (this fact is used in the previous approximation lemmas in [40, 42]). However, in our proof of Lemma 4.7 (with its strong quantitative aspects), we will need to consider quasirandomness and triangle-typicality with quite different parameters.

In the next two lemmas, we show that for *any* Latin square $L \in \mathcal{L}_n$, if we take a random ordering of a random subset of L, then the resulting random object is likely to satisfy certain triangle-typicality and quasirandomness properties. This will allow us to assume that various graphs we encounter later in the proof are triangle-typical/quasirandom, which will be necessary for our enumeration techniques.

Lemma 4.13. *Fix a constant $\alpha \in (0, 1)$. Consider any $L \in \mathcal{L}_n$ and let $\vec{P}_{\alpha n^2}(L) \in \vec{\mathcal{P}}_{n,\alpha n^2}$ be a random ordering of a random set of αn^2 hyperedges of L. Then, with $\varepsilon = n^{-1} \log n$, we have*

$$\mathbb{P}[\vec{P}_{\alpha n^2}(L) \notin \vec{\mathcal{P}}_{n,\alpha n^2}^{\triangle:\varepsilon}] \leq \exp(-\Omega(\log^2 n)).$$

Proof. Fix some $m \leq \alpha n^2$. We will show that $G(\vec{\mathbf{P}}_m(L))$ is ε-triangle-typical with probability at least $1 - \exp(-\Omega(\log^2 n))$. The statement of the lemma will then follow by a union bound over all $m \leq \alpha n^2$.

Note that $\vec{\mathbf{P}}_m(L)$ consists of a uniformly random set of m hyperedges of L. Let $\mathbf{B}_p(L)$ be a random subset of the hyperedges of L where each hyperedge is taken independently at random with probability $p = m/n^2$. By "Pittel's inequality" (see e.g. page 17 in [32]), for any property $\mathcal{Q}$ of unordered partial Latin squares,
$$\mathbb{P}[\vec{\mathbf{P}}_m(L) \notin \mathcal{Q}] \leq 3\sqrt{m}\,\mathbb{P}[\mathbf{B}_p(L) \notin \mathcal{Q}].$$
Thus, letting $\mathbf{X}$ be the number of triangles in $G(\vec{\mathbf{P}}_m(L))$, it suffices to show that $\mathbb{P}[\mathbf{X} \notin (1\pm\varepsilon)n^3(1-p)^3] \leq \exp(-\Omega(\log^2 n))$.

For any $v_1 \in V_1, v_2 \in V_2, v_3 \in V_3$ such that $v_1 v_2 v_3 \in L$, the probability that $v_1 v_2 v_3$ is a triangle in $G(\mathbf{B}_p(L))$ is precisely $1 - p = 1 - m/n^2$. On the other hand, if $v_1 v_2 v_3 \notin L$, the probability $v_1 v_2 v_3$ is a triangle in $G(\mathbf{B}_p(L))$ is $(1-p)^3$, since each pair of vertices among v_1, v_2, v_3 is in a distinct hyperedge of L, and all these hyperedges need to be absent from $\mathbf{B}_p(L)$ in order for $v_1 v_2 v_3$ to be in $G(\mathbf{B}_p(L))$. Thus,
$$\mu := \mathbb{E}[\mathbf{X}] = n^2(1-p) + (n^3 - n^2)(1-p)^3 = \Theta(n^3).$$
There are n^2 hyperedges in L, and adding or removing a hyperedge from $\mathbf{B}_p(L)$ can affect $\mathbf{X}$ by at most $1 + 3(n-1)$, since each pair of vertices in e can form a triangle with $n-1$ other vertices in $G(\mathbf{B}_p(L))$. Thus, by Theorem 3.1, we have
$$\mathbb{P}[\mathbf{X} \notin (1 \pm \varepsilon)n^3(1-p)^3] \leq \mathbb{P}[|\mathbf{X} - \mu| \geq \varepsilon\mu/2]$$
$$\leq 2\exp\left(-\frac{\varepsilon^2\mu^2/4}{2n^2p\big(1+3(n-1)\big)^2 + 2\big(1+3(n-1)\big)\varepsilon\mu/6}\right)$$
$$= \exp\big(-\Omega(\log^2 n)\big),$$
where in the first inequality we used that $\mu - n^3(1-p)^3 = o\big(\varepsilon n^3(1-p)^3\big)$. $\qquad\square$

Lemma 4.14. *Fix a constant $\alpha \in (0,1)$. Consider any $L \in \mathcal{L}_n$ and let $\mathbf{P}_{\alpha n^2}(L) \in \mathcal{P}_{n,\alpha n^2}$ be a random subset of αn^2 hyperedges of L. Let $\gamma = \omega(n^{-1/2}\log^{1/2} n)$. Then*
$$\mathbb{P}[\mathbf{P}_{\alpha n^2}(L) \notin \mathcal{P}^\gamma_{n,\alpha n^2}] \leq \exp(-\Omega(\gamma^2 n)).$$

Proof. Let $m := \alpha n^2$. We will show that $G(\mathbf{P}_m(L))$ is γ-quasirandom with probability at least $1 - \exp(-\Omega(\gamma^2 n))$.

As in the proof of Lemma 4.13, let $\mathbf{B}_p(L)$ be a random subset of the hyperedges of L where each hyperedge is taken independently at random with probability $p = m/n^2 = \alpha$. Let $\mathbf{S}_{u,v}$ be the set of common neighbours of u and v, with respect to the graph $G(\mathbf{B}_p(L))$. It suffices to show that for every u, v that are in different parts of $G(\mathbf{B}_p(L))$, it holds that $\mathbb{P}[|\mathbf{S}_{u,v}| \notin (1 \pm \gamma)n(1-\alpha)^2] \leq \exp(-\Omega(\gamma^2 n))$. A union bound over all pairs of vertices u, v and Pittel's inequality would then complete the proof.

Fix some u, v in different parts of $G(\mathbf{B}_p(L))$. Let w be the unique vertex such that uvw is a hyperedge of L. Then $\mathbb{P}[w \in \mathbf{S}_{u,v}] = 1 - p$. On the other hand, for all $w' \neq w$ that are in the same part as w, we have $\mathbb{P}[w' \in \mathbf{S}_{u,v}] = (1-p)^2$ (since w' participates in some hyperedge together with u and in another hyperedge together with v, and both of these need to not be selected in $\mathbf{B}_p(L)$ to have $w' \in \mathbf{S}_{u,v}$). Thus, for $\mathbf{X} = |\mathbf{S}_{u,v}|$,
$$\mu := \mathbb{E}[\mathbf{X}] = (1-p) + (n-1)(1-p)^2 = \Theta(n).$$
There are $2n - 1$ hyperedges in L that affect $\mathbf{X}$, and adding or removing such a hyperedge from $\mathbf{B}_p(L)$ can affect $\mathbf{X}$ by at most 1. Thus, by Theorem 3.1, we have
$$\mathbb{P}[\mathbf{X} \notin (1 \pm \gamma)n(1-p)^2] \leq \mathbb{P}[|\mathbf{X} - \mu| \geq \gamma\mu/2] \leq \exp\left(-\frac{\gamma^2\mu^2/4}{2(2n-1) + \gamma\mu/3}\right) = \exp(-\Omega(\gamma^2 n)),$$
where in the first inequality we used that $\mu - n(1-p)^2 = (1-p)p = o(\gamma\mu)$. $\qquad\square$

We next prove an upper bound on the number of completions of a partial Latin square, under triangle-typicality and quasirandomness assumptions.

Theorem 4.15. *Fix a constant $\alpha \in (0,1)$, let $\gamma = o(1)$ and let $\varepsilon = n^{-1}\log n$. For any partial Latin square $P \in \mathcal{P}^\gamma_{n,\alpha n^2} \cap \mathcal{P}^{\triangle:\varepsilon}_{n,\alpha n^2}$, the number $|\mathcal{L}_n(P)|$ of completions of P to an $n \times n$ Latin square satisfies*
$$|\mathcal{L}_n(P)| \leq \left(\frac{(1-\alpha)^2 n}{e^2}\right)^{(1-\alpha)n^2} \cdot \exp\left(\frac{n\log^2 n}{1 - \alpha - o(1)}\right).$$

Proof. For a row x (respectively, column y), we write r_x (respectively, c_y) for the number of empty cells in that row, in P. Then for an empty cell $e = (x, y)$, we write $Q_{x,y}$ for the set of symbols that do not appear in the row or column of e. Note that $Q_{x,y}$ is precisely the set of common neighbours of x and y in $G(P)$. Then, let $\mathbf{L}(P)$ be a uniformly random completion of P. We will estimate the entropy $H[\mathbf{L}(P)] = \log|\mathcal{L}_n(P)|$ of $\mathbf{L}(P)$.

Let $E \subseteq [n]^2$ be the set of empty cells in P. For each $e = (x, y) \in E$, let $\mathbf{z}_e$ be the symbol in cell e in $\mathbf{L}(P)$. So the sequence $(\mathbf{z}_e)_{e \in E}$ determines $\mathbf{L}(P)$. For any total ordering $\prec$ on E, we have

$$H[\mathbf{L}(P)] = \sum_{e \in E} H\big[\mathbf{z}_e \,\big|\, (\mathbf{z}_{e'} : e' \prec e)\big]. \tag{4.1}$$

Now, consider a pair of sequences $\mu = (\mu_x)_x, \nu = (\nu_y)_y \in [0,1]^n$ (of real numbers in the range $[0,1]$, which we assume to be all distinct from each other). Let $\lambda = (\lambda_{x,y})_{x,y}$ be the array defined by $\lambda_{x,y} = (\mu_x, \nu_y)$. We can use λ to define a total order $\prec_\lambda$ on E, via reverse lexicographic order on the pairs $\lambda_{x,y}$. To be precise, write $(x', y') \prec_\lambda (x, y)$ when $\mu_{x'} > \mu_x$ or when $x' = x$ and $\nu_{y'} > \nu_y$. Let $\mathbf{R}_e(\lambda)$ be an upper bound on the conditional support size

$$\Big|\operatorname{supp}\big(\mathbf{z}_e \,\big|\, (\mathbf{z}_{e'} : e' \prec_\lambda e)\big)\Big|$$

defined as follows. $\mathbf{R}_e(\lambda)$ for $e = (x, y)$ is 1 plus the number of symbols $z \in Q_e \setminus \mathbf{z}_e$ for which $\mu_{x_z} < \mu_x$ and $\nu_{y_z} < \nu_y$, where x_z and y_z are the row and column such that cells (x_z, y) and (x, y_z) contain symbol z in $\mathbf{L}(P)$.

Since $\mathbf{R}_e(\lambda)$ is an upper bound on $|\operatorname{supp}(\mathbf{z}_e \,|\, (\mathbf{z}_{e'} : e' \prec_\lambda e))|$, we have

$$H\big[\mathbf{z}_e \,\big|\, (\mathbf{z}_{e'} : e' \prec_\lambda e)\big] \leq \mathbb{E}[\log \mathbf{R}_e(\lambda)]. \tag{4.2}$$

It follows from (4.1) and (4.2) that

$$H[\mathbf{L}(P)] \leq \sum_{e \in E} \mathbb{E}[\log \mathbf{R}_e(\lambda)].$$

This is true for any fixed λ, so it is also true if λ is chosen randomly: let $\boldsymbol{\lambda} \in \left([0,1]^2\right)^{n^2}$ be obtained via $2n$ independent random variables $\boldsymbol{\mu}_x, \boldsymbol{\nu}_y$ each uniformly distributed in $[0,1]$ (with probability 1 they are all distinct). Then

$$H[\mathbf{L}(P)] \leq \sum_{e \in E} \mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda})].$$

Next, for any completion L of P, any $e = (x, y)$ and any $\lambda_e = (\mu_x, \nu_y) \in [0,1]^2$, let

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

(Note that $\boldsymbol{\lambda}_e = \lambda_e$ occurs with probability zero, so formally we should condition on $\boldsymbol{\lambda}_e = \lambda_e \pm d\lambda_e$ and take limits in what follows, but there are no continuity issues so we will ignore this detail). By the definition of $\mathbf{R}_e(\boldsymbol{\lambda})$ and linearity of expectation, we have

$$R_e^{L, \lambda_e} = 1 + (|Q_e| - 1)\mu_x \nu_y.$$

By Jensen's inequality,

$$\mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda}) \,|\, \mathbf{L}(P) = L, \boldsymbol{\lambda}_e = \lambda_e] \leq \log R_e^{L, \lambda_e},$$

so

$$\mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda}) \,|\, \mathbf{L}(P) = L] \leq \mathbb{E}\Big[\log R_e^{L, \boldsymbol{\lambda}_e}\Big]$$

$$= \int_{[0,1]^2} \log\big(1 + (|Q_e| - 1)\mu_x \nu_y\big)\, d\lambda_e.$$

For $C > 0$ we compute

$$\int_0^1 \int_0^1 \log(1 + Cts)\, dt\, ds = \log(1 + C) - 2 + \frac{\log(1 + C) - \operatorname{Li}_2(-C)}{C},$$

where

$$\operatorname{Li}_2(C) = -\int_0^C \frac{\log(1 - t)}{t}\, dt$$

is an evaluation of the polylogarithm function, so, as $C \to \infty$,

$$-\operatorname{Li}_2(-C) = \int_0^{-C} \frac{\log(1 - t)}{t}\, dt = \int_0^C \frac{\log(1 + t)}{t}\, dt$$

$$\leq \int_0^1 \frac{\log(1+t)}{t}\,\mathrm{d}t + \sum_{t=1}^{C-1} \frac{\log(2+t)}{t} \leq \log^2(C+1) + O(1).$$

Therefore, substituting $C = |Q_e| - 1$, we have

$$\mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda}) \mid \mathbf{L}(P) = L] \leq \log|Q_e| - 2 + \frac{\log|Q_e| + \log^2|Q_e| + O(1)}{|Q_e|} = \log|Q_e| - 2 + \frac{(1+o(1))\log^2|Q_e|}{|Q_e|}.$$

Now recall that $P \in \mathcal{P}_{n,\alpha n^2}^\gamma$, and that $|Q_e| = |Q_{x,y}|$ is precisely the number of common neighbours of x and y in $G(P)$. Since $G(P)$ is γ-quasirandom, we have $|Q_e| = (1\pm\gamma)n(1-\alpha)^2 = n(1-\alpha-o(1))^2$. Thus, using $|Q_e| \leq n$, and noting that the expression above does not depend on the choice of L, we get

$$\mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda})] \leq \log|Q_e| - 2 + \frac{\log^2 n}{(1-\alpha-o(1))^2 n},$$

Next, since $P \in \mathcal{P}_{n,\alpha n^2}^{\triangle:\varepsilon}$ with $\varepsilon := n^{-1}\log n$, we have that the number of triangles in $G(P)$ is

$$\sum_{e \in E} |Q_e| = (1\pm\varepsilon)n^3(1-\alpha)^3.$$

Since the natural logarithm is a concave function, Jensen's inequality yields

$$\sum_{e \in E} \log|Q_e| \leq |E|\log\left(\frac{\sum_{e\in E}|Q_e|}{|E|}\right) \leq |E|\log\left((1+\varepsilon)n(1-\alpha)^2\right).$$

We now get

$$\log|\mathcal{L}_n(P)| = H[\mathbf{L}(P)] \leq \sum_{e\in E} \mathbb{E}[\log \mathbf{R}_e(\boldsymbol{\lambda})] \leq \sum_{e\in E} \log|Q_e| + |E|\left(\frac{\log^2 n}{(1-\alpha-o(1))^2 n} - 2\right)$$

$$\leq n^2(1-\alpha)\left(\log\left((1+\varepsilon)(1-\alpha)^2 n\right) - 2\right) + \frac{n\log^2 n}{1-\alpha-o(1)}.$$

Thus,

$$|\mathcal{L}_n(P)| \leq \left(\frac{(1+\varepsilon)(1-\alpha)^2 n}{e^2}\right)^{(1-\alpha)n^2} \cdot \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)$$

$$\leq \left(\frac{(1-\alpha)^2 n}{e^2}\right)^{(1-\alpha)n^2} \cdot \exp\left(\varepsilon(1-\alpha)n^2 + \frac{n\log^2 n}{1-\alpha-o(1)}\right)$$

$$= \left(\frac{(1-\alpha)^2 n}{e^2}\right)^{(1-\alpha)n^2} \cdot \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right),$$

recalling that $\varepsilon = n^{-1}\log n$ and so $\exp(\varepsilon(1-\alpha)n^2) = \exp(o(n\log^2 n))$. $\qquad\square$

The next ingredient we will need is a lower bound on the total number of $n \times n$ Latin squares. The current state of the art is the following bound (obtained via a celebrated result of Egorychev–Falikman [21, 22] on permanents of doubly stochastic matrices).

Lemma 4.16 ([56, Theorem 17.2]). *The number of $n \times n$ Latin squares is at least*

$$|\mathcal{L}_n| \geq \left(\frac{n}{e^2}\right)^{n^2} \cdot \exp(-O(n\log n)).$$

Finally, we need the following lemma, which shows that each triangle-typical ordered partial Latin square is roughly equally likely to be the outcome of the triangle removal process.

Lemma 4.17. *Fix a constant $\alpha \in (0,1)$. Consider $\vec{P} \in \vec{\mathcal{P}}_{n,\alpha n^2}^{\triangle:\varepsilon}$ for some $\varepsilon > 0$. Then for $\vec{\mathbf{R}} \sim$* $\mathrm{TRP}(n,m)$, *we have*

$$\mathbb{P}[\vec{\mathbf{R}} = \vec{P}] = (1\pm 2\varepsilon)^{\alpha n^2} \exp\left(O(\log n)\right)\left(\frac{e}{n}\right)^{3\alpha n^2}(1-\alpha)^{3n^2(1-\alpha)}.$$

Proof. Let $N = n^2$. Since $\vec{P} \in \vec{\mathcal{P}}_{n,\alpha N}^{\triangle:\varepsilon}$,

$$\mathbb{P}[\vec{\mathbf{R}} = \vec{P}] = \prod_{i=0}^{\alpha N - 1} \frac{1}{(1 \pm \varepsilon)(1 - i/N)^3 n^3} = (1 \pm 2\varepsilon)^{\alpha N} n^{-3\alpha N} \exp\left(-3 \sum_{i=0}^{\alpha N - 1} \log\left(1 - \frac{i}{N}\right) \right).$$

Note that

$$\sum_{i=0}^{\alpha N - 1} \frac{1}{N} \log\left(1 - \frac{i+1}{N}\right) \leq \int_0^\alpha \log(1-t) \mathrm{d}t \leq \sum_{i=0}^{\alpha N - 1} \frac{1}{N} \log\left(1 - \frac{i}{N}\right).$$

We have

$$\sum_{i=0}^{\alpha N - 1} \left(\log\left(1 - \frac{i}{N}\right) - \log\left(1 - \frac{i+1}{N}\right) \right) = \sum_{i=0}^{\alpha N - 1} \log\left(1 + \frac{1}{N - (i+1)}\right) \leq \sum_{i=0}^{\alpha N - 1} \frac{1}{N - (i+1)} = O(\log n).$$

Thus, using the fact that $\int \log x = x(\log x - 1)$, we get

$$3 \sum_{i=0}^{\alpha N - 1} \log\left(1 - \frac{i}{N}\right) = 3N \int_0^\alpha \log(1-t) \, \mathrm{d}t \pm O(\log n)$$

$$= 3N \int_{1-\alpha}^1 \log s \, \mathrm{d}s \pm O(\log n)$$

$$= 3N \Big(1(\log 1 - 1) - (1-\alpha)\big(\log(1-\alpha) - 1\big) \Big) \pm O(\log n)$$

$$= 3N \big(-\alpha - (1-\alpha)\log(1-\alpha) \big) \pm O(\log n),$$

so

$$\exp\left(-3 \sum_{i=0}^{\alpha N - 1} \log\left(1 - \frac{i}{N}\right) \right) = e^{3N\alpha}(1-\alpha)^{3N(1-\alpha)} \exp\big(O(\log n)\big).$$

Therefore,

$$\mathbb{P}[\vec{\mathbf{R}} = \vec{P}] = (1 \pm 2\varepsilon)^{\alpha N} \exp\big(O(\log n)\big) \left(\frac{e}{n}\right)^{3\alpha N} (1-\alpha)^{3N(1-\alpha)}. \qquad \square$$

We are now ready to prove Lemma 4.7.

Proof of Lemma 4.7. Let $N := n^2$ and $\varepsilon := n^{-1} \log n$ and $\gamma := n^{-1/4}$. By the fact that $\vec{\mathcal{U}}$ is (ρ, m)-inherited from $\mathcal{T}$ and by Lemmas 4.13 and 4.14, we have that for every Latin square $L \in \mathcal{T}$, a random ordering $\vec{\mathbf{P}}_m(L)$ of a random set of m hyperedges of L satisfies

$$\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}} \cap \vec{\mathcal{P}}_{n,\alpha N}^{\triangle:\varepsilon} \cap \mathcal{P}_{n,\alpha N}^\gamma] \geq \rho - o(1),$$

where we say that an ordered partial Latin square is in $\mathcal{P}_{n,\alpha N}^\gamma$ if its underlying unordered partial Latin square is in $\mathcal{P}_{n,\alpha N}^\gamma$. Thus, letting $\mathcal{S} := \vec{\mathcal{U}} \cap \vec{\mathcal{P}}_{n,\alpha N}^{\triangle:\varepsilon} \cap \mathcal{P}_{n,\alpha N}^\gamma$,

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}] = \sum_{L \in \mathcal{T}} \frac{1}{|\mathcal{L}_n|} \leq \sum_{L \in \mathcal{T}} \frac{1}{|\mathcal{L}_n|} \cdot \frac{\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \mathcal{S}]}{\rho - o(1)} = \frac{1}{\rho - o(1)} \cdot \frac{1}{|\mathcal{L}_n|} \sum_{L \in \mathcal{T}} \sum_{\vec{P} \in \mathcal{S}} \mathbb{P}[\vec{\mathbf{P}}_m(L) = \vec{P}].$$

Changing the order of the two sums, we get

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}] \leq \frac{1}{\rho - o(1)} \sum_{\vec{P} \in \mathcal{S}} \mathbb{P}[\vec{\mathbf{P}}_m(\mathbf{L}) = \vec{P}],$$

where $\vec{\mathbf{P}}_m(\mathbf{L})$ is a uniformly random ordering of a uniformly random subset of m hyperedges of $\mathbf{L}$. Letting $|\mathcal{L}_n(\vec{P})|$ denote the number of completions of the underlying unordered partial Latin square of $\vec{P}$, we have

$$\mathbb{P}[\vec{\mathbf{P}}_m(\mathbf{L}) = \vec{P}] = |\mathcal{L}_n(\vec{P})| \cdot \frac{1}{|\mathcal{L}_n|} \cdot \frac{(N-m)!}{N!},$$

since for each Latin square, there are $N(N-1)\ldots(N-m+1)$ ways to select an (ordered) sequence of m hyperedges. Therefore, by Theorem 4.15 and Lemma 4.16,

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}] \leq \sum_{\vec{P} \in \mathcal{S}} (\rho^{-1} + o(1)) \cdot \frac{|\mathcal{L}_n(\vec{P})|}{|\mathcal{L}_n|} \cdot \frac{(N-m)!}{N!}$$

$$\leq \sum_{\vec{P}\in\mathcal{S}} (\rho^{-1}+o(1))\cdot\left(\frac{n(1-\alpha)^2}{e^2}\right)^{N-|\vec{P}|}\exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\cdot n^{O(n)}\left(\frac{e^2}{n}\right)^{n^2}\cdot\frac{(N-m)!}{N!}$$

$$\leq \sum_{\vec{P}\in\mathcal{S}} \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\cdot\left(\frac{e^2}{n}\right)^{\alpha N}\cdot(1-\alpha)^{2N(1-\alpha)}\cdot\frac{(N-m)!}{N!}.$$

By Stirling's approximation, we deduce

$$\mathbb{P}[\mathbf{L}\in\mathcal{T}] \leq \sum_{\vec{P}\in\mathcal{S}} \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\cdot\left(\frac{e^2}{n}\right)^{\alpha N}\cdot(1-\alpha)^{2N(1-\alpha)}\cdot\frac{N^{N(1-\alpha)}(1-\alpha)^{N(1-\alpha)}}{e^{N(1-\alpha)}}\cdot\frac{e^N}{N^N}$$

$$= \sum_{\vec{P}\in\mathcal{S}} \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\left(\frac{e}{n}\right)^{3\alpha N}(1-\alpha)^{3N(1-\alpha)}.$$

On the other hand, for each $\vec{P}\in\vec{\mathcal{U}}\cap\vec{\mathcal{P}}_{n,\alpha N}^{\triangle:\varepsilon}\cap\mathcal{P}_{n,\alpha N}^{\gamma}$, by Lemma 4.17 we have

$$\mathbb{P}[\vec{\mathbf{R}}=\vec{P}] \geq (1-2\varepsilon)^{\alpha N}e^{-\Theta(\log n)}\left(\frac{e}{n}\right)^{3\alpha N}(1-\alpha)^{3N(1-\alpha)}$$

$$\geq \exp\left(-(2+o(1))\alpha n\log n\right)\left(\frac{e}{n}\right)^{3\alpha N}(1-\alpha)^{3N(1-\alpha)}.$$

Therefore,

$$\mathbb{P}[\mathbf{L}\in\mathcal{T}] \leq \sum_{\vec{P}\in\mathcal{S}} \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\mathbb{P}[\vec{\mathbf{R}}=\vec{P}] \leq \exp\left(\frac{n\log^2 n}{1-\alpha-o(1)}\right)\mathbb{P}[\vec{\mathbf{R}}\in\vec{\mathcal{U}}]. \qquad \square$$

5. Approximating the triangle removal process

In this section, we collect some simple lemmas that allow us to analyse the triangle removal process via Erdős–Rényi (binomial) random hypergraphs. The latter are defined in terms of independent choices, and are therefore much easier to analyse.

We then use these lemmas to deduce a user-friendly corollary (Corollary 5.5) of Lemma 4.7.

Definition 5.1. Let $K_{n,n,n}^{(3)}$ be the complete 3-uniform 3-partite hypergraph on $n+n+n$ vertices, and let $\mathbb{G}^{(3)}(n,p)$ be the probability distribution on subgraphs of $K_{n,n,n}^{(3)}$ obtained by including each possible hyperedge with probability p independently.

Note that $\mathbf{G}\sim\mathbb{G}^{(3)}(n,p)$ may not be a partial Latin square, as it may contain two hyperedges intersecting in more than one vertex. Call this a *conflict*.

Now, our first lemma is similar to [24, Lemma 2.6]. It approximates the triangle removal process with a binomial random hypergraph, for properties $\mathcal{U}$ that are *monotone increasing* in the sense that $P\in\mathcal{U}$ and $P\subseteq P'$ then $P'\in\mathcal{U}$ (i.e. "adding additional hyperedges doesn't hurt"). For multiple-exposure arguments, it will be important that this lemma can handle the triangle removal process started from a proper subgraph of $K_{n,n,n}$.

Definition 5.2. For a partial Latin square P, let $\mathrm{TRP}(P,m)$ be the partial Latin square distribution obtained with m steps of the triangle removal process starting from $G(P)$. Thus, an outcome of $\mathrm{TRP}(P,m)$ always has m hyperedges, and none of these hyperedges conflict with P (unless we run out of triangles and get the outcome "$\perp$")

Lemma 5.3. *Fix constants* γ,α_1,α_2 *such that* $\gamma>0$, *and* $\alpha_1,\alpha_2\in[0,1]$ *are sufficiently small with respect to* γ. *Define* $m_1=\alpha_1 n^2$ *and* $m_2=\alpha_2 n^2$ *and* $p=(1+\gamma)\alpha_2/n$. *Let* $P\in\mathcal{P}_{n,m_1}$ *be a partial Latin square, and consider a monotone increasing property* $\mathcal{U}$ *of subhypergraphs of* $K_{n,n,n}^{(3)}$. *Let* $\mathbf{R}\sim\mathrm{TRP}(P,m_2)$ *and* $\mathbf{G}\sim\mathbb{G}^{(3)}(n,p)$. *Then*

$$\mathbb{P}[\mathbf{R}\in\mathcal{U}] \leq \mathbb{P}[\mathbf{G}\in\mathcal{U}] + \exp(-\Omega(n^2)).$$

Proof. Let $\mathbf{G}^*$ be obtained from $\mathbf{G}$ by deleting all hyperedges involved in conflicts. We say that the hyperedges of $\mathbf{G}^*$ are *isolated* in $\mathbf{G}$. Note that we can couple $(\mathbf{G}^*,\mathbf{G})$ and $\mathbf{R}$ such that $\mathbf{R}\subseteq\mathbf{G}$ as long as $\mathbf{G}^*$ has at least $m_2=\alpha_2 n^2$ hyperedges (to see this, randomly order the hyperedges in $\mathbf{G}$ and run the triangle removal process on that ordering).

Let $\mathbf{X}$ be the number of hyperedges in $\mathbf{G}^*$ (i.e., the number of isolated hyperedges in $\mathbf{G}$). Note that each hyperedge in P conflicts with no more than $3n$ of the n^3 potential hyperedges in $\mathbf{G}$. Thus, there

are at least $(1 - 3\alpha_1)n^3$ potential hyperedges that do not conflict with P. Each of these is present and isolated in $\mathbf{G}$ with probability at least $p(1-p)^{3n} \geq (1+\gamma)\alpha_2 \exp(-6(1+\gamma)\alpha_2)/n$. Thus, we have

$$\mathbb{E}[\mathbf{X}] \geq (1 - 3\alpha_1)n^3 \cdot (1+\gamma)\alpha_2 \exp(-6(1+\gamma)\alpha_2)/n \geq (1+\gamma/2)\alpha_2 n^2.$$

Adding a hyperedge to $\mathbf{G}$ can increase $\mathbf{X}$ by at most 1, and removing a hyperedge from $\mathbf{G}$ can increase $\mathbf{X}$ by at most 3, by making three hyperedges isolated. Thus, by Theorem 3.1,

$$\mathbb{P}[\mathbf{X} < \alpha_2 n^2] \leq \mathbb{P}[|\mathbf{X} - \mathbb{E}[\mathbf{X}]| \geq \gamma\alpha_2 n^2/2] \leq \exp(-\Omega(n^2)).$$

Therefore,

$$\mathbb{P}[\mathbf{R} \in \mathcal{U}] \leq \mathbb{P}[\mathbf{G} \in \mathcal{U}] + \mathbb{P}[\mathbf{X} < \alpha_2 n^2] = \mathbb{P}[\mathbf{G} \in \mathcal{U}] + \exp(-\Omega(n^2)). \qquad \square$$

The next lemma appears as [42, Lemma 5.2], and handles properties $\mathcal{U}$ that are monotone decreasing in the sense that if $P \in \mathcal{U}$ and $P' \subseteq P$ then $P' \in \mathcal{U}$ (i.e., "removing hyperedges doesn't hurt").

Lemma 5.4 ([42, Lemma 5.2]). *Fix a constant $\alpha \in (0, 1)$. Let $\mathcal{U} \subseteq \mathcal{P}_n$ be a monotone decreasing property of partial Latin squares, let[10] $\mathbf{R} \sim \mathrm{TRP}(n, \alpha n^2)$, let $\mathbf{G} \sim \mathbb{G}^{(3)}(n, \alpha/n)$, and let $\mathbf{G}^*$ be obtained from $\mathbf{G}$ by deleting all hyperedges involved in conflicts. Then*

$$\mathbb{P}[\mathbf{R} \in \mathcal{U} \text{ or } \mathbf{R} = \perp] = O(\mathbb{P}[\mathbf{G}^* \in \mathcal{U}]).$$

For completeness, we reproduce the simple proof of Lemma 5.4 (from [42]).

Proof. We can couple $\mathbf{R}$ and $\mathbf{G}^*$ such that, if $e(\mathbf{G}) \leq \alpha n^2$, then either $\mathbf{G}^* \subseteq \mathbf{R}$ or $\mathbf{R} = \perp$ (to see this, randomly order the hyperedges of $\mathbf{G}$ and run the triangle removal process on that ordering).

Since $\mathcal{U}$ is monotone decreasing, we have

$$\mathbb{P}[\mathbf{R} \in \mathcal{U} \text{ or } \mathbf{R} = \perp] \leq \mathbb{P}[\mathbf{G}^* \in \mathcal{U} \mid e(\mathbf{G}) \leq \alpha n^2].$$

Since $e(\mathbf{G})$ is binomial with mean αn^2, we have that $\mathbb{P}[e(\mathbf{G}) \leq \alpha n^2] = \Omega(1)$. Thus,

$$\mathbb{P}[\mathbf{R} \in \mathcal{U} \text{ or } \mathbf{R} = \perp] \leq \mathbb{P}[\mathbf{G}^* \in \mathcal{U}]/\mathbb{P}[e(\mathbf{G}) \leq \alpha n^2] = O(\mathbb{P}[\mathbf{G}^* \in \mathcal{U}]). \qquad \square$$

To illustrate the tools from this section, we conclude this section by stating a corollary of Lemma 4.7, which is an easy-to-apply tool for studying random Latin squares. We do not actually use this corollary in the present paper (due to some technical issues related to multiple-exposure arguments, we need the full power of Lemma 4.7), but we hope that it may be convenient in future work on random Latin squares.

Note that we have defined (ρ, m)-inheritedness (in Definition 4.3) only for properties of ordered partial Latin squares, but in what follows we extend this definition to unordered partial Latin squares in the natural way.

Corollary 5.5. *Fix constants $\alpha \in (0, 1)$ and $\rho \in (0, 1)$, and let $m = \alpha n^2$. Let $\mathcal{T}$ be a property of $n \times n$ Latin squares and let $\mathcal{U}$ be a property of subhypergraphs of $K_{n,n,n}^{(3)}$, such that $\mathcal{U}$ is (ρ, m)-inherited from $\mathcal{T}$. Let $\mathbf{L}$ be a uniformly random $n \times n$ Latin square.*

(1) If $\mathcal{U}$ is a monotone decreasing *property, we have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}] \leq \exp\big(O(n \log^2 n)\big)\mathbb{P}[\mathbf{G}^* \in \mathcal{U}],$$

where $\mathbf{G} \sim \mathbb{G}^{(3)}(n, \alpha/n)$, and $\mathbf{G}^*$ is obtained from $\mathbf{G}$ by deleting all hyperedges involved in conflicts.

(2) If $\mathcal{U}$ is a monotone increasing *property, then for any constant $\gamma > 0$, if α is sufficiently small with respect to γ, we have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}] \leq \exp\big(O(n \log^2 n)\big)\mathbb{P}[\mathbf{G} \in \mathcal{U}] + \exp(-\Omega(n^2)),$$

where $\mathbf{G} \sim \mathbb{G}^{(3)}(n, (1+\gamma)\alpha/n)$.

Part (1) is obtained by combining Lemma 4.7 with Lemma 5.3 (with $P = \emptyset$ and $\alpha_1 = 0$ and $\alpha_2 = \alpha$), and Part (2) is obtained by combining Lemma 4.7 with Lemma 5.4.

---

[10]We previously defined $\mathrm{TRP}(n, m)$ as a distribution on ordered partial Latin squares, but here we are somewhat abusively viewing it as a distribution on unordered partial Latin squares (we simply ignore the ordering).

6. The master theorem for parity distributions

All parts of Theorem 1.3 will be deduced from a master theorem which describes the distribution of $(N_{\mathrm{row}}(\mathbf{L}), N_{\mathrm{col}}(\mathbf{L}), N_{\mathrm{sym}}(\mathbf{L}))$ in somewhat technical terms. To state this theorem, we need some terminology to describe certain types of distributions that arise naturally from random intercalate switches.

Definition 6.1. Consider a random vector $(c_1 + \mathbf{B}_1, c_2 + \mathbf{B}_2, c_3 + \mathbf{B}_3)$, where $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3$ are independent binomial random variables satisfying $\mathbf{B}_i \sim \mathrm{Bin}(n_i, 1/2)$. We say that this random vector is (n, d)-*near-binomial* if $n - d \le n_i \le n$ and $c_i \le d$ for each i.

The reader should think of d as a "defect parameter": if d is small, then (n, d)-near-binomial random vectors are approximately distributed like three independent copies of $\mathrm{Bin}(n, 1/2)$. We remark that if we want this approximation to have $o(1)$ total variation error, then we need $d = o(\sqrt{n})$. Indeed, for each i we need $\mathbb{E}[\mathbf{X}_i] - n/2$ to be negligible compared to the fluctuations of $\mathrm{Bin}(n, 1/2)$ (which have order of magnitude $\sqrt{n}$).

Definition 6.2. Say that a random vector $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) \in \mathbb{N}^3$ is (n, d)-*2-near-binomial* if it can be obtained from an (n, d)-near-binomial random vector by conditioning on a particular parity (even or odd) for each of $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$.

Definition 6.3. Say that a random vector $\vec{\mathbf{X}}$ is a *2-near-binomial mixture* with parameters (n, d, ε) if one can define an auxiliary random object $\mathbf{Q}$ (on the same probability space as $\vec{\mathbf{X}}$), and an event $\mathcal{A}$ depending only on $\mathbf{Q}$, such that:

- $\mathbb{P}[\mathcal{A}] \ge 1 - \varepsilon$, and
- For any of the possible outcomes Q of $\mathbf{Q}$ satisfying $\mathcal{A}$, the conditional distribution of $\vec{\mathbf{X}}$ given $\mathbf{Q} = Q$ is (n, d)-2-near-binomial.

If d is small, then the above definition says that $\vec{\mathbf{X}}$ is closely approximated by a sequence of three independent $\mathrm{Bin}(n, 1/2)$ random variables, except for possible biases mod 2. Note that the distribution μ^* featuring in Theorem 1.3 is a 2-near-binomial mixture with parameters $(n, 0, 0)$.

Now, our main technical theorem is as follows.

Theorem 6.4. *Let $\mathbf{L}$ be a uniformly random $n \times n$ Latin square. Then $(N_{\mathrm{row}}(\mathbf{L}), N_{\mathrm{col}}(\mathbf{L}), N_{\mathrm{sym}}(\mathbf{L}))$ is a 2-near-binomial mixture with parameters*

$$\left(n, \; O(\log^{11} n), \; \exp(-\omega(n \log^2 n)) \right).$$

6.1. Corollaries of the master theorem.

In this subsection we briefly explain how to deduce all parts of Theorem 1.3 from our master theorem (Theorem 6.4). This is all quite routine. Note that Theorem 1.2 is an immediate consequence of Theorem 1.3(4).

We start with basic consequences of the de Moivre–Laplace local central limit theorem and a large deviation principle for binomial distributions.

Lemma 6.5. *Let $\vec{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ be a $(n, o(\sqrt{n}))$-2-near-binomial random vector.*

(i) *Recall that $(\mathbf{X}_1 \bmod 2, \mathbf{X}_2 \bmod 2, \mathbf{X}_3 \bmod 2)$ always takes a common value $\vec{v} \in \{0, 1\}^3$. For any $\vec{x} \in \mathbb{Z}^3$ satisfying $\vec{x} = \vec{v} \bmod 2$, we have*

$$\mathbb{P}[\vec{\mathbf{X}} = \vec{x}] = 8 \cdot \frac{1}{(2\pi(n/4))^{3/2}} \exp\left(-\frac{(x_1 - n/2)^2 + (x_2 - n/2)^2 + (x_3 - n/2)^2}{2(n/4)} \right) + o(n^{-3/2}).$$

(ii) *Let $H_2 : \alpha \mapsto -\alpha \log_2 \alpha - (1 - \alpha) \log_2(1 - \alpha)$ be the base-2 binary entropy function, and let $I(x_1, x_2, x_3) = 3 - H_2(x_1) - H_2(x_2) - H_2(x_3)$. We have*

$$-\inf_{\vec{x} \in E^\circ} I(\vec{x}) \le \liminf n^{-1} \log_2 \mathbb{P}[n^{-1} \vec{\mathbf{X}} \in E] \le \limsup n^{-1} \log_2 \mathbb{P}[n^{-1} \vec{\mathbf{X}} \in E] \le -\inf_{\vec{x} \in \overline{E}} I(\vec{x})$$

for all Borel $E \subseteq \mathbb{R}^3$.

Proof. Recall that $\vec{\mathbf{X}}$ is obtained from a (n, d)-near-binomial random vector $\vec{\mathbf{Y}}$ by conditioning on the event that $\vec{\mathbf{Y}} = \vec{v} \bmod 2$ (which occurs with probability exactly[11] $1/8$). Now, the de Moivre–Laplace local central limit theorem (see e.g. [23, Chapter VII, Section 3, Theorem 1]) tells us that

$$\mathbb{P}[\vec{\mathbf{Y}} = \vec{x}] = \prod_{i=1}^{3} \left(\frac{1}{\sqrt{2\pi(n_i/4)}} \exp\left(-\frac{(x_i' - n_i/2)^2}{2(n_i/4)} \right) + o(n^{-1/2}) \right)$$

---

[11]Here we are using that $\mathrm{Bin}(n, 1/2)$ is equally likely to be even and odd, which is easy to prove e.g. by induction.

for some $n_1, n_2, n_3 = n - o(\sqrt{n})$, and some x_1', x_2', x_3' satisfying $x_i' = x_i - o(\sqrt{n})$. It is easy to see that the error terms in the n_i and x_i' can be absorbed into the main error term (i.e., it makes no difference to replace n_1, n_2, n_3 with n, and replace each x_i' with x_i). The desired result follows, noting that conditioning on an event that occurs with probability $1/8$ introduces a factor of 8.

We can similarly deduce (ii) from a large deviation principle for binomial distributions (e.g. Sanov's theorem, see [20, Theorem 2.1.10][12]). Actually, this is even simpler than (i), because in this case the conditioning mod 2 has a negligible impact. $\square$

Now we are ready to deduce Theorem 1.3.

Proof of Theorem 1.3. We start with Theorem 1.3(3) (the local central limit theorem). Consider any $\vec{x} \in \mathbb{Z}^3$ satisfying $x_1 + x_2 + x_3 = f(n) \pmod 2$. By Theorem 6.4, together with Lemma 6.5(i) and the definition of a 2-near-binomial mixture, $\mathbb{P}[\vec{X} = \vec{x}]$ is of the form

$$\mathbb{P}[\vec{X} = \vec{x} \bmod 2] \cdot \left(8 \cdot \frac{1}{(2\pi(n/4))^{3/2}} \exp\left(\frac{(x_1 - n/2)^2 + (x_2 - n/2)^2 + (x_3 - n/2)^2}{2(n/4)} \right) + o(n^{-3/2}) \right)$$
$$+ \exp(-\omega(n \log^2 n)).$$

By the Cavenagh–Wanless theorem [18] (mentioned in the introduction), we have $\mathbb{P}[\vec{X} = \vec{x} \bmod 2] = 1/4 + o(1)$, and Theorem 1.3(3) follows.

We can then deduce Theorem 1.3(4) (total variation convergence) by summing over all $\vec{x}$ (using e.g. Chebyshev's inequality to control the tails). Indeed, let $\vec{Z} \sim \mu^*$. For any N, the variance of $\mathrm{Bin}(N, 1/2)$ is $N/4 = O(N)$, so (recalling the definition of a 2-near-binomial mixture and noting that the means of all relevant binomials are within $O((\log n)^{11}) = O(\sqrt{n})$ of each other), the variance of each component of $\vec{X}$ and $\vec{Z}$ is $O(n)$. Then, for any $K \geq 1$, taking $\vec{Z} \sim \mu^*$ and writing $(n/2)\vec{1} = (n/2, n/2, n/2)$, we have

$$d_{\mathrm{TV}}(\vec{X}, \mu^*) = \frac{1}{2} \sum_{\vec{x} \in \mathbb{Z}^3} \left| \mathbb{P}[\vec{X} = \vec{x}] - \mathbb{P}[\vec{Z} = \vec{x}] \right|$$
$$\leq O((K\sqrt{n})^3) \cdot o(n^{-3/2}) + \mathbb{P}\left[\|\vec{X} - (n/2)\vec{1}\|_\infty \geq K\sqrt{n} \right] + \mathbb{P}\left[\|\vec{Z} - (n/2)\vec{1}\|_\infty \geq K\sqrt{n} \right]$$
$$\leq o(1) \cdot K^3 + O(1/K^2).$$

Taking $K \to \infty$ sufficiently slowly, Theorem 1.3(4) follows.

It is then a near-trivial matter to deduce Theorem 1.3(1–2) (the central limit theorem and law of large numbers), since total variation convergence is much stronger than convergence in distribution (see e.g. [54, Proposition 1.2]).

Finally, Theorem 1.3(5) (the large deviation principle) follows from Theorem 6.4 and Lemma 6.5(ii). Indeed, Theorem 6.4 says that $\vec{X}$ is a mixture of $(n, o(\sqrt{n}))$-2-near-binomial distributions except with an exceptional probability $\exp(-\omega(n \log^2 n))$ (which is negligible). $\square$

Remark 6.6. In the above proof, it was convenient to use the Cavenagh–Wanless theorem as a black box. However, we remark that given the machinery we developed in Sections 4 and 5, it would be a very simple matter to re-prove the Cavenagh–Wanless theorem. Indeed, using this machinery it is easy to prove that $\mathbf{L}$ is very likely to have a 2×3 Latin subrectangle (in the first half of the rows, columns and symbols of $\mathbf{L}$, say), and a 3×2 Latin subrectangle (in the second half of the rows, columns and symbols of $\mathbf{L}$, say), and switching on these two Latin subrectangles mixes between all four different possibilities for $(N_{\mathrm{row}}(\mathbf{L}), N_{\mathrm{col}}(\mathbf{L}), N_{\mathrm{sym}}(\mathbf{L})) \bmod 2$.

7. STABLE INTERCALATES

An *intercalate* in a Latin square is a 2×2 Latin subsquare. To *switch* an intercalate with symbols a and b means to replace both occurrences of a with b and vice versa. The result of this operation is always another Latin square.

As discussed in Section 2, we will prove Theorem 6.4 via random intercalate switches. It is very important that if we start with a uniformly random Latin square, and perform our random switches, the resulting Latin square is still uniformly random. To ensure this, we need to restrict our attention to intercalates with a certain *stability* property. We define this property in such a way that switching stable

---

[12]A simple approximation argument (approximating general open and closed sets with product sets) can be used to combine large deviation principles for three independent random variables, see e.g. [48, Lemma 2.8].

intercalates can never create a new intercalate or destroy an existing intercalate (and, in particular, if we only ever switch stable intercalates, then the set of available stable intercalate switches never changes).

Specifically, we define an intercalate to be *stable* if it does not intersect other intercalates, and if switching it and/or some other intercalates cannot make it intersect another intercalate, as follows.

Definition 7.1. Let P be a partial Latin square. For an intercalate A, we denote by $\bar{A}$ the outcome of switching A.

- We say that an intercalate in P is *isolated* if it does not share an entry with any other intercalate.
- Consider a set of isolated intercalates $\{A_1, \ldots, A_t\}$ in P. We say that this is a *critical* set of intercalates if it is possible to switch a subset of $A_1, \ldots, A_t$ (resulting in intercalates $B_1, \ldots, B_t$ where $B_i = A_i$ or $B_i = \bar{A}_i$ for each $i \in [t]$) to create a new intercalate that shares an entry with each of $B_1, \ldots, B_t$.
- We say that an isolated intercalate in P is *stable* if it is not contained in a critical set of intercalates.

Observe that a critical set $\{A_1, \ldots, A_t\}$ may be a subset of another critical set $\{A_1, \ldots, A_{t+1}\}$ (e.g., if switching a subset of $A_1, \ldots, A_t$ creates a new intercalate that intersects A_{t+1}). Any critical set $\mathcal{D}$ has size at most 4, since its members $A_1, \ldots, A_t$ are pairwise entry-disjoint, and for each A_i, either A_i or $\bar{A}_i$ shares an entry with the new intercalate that can be produced by switching a subset of $\mathcal{D}$.

Now, in the following lemma we record the fact that switching stable intercalates leaves the set of stable intercalates switches unchanged.

Definition 7.2. For an intercalate A with rows r_1, r_2, columns c_1, c_2, and symbols s_1, s_2, define the tuple $\sigma(A) = (\{r_1, r_2\}, \{c_1, c_2\}, \{s_1, s_2\})$ (i.e., this records the rows, columns and symbols of the intercalate, without recording which of the two possible "states" it is in). For a (partial) Latin square P, let $\mathcal{S}(P)$ be the set of all tuples $\sigma(A)$ for all stable intercalates A in P (i.e., this records the intercalate switches we can perform, without recording which state each of those intercalates is in).

Lemma 7.3. *Let P_1 be a Latin square, and suppose that we can switch a single stable intercalate in P_1 to obtain a second Latin square P_2. Then $\mathcal{S}(P_1) = \mathcal{S}(P_2)$.*

Proof. Let A be the stable intercalate in P_1 which is switched to obtain P_2. Then $\sigma(A) \in \mathcal{S}(P_1)$. We start with a simple observation which will be used multiple times.

Claim 7.4. *If an intercalate $B \neq \bar{A}$ is isolated in P_2, it is also present and isolated in P_1.*

Proof of claim. Since B shares no entry with $\bar{A}$ in P_2, we see that B is also present in P_1. Furthermore, if there were any intercalate C in P_1 that shares an entry with B, that intercalate C would not be present in P_2, so it would have to share an entry with A in P_1. But that would imply A is not isolated in P_1, which would contradict our assumption that A is stable in P_1. ∎

Step 1: reversibility of the A switch. We will first show $\sigma(A) \in \mathcal{S}(P_2)$, i.e., we could equally well have switched $\bar{A}$ in P_2 to obtain P_1. This amounts to showing that $\bar{A}$ is isolated in P_2, and not contained in a critical set of intercalates.

First, note that if $\bar{A}$ were not isolated in P_2, then there would be some intercalate B in P_2 that shares an entry with $\bar{A}$. Since that entry would not be in P_1, we would have that B is not in P_1, implying that $\{A\}$ is a critical set in P_1, contradicting the stableness of A in P_1. Thus, $\bar{A}$ is isolated in P_2.

Now, suppose for contradiction that $\bar{A}$ is contained in a critical set $\{\bar{A}, A_2, \ldots, A_t\}$ of isolated intercalates in P_2. If $t = 1$, this would contradict the isolatedness of either A in P_1 or $\bar{A}$ in P_2; thus $t \geq 2$. Since the intercalates $A_2, \ldots, A_t$ are isolated in P_2, they are also present and isolated in P_1, by Claim 7.4. But then $\{A, A_2, \ldots, A_t\}$ is a critical set in P_1: if switching $S \subseteq \{\bar{A}, A_2, \ldots, A_t\}$ creates a new intercalate in P_2, then switching $\bar{S} = S \cup \{A\}$ (if $\bar{A} \notin S$) or $\bar{S} = S \setminus \{\bar{\bar{A}}\}$ (if $\bar{A} \in S$) creates a new intercalate in P_1. This contradicts the stableness of A in P_1. We conclude that $\bar{A}$ is stable in P_2, so $\sigma(A) \in \mathcal{S}(P_2)$.

Step 2: the intercalates other than A. Since $\sigma(A) \in \mathcal{S}(P_1) \cap \mathcal{S}(P_2)$, by symmetry it is enough to show that for every intercalate B that is stable in P_1, we have $\sigma(B) \in \mathcal{S}(P_2)$. Let $B \neq A$ be a stable intercalate in P_1. Then B must be present and isolated in P_2, by Claim 7.4 and the symmetry between P_1 and P_2.

Suppose for the purpose of contradiction that B is not stable in P_2. Then B is contained in a critical set $\mathcal{Z} = \{B, A_2, \ldots, A_t\}$ of isolated intercalates in P_2. Since the intercalates in $\{A_2, \ldots, A_t\} \setminus \{\bar{A}\}$ are isolated in P_2 and distinct from $\bar{A}$, they are also present and isolated in P_1. Now, either $\mathcal{Z} \setminus \{\bar{A}\} \cup \{A\}$ or $\mathcal{Z} \setminus \{\bar{A}\}$ is a critical set in P_1. In particular, if the new intercalate that can be created in P_2 by switching a subset of $\mathcal{Z}$ shares an entry with either A or $\bar{A}$, then $\mathcal{Z} \setminus \{\bar{A}\} \cup \{A\}$ is a critical set in P_1. Otherwise,

$\mathcal{Z} \setminus \{\bar{A}\}$ is a critical set in P_1. In either case, B is contained in a critical set in P_1, contradicting that B is stable in P_1. Thus, B must be stable in P_2, so $\sigma(B) \in \mathcal{S}(P_2)$, concluding the proof. $\qquad \square$

8. Deducing the master theorem from stable intercalate information

As discussed in Section 2, the main challenge in this paper is to show that random Latin squares typically have very rich constellations of stable intercalates. These intercalates can then be randomly switched to prove our master theorem on parity distributions (Theorem 6.4). In this section, we state our key lemma about stable intercalates in random Latin squares (Lemma 8.3), and show how to use this key lemma (together with some linear-algebraic arguments) to prove Theorem 6.4. After this section, the rest of the paper will be devoted to the proof of Lemma 8.3.

8.1. **The key lemma on stable intercalates.** Roughly speaking, we need the property that for all "reasonably large" sets of rows, columns and symbols, we can find stable intercalates compatible with those choices. This will imply that the stable intercalates are so well-spread throughout the rows, columns and symbols that randomly switching them will thoroughly mix up the row, column and symbol parities.

Definition 8.1. Let P be an $n \times n$ (partial) Latin square, and consider some $1 \leq \ell \leq n$. We say that P is an (ℓ, β)-*intercalate-expander* if the following property holds. For any sets of rows $\mathscr{R}, \mathscr{R}^*$, any sets of columns $\mathscr{C}, \mathscr{C}^*$, and any sets of symbols $\mathscr{S}, \mathscr{S}^*$, such that five out of these six sets have size βn and the last one has size ℓ, there is a stable intercalate in P with one row in $\mathscr{R}$, the other row in $\mathscr{R}^*$, one column in $\mathscr{C}$, the other column in $\mathscr{C}^*$, one symbol in $\mathscr{S}$, and the other symbol in $\mathscr{S}^*$.

Note that the sets of rows, columns, and symbols above are not necessarily disjoint. Thus, the stable intercalate promised by Definition 8.1 may for instance have both its rows in $\mathscr{R} \cap \mathscr{R}^*$.

Random Latin squares themselves are unlikely to be good intercalate-expanders, because they have too many intercalates (that are too likely to intersect each other and form critical sets of intercalates). We are able to overcome this with a *sparsification* technique: instead of considering all the stable intercalates in a Latin square L, we consider only the stable intercalates in some sparsified *partial* Latin square (and we make sure to use the same "sparsification template" T for all Latin squares). To formalise this, we need another definition.

Definition 8.2. An $n \times n$ *template* is a subset $T \subseteq [n]^2$, which we interpret as a set of row/column pairs. For a Latin square $L \in \mathcal{L}_n$, we write $T \cap L \in \mathcal{P}_n$ for the partial Latin square containing just the entries of L in positions specified by T.

Now, our key lemma is as follows.

Lemma 8.3. *Let $\beta > 0$ be a constant. There exists a template $T \subseteq [n]^2$ such that the following holds. Let $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$ be a uniformly random Latin square (interpreted as a subgraph of $K_{n,n,n}^{(3)}$) and let $\ell = \log^{11} n$. Then*

$$\mathbb{P}[T \cap \mathbf{L} \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \geq 1 - \exp(-\omega(n \log^2 n)).$$

8.2. **Deducing the master theorem on parity distributions.** To deduce Theorem 6.4 from Lemma 8.3, we will need some more notation and preliminaries.

Definition 8.4. Consider an $n \times n$ Latin square L. For a set of rows R and a set of intercalates $\mathcal{I}$ in L, let $M_{R,\mathcal{I}} \in \{0,1\}^{|R| \times |\mathcal{I}|}$ be the "incidence matrix" defined as follows:

$$M_{R,\mathcal{I}}(r, I) = \begin{cases} 1 & \text{if row } r \text{ participates in intercalate } I \\ 0 & \text{otherwise.} \end{cases}$$

We can similarly define $M_{C,\mathcal{I}}$ or $M_{S,\mathcal{I}}$ for a set of columns C or a set of symbols S.

Lemma 8.5. *Consider an $n \times n$ partial Latin square P that is an $(\ell, 1/10)$-intercalate-expander, where n is sufficiently large and $1 \leq \ell = o(\sqrt{n})$. Then we can find sets R, C, S of rows, columns and symbols, satisfying $|R|, |C|, |S| \geq n - \ell$, with the following property.*

If we let $\mathcal{I}(R, C, S)$ be the set of all stable intercalates in P which are contained in the rows in R and the columns in C, only using the symbols in S, then the matrix

$$\begin{pmatrix} M_{R,\mathcal{I}(R,C,S)} \\ M_{C,\mathcal{I}(R,C,S)} \\ M_{S,\mathcal{I}(R,C,S)} \end{pmatrix} \in \{0,1\}^{(|R|+|C|+|S|) \times |\mathcal{I}(R,C,S)|}$$

has rank exactly $|R| + |C| + |S| - 3$ over $\mathbb{F}_2$. The left kernel vectors of this matrix are those vectors of the form $(\vec{x}_R \ \vec{x}_C \ \vec{x}_S) \in \mathbb{F}_2^{|R|+|C|+|S|}$ where $\vec{x}_Z \in \mathbb{F}_2^{|Z|}$ is the all-zero vector $(0, \dots, 0)$ or the all-one vector $(1, \dots, 1)$ for each $Z \in \{R, C, S\}$.

Proof. First, we translate the desired statement into more combinatorial language. Consider the auxiliary 6-uniform hypergraph $\mathcal{Q}(R, C, S)$ on the vertex set[13] $R \sqcup C \sqcup S$, where for every intercalate in $\mathcal{I}(R, C, S)$ we put a hyperedge consisting of the two rows, two columns and two symbols of that intercalate. We want to choose R, C, S to have the property that for any choices of vertex subsets $R' \subseteq R$ and $C' \subseteq C$ and $S' \subseteq S$ which are not all "trivial" (i.e., at least one of these three subsets is nonempty and proper), there is some hyperedge of $\mathcal{Q}(R, C, S)$ which intersects $R' \sqcup C' \sqcup S'$ in an odd number of vertices.

Now, we specify R, C, S as follows. Let R be the set of all rows r for which there exist at least 6ℓ different stable intercalates in P which all involve r and which otherwise involve pairwise disjoint rows, columns and symbols. Define C, S similarly.

Claim. *Each of R, C, S has size greater than $n - \ell$.*

Proof of claim. Suppose for contradiction (and without loss of generality) that there is a set R_0 of ℓ rows which are not in R. Then we can use the intercalate-expansion property to greedily find $n/10$ stable intercalates in P which all involve disjoint rows, columns and symbols (except for their rows in R_0). So, there must be some $r \in R_0$ involved in $(n/10)/|R_0| \geq 6\ell$ stable intercalates which involve disjoint rows, columns and symbols (except that they all involve r). This contradicts the choice of R_0. ∎

Recall that the defining properties of R, C, S were that they are included in many (at least 6ℓ) "externally disjoint" stable intercalates in P. The above claim shows that almost all of those intercalates lie in $\mathcal{I}(R, C, S)$, and therefore correspond to hyperedges of $\mathcal{Q}(R, C, S)$. Specifically, each "bad" row/column/symbol (which is not in R, C, S) kills at most one of our externally disjoint stable intercalates, so each vertex of $\mathcal{Q}(R, C, S)$ is contained in more than $6\ell - 3\ell \geq 3\ell$ externally disjoint 6-uniform hyperedges.

Now, consider vertex subsets $R' \subseteq R$ and $C' \subseteq C$ and $S' \subseteq S$ which are not all trivial (at least one of these subsets is nonempty and proper). We need to prove there is some hyperedge intersecting $R' \sqcup C' \sqcup S'$ in an odd number of vertices.

Case 1: First, suppose that for each $Z \in \{R', C', S'\}$, either Z or its complement has size less than ℓ. This gives us three "small sets" of size less than ℓ. At least one of our small sets must be non-empty; let v be a vertex in that set. Recall that v is contained in at least 3ℓ externally disjoint hyperedges of $\mathcal{Q}(R, C, S)$, so at least one of these hyperedges (externally) avoids our three small sets. Any such hyperedge has odd intersection with $R \sqcup C \sqcup S$.

Case 2: If Case 1 does not occur, we may assume without loss of generality that R' and $R \setminus R'$ both have size at least ℓ. Then we apply the intercalate-expansion property directly. We take the smaller of R' and $R \setminus R'$ as $\mathscr{R}^*$ and the larger as $\mathscr{R}$; we take the larger of C' and $C \setminus C'$ as $\mathscr{C} = \mathscr{C}^*$; and we take the larger of S' and $S \setminus S'$ as $\mathscr{S} = \mathscr{S}^*$. This again gives us a stable intercalate in P (corresponding to an edge of $\mathcal{Q}(R, C, S)$) with the desired odd intersection. □

Now we can prove Theorem 6.4.

Proof of Theorem 6.4. Let T be as given by Lemma 8.3, and let $\ell = \log^{11} n$. For a Latin square $L \in \mathcal{L}_n$, we say an intercalate in L is *T-stable* if it is present and stable in $T \cap L$. Recall the notation $\mathcal{S}(P)$ from Definition 7.2 (recording the set of stable intercalates in a partial Latin square P, but without recording which of the two possible "states" each intercalate is in).

We consider an auxiliary multigraph $\mathcal{G}$, whose vertices are the Latin squares $L \in \mathcal{L}_n$, and where there is an edge between two Latin squares L_1, L_2 if it is possible to switch some of the T-stable intercalates in L_1 to transform it into L_2. So, for every $L \in \mathcal{L}_n$, the degree of L in $\mathcal{G}$ is precisely $2^{|\mathcal{S}(T \cap L)|}$ (including one loop edge from L to itself). In fact, by Lemma 7.3, the connected component of L is a clique (with loops) of order $2^{|\mathcal{S}(T \cap L)|}$, where each vertex $L' \in \mathcal{L}_n$ in this clique has $\mathcal{S}(T \cap L') = \mathcal{S}(T \cap L)$. In particular, each component is regular, so the uniform distribution $\mathrm{Unif}(\mathcal{L}_n)$ is a stationary distribution for the random walk on $\mathcal{G}$.

Therefore, if we first sample $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$, and then we "re-randomise" $\mathbf{L}$ by randomly switching each T-stable intercalate with probability $1/2$ independently (this is the same as stepping to a random

---

[13]Here we use "$\sqcup$" to indicate a *disjoint* union.

neighbour of $\mathbf{L}$ in $\mathcal{G}$), then the resulting random Latin square $\mathbf{L}'$ still has the same uniform distribution as $\mathbf{L}$.

Now, for the rest of the proof, our goal is to show that $\vec{\mathbf{X}}' = (N_{\mathrm{row}}(\mathbf{L}'), N_{\mathrm{col}}(\mathbf{L}'), N_{\mathrm{sym}}(\mathbf{L}'))$ is a 2-near-binomial mixture with parameters $(n, \ell, \exp(-\omega(n \log^2 n)))$. Recalling Definition 6.3, this means we need to show how to reveal certain information about $\mathbf{L}$ and $\mathbf{L}'$, in such a way that, conditional on a typical outcome of this revealed information, $\vec{\mathbf{X}}'$ is a (n, ℓ)-2-near-binomial random vector (where here "typical" refers to an event which occurs with probability $1 - \exp(-\omega(n \log^2 n))$).

First, reveal an outcome of $\mathbf{L}$ such that $T \cap \mathbf{L}$ is an $(\ell, 1/10)$-intercalate-expander. By Lemma 8.3, this occurs with probability $1 - \exp(-\omega(n \log^2 n))$. From now on, we will view $\mathbf{L}$ as a non-random object (i.e., all probabilities will implicitly be with respect to the conditional probability space given our outcome of $\mathbf{L}$). Let R, C, S be the sets of rows, columns and symbols from Lemma 8.5, and let $\mathcal{I} = \mathcal{I}(R, C, S)$ be as in the statement of Lemma 8.5. Recall that $\mathbf{L}'$ is obtained from $\mathbf{L}$ via random T-stable intercalate switches; reveal any outcome of these random switches for T-stable intercalates which are *not* in $\mathcal{I}$. We will show that, in the resulting conditional probability space (which can be described in terms of $|\mathcal{I}|$ random coin flips), $\vec{\mathbf{X}}'$ has a 2-near-binomial distribution with parameters (n, ℓ).

To see this, let $\vec{\mathbf{x}}_1 \in \{0, 1\}^n, \vec{\mathbf{x}}_2 \in \{0, 1\}^n, \vec{\mathbf{x}}_3 \in \{0, 1\}^n$ be the sequences of parities of rows, columns and symbols in $\mathbf{L}'$. In our conditional probability space, the only parts of $\vec{\mathbf{x}}_1, \vec{\mathbf{x}}_2, \vec{\mathbf{x}}_3$ that remain random are the subsequences $\vec{\mathbf{x}}_R \in \{0, 1\}^R, \vec{\mathbf{x}}_C \in \{0, 1\}^C, \vec{\mathbf{x}}_S \in \{0, 1\}^S$ corresponding to our identified sets of rows, columns and symbols. We can describe their joint distribution as

$$(\vec{\mathbf{x}}_R \ \vec{\mathbf{x}}_C \ \vec{\mathbf{x}}_S) = (\vec{y}_R \ \vec{y}_C \ \vec{y}_S) + \begin{pmatrix} M_{R,\mathcal{I}} \\ M_{C,\mathcal{I}} \\ M_{S,\mathcal{I}} \end{pmatrix} \vec{\mathbf{r}},$$

where $\vec{y}_R, \vec{y}_C, \vec{y}_S$ describe the parities of our identified rows, columns and symbols in $\mathbf{L}$ (which we are viewing as non-random), $\vec{\mathbf{r}} \in \{0, 1\}^{|\mathcal{I}|}$ is a uniformly random zero-one sequence of length $|\mathcal{I}|$, and arithmetic is mod 2.

Now, recall the structure of the left kernel vectors from Lemma 8.5. We see that $(\vec{\mathbf{x}}_R \ \vec{\mathbf{x}}_C \ \vec{\mathbf{x}}_S)$ is a uniformly random element of $\{0, 1\}^R \times \{0, 1\}^C \times \{0, 1\}^S$, except that the values of $\vec{1} \cdot \vec{\mathbf{x}}_R$, $\vec{1} \cdot \vec{\mathbf{x}}_C$ and $\vec{1} \cdot \vec{\mathbf{x}}_S$ are constrained (i.e., the parities of the rows, columns and symbols in R, C, S are uniformly random, except that the number of odd rows, number of odd columns, and number of odd symbols are constrained mod 2). This implies the desired result. $\qquad\square$

9. Setup for the proof of the intercalate-expander lemma

In this section, we set the stage for the proof of Lemma 8.3. We first rephrase it by slightly changing the nature and role of the template. After that, we state two main lemmas, and show how to deduce the (rephrased) intercalate-expander lemma from them (we will then spend the following sections of the paper proving these two lemmas).

Recall the sparsifying role of the template T in Lemma 8.3. To prove Lemma 8.3, we will take a slightly different view on T. Namely, for most of the proof it will be convenient to instead work with a template *hypergraph* H (which specifies a set of row/column/symbol triples, instead of a set of row/column pairs). Our template hypergraph will be obtained by sampling $\mathbb{G}^{(3)}(n, \varepsilon)$ with an appropriate ε, as specified in the following definition.

Definition 9.1. Throughout this and the following sections, we fix a constant $\beta > 0$ (cf. Lemma 8.3). We also let $\varepsilon = \eta \log^{-1} n$ for some $\eta = o(1)$ that slowly goes to 0 (concretely, take $\eta = 1/\log\log n$) and let $\ell = \log^{11} n$.

We now state a version of the intercalate-expander lemma with a random template hypergraph.

Lemma 9.2. *Fix a constant $\beta > 0$. Consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$ (both interpreted as subgraphs of $K_{n,n,n}^{(3)}$). Then*

$$\mathbb{P}[\mathbf{H} \cap \mathbf{L} \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \geq 1 - \exp(-\omega(n \log^2 n)).$$

Remark 9.3. Lemma 9.2 trivially implies that there is a hypergraph $H \subseteq K_{n,n,n}^{(3)}$ such that

$$\mathbb{P}[H \cap \mathbf{L} \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \geq 1 - \exp(-\omega(n \log^2 n)).$$

However, this would not have sufficed to prove Theorem 6.4 (we would have no guarantee that if we switch an intercalate which lies in H, the resulting switched intercalate still lies in H). The slightly different notion of a template in Definition 8.2 was chosen to avoid this issue.

We first show the (simple) deduction of Lemma 8.3 from Lemma 9.2. For this, we record the following simple fact, which we repeatedly use throughout the rest of the paper.

Fact 9.4. Suppose $\mathbf{H}_1 \in \mathcal{H}_1$ and $\mathbf{H}_2 \in \mathcal{H}_2$ are independent random objects, let $\mathcal{P}^{(2)} \subseteq \mathcal{H}_1 \times \mathcal{H}_2$, and for any $H_1 \in \mathcal{H}_1$ and $p \in [0,1]$ let $\mathcal{P}(p)$ be the set of all $H_1 \in \mathcal{H}_1$ such that $\mathbb{P}[(H_1, \mathbf{H}_2) \in \mathcal{P}^{(2)}] > p$. Then

(1) if $\mathbb{P}[(\mathbf{H}_1, \mathbf{H}_2) \in \mathcal{P}^{(2)}] \le p$, then $\mathbb{P}[\mathbf{H}_1 \in \mathcal{P}(\sqrt{p})] \le \sqrt{p}$.

(2) If $\mathbb{P}[\mathbf{H}_1 \in \mathcal{P}(p_2)] \le p_1$, then $\mathbb{P}[(\mathbf{H}_1, \mathbf{H}_2) \in \mathcal{P}^{(2)}] \le p_1 + p_2$.

Proof of Lemma 8.3. Say that $L \in \mathcal{L}_n$ is *good* if

$$\mathbb{P}[\mathbf{H} \cap L \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \ge 1 - \exp(-\omega(n \log^2 n)).$$

By Lemma 9.2 and Fact 9.4(1), $\mathbf{L}$ is good with probability $1 - \exp(-\omega(n \log^2 n))$.

Now, let $\mathbf{T} \subseteq [n]^2$ be a random subset of row/column pairs, where each is included with probability ε independently. Note that for each fixed outcome L of $\mathbf{L}$, the two random partial Latin squares $\mathbf{H} \cap L$ and $\mathbf{T} \cap L$ have exactly the same distribution. Therefore, for each good L,

$$\mathbb{P}[\mathbf{T} \cap L \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \ge 1 - \exp(-\omega(n \log^2 n)),$$

and Fact 9.4(2) yields

$$\mathbb{P}[\mathbf{T} \cap \mathbf{L} \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \ge 1 - \exp(-\omega(n \log^2 n)).$$

Applying Fact 9.4(1) again, we see that almost all outcomes T of $\mathbf{T}$ are suitable for the conclusion of Lemma 8.3 (namely, $\mathbf{T}$ satisfies the required property with probability $1 - \exp(-\omega(n \log^2 n))$). $\square$

We now move on to reducing Lemma 9.2 to two main lemmas. Recall that the property of being an intercalate-expander says that (with respect to any appropriate 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$) there is at least one stable intercalate consistent with our 6-tuple. Our first lemma (Lemma 9.11) will say that there are likely to be *many* intercalates consistent with every appropriate 6-tuple (and says nothing about isolatedness or stability). Our second lemma (Lemma 9.13) will say that there is likely to be a much smaller number of entries which participate in "bad sets" of intercalates which violate isolatedness or stability (i.e., they participate in a critical set of intercalates or a pair of intersecting intercalates).

Remark 9.5. It is crucially important here that we consider the *number of entries* in bad sets of intercalates (as opposed to the number of bad sets of intercalates themselves). This is due to the "infamous upper tail" problem which arises when trying to estimate the upper tail of subgraph counts in random graphs and hypergraphs. Due to clustering phenomena, these upper tails are 'fatter' than their lower counterparts (we need to pay very close attention to tail bounds due to the exponential error terms in Lemma 4.7).

Before stating our two lemmas, we need some preparations.

Definition 9.6. For sets $\mathscr{R}$ and $\mathscr{R}^*$ of rows, sets $\mathscr{C}$ and $\mathscr{C}^*$ of columns and sets $\mathscr{S}$ and $\mathscr{S}^*$ of symbols, we say that an intercalate is $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-*split* (or just *split* if the parameters are clear from context) if it has one row in $\mathscr{R}$, the other row in $\mathscr{R}^*$, one column in $\mathscr{C}$, the other column in $\mathscr{C}^*$, one symbol in $\mathscr{S}$, and the other symbol in $\mathscr{S}^*$.

By symmetry between the rows, columns, and symbols of a Latin square, in the context of Definition 8.1 it suffices to consider the case where $|\mathscr{R}^*| = \ell$ and $|\mathscr{R}|, |\mathscr{C}|, |\mathscr{C}^*|, |\mathscr{S}|, |\mathscr{S}^*| \ge \beta n$.

Definition 9.7. We say a 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$ is (ℓ, β)-*permissible* if we have $|\mathscr{R}^*| = \ell$ and $|\mathscr{R}|, |\mathscr{C}|, |\mathscr{C}^*|, |\mathscr{S}|, |\mathscr{S}^*| \ge \beta n$.

Our goal is then to show that, with very high probability, all (ℓ, β)-permissible 6-tuples have a split stable intercalate.

Definition 9.8. Let P be a partial Latin square, let $\{A_1, \ldots, A_t\}$ be a critical set of isolated intercalates in P, and let A' be an intercalate that can be created by switching a subset of $\{A_1, \ldots, A_t\}$. Then we say the entries in $A_1 \cup \cdots \cup A_t \cup (A' \cap P)$ comprise a *critical configuration* with respect to P.

We say that a set of entries in a partial Latin square P is a *bad configuration* with respect to P either if it is a critical configuration with respect to P, or if it comprises two intercalates that intersect in one entry. (We will from now on omit "with respect to P" if P is clear from the context.)

We say that a bad configuration is $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-*split* if at least one intercalate in it (that is, one of the at most four intercalates in the critical set of intercalates or one of the two intersecting intercalates, depending on the type of bad configuration) is $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split.

Remark 9.9. We will sometimes need to talk about sets of entries which make up some arrangement of intercalates, and we will sometimes need to talk about sets of intercalates themselves. To try to assist the reader to keep track of this distinction, we will reserve the word "configuration" for a set of entries in a partial Latin square (i.e. a set of hyperedges, in the hypergraph perspective in Fact 4.1).

Note that pairs of intercalates in a partial Latin square can intersect in at most one entry, so if an intercalate in a partial Latin square does not intersect any other intercalate in one entry, then it is isolated. Thus, an intercalate in a (partial) Latin square is stable if and only if it is not a subset of any bad configuration.

Instead of reasoning directly about the total number of split intercalates, it will be easier to show lower bounds on the size of the maximum family of disjoint split intercalates (as this random variable behaves better with respect to Theorem 3.1). The next lemma states that if we consider a random Latin square intersected with a random sparse template hypergraph, then there is very likely to be a large family of disjoint split intercalates.

Definition 9.10. For $c > 0$ and a tripartite hypergraph $H \subseteq K_{n,n,n}^{(3)}$, let $\mathcal{T}^{\mathrm{int}}(H, c) \subseteq \mathcal{L}_n$ be the set of Latin squares L such that for some (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$, the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $H \cap L$ has size smaller than $c\varepsilon^4 n\ell$.

Lemma 9.11. *Fix a constant $c > 0$ which is sufficiently small with respect to β. Then, consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{int}}(\mathbf{H}, c)] \leq \exp(-\omega(n \log^2 n)).$$

Next, the following lemma gives an upper bound on the number of entries in split bad configurations.

Definition 9.12. We say that an entry in a row in $\mathscr{R}^*$ is *covered* by some split bad configuration if it belongs to a split intercalate in it. We say an intercalate or a bad configuration is $\mathscr{R}^*$-split if it is $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split for some (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$ containing $\mathscr{R}^*$. For $C > 0$ and a hypergraph $H \subseteq K_{n,n,n}^{(3)}$, let $\mathcal{T}^{\mathrm{bad}}(H, C) \subseteq \mathcal{L}_n$ be the set of Latin squares L such that for some set of rows $\mathscr{R}^*$ of size ℓ, the number of entries in rows in $\mathscr{R}^*$ covered by $\mathscr{R}^*$-split bad configurations in $H \cap L$ is more than $C\varepsilon^5 n\ell$.

Lemma 9.13. *There is an absolute constant $C > 0$ such that the following holds. Consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C)] \leq \exp(-\omega(n \log^2 n)).$$

We can now deduce Lemma 9.2 from Lemmas 9.11 and 9.13.

Proof of Lemma 9.2. By the union bound, Lemma 9.11, and Lemma 9.13, there are $C, c > 0$ such that we have

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{int}}(\mathbf{H}, c) \cup \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C)] \leq \exp(-\omega(n \log^2 n)).$$

Note that if $\mathbf{L} \notin \mathcal{T}^{\mathrm{int}}(\mathbf{H}, c) \cup \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C)$, then for every (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$, there are many $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $\mathbf{H} \cap \mathbf{L}$, and only a small fraction of them can have entries covered by $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split bad configurations. Therefore, at least one of them must be stable. By symmetry between rows, columns, and symbols, and a union bound, we see that we have

$$\mathbb{P}[\mathbf{H} \cap \mathbf{L} \text{ is an } (\ell, \beta)\text{-intercalate-expander}] \geq 1 - \exp(-\omega(n \log^2 n)).$$

That is to say, almost all outcomes $H \cap L$ of $\mathbf{H} \cap \mathbf{L}$ satisfy the conclusion of Definition 8.1. □

It remains to prove Lemmas 9.11 and 9.13, which we will do in Sections 10 and 11, respectively. Both proofs will make crucial use of Lemma 4.7 (to deduce results about random Latin squares from very-high-probability results about the triangle removal process). Given Lemma 4.7, the proof of Lemma 9.11 is a rather quick (though somewhat delicate) consequence of Theorem 3.1. The proof of Lemma 9.13 is much longer, and requires several additional ideas.

10. Existence of many disjoint intercalates

As already mentioned, to prove Lemma 9.11, we employ Lemma 4.7 and work in the setting of the triangle removal process. To do that, we define a property of ordered partial Latin squares that satisfies suitable inheritance properties with respect to $\mathcal{T}^{\mathrm{int}}(H, c)$ (recall Definition 4.3).

Definition 10.1. Recall the definitions of ℓ and ε from Definition 9.1. For $c > 0$ and $H \subseteq K_{n,n,n}^{(3)}$, let $\vec{\mathcal{U}}^{\mathrm{int}}(H,c) \subseteq \vec{\mathcal{P}}_n$ be the set of ordered partial Latin squares $\vec{P}$ such that for some (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$, the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $H \cap \vec{P}$ has size less than $c\varepsilon^4 n\ell$.

Now, the following lemma is a version of Lemma 9.11 for the triangle removal process.

Lemma 10.2. *Fix constants $\alpha, c > 0$, such that α is sufficiently small and c is sufficiently small in terms of α and β. Consider independent random hypergraphs $\vec{\mathbf{R}} \sim \mathrm{TRP}(n, \alpha n^2)$ and $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$. Then*
$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{int}}(\mathbf{H}, c)] \le \exp(-\omega(n \log^2 n)).$$

We first deduce Lemma 9.11 from Lemma 10.2 using Lemma 4.7.

Proof of Lemma 9.11. Let α, c be small enough for Lemma 10.2 and let $\rho = 1$. By Lemma 10.2 and Fact 9.4(1), with probability $1 - \exp(-\omega(n \log^2 n)$ our random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ is such that $\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{int}}(\mathbf{H}, c) \mid \mathbf{H}] \le \exp(-\omega(n \log^2 n))$. Let H be an outcome of $\mathbf{H}$ for which this holds. Note that $\vec{\mathcal{U}}^{\mathrm{int}}(H, c)$ is $(\rho, \alpha n^2)$-inherited from $\mathcal{T}^{\mathrm{int}}(H, c)$: indeed, if $L \in \mathcal{T}^{\mathrm{int}}(H, c)$, then there is some (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$ such that the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $H \cap L$ has size smaller than $c\varepsilon^4 n\ell$, so for any $P \subseteq L$, the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $H \cap P$ also has size smaller than $c\varepsilon^4 n\ell$. Thus, by Lemma 4.7,
$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{int}}(H, c)] \le \exp(2n \log^2 n)\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{int}}(H, c)] \le \exp(-\omega(n \log^2 n)).$$
Recalling our choice of H, Fact 9.4(2) then implies that
$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{int}}(\mathbf{H}, c)] \le \exp(-\omega(n \log^2 n)),$$
as desired. $\qquad\square$

Proof of Lemma 10.2. We will prove the desired statement with $c = 10^{-5}\beta^5\alpha^4$.

Note that the property $\vec{\mathcal{U}}^{\mathrm{int}}(H, c)$ is monotone decreasing. Let $\mathbf{G} \sim \mathbb{G}^{(3)}(n, \alpha/n)$, and let $\mathbf{G}^*$ be obtained from $\mathbf{G}$ by deleting all hyperedges involved in conflicts (recall that a conflict is a pair of hyperedges that intersect in more than one vertex). By Lemma 5.4, for each $H \subseteq K_{n,n,n}^{(3)}$ we have
$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{int}}(H, c)] = O(\mathbb{P}[\mathbf{G}^* \in \vec{\mathcal{U}}^{\mathrm{int}}(H, c)]),$$
so for our random hypergraph $\mathbf{H}$ we have
$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{int}}(\mathbf{H}, c)] = O(\mathbb{P}[\mathbf{G}^* \in \vec{\mathcal{U}}^{\mathrm{int}}(\mathbf{H}, c)]).$$
Thus, from now on we work in $\mathbf{G}^*$. Fix an (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$. We will show that the probability that the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $\mathbf{H} \cap \mathbf{G}^*$ has size smaller than $c\varepsilon^4 n\ell$ is at most $\exp(-\omega(n \log^2 n))$.

Our main tool will be Theorem 3.1, but to effectively apply this concentration inequality we need to be quite careful with the way we interpret $\mathbf{H} \cap \mathbf{G}^*$ as a function of independent choices. Specifically, for each potential hyperedge e, let $\mathbf{g}_e^{\mathbf{h}}$ and $\mathbf{g}_e$ be two independent Bernoulli random variables with $\mathbb{P}[\mathbf{g}_e^{\mathbf{h}} = 1] = \alpha\varepsilon/n$ and $\mathbb{P}[\mathbf{g}_e = 1] = \alpha(1-\varepsilon)/(n - \alpha\varepsilon)$. We generate $\mathbf{G}$ and $\mathbf{H} \cap \mathbf{G}$ by sampling $\mathbf{g}_e^{\mathbf{h}}$ and $\mathbf{g}_e$ for each potential edge e, setting $e \in \mathbf{H} \cap \mathbf{G}$ if $\mathbf{g}_e^{\mathbf{h}} = 1$, and setting $e \in \mathbf{G}$ if $\mathbf{g}_e^{\mathbf{h}} = 1$ or $\mathbf{g}_e = 1$ (or both). To see that this correctly describes the joint distribution of $\mathbf{G}$ and $\mathbf{H} \cap \mathbf{G}$, note that
$$\mathbb{P}[e \in \mathbf{G}] = 1 - (1 - \mathbb{P}[\mathbf{g}_e^{\mathbf{h}} = 1])(1 - \mathbb{P}[\mathbf{g}_e = 1]) = 1 - \left(1 - \frac{\alpha\varepsilon}{n}\right)\left(1 - \frac{\alpha(1-\varepsilon)}{n - \alpha\varepsilon}\right)$$
$$= \frac{\alpha\varepsilon}{n} + \frac{\alpha(1-\varepsilon)}{n - \alpha\varepsilon} - \frac{\alpha^2\varepsilon(1-\varepsilon)}{n(n - \alpha\varepsilon)} = \frac{\alpha}{n}.$$

Now, the plan is to reveal $\mathbf{G}$ and $\mathbf{H}$ (via the random variables $\mathbf{g}_e^{\mathbf{h}}$ and $\mathbf{g}_e$) in two phases. First, we reveal all entries in rows outside of $\mathscr{R}^*$, and then we reveal entries in rows in $\mathscr{R}^*$.

Recall that a conflict is a pair of hyperedges that intersect in more than one vertex. We say a pair of entries $\{(r, c_1, s_1), (r, c_2, s_2)\}$ in $\mathbf{H} \cap \mathbf{G}$ on the same row $r \notin \mathscr{R}^*$ has a *second-order conflict* if s_1 appears in column c_2 or if s_2 appears in column c_1 on a row outside of $\mathscr{R}^*$, with respect to $\mathbf{G}$. For $r \in \mathscr{R} \setminus \mathscr{R}^*$, $c_1 \in \mathscr{C}$, $c_2 \in \mathscr{C}^*$, $s_1 \in \mathscr{S}$ and $s_2 \in \mathscr{S}^*$, we say a pair $\{(r, c_1, s_1), (r, c_2, s_2)\}$ is a *half-intercalate* if both its hyperedges are present in $\mathbf{H} \cap \mathbf{G}$ and have no conflicts with entries of $\mathbf{G}$ outside of $\mathscr{R}^*$, and the pair has no second-order conflicts (which would make it impossible to complete it to an intercalate in $\mathbf{G}^*$).

The idea is that in the first phase, by revealing all the entries in the rows outside $\mathscr{R}^*$, we reveal all the half-intercalates. In the second phase, each half-intercalate $\{(r, c_1, s_1), (r, c_2, s_2)\}$ can then be completed to a split intercalate in $\mathbf{H} \cap \mathbf{G}^*$ via any pair of entries $(r', c_1, s_1), (r', c_2, s_2)$ (with $r' \in \mathscr{R}^*$) in $\mathbf{G}$, as long as the four entries involved in the intercalate do not conflict with other entries in $\mathbf{G}$ on rows in $\mathscr{R}^*$.

Let $\mathbf{X}$ be the number of half-intercalates. We will apply Theorem 3.1 to show concentration of $\mathbf{X}$. First note that

$$\mathbb{E}[\mathbf{X}] \geq \binom{\beta n}{2}^2 (\beta n - \ell)(\varepsilon \alpha/n)^2 (1 - \alpha/n)^{8n-8} \geq 10^{-1} \varepsilon^2 \alpha^2 (\beta n - \ell) \beta^4 n^2 e^{-16\alpha} \geq 10^{-2} \varepsilon^2 \alpha^2 \beta^5 n^3.$$

To see this, recall that $|\mathscr{R} \setminus \mathscr{R}^*| \geq \beta n - \ell$ and each of $\mathscr{R}, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*$ have size at least βn; also, each pair of entries is present in $\mathbf{H} \cap \mathbf{G}$ with probability $(\varepsilon \alpha/n)^2$, it has no conflicts with other entries in $\mathbf{G}$ with probability at least $(1 - \alpha/n)^{6n-6}$, and it has no second-order conflicts with probability at least $(1 - \alpha/n)^{2n-2}$.

Recall that we work with the independent random variables $\mathbf{g}_e^{\mathbf{h}}$ and $\mathbf{g}_e$ for each of the $O(n^3)$ many potential hyperedges e outside the rows in $\mathscr{R}^*$, and each of these random variables is 1 with probability $O(1/n)$. The appearance of a hyperedge in $\mathbf{H} \cap \mathbf{G}$ can increase $\mathbf{X}$ by at most $n - 1$, since each entry can be in at most $n - 1$ half-intercalates. The appearance of a hyperedge in $\mathbf{G}$ can decrease $\mathbf{X}$ by at most $4n$, since it can destroy at most 3 hyperedges in $\mathbf{G}^*$ due to conflicts (and each of them could be in at most $n - 1$ pairs counted by $\mathbf{X}$), and can introduce at most n second-order conflicts. Thus, the effect of changing any individual random variable (of the form $\mathbf{g}_e^{\mathbf{h}}$ or $\mathbf{g}_e$) is $O(n)$. By Theorem 3.1, with probability $1 - \exp(-\Omega(\varepsilon^4 n^2))$, we have $\mathbf{X} \geq 10^{-3} \varepsilon^2 \alpha^2 \beta^5 n^3$. For the rest of the proof, we condition on an outcome of the first phase such that this is the case (i.e., in the rest of the proof, all probabilistic considerations are implicitly with respect to this conditional probability space).

Now, in the second phase we reveal $\mathbf{G}$ and $\mathbf{H} \cap \mathbf{G}$ in rows in $\mathscr{R}^*$, and estimate the size $\mathbf{Z}$ of the maximum family of disjoint $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$-split intercalates in $\mathbf{H} \cap \mathbf{G}^*$. Note that each half-intercalate counted by $\mathbf{X}$ forms an intercalate with two other entries on row $r \in \mathscr{R}^*$ if and only if these two other entries are present in $\mathbf{H} \cap \mathbf{G}$ and no entry in $\mathbf{G}$ in a row in $\mathscr{R}^*$ conflicts with any of the four entries of the intercalate. We can compute a lower bound on $\mathbb{E}[\mathbf{Z}]$ by subtracting the expected number of pairs of intersecting split intercalates from the expected number of split intercalates. Indeed,

$$\mathbb{E}[\mathbf{Z}] \geq \mathbf{X} \ell (\varepsilon \alpha/n)^2 (1 - \alpha/n)^{8n-8} - \mathbf{X} n \ell^2 (\varepsilon \alpha/n)^4 - \mathbf{X} n \ell (\varepsilon \alpha/n)^3 \geq \mathbf{X} \frac{\ell \alpha^2 \varepsilon^2}{2n^2} \geq 10^{-4} \alpha^4 \beta^5 \varepsilon^4 n \ell,$$

where the term $(1 - \alpha/n)^{8n-8}$ accounts for the probability that some entry in a fixed potential split intercalate conflicts with some entry in a row in $\mathscr{R}^*$. The second term bounds the expected number of pairs of split intercalates that intersect in an entry outside of $\mathscr{R}^*$, since for a fixed half-intercalate, there are at most n choices for the other column of the other intercalate and at most ℓ choices for the other row of each intercalate, and for each such combination of choices, four entries need to be present in rows in $\mathscr{R}^*$. The third term bounds the expected number of pairs of split intercalates that intersect in an entry in $\mathscr{R}^*$, since for a fixed half-intercalate, there are at most ℓ choices for the common to both intercalates row in $\mathscr{R}^*$ and at most n choices for the other row of the other intercalate, which is outside of $\mathscr{R}^*$ (and this determines the entries of both intercalates completely). We again apply Theorem 3.1 with the remaining unrevealed random variables $\mathbf{g}_e^{\mathbf{h}}$ and $\mathbf{g}_e$. Since we are interested in a family of disjoint intercalates, the appearance of a hyperedge can increase $\mathbf{Z}$ by at most 1 and can decrease it by at most 3 (as it can have a conflict with at most 3 edges in $\mathbf{G}^*$). Crucially, there are now only $O(n^2 \ell)$ many unrevealed random variables. Thus, by Theorem 3.1, we have that $\mathbf{Z} \geq 10^{-5} \alpha^4 \beta^5 \varepsilon^4 n \ell$ with probability at least $1 - \exp(-\Omega_{\alpha,\beta}(\varepsilon^8 n \ell))$.

We conclude the proof with a union bound over all (at most 2^{6n}) choices for an (ℓ, β)-permissible 6-tuple $(\mathscr{R}, \mathscr{R}^*, \mathscr{C}, \mathscr{C}^*, \mathscr{S}, \mathscr{S}^*)$. $\qquad\square$

11. Upper-bounding the number of entries in bad configurations

We now turn to the proof of Lemma 9.13. For any set of ℓ rows $\mathscr{R}^*$, our goal is to upper bound the probability that there are many entries in rows in $\mathscr{R}^*$ covered by $\mathscr{R}^*$-split bad configurations.

11.1. **Setup.** As in the proof of Lemma 10.2, the special role of the rows in $\mathscr{R}^*$ requires us to consider two "phases". In the first phase we reveal a subset of the entries of our random Latin square, and study configurations of entries which are "almost" bad configurations (they are just missing some entries from $\mathscr{R}^*$). In the second phase we reveal the rest of our random Latin square, and we study which of our

"almost bad configurations" give rise to actual bad configurations. However, this breakdown into phases is much more complicated than for Lemma 10.2 (for several different reasons), and requires some setup.

Definition 11.1. Given a $\mathscr{R}^*$-split bad configuration F, we say that the *special entries* of F are its two entries which belong to rows in $\mathscr{R}^*$. (Technically, it is possible that there is more than one split intercalate in a bad configuration, or that the bad configuration contains an intercalate with both its rows in $\mathscr{R}^*$. In these cases we imagine multiple "copies" of the bad configuration, each time making a different choice for the two special entries). Also, we say that the intercalate in F which contains the special entries is the *special intercalate*.

Definition 11.2. Given a partial Latin square P and a set of rows $\mathscr{R}^*$, consider a pair $\{e_1, e_2\}$ of row/column/symbol triples, belonging to some row in $\mathscr{R}^*$, which are not already present in P but which can be added to P without causing conflicts (i.e., $P \cup \{e_1, e_2\}$ is a partial Latin square). Say that $\{e_1, e_2\}$ is a *threatened pair* for P if there is some split bad configuration F in $P \cup \{e_1, e_2\}$ which has e_1 and e_2 as special entries. In this case, we say $F \cap P$ is a *threat configuration* in P.

Informally speaking, threat configurations are the sets of entries which are "in danger of being completed to a split bad configuration", only requiring the addition of their two special entries. Each of the possible ways to complete a threat configuration to a split bad configuration is described by a threatened pair. A single threatened pair may complete multiple different threat configurations to split bad configurations, and a single threat configuration may be completeable by multiple different threatened pairs (to different split bad configurations).

Now, given the above terminology, we can be a bit more concrete about the high-level plan to prove Lemma 9.13. In a first phase, we reveal part of our random Latin square and upper-bound the number of threatened pairs in it (we do *not* upper bound the number of threat configurations themselves, cf. Remark 9.5). Then, in a second phase, we reveal the rest of the Latin square and upper bound the number of threatened pairs from the first phase that actually appear as entries in the second phase. By an averaging argument (over all the ways to split into two phases) we are able to deduce the desired bound on the number of entries in split bad configurations.

For the first phase, we consider the following property.

Definition 11.3. Recall the definitions of ℓ and ε from Definition 9.1. For $H \subseteq K_{n,n,n}^{(3)}$ and $C, \alpha, \phi > 0$, let $\mathcal{T}^{\mathrm{threat}}(H, C, \alpha, \phi) \subseteq \mathcal{L}_n$ be the set of Latin squares L such that the following holds. If we take a random subset $\mathbf{P}_{\alpha n^2}(L) \in \mathcal{P}_{n, \alpha n^2}$ of αn^2 entries in L, then with probability at least ϕ, for some set of rows $\mathscr{R}^*$ of size $|\mathscr{R}^*| = \ell$, the number of threatened pairs for $H \cap \mathbf{P}_{\alpha n^2}(L)$ is more than $C \varepsilon^3 n^3 \ell$.

Note that here we are defining a property of Latin squares in terms of how their random subsets of entries behave. This may seem rather unwieldy, but it is very convenient for the ρ-inheritedness assumption in Lemma 4.7 (when it comes time to apply Lemma 4.7, the ρ-inheritedness assumption will hold basically by definition). The issue with more straightforward properties is that bad configurations and threatened pairs do not behave well under subsampling entries, and therefore ρ-inheritedness would be quite difficult to prove. (The issue is that a bad configuration is defined in terms of *isolated* intercalates, and deleting entries can cause more intercalates to be isolated[14]).

Now, the following two lemmas correspond to the two "phases" informally discussed above.

Lemma 11.4. *There is an absolute constant $C' > 0$ such that the following holds. Fix constants $\alpha, \phi > 0$, and consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{threat}}(\mathbf{H}, C', \alpha, \phi)] \leq \exp(-\omega(n \log^2 n)).$$

Lemma 11.5. *Fix constants $\alpha, C', C, \phi > 0$, such that α is sufficiently small, C is sufficiently large in terms of α and C', and ϕ is sufficiently small in terms of α. Consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \mathrm{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C) \setminus \mathcal{T}^{\mathrm{threat}}(\mathbf{H}, C', \alpha, \phi)] \leq \exp(-\omega(n \log^2 n)).$$

Lemmas 11.4 and 11.5 are both proved using Lemma 4.7, comparing random Latin squares to the triangle removal process. In addition, the proof of Lemma 11.4 also involves a switching argument directly

---

[14] The reader may then wonder why we insist on isolatedness in the definition of a bad configuration. The reason is that it is very convenient to restrict our attention to isolated intercalates in a switching argument which will appear later in this section.

on random Latin squares (to tame the complexity of all the different possibilities for the structure of a threat configuration).

We conclude this subsection with the brief deduction of Lemma 9.13 from Lemmas 11.4 and 11.5.

Proof of Lemma 9.13. Let C' be as in Lemma 11.4, and then let α, C, ϕ be as in Lemma 11.5 (for the same value of C'). Then we have

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C)] \leq \mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{bad}}(\mathbf{H}, C) \setminus \mathcal{T}^{\mathrm{threat}}(\mathbf{H}, C', \alpha, \phi)] + \mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{threat}}(\mathbf{H}, C', \alpha, \phi)]$$

$$\leq \exp(-\omega(n \log^2 n)) + \exp(-\omega(n \log^2 n))$$

by Lemmas 11.4 and 11.5. $\qquad\qquad\square$

In the next three subsections, we will prove Lemmas 11.4 and 11.5 (in reverse order).

11.2. **Few bad configurations or many threatened pairs.** In this subsection we prove Lemma 11.5. As previously mentioned, we will work with the triangle removal process, via Lemma 4.7. In contrast to our previous application of Lemma 4.7 in Section 10, this time the order of the edges will be important.

Definition 11.6. Let $\vec{P} \in \vec{\mathcal{P}}_n$ be an ordered partial Latin square, and write $e_1, \ldots, e_m$ for its entries (in order). For $\iota, \kappa \in [0, 1]$, we write $\vec{P}[\iota, \kappa] \in \mathcal{P}_{n,(\kappa-\iota)m}$ for the partial Latin square consisting of the edges e_i with $\iota m < i \leq \kappa m$.

Recall the definitions of ℓ and ε from Definition 9.1.

Definition 11.7. Let $\alpha \in (0, 1)$, $C > 0$, and $H \subseteq K^{(3)}_{n,n,n}$. Let $\vec{\mathcal{U}}^{\mathrm{threat}}(H, C) \subseteq \vec{\mathcal{P}}_n$ be the set of partial ordered Latin squares $\vec{P}$ such that for some set of rows $\mathscr{R}^*$ of size $\ell = |\mathscr{R}^*|$, the number of threatened pairs for $H \cap \vec{P}[0, 1/2]$ is more than $C\varepsilon^3 n^3 \ell$.

Definition 11.8. We call a bad configuration *consistent with* $\vec{P}$ if its corresponding threat configuration is in $\vec{P}[0, 1/2]$, one of its special entries is in $\vec{P}[1/2, 3/4]$, and its other special entry is in $\vec{P}[3/4, 1]$. Let $\vec{\mathcal{U}}^{\mathrm{bad}}(H, C) \subseteq \vec{\mathcal{P}}_n$ be the set of partial ordered Latin squares $\vec{P}$ such that for some set of rows $\mathscr{R}^*$ of size $\ell = |\mathscr{R}^*|$, there are more than $C\varepsilon^5 n\ell$ entries in $\vec{P}[3/4, 1]$ that are in rows in $\mathscr{R}^*$ and are covered by $\mathscr{R}^*$-split bad configurations in $H \cap \vec{P}$ that are consistent with $\vec{P}$.

The following lemma is a version of Lemma 11.5 for the triangle removal process.

Lemma 11.9. *Fix constants α, C', such that α is sufficiently small. Consider independent random hypergraphs $\vec{\mathbf{R}} \sim \mathrm{TRP}(n, \alpha n^2)$ and $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$. Then*

$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{bad}}(\mathbf{H}, C') \setminus \vec{\mathcal{U}}^{\mathrm{threat}}(\mathbf{H}, C')] \leq \exp(-\omega(n \log^2 n)).$$

Before proving Lemma 11.9, we show how it implies Lemma 11.5, using Lemma 4.7 and an averaging argument.

Proof of Lemma 11.5. Let $m = 2\alpha n^2$ and $\rho = \alpha^{16}/2^4$. We assume that $\phi \leq \rho$ and $\rho C \geq C'$.

By Lemma 11.9 and Fact 9.4(1), with probability $1 - \exp(-\omega(n \log^2 n))$ over the randomness of $\mathbf{H}$ we have

$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\mathrm{bad}}(\mathbf{H}, C') \setminus \vec{\mathcal{U}}^{\mathrm{threat}}(\mathbf{H}, C') \,|\, \mathbf{H}] \leq \exp(-\omega(n \log^2 n)).$$

The desired result would therefore follow from Lemma 4.7 and Fact 9.4(2), if we knew that $\vec{\mathcal{U}}^{\mathrm{bad}}(H, C') \setminus \vec{\mathcal{U}}^{\mathrm{threat}}(H, C')$ is (ρ, m)-inherited from $\mathcal{T}^{\mathrm{bad}}(H, C) \setminus \mathcal{T}^{\mathrm{threat}}(H, C', \alpha, \phi)$ for all outcomes H of $\mathbf{H}$. For the rest of the proof, our goal is to prove that this is the case.

Suppose $L \in \mathcal{T}^{\mathrm{bad}}(H, C) \setminus \mathcal{T}^{\mathrm{threat}}(H, C', \alpha, \phi)$, and let $\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{P}}_{n,m}$ be a random ordering of m random entries of L. We need to prove that $\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}^{\mathrm{bad}}(H, C') \setminus \vec{\mathcal{U}}^{\mathrm{threat}}(H, C')$ with probability at least ρ.

By the definition of $\mathcal{T}^{\mathrm{threat}}(H, C', \alpha, \phi)$ and the fact that $\vec{\mathbf{P}}_m(L)[0, 1/2]$ is a uniformly random set of αn^2 entries of L, first note that $\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}^{\mathrm{threat}}(H, C')] \leq \phi$. So, it suffices to show that $\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}^{\mathrm{bad}}(H, C')$ with probability at least $\rho + \phi$.

Since $L \in \mathcal{T}^{\mathrm{bad}}(H, C)$, there is some $\mathscr{R}^*$ such that there are $X \geq C\varepsilon^5 n\ell$ many entries in rows in $\mathscr{R}^*$ covered by $\mathscr{R}^*$-split bad configurations in $H \cap L$. Let $Y \leq X$ be the number of such entries in $\vec{\mathbf{P}}_m(L)[3/4, 1]$ which are covered by $\mathscr{R}^*$-split bad configurations in $H \cap \vec{\mathbf{P}}_m(L)$ consistent with $\vec{\mathbf{P}}_m(L)$. For each covered entry (r, c, s) in $H \cap L$ counted by X, the probability that its associated split bad

configuration with $s \leq 16$ entries is present in $\vec{\mathbf{P}}_m(L)$ is $\binom{n^2-s}{m-s}/\binom{n^2}{m} = (1+o(1))(2\alpha)^s$. Given that it is present in $\vec{\mathbf{P}}_m(L)$, the probability that it is consistent with $\vec{\mathbf{P}}_m(L)$ and that $(r,c,s) \in \vec{\mathbf{P}}_m(L)[3/4,1]$ is

$$\frac{m/4}{m} \cdot \frac{m/4}{m-1} \cdot \binom{m/2}{s-2}/\binom{m-2}{s-2} = (1+o(1))(1/4)^2(1/2)^{s-2}.$$

Note that each isolated intercalate in $H \cap L$ whose entries are in $\vec{\mathbf{P}}_m(L)$ is also isolated in $H \cap \vec{\mathbf{P}}_m(L)$, so each bad configuration in $H \cap L$ that is present in $H \cap \vec{\mathbf{P}}_m(L)$ is also a bad configuration in $H \cap \vec{\mathbf{P}}_m(L)$. Therefore

$$\mathbb{E}[\mathbf{Y}] \geq (1+o(1))X(2\alpha)^s(1/4)^2(1/2)^{s-2} \geq X\alpha^{16}/4.$$

Recall our assumptions on ϕ, C from the start of the proof. Let $\delta = \alpha^{16}/8 \geq \rho + \phi$ and note that $\delta X \geq C'\varepsilon^5 n\ell$. By Markov's inequality, $\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}^{\mathrm{bad}}(H, C')]$ is at least

$$\mathbb{P}[\mathbf{Y} \geq \delta X] = 1 - \mathbb{P}[X - \mathbf{Y} > (1-\delta)X] \geq 1 - \frac{\mathbb{E}[X-\mathbf{Y}]}{(1-\delta)X} \geq 1 - \frac{1-2\delta}{1-\delta} \geq \rho + \phi,$$

as desired. $\qquad\square$

Now we prove Lemma 11.9.

Proof of Lemma 11.9. First, we reveal $\mathbf{H} \cap \vec{\mathbf{R}}[0, 1/2]$. We assume that for each choice of $\mathscr{R}^*$ of size ℓ, the number of threatened pairs for $\mathbf{H} \cap \vec{\mathbf{R}}[0, 1/2]$ is at most $C'\varepsilon^3 n^3 \ell$ (otherwise $\vec{\mathcal{U}}^{\mathrm{threat}}(\mathbf{H}, C')$ holds, and there is nothing to prove). From now on, when we talk about threatened pairs in this proof, we mean threatened pairs for $\mathbf{H} \cap \vec{\mathbf{R}}[0, 1/2]$.

Given revealed information, our goal is to show that with probability $1 - \exp(-\omega(n\log^2 n))$, for every $\mathscr{R}^*$ with $|\mathscr{R}^*| = \ell$, there are at most $C'\varepsilon^5 n\ell$ entries of $\vec{\mathbf{R}}[3/4, 1]$ that are in rows in $\mathscr{R}^*$ and are covered by $\mathscr{R}^*$-split bad configurations in $\mathbf{H} \cap \vec{\mathbf{R}}$ that are consistent with $\vec{\mathbf{R}}$. Actually, it suffices to show this for a fixed choice of $\mathscr{R}^*$ (a union bound over choices of $\mathscr{R}^*$ has a negligible impact). So, fix some $\mathscr{R}^*$ with $|\mathscr{R}^*| = \ell$.

The next step is to consider $\mathbf{H} \cap \vec{\mathbf{R}}[1/2, 3/4]$. We say a threatened pair *survives* if one of its entries is in $\mathbf{H} \cap \vec{\mathbf{R}}[1/2, 3/4]$.

Claim. *Given revealed information, the number of surviving threatened pairs is at most $C'\varepsilon^4 n^2 \ell$ with probability $1 - \exp(-\omega(n\log^2 n))$.*

Proof of claim. Let $\mathcal{U}$ be the set of $G \subseteq K_{n,n,n}^{(3)}$ which intersect more than $C'\varepsilon^4 n^2 \ell$ threatened pairs. Then, for each possible outcome H of $\mathbf{H}$, the set $\{G \subseteq K_{n,n,n}^{(3)} : H \cap G \in \mathcal{U}\}$ is a monotone increasing property of subgraphs of $K_{n,n,n}^{(3)}$, so by Lemma 5.3, the probability that there are more than $C'\varepsilon^4 n^2 \ell$ surviving threatened pairs is

$$\mathbb{P}[\mathbf{H} \cap \vec{\mathbf{R}}[1/2, 3/4] \in \mathcal{U}] \leq \mathbb{P}[\mathbf{H} \cap \mathbf{G}' \in \mathcal{U}] + \exp(-\Omega(n^2)) = \mathbb{P}[\mathbf{G} \in \mathcal{U}] + \exp(-\Omega(n^2)),$$

where $\mathbf{G}' \sim \mathbb{G}^{(3)}(n, (\alpha/2)/n)$ and $\mathbf{G} \sim \mathbb{G}^{(3)}(n, (\varepsilon\alpha/2)/n)$.

From here, we just need to apply a concentration inequality in $\mathbf{G}$, to upper bound $\mathbb{P}[\mathbf{G} \in \mathcal{U}]$. For each row/column/symbol triple e, let w_e be the number of threatened pairs that e is in, and let $\mathbf{X}$ be the sum of w_e over all hyperedges e of $\mathbf{G}$. Note that (for small $\alpha > 0$) we have

$$\mathbb{E}[\mathbf{X}] \leq 2 \cdot C'\varepsilon^3 n^3 \ell \cdot \frac{\varepsilon\alpha/2}{n} \leq \frac{C'\varepsilon^4 n^2 \ell}{2}.$$

Then, for any row/column/symbol triple e, consider each of the threatened pairs that e is in. Each of these threatened pairs, if added to $\vec{\mathbf{R}}[1/2, 1]$, would complete a split bad configuration in which e would be in a different intercalate. Since a single entry can be in at most n many intercalates (in any Latin square), it follows that e is in at most n threatened pairs; that is, $w_e \leq n$.

We can therefore apply Corollary 3.2 with $\Delta = n$ and $\delta = 1$, to obtain

$$\mathbb{P}[\mathbf{G} \in \mathcal{U}] \leq \mathbb{P}\left[\mathbf{X} \geq \mathbb{E}[\mathbf{X}] + \frac{C'\varepsilon^4 n^2 \ell}{2}\right] \leq \exp(-\Omega(\varepsilon^4 n\ell)),$$

and the desired result follows (recalling the relationship between ℓ and ε in Definition 9.1). $\qquad\blacksquare$

Now, reveal $\mathbf{H} \cap \vec{\mathbf{R}}[1/2, 3/4]$; by the above claim we may assume that our revealed outcome is such that there are at most $C'\varepsilon^4 n^2 \ell$ surviving threatened pairs. Then, the final step is to consider $\mathbf{H} \cap \vec{\mathbf{R}}[3/4, 1]$.

Each surviving threatened pair's presence in $\mathbf{H} \cap \vec{\mathbf{R}}$ is determined by whether its second entry is in $\mathbf{H} \cap \vec{\mathbf{R}}[3/4, 1]$. Say a row/column/symbol triple e is a *threatened entry* if it is the second entry of some surviving threatened pair.

Claim. *Given revealed information, there are at most $C'\varepsilon^5 n\ell$ threatened entries present in $\mathbf{H} \cap \vec{\mathbf{R}}[3/4, 1]$, with probability $1 - \exp(-\omega(n \log^2 n))$.*

Proof of claim. We proceed very similarly to the last claim (in fact, the situation is even simpler). Let $\mathbf{G} \sim \mathbb{G}^{(3)}(n, (\alpha\varepsilon/2)/n)$, and let $\mathbf{Y}$ be the number of threatened entries present in $\mathbf{G}$. Using Lemma 5.3 as before, the probability that there are more than $C'\varepsilon^5 n\ell$ threatened entries present in $\mathbf{H} \cap \vec{\mathbf{R}}[3/4, 1]$ is at most $\mathbb{P}[\mathbf{Y} > C'\varepsilon^5 n\ell] + \exp(-\Omega(n^2))$.

We have
$$\mathbb{E}[\mathbf{Y}] \leq C'\varepsilon^4 n^2 \ell \cdot \frac{\alpha\varepsilon/2}{n} \leq \frac{C'\varepsilon^5 n\ell}{2},$$
so by Corollary 3.2 (with $\Delta = 1$), we have $\mathbb{P}[\mathbf{Y} > C'\varepsilon^5 n\ell] \leq \exp(-\Omega(\varepsilon^5 n\ell))$; the desired result follows. $\blacksquare$

Now, if e is an entry of $\vec{\mathbf{R}}[3/4, 1]$ that is in a row in $\mathscr{R}^*$ and is covered by a split bad configuration in $\mathbf{H} \cap \vec{\mathbf{R}}$ consistent with $\vec{\mathbf{R}}$, then e was a threatened entry which happened to be present in $\mathbf{H} \cap \vec{\mathbf{R}}[3/4, 1]$ (the converse may not hold, as bad configurations in subsets of $\mathbf{H} \cap \vec{\mathbf{R}}$ are not always bad configurations in $\mathbf{H} \cap \vec{\mathbf{R}}$ itself). So, the desired result follows from the above claim. $\square$

11.3. **Upper-bounding the number of threatened pairs: setup and switching.** It remains to prove Lemma 11.4, which amounts to a careful study of threat configurations in a random subset of a random Latin square.

Recall that threat configurations are sets of entries that are "two entries away" from a bad configuration, and there are a very large number of possibilities for the structure of a bad configuration (e.g., critical configurations can have different numbers of intercalates, they can require different subsets of intercalates to be switched to create a new intercalate, and they can have different rows in $\mathscr{R}^*$). To tame this complexity, one crucial observation is that for every bad configuration in a partial Latin square, there is some way to switch isolated intercalates to obtain a pair of intersecting intercalates. So, with a switching argument (unrelated to the main switching argument in Section 8), we can reduce the number of cases significantly: it suffices to consider bad configurations that correspond to pairs of intersecting intercalates, and bad configurations that are "one switch away" from a pair of intersecting intercalates.

Definition 11.10. Given a partial Latin square $P \in \mathcal{P}_n$ and a set of rows $\mathscr{R}^*$, we define a *basic split bad configuration* to be a set F of seven entries in P, such that F contains an intercalate A with a row in $\mathscr{R}^*$, and either F or $(F \setminus A) \cup \bar{A}$ is a union of two intersecting intercalates.

Then, say that a pair $\{e_1, e_2\}$, belonging to some row in $\mathscr{R}^*$, is a *basic threatened pair* for P if there is some basic split bad configuration F in $P \cup \{e_1, e_2\}$ with $e_1, e_2 \in F$. In this case, we say $F \setminus \{e_1, e_2\}$ is a *basic threat configuration* in P.

There are four cases for the structure of a basic threatened pair/basic threat configuration; see Figure 1. Note that we make no isolatedness assumptions in the above definition, so technically it is possible that a basic split bad configuration is not actually a split bad configuration (or that a basic threatened pair is not actually a threatened pair, or that a basic threat configuration is not actually a threat configuration). However, we still borrow the same terminology (e.g., we may talk about the special entries of a basic split bad configuration).

	c_1	c_2	c_3
r_1^*	s_1	s_2	
r_2	s_2	s_1	s_3
r_3		s_3	s_2

	c_1	c_2	c_3
r_1^*	s_2	s_1	
r_2	s_1	s_2	s_3
r_3		s_3	s_2

	c_1	c_2	c_3
r_1	s_1	s_2	
r_2^*	s_2	s_1	s_3
r_3		s_3	s_2

	c_1	c_2	c_3
r_1	s_2	s_1	
r_2^*	s_1	s_2	s_3
r_3		s_3	s_2

Figure 1. The four possibilities for the structure of a basic threat configuration, each illustrated with a basic threatened pair in orange. In each case, rows in $\mathscr{R}^*$ are marked with a star. (Rows not marked with a star may or may not be in $\mathscr{R}^*$.)

Fact 11.11. Consider any partial Latin square $R \in \mathcal{P}_n$ and set of rows $\mathscr{R}^*$, and let $\mathbf{R}'$ be obtained from R by switching each isolated intercalate with probability $1/2$, independently. Let $\{e_1, e_2\}$ be any threatened pair for R. Then, $\{e_1, e_2\}$ is a basic threatened pair for $\mathbf{R}'$, with probability at least $(1/2)^5$.

Proof. There is some threat configuration F_0 in R corresponding to the threatened pair $\{e_1, e_2\}$. By the definition of a split bad configuration, in $R \cup \{e_1, e_2\}$, there is a way to switch isolated intercalates to introduce a pair of intersecting intercalates containing $\{e_1, e_2\}$. If we do the same switches in R (except possibly the switch of the special intercalate, which may not actually exist in R), we make $\{e_1, e_2\}$ a basic threatened pair. See Figure 2 for an example.

Now, F_0 intersects at most five isolated intercalates in R (it is clear that F_0 contains at most three isolated intercalates, but the other entries in F_0 could be contained in totally different isolated intercalates). The desired result follows. $\square$

We next record a lemma saying that there are typically not many entries covered by basic threatened pairs. Recall the definitions of ℓ and ε from Definition 9.1.

Definition 11.12. For $C > 0$ and a hypergraph $H \subseteq K_{n,n,n}^{(3)}$, let $\mathcal{T}^{\text{basic}}(H, C) \subseteq \mathcal{L}_n$ be the set of Latin squares L such that for some set of rows $\mathscr{R}^*$ of size $|\mathscr{R}^*| = \ell$, the number of basic threatened pairs in $H \cap L$ is more than $C\varepsilon^3 n^3 \ell$.

Lemma 11.13. *There is an absolute constant $C > 0$ such that the following holds. Consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \text{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\text{basic}}(\mathbf{H}, C)] \leq \exp(-\Omega(\varepsilon^3 n^2)).$$

We defer the proof of Lemma 11.13 to Section 11.4. First, we show how to use a switching argument (with Fact 11.11) to deduce Lemma 11.4. Note that, as we will see, this deduction crucially requires that the error probability in Lemma 11.13 is $\exp(-\Omega(\varepsilon^3 n^2))$ rather than our usual $\exp(-\omega(n \log^2 n))$. We also need an upper bound on the size of the largest family of pairwise disjoint intercalates, as follows.

Definition 11.14. Let $\mathcal{T}^{\text{int}}_{\text{upper}}(H) \subseteq \mathcal{L}_n$ be the set of Latin squares L for which there is a family of more than $80\varepsilon^4 n^2$ disjoint intercalates in $H \cap L$.

Lemma 11.15. *Consider a random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ and an independent random Latin square $\mathbf{L} \sim \text{Unif}(\mathcal{L}_n)$. We have*

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\text{int}}_{\text{upper}}(\mathbf{H})] \leq \exp(-\omega(n \log^2 n)).$$

We prove Lemma 11.15 with an averaging argument and Lemma 4.7 (this is a subset of the ideas we have already seen in Sections 10 and 11.2). Indeed, the following lemma is a version of Lemma 11.15 for the triangle removal process.

Definition 11.16. Let $\vec{\mathcal{U}}^{\text{int}}_{\text{upper}}(H, \alpha) \subseteq \vec{\mathcal{P}}_n$ be the set of partial ordered Latin squares $\vec{P}$ for which there is a family of more than $20\alpha^4 \varepsilon^4 n^2$ disjoint intercalates in $H \cap \vec{P}$.

Lemma 11.17. *Let $\alpha \in (0, 1)$ be a sufficiently small constant. Consider independent random hypergraphs $\vec{\mathbf{R}} \sim \text{TRP}(n, \alpha n^2)$ and $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$. Then we have*

$$\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}^{\text{int}}_{\text{upper}}(\mathbf{H}, \alpha)] \leq \exp(-\omega(n \log^2 n)).$$

Figure 2. On the left, an example of a threat configuration in a partial Latin square (with an example threatened pair in **orange**). If we switch the three highlighted intercalates, then we obtain the partial Latin square on the right, which contains a basic threat configuration (highlighted). The threatened pair on the left has now become a basic threatened pair on the right.

Proof of Lemma 11.15 given Lemma 11.17. Let $\alpha > 0$ be small enough for Lemma 11.17, let $m = \alpha n^2$ and let $\rho = \alpha^4/4$.

By Lemma 11.17 and Fact 9.4(1), with probability $1 - \exp(-\omega(n\log^2 n))$ our random hypergraph $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$ satisfies $\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(\mathbf{H}, \alpha) \,|\, \mathbf{H}] \leq \exp(-\omega(n\log^2 n))$. Let H be such an outcome of $\mathbf{H}$.

To show $\vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)$ is (ρ, m)-inherited from $\mathcal{T}_{\text{upper}}^{\text{int}}(H)$, let $L \in \mathcal{T}_{\text{upper}}^{\text{int}}(H)$, so there is a family of $X \geq 80\varepsilon^4 n^2$ disjoint intercalates in $H \cap L$. Let $\mathbf{Y}$ be the number of these intercalates which lie in a random ordered m-subset $\vec{\mathbf{P}}_m(L)$ of L. We have

$$\mathbb{E}[\mathbf{Y}] = X \cdot (1 + o(1))\alpha^4 \geq \alpha^4 X/2.$$

By Markov's inequality,

$$\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)] \geq \mathbb{P}\left[\mathbf{Y} \geq \frac{\alpha^4}{4} X\right] = 1 - \mathbb{P}\left[X - \mathbf{Y} > \left(1 - \frac{\alpha^4}{4}\right) X\right] \geq 1 - \frac{1 - \alpha^4/2}{1 - \alpha^4/4} \geq \rho.$$

That is to say, $\vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)$ is (ρ, m)-inherited from $\mathcal{T}_{\text{upper}}^{\text{int}}(H)$. Thus, by Lemma 4.7,

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}_{\text{upper}}^{\text{int}}(H)] \leq \exp(2n\log^2 n)\mathbb{P}[\vec{\mathbf{R}} \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)] = \exp(-\omega(n\log^2 n)).$$

Recalling our choice of H, Fact 9.4(2) concludes the proof. $\qquad\square$

Proof of Lemma 11.17. For each $H \subseteq K_{n,n,n}^{(3)}$, note that $\vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)$ is a monotone increasing property. So, by Lemma 5.3 (applied to each possible outcome H of $\mathbf{H}$), we have

$$\mathbb{P}[\mathbf{H} \cap \vec{\mathbf{R}} \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)] \leq \mathbb{P}[\mathbf{H} \cap \mathbf{G}' \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)] + \exp(-\Omega(n^2)) = \mathbb{P}[\mathbf{G} \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)] + \exp(-\Omega(n^2)),$$

where $\mathbf{G}' \sim \mathbb{G}^{(3)}(n, 2\alpha/n)$ and $\mathbf{G} \sim \mathbb{G}^{(3)}(n, 2\varepsilon\alpha/n)$. Let $\mathbf{X}$ be the size of the maximum family of disjoint intercalates in $\mathbf{G}$. We have

$$\mathbb{E}[\mathbf{X}] \leq n^6 (2\alpha\varepsilon/n)^4 = 16\alpha^4\varepsilon^4 n^2,$$

and changing an edge of $\mathbf{G}$ changes $\mathbf{X}$ by at most 1 (since an entry can be part of at most one intercalate in a maximum disjoint family). By Theorem 3.1, $\mathbb{P}[\mathbf{G} \in \vec{\mathcal{U}}_{\text{upper}}^{\text{int}}(H, \alpha)]$ is at most

$$\mathbb{P}[\mathbf{X} > 20\alpha^4\varepsilon^4 n^2] \leq \exp\left(-\Omega\left(\frac{\varepsilon^8 n^4}{n^3 \cdot \varepsilon/n + \varepsilon^4 n^2}\right)\right) = \exp(-\Omega(\varepsilon^7 n^2)) = \exp(-\omega(n\log^2 n)),$$

and the desired result follows. $\qquad\square$

We now deduce Lemma 11.4 from Lemmas 11.13 and 11.15 and Fact 11.11.

Proof of Lemma 11.4. Let C be as in Lemma 11.13, and let $C' = 100C$.

By Fact 9.4(1) and Lemmas 11.13 and 11.15, with probability $1 - \exp(-\omega(n\log^2 n))$ over the randomness of $\mathbf{H}$, we have

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\text{basic}}(\mathbf{H}, C) \,|\, \mathbf{H}] \leq \exp(-\Omega(\varepsilon^3 n^2)), \quad \mathbb{P}[\mathbf{L} \in \mathcal{T}_{\text{upper}}^{\text{int}}(\mathbf{H}) \,|\, \mathbf{H}] \leq \exp(-\omega(n\log^2 n)).$$

Fix such an outcome $H \in K_{n,n,n}^{(3)}$ of $\mathbf{H}$, and let $\mathcal{S} = \mathcal{T}^{\text{threat}}(H, C', \alpha, \phi) \setminus \mathcal{T}_{\text{upper}}^{\text{int}}(H)$. By Fact 9.4(2), it suffices to show that

$$\mathbb{P}[\mathbf{L} \in \mathcal{S}] \leq \exp(-\omega(n\log^2 n)).$$

Let $\mathbf{P} \in \mathcal{P}_{n,\alpha n^2}$ be a uniformly random subset of αn^2 edges of $\mathbf{L}$ and let $\mathbf{P}'$ be obtained from $\mathbf{P}$ by switching each isolated intercalate in $H \cap \mathbf{P}$ with probability $1/2$ independently. Let $\mathbf{L}' = \mathbf{P}' \cup (\mathbf{L} \setminus \mathbf{P})$.

Let $\mathbf{X}$ be the maximum over all choices of $\mathscr{R}^*$ of the number of threatened pairs in $H \cap \mathbf{P}$, and let $\mathbf{Y} \leq \mathbf{X}$ be the maximum over all choices of $\mathscr{R}^*$ of the number of threatened pairs in $H \cap \mathbf{P}$ which become basic threatened pairs in $H \cap \mathbf{P}'$. By Fact 11.11 with $\mathbf{R} = H \cap \mathbf{P}$ (and the fact that an expected maximum is at least the maximum expectation) we have $\mathbb{E}[\mathbf{Y} \,|\, \mathbf{X}] \geq (1/32)\mathbf{X}$, so by Markov's inequality

$$\mathbb{P}\left[\mathbf{Y} \geq \frac{1}{100}\mathbf{X} \,\Big|\, \mathbf{X}\right] = 1 - \mathbb{P}\left[\mathbf{X} - \mathbf{Y} > \frac{99}{100}\mathbf{X} \,\Big|\, \mathbf{X}\right] \geq 1 - \frac{31/32}{99/100} \geq \frac{1}{100}.$$

Also, by the definition of $\mathcal{T}^{\text{threat}}(H, C', \alpha, \phi) \supseteq \mathcal{S}$, we have

$$\mathbb{P}[\mathbf{X} \geq C'\varepsilon^3 n^3 \ell \,|\, \mathbf{L} \in \mathcal{S}] \geq \phi.$$

Recalling that $C' = 100C$, and letting $\mathcal{T} = \mathcal{T}^{\text{basic}}(H, C)$, we deduce that $\mathbb{P}[\mathbf{L}' \in \mathcal{T} \,|\, \mathbf{L} \in \mathcal{S}] \geq \phi/100$ or in other words

$$\mathbb{P}[\mathbf{L} \in \mathcal{S}] \leq \frac{100}{\phi}\mathbb{P}[\mathbf{L} \in \mathcal{S} \text{ and } \mathbf{L}' \in \mathcal{T}].$$

So, it suffices to prove that $\mathbb{P}[\mathbf{L} \in \mathcal{S}$ and $\mathbf{L}' \in \mathcal{T}] \leq \exp(-\omega(n \log^2 n))$.

We introduce the notation $L_1 \to L_2$ to mean that we can obtain L_2 from L_1 by switching some disjoint intercalates in $H \cap L_1$. Recall that $\mathbf{L}'$ is obtained from $\mathbf{L}$ by switching some intercalates which are isolated in $H \cap \mathbf{P}$, and these intercalates are certainly disjoint, so for any of the possible outcomes (L_1, L_2) of $(\mathbf{L}, \mathbf{L}')$ we have $L_1 \to L_2$. We therefore have

$$\mathbb{P}[\mathbf{L} \in \mathcal{S} \text{ and } \mathbf{L}' \in \mathcal{T}] \leq \sum_{L_2 \in \mathcal{T}} \sum_{L_1 \in \mathcal{S}, L_1 \to L_2} \mathbb{P}[\mathbf{L} = L_1] \mathbb{P}[\mathbf{L}' = L_2 \mid \mathbf{L} = L_1]$$

$$\leq \sum_{L_2 \in \mathcal{T}} \mathbb{P}[\mathbf{L} = L_2] \sum_{L_1 \in \mathcal{S}, L_1 \to L_2} 1$$

$$\leq \sum_{L_2 \in \mathcal{T}} \mathbb{P}[\mathbf{L} = L_2] \binom{n^4 + 80\varepsilon^4 n^2}{80\varepsilon^4 n^2}$$

$$\leq \mathbb{P}[\mathbf{L} \in \mathcal{T}] \exp(O(80\varepsilon^4 n^2 \log n))$$

$$\leq \exp(-\Omega(\varepsilon^3 n^2)) \cdot \exp(O(80\varepsilon^4 n^2 \log n))$$

$$\leq \exp(-\omega(n \log^2 n)),$$

as desired. Here, in the second inequality we used the fact that $\mathbb{P}[\mathbf{L} = L_1] = \mathbb{P}[\mathbf{L} = L_2]$ (since $\mathbf{L}$ has the same probability of being equal to any Latin square). In the third inequality we used the fact that every Latin square has at most n^4 intercalates, and the difference between L_1 and L_2 is described by at most $80\varepsilon^4 n^2$ disjoint intercalate switches (note that $L_1 \in \mathcal{S}$ implies that $L_1 \notin \mathcal{T}_{\text{upper}}^{\text{int}}(H)$, i.e., $H \cap L_1$ has at most $80\varepsilon^4 n^2$ disjoint intercalates). Given a set of size N, the number of subsets of size at most k is bounded by $\binom{N+k}{k}$. The fifth inequality is by our choice of H, and the final one uses that $\varepsilon = o(1/\log n)$ (recall Definition 9.1). It is vital here that the error probability in Lemma 11.13 is $\exp(-\Omega(\varepsilon^3 n^2))$. $\qquad\square$

11.4. Upper-bounding the number of basic threatened pairs.

It remains to prove Lemma 11.13. This does not really involve any new ideas. Specifically, the approach is to carefully consider how basic threat configurations can emerge through the triangle removal process, using Lemma 4.7 to relate this to a random Latin square. There are four different cases for the structure of a basic threat configuration (cf. Figure 1), each of which needs to be treated in a slightly different way.

We start by giving each of the four cases in Figure 1 a name.

Definition 11.18. Consider the four cases for the structure of a basic threat configuration, as follows[15].

	c_1	c_2	c_3			c_1	c_2	c_3			c_1	c_2	c_3			c_1	c_2	c_3
r_1^*					r_1^*					r_1^*			s_3		r_1^*			s_3
r_2	s_2	s_1	s_3		r_2	s_1	s_2	s_3		r_2	s_1	s_2			r_2	s_2	s_1	
r_3		s_3	s_2		r_3		s_3	s_2		r_3		s_3	s_2		r_3		s_3	s_2

These four cases describe four partial Latin squares $Q_1, Q_2, Q_3, Q_4 \in \mathcal{P}_3$ (from left to right). Each Q_t has three rows, three columns and three symbols (though in Q_1 and Q_2, one of the rows does not have any entries in it).

For each type t, and any partial Latin square $P \in \mathcal{P}_n$, an *embedding* from Q_t into P is an injective map from the 9 vertices of Q_t into the vertices of P (where row-vertices are mapped to row-vertices, column-vertices are mapped to column-vertices, and symbol-vertices are mapped to symbol-vertices), such that the image of every hyperedge of Q_t is a hyperedge of P_t.

Then, for a partial Latin square $P \in \mathcal{P}_n$, note that a basic threatened pair (together with an associated basic threat configuration) is specified by an embedding of some Q_t into P, where we demand that r_1^* is mapped into $\mathscr{R}^*$. In this case we say that the basic threatened pair has *type t*.

Note that a basic threatened pair can have multiple types (if it has multiple associated basic threat configurations).

Definition 11.19. For some $t \in \{1, 2, 3, 4\}$, let $\pi = (e_1, \dots, e_5)$ be an ordering of the entries of Q_t, and let $\phi(Q_t)$ be an embedding of Q_t into an ordered partial Latin square $\vec{P} \in \vec{\mathcal{P}}_n$. We say that $\phi(Q_t)$ is *π-consistent* with $\vec{P}$ if $\phi(e_i) \in \vec{P}[(i-1)/5, i/5]$ for all $i \in \{1, \dots, 5\}$. We say that a basic threatened pair of type t is *π-consistent* with $\vec{P}$ if it is associated with some π-consistent embedding $\phi(Q_t)$.

Now, we state a version of Lemma 11.13 for the triangle removal process. Recall the definitions of ℓ and ε from Definition 9.1.

---

[15] These are exactly as in Figure 1, but in the last two cases we have swapped the first and second rows.

Definition 11.20. For a hypergraph H, a type $t \in \{1, 2, 3, 4\}$ and an ordering π of the entries of Q_t, let $\vec{\mathcal{U}}^{\mathrm{basic}}(H, t, \pi) \subseteq \vec{\mathcal{P}}_n$ be the set of partial ordered Latin squares $\vec{P}$ such that for some set of rows $\mathscr{R}^*$ of size $|\mathscr{R}^*| = \ell$, the number of type-t basic threatened pairs in $H \cap \vec{P}$ which are π-consistent with $\vec{P}$ is more than $\varepsilon^3 n^3 \ell$.

Lemma 11.21. *Fix a sufficiently small constant $\alpha > 0$, and any $t \in \{1, 2, 3, 4\}$. There is an ordering π of Q_t such that for independent random hypergraphs $\vec{R} \sim \mathrm{TRP}(n, \alpha n^2)$ and $\mathbf{H} \sim \mathbb{G}^{(3)}(n, \varepsilon)$, we have*

$$\mathbb{P}[\vec{R} \in \vec{\mathcal{U}}^{\mathrm{basic}}(\mathbf{H}, t, \pi)] \leq \exp(-\Omega(\varepsilon^3 n^2)).$$

Before proving Lemma 11.21, we show how to deduce Lemma 11.13.

Proof of Lemma 11.13. Let $\alpha \in (0, 1)$ be small enough for Lemma 11.21, let $m = \alpha n^2$ and let $\rho = \alpha^5/5^7$, and assume $C \geq 4(5^7/\alpha^5)$. For each t, let π_t be the ordering in Lemma 11.21.

By Lemma 11.21 and Fact 9.4(1), with probability $1 - \exp(-\Omega(\varepsilon^3 n^2))$ over the randomness of $\mathbf{H}$, we have that for each $t \in \{1, 2, 3, 4\}$,

$$\mathbb{P}[\vec{R} \in \vec{\mathcal{U}}^{\mathrm{basic}}(\mathbf{H}, t, \pi_t) \,|\, \mathbf{H}] \leq \exp(-\Omega(\varepsilon^3 n^2)). \tag{11.1}$$

Fix such an outcome $H \subseteq K_{n,n,n}^{(3)}$ of $\mathbf{H}$.

Let $\vec{\mathcal{U}}^{\mathrm{basic}}(H) = \bigcup_{t=1}^{4} \vec{\mathcal{U}}^{\mathrm{basic}}(H, t, \pi_t)$. Given an ordered partial Latin square $\vec{P}$ and a set of rows $\mathscr{R}^*$, we say a type-t basic threatened pair is *consistent* if it is π_t-consistent.

To show $\vec{\mathcal{U}}^{\mathrm{basic}}(H)$ is (ρ, m)-inherited from $\mathcal{T}^{\mathrm{basic}}(H, C)$, let $L \in \mathcal{T}^{\mathrm{basic}}(H, C)$, so there is $\mathscr{R}^*$ of size $|\mathscr{R}^*| = \ell$ for which there are $X > C\varepsilon^3 n^3 \ell$ many basic threatened pairs in $H \cap L$. Let $\mathbf{Y}$ be the number of basic threatened pairs in $H \cap L$ which are present and consistent in an ordered random m-subset $\vec{\mathbf{P}}_m(L)$ of L. We have

$$\mathbb{E}[\mathbf{Y}] = X \cdot (1 + o(1))\alpha^5 (1/5)^5 \geq \frac{\alpha^5}{5^6} X.$$

Now, if $\mathbf{Y} \geq (\alpha^5/5^7)X > 4\varepsilon^3 n^3 \ell$ then $\vec{\mathcal{U}}^{\mathrm{basic}}(H, t, \pi_t)$ holds for some t. Using Markov's inequality, it follows that

$$\mathbb{P}[\vec{\mathbf{P}}_m(L) \in \vec{\mathcal{U}}^{\mathrm{basic}}(H)] \geq \mathbb{P}\left[\mathbf{Y} \geq \frac{\alpha^5}{5^7} X\right] = 1 - \mathbb{P}\left[X - \mathbf{Y} > \left(1 - \frac{\alpha^5}{5^7}\right)X\right] \geq 1 - \frac{1 - (\alpha^5/5^6)}{1 - (\alpha^5/5^7)} \geq \rho,$$

which means $\vec{\mathcal{U}}^{\mathrm{basic}}(H)$ is (ρ, m)-inherited from $\mathcal{T}^{\mathrm{basic}}(H, C)$.

Thus, we have

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{basic}}(H, C)] \leq \exp(2n \log^2 n)\mathbb{P}[\vec{R} \in \vec{\mathcal{U}}^{\mathrm{basic}}(H)] \leq \exp(-\Omega(\varepsilon^3 n^2)),$$

where the first inequality follows by Lemma 4.7, and the second inequality follows by (11.1).

Recalling our choice of H, Fact 9.4(2) implies that

$$\mathbb{P}[\mathbf{L} \in \mathcal{T}^{\mathrm{basic}}(\mathbf{H}, C)] \leq \exp(-\Omega(\varepsilon^3 n^2)),$$

as desired. $\qquad\square$

Now we prove Lemma 11.21, separately considering types 1 and 2, and types 3 and 4.

Proof of Lemma 11.21 for types $t = 1$ and $t = 2$. Recall that Q_1 and Q_2 have the following forms.

	c_1	c_2	c_3
r_1^*			
r_2	s_2	s_1	s_3
r_3		s_3	s_2

	c_1	c_2	c_3
r_1^*			
r_2	s_1	s_2	s_3
r_3		s_3	s_2

If $t = 1$, we take our ordering π to be

$$e_1 = (r_3, c_3, s_2), \; e_2 = (r_2, c_3, s_3), \; e_3 = (r_3, c_2, s_3), \; e_4 = (r_2, c_2, s_1), \; e_5 = (r_2, c_1, s_2).$$

If $t = 2$, we take our ordering π to be

$$e_1 = (r_3, c_3, s_2), \; e_2 = (r_2, c_3, s_3), \; e_3 = (r_3, c_2, s_3), \; e_4 = (r_2, c_2, s_2), \; e_5 = (r_2, c_1, s_1).$$

Fix $t \in \{1, 2\}$ and let $\mathscr{R}^*$ be a set of rows with $|\mathscr{R}^*| = \ell$. Among all the π-consistent embeddings $\phi(Q_t)$ of Q_t into $\mathbf{H} \cap \vec{R}$ which satisfy $\phi(r_1^*) \in \mathscr{R}^*$, let $\mathbf{N}$ be the number of possibilities for $(\phi(r_1^*), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2))$. Then, $\mathbf{N}$ is an upper bound on the number of type-t basic threatened pairs in $\mathbf{H} \cap \vec{R}$ which are π-consistent with $\vec{R}$.

We first reveal $\vec{\mathbf{R}}[0,1/5]$, which has $\alpha n^2/5$ entries. Each of them is present in $\mathbf{H}$ with probability ε independently. Thus, by a Chernoff bound, with probability $1 - \exp(-\Omega(\varepsilon n^2))$, there are at most εn^2 many viable choices for $\phi(e_1)$ (and therefore $(\phi(r_3), \phi(c_3), \phi(s_2))$). Assume this is the case.

Next, we reveal $\mathbf{H} \cap \vec{\mathbf{R}}[1/5, 3/5]$. For each choice of $(\phi(r_3), \phi(c_3), \phi(s_2))$, there are at most n choices for $\phi(e_2)$ in column $\phi(c_3)$ of $\vec{\mathbf{R}}[1/5, 2/5]$. This determines $\phi(r_2)$ and $\phi(s_3)$, and then there is at most one occurrence of $\phi(s_3)$ in row $\phi(r_3)$ of $\vec{\mathbf{R}}[2/5, 3/5]$, determining $\phi(c_2)$. This makes a total of at most εn^3 choices for $(\phi(r_2), \phi(r_3), \phi(c_2), \phi(c_3), \phi(s_2), \phi(s_3))$.

Let S be the set of all possible choices for $(\phi(r_2), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2))$ at this stage, so we have $|S| \le \varepsilon n^5$ (there are at most n^2 choices for $(\phi(c_1), \phi(s_1))$). Let $\mathbf{S}_4 \subseteq S$ be the set of all $(\phi(r_2), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2)) \in S$ for which $\phi(e_4)$ is an edge of $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$, and let $\mathbf{S}_5 \subseteq \mathbf{S}_4$ be the set of all such tuples for which $\phi(e_5)$ is an edge of $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$. Note that $\mathbf{N} \le \ell |\mathbf{S}_5|$ (there are at most ℓ choices for $\phi(r_1^*) \in \mathscr{R}^*$).

We next consider $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$. We claim that $|\mathbf{S}_4| \le \varepsilon^2 n^4$ with probability $1 - \exp(-\Omega(\varepsilon^2 n^2))$. To see this, first note that, by Lemma 5.3, it suffices to prove an analogous bound with $\mathbf{H} \cap \mathbf{G} \sim \mathbb{G}^{(3)}(n, (2\varepsilon\alpha/5)/n)$ in place of $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$, where $\mathbf{G} \sim \mathbb{G}^{(3)}(n, 2(\alpha/5)/n)$. But this is a direct consequence of Corollary 3.2, since each potential entry $\phi(e_4)$ is present in $\mathbf{H} \cap \mathbf{G}$ with probability at most $2\varepsilon\alpha/(5n)$, and appears in at most n^2 tuples in S.

Reveal an outcome of $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$ satisfying $|\mathbf{S}_4| \le \varepsilon^2 n^4$. We next consider $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$: we claim that $|\mathbf{S}_5| \le \varepsilon^3 n^3$ with probability $1 - \exp(-\Omega(\varepsilon^3 n^2))$. The proof is basically as before: first note that it suffices to prove an analogous bound with $\mathbf{H} \cap \mathbf{G} \sim \mathbb{G}^{(3)}(n, (2\varepsilon\alpha/5)/n)$ in place of $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$, and then the desired result follows from Corollary 3.2 (noting that each potential entry $\phi(e_5)$ appears in at most n tuples in $\mathbf{S}_4$, since such a tuple is determined by a choice of $\phi(e_4)$ in the same row).

We have proved that the number of type-t basic threatened pairs in $\mathbf{H} \cap \vec{\mathbf{R}}$ which are π-consistent with $\vec{\mathbf{R}}$ is at most $\mathbf{N} \le \ell |\mathbf{S}_5| \le \varepsilon^3 n^3 \ell$, with probability at least $1 - \exp(-\Omega(\varepsilon^3 n^2))$. A union bound over all (at most 2^n) choices of $\mathscr{R}^*$ finishes the proof. $\qquad\square$

Proof of Lemma 11.21 for types $t = 3$ and $t = 4$. Recall that Q_3 and Q_4 have the following forms.

	c_1	c_2	c_3			c_1	c_2	c_3
r_1^*			s_3		r_1^*			s_3
r_2	s_1	s_2			r_2	s_2	s_1	
r_3		s_3	s_2		r_3		s_3	s_2

If $t = 3$, we take our ordering π to be

$$e_1 = (r_3, c_3, s_2), \; e_2 = (r_1^*, c_3, s_3), \; e_3 = (r_3, c_2, s_3), \; e_4 = (r_2, c_2, s_2), \; e_5 = (r_2, c_1, s_1).$$

If $t = 4$, we take our ordering π to be

$$e_1 = (r_3, c_3, s_2), \; e_2 = (r_1^*, c_3, s_3), \; e_3 = (r_3, c_2, s_3), \; e_4 = (r_2, c_2, s_1), \; e_5 = (r_2, c_1, s_2).$$

Fix $t \in \{3, 4\}$ and let $\mathscr{R}^*$ be a set of rows with $|\mathscr{R}^*| = \ell$. Among all the π-consistent embeddings $\phi(Q_t)$ of Q_t into $\mathbf{H} \cap \vec{\mathbf{R}}$ which satisfy $\phi(r_1^*) \in \mathscr{R}^*$, let $\mathbf{N}$ be the number of possibilities for $(\phi(r_1^*), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2))$. Then, $\mathbf{N}$ is an upper bound on the number of type-t basic threatened pairs in $\mathbf{H} \cap \vec{\mathbf{R}}$ which are π-consistent with $\vec{\mathbf{R}}$.

The first part is exactly the same as the previous proof: we first reveal $\vec{\mathbf{R}}[0, 1/5]$, which has $\alpha n^2/5$ entries. By a Chernoff bound over the randomness of $\mathbf{H}$, with probability $1 - \exp(-\Omega(\varepsilon n^2))$, there are at most εn^2 many viable choices for $\phi(e_1)$ (and therefore $(\phi(r_3), \phi(c_3), \phi(s_2))$). Assume this is the case.

Next, we reveal $\mathbf{H} \cap \vec{\mathbf{R}}[1/5, 3/5]$. The numbers are slightly different than in the last proof: for each choice of $(\phi(r_3), \phi(c_3), \phi(s_2))$, there are at most ℓ choices for $\phi(e_2)$ in column $\phi(c_3)$ of $\vec{\mathbf{R}}[1/5, 2/5]$ (in a row in $\mathscr{R}^*$). This determines $\phi(r_1^*)$ and $\phi(s_3)$, and then there is at most one occurrence of $\phi(s_3)$ in row $\phi(r_3)$ of $\vec{\mathbf{R}}[2/5, 3/5]$. This makes a total of at most $\varepsilon n^2 \ell$ choices for $(\phi(r_1^*), \phi(r_3), \phi(c_2), \phi(c_3), \phi(s_2), \phi(s_3))$.

Let S be the set of all possible choices for $(\phi(r_1^*), \phi(r_2), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2))$ at this stage, so we have $|S| \le \varepsilon n^5 \ell$ (there are at most n^3 choices for $(\phi(r_2), \phi(c_1), \phi(s_1))$). Let $\mathbf{S}_4 \subseteq S$ be the set of all $(\phi(r_1^*), \phi(r_2), \phi(c_1), \phi(c_2), \phi(s_1), \phi(s_2)) \in S$ for which $\phi(e_4)$ is an edge of $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$, and let $\mathbf{S}_5 \subseteq \mathbf{S}_4$ be the set of all such tuples for which $\phi(e_5)$ is an edge of $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$. Note that $\mathbf{N} \le |\mathbf{S}_5|$.

The end of the proof is similar to before. We next consider $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$: we claim that $|\mathbf{S}_4| \le \varepsilon^2 n^4 \ell$ with probability $1 - \exp(-\Omega(\varepsilon^2 n^2))$. To see this, we first note that it suffices to prove an analogous bound

with $\mathbf{H} \cap \mathbf{G} \sim \mathbb{G}^{(3)}(n, (2\varepsilon\alpha/5)/n)$ in place of $\mathbf{H} \cap \vec{\mathbf{R}}[3/5, 4/5]$, and then the desired result follows from Corollary 3.2, noting that each potential entry $\phi(e_4)$ appears in at most $n^2\ell$ tuples in S.

Finally we consider $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$: we claim that $|\mathbf{S}_5| \leq \varepsilon^3 n^3 \ell$ with probability $1 - \exp(-\Omega(\varepsilon^3 n^2))$. Again, we first note that it suffices to prove an analogous bound with $\mathbf{H} \cap \mathbf{G} \sim \mathbb{G}^{(3)}(n, (2\varepsilon\alpha/5)/n)$ in place of $\mathbf{H} \cap \vec{\mathbf{R}}[4/5, 1]$, and then the desired result follows from Corollary 3.2 (noting that each potential entry $\phi(e_5)$ appears in at most $n\ell$ tuples in $\mathbf{S}_4$, since such a tuple is determined by a choice of $\phi(e_4)$ in the same row, and a choice of $\phi(r_1^*)$).

We then finish with a union bound over choices of $\mathscr{R}^*$, as before. $\qquad\square$

12. Concluding remarks

In this paper we have proved (a generalisation of) Cameron's conjecture, describing the joint distribution of the number of odd row/column/symbol permutations in a random $n \times n$ Latin square. There are various directions for further research, as follows.

- Can quantitative aspects be improved? For example, in Theorem 1.3(4), we see that the total variation error of our approximation is $o(1)$. Actually, our proof gives an error of $n^{-1/2+o(1)}$, and it seems this can be improved to $n^{-1+o(1)}$ with some additional calculations using the fact that $\mathbb{E}[N_{\mathrm{row}}(\mathbf{L})] = \mathbb{E}[N_{\mathrm{col}}(\mathbf{L})] = \mathbb{E}[N_{\mathrm{sym}}(\mathbf{L})] = n/2$. However, it seems plausible that the true error of the approximation might be super-exponentially small.
- What about constrained distributions on Latin squares (e.g. symmetric Latin squares)? We expect that similar results should hold, but the available enumeration estimates are much weaker. It might be possible to prove analogues of Theorem 1.3(1) and Theorem 1.3(5) (using the ideas described in Section 2.4, which have relatively weak quantitative requirements), but new ideas would be required for the other parts of Theorem 1.3.
- What about analogous questions for higher-dimensional analogues of Latin squares (sometimes called *high dimensional permutations*, see e.g. [46])? Here it seems completely new ideas would be required, because (we predict that) small switchings become rarer as the dimension increases.

References

[1] Y. Alimohammadi, P. Diaconis, M. Roghani, and A. Saberi, *Sequential importance sampling for estimating expectations over the space of perfect matchings*, Ann. Appl. Probab. **33** (2023), no. 2, 799–833.

[2] J. Allsop and I. M. Wanless, *Subsquares in random Latin rectangles*, arXiv:2409.08446.

[3] L. Alpoge, *Square-root cancellation for the signs of Latin squares*, Combinatorica **37** (2017), no. 2, 137–142.

[4] B. Bollobás, *The chromatic number of random graphs*, Combinatorica **8** (1988), no. 1, 49–55.

[5] C. Bowtell and R. Montgomery, *Almost every Latin square has a decomposition into transversals*, arXiv:2501.05438.

[6] L. M. Brègman, *Certain properties of nonnegative matrices and their permanents*, Dokl. Akad. Nauk SSSR **211** (1973), 27–30.

[7] P. J. Cameron, *Permutations*, Paul Erdős and his mathematics, II (Budapest, 1999), Bolyai Soc. Math. Stud., vol. 11, János Bolyai Math. Soc., Budapest, 2002, pp. 205–239.

[8] P. J. Cameron, *Asymmetric Latin squares, Steiner triple systems, and edge-parallelisms*, arXiv:1507.02190.

[9] P. J. Cameron, *Almost all quasigroups have rank* 2, Discrete Math. **106**/**107** (1992), 111–115.

[10] P. J. Cameron, *BCC problem list*, https://webspace.maths.qmul.ac.uk/p.j.cameron/bcc/allprobs.pdf, 2001.

[11] P. J. Cameron, *Random Latin squares*, slides for a presentation at the RAND-APX meeting at Oxford in 2003, https://webspace.maths.qmul.ac.uk/p.j.cameron/slides/rand-apx.pdf, 2003.

[12] P. J. Cameron, *Problems from "Permutations"*, https://webspace.maths.qmul.ac.uk/p.j.cameron/permgps/permutations.html, 2005.

[13] P. J. Cameron, *Peter Cameron's BCC problems*, https://webspace.maths.qmul.ac.uk/p.j.cameron/pjcatbcc.html, 2007.

[14] P. J. Cameron, *Problems*, https://webspace.maths.qmul.ac.uk/p.j.cameron/oldprob.html, 2013.

[15] P. J. Cameron, *A niggling problem*, posted on Peter Cameron's blog, https://cameroncounts.wordpress.com/2015/01/24/a-niggling-problem/, 2015.

[16] N. Cavenagh and R. Ramadurai, *On the distances between Latin squares and the smallest defining set size*, J. Combin. Des. **25** (2017), no. 4, 147–158.

[17] 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.

[18] N. J. Cavenagh and I. M. Wanless, *There are asymptotically the same number of Latin squares of each parity*, Bull. Aust. Math. Soc. **94** (2016), no. 2, 187–194.

[19] M. Delcourt and L. Postle, *Refined absorption: A new proof of the existence conjecture*, arXiv:2402.17855.

[20] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Stochastic Modelling and Applied Probability, vol. 38, Springer-Verlag, Berlin, 2010, Corrected reprint of the second (1998) edition.

[21] G. P. Egorychev, *The solution of van der Waerden's problem for permanents*, Adv. in Math. **42** (1981), no. 3, 299–305.

[22] 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.

[23] W. Feller, *An introduction to probability theory and its applications. Vol. I*, third ed., John Wiley & Sons, Inc., New York-London-Sydney, 1968.

[24] A. Ferber and M. Kwan, *Almost all Steiner triple systems are almost resolvable*, Forum of Mathematics, Sigma, vol. 8, Cambridge University Press, 2020, p. e39.

[25] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probability **3** (1975), 100–118.

[26] B. Friedman and S. McGuinness, *The Alon-Tarsi conjecture: a perspective on the main results*, Discrete Math. **342** (2019), no. 8, 2234–2253.

[27] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F*, Mem. Amer. Math. Soc. **284** (2023), no. 1406, v+131.

[28] S. Gould and T. Kelly, *Hamilton transversals in random Latin squares*, Random Structures Algorithms **62** (2023), no. 2, 450–478.

[29] 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.

[30] R. Huang and G.-C. Rota, *On the relations of various conjectures on Latin squares and straightening coefficients*, Discrete Math. **128** (1994), no. 1-3, 225–236.

[31] M. T. Jacobson and P. Matthews, *Generating uniformly distributed random Latin squares*, J. Combin. Des. **4** (1996), no. 6, 405–437.

[32] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[33] S. Janson and A. Ruciński, *The infamous upper tail*, Random Structures Algorithms **20** (2002), no. 3, 317–342, Probabilistic methods in combinatorial optimization.

[34] J. C. M. Janssen, *On even and odd Latin squares*, J. Combin. Theory Ser. A **69** (1995), no. 1, 173–181.

[35] 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 Erdös.

[36] P. Keevash, *The existence of designs*, arXiv:1401.3665.

[37] P. Keevash, *The existence of designs II*, arXiv:1802.05900.

[38] P. Keevash, *A short proof of the existence of designs*, arXiv:2411.18291.

[39] P. Keevash, *Counting designs*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 4, 903–927.

[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, A. Sah, and M. Sawhney, *Note on random latin squares and the triangle removal process*, arXiv:2109.15201.

[42] M. Kwan, A. Sah, and M. Sawhney, *Large deviations in random Latin squares*, Bull. Lond. Math. Soc. **54** (2022), no. 4, 1420–1438.

[43] M. Kwan, A. Sah, M. Sawhney, and M. Simkin, *Substructures in Latin squares*, Israel J. Math. **256** (2023), no. 2, 363–416.

[44] M. Kwan and B. Sudakov, *Intercalates and discrepancy in random Latin squares*, Random Structures Algorithms **52** (2018), no. 2, 181–196.

[45] N. Linial and Z. Luria, *An upper bound on the number of Steiner triple systems*, Random Structures Algorithms **43** (2013), no. 4, 399–406.

[46] N. Linial and Z. Luria, *Discrepancy of high-dimensional permutations*, Discrete Anal. (2016), Paper No. 11, 8.

[47] N. Linial and M. Simkin, *Monotone subsequences in high-dimensional permutations*, Combin. Probab. Comput. **27** (2018), no. 1, 69–83.

[48] J. Lynch and J. Sethuraman, *Large deviations for processes with independent increments*, Ann. Probab. **15** (1987), no. 2, 610–627.

[49] C. McDiarmid, *On the method of bounded differences*, Surveys in combinatorics, 1989 (Norwich, 1989), London Math. Soc. Lecture Note Ser., vol. 141, Cambridge Univ. Press, Cambridge, 1989, pp. 148–188.

[50] B. D. McKay and I. M. Wanless, *Most Latin squares have many subsquares*, J. Combin. Theory Ser. A **86** (1999), no. 2, 322–347.

[51] A. O. Pittenger, *Mappings of Latin squares*, Linear Algebra Appl. **261** (1997), 251–268.

[52] J. Radhakrishnan, *An entropy proof of Bregman's theorem*, J. Combin. Theory Ser. A **77** (1997), no. 1, 161–164.

[53] V. Rödl and L. Thoma, *Asymptotic packing and the random greedy algorithm*, Random Structures Algorithms **8** (1996), no. 3, 161–177.

[54] N. Ross, *Fundamentals of Stein's method*, Probab. Surv. **8** (2011), 210–293.

[55] J. Spencer, *Asymptotic packing via a branching process*, Random Structures Algorithms **7** (1995), no. 2, 167–172.

[56] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, 2nd ed., Cambridge University Press, Cambridge, 2001.

[57] L. Warnke, *On the method of typical bounded differences*, Combinatorics, Probability and Computing **25** (2016), no. 2, 269–299.

[58] P. Zappa, *Triplets of Latin squares*, Boll. Un. Mat. Ital. A (7) **10** (1996), no. 1, 63–69.

Institute of Science and Technology Austria (ISTA). Am Campus 1, 3400 Klosterneuburg, Austria
Email address: matthew.kwan@ist.ac.at

Institute of Science and Technology Austria (ISTA). Am Campus 1, 3400 Klosterneuburg, Austria
Email address: kalina.petrova@ist.ac.at

Department of Mathematics, Columbia University, New York, NY 10027
Email address: m.sawhney@columbia.edu