Journal of Combinatorial Theory, Series A **86**, 323–347 (1999)
Article ID jcta.1998.2947, available online at http://www.idealibrary.com on IDEAL®

Most Latin Squares Have Many Subsquares

B. D. McKay and I. M. Wanless*

Department of Computer Science, Australian National University, Canberra, ACT 0200, Australia
E-mail: bdm@cs.anu.edu.au; ianw@ms.unimelb.edu.au

Communicated by the Managing Editors

Received January 30, 1998

A $k \times n$ Latin rectangle is a $k \times n$ matrix of entries from $\{1, 2, ..., n\}$ such that no symbol occurs twice in any row or column. An intercalate is a 2×2 Latin subrectangle. Let $N(R)$ be the number of intercalates in R, a randomly chosen $k \times n$ Latin rectangle. We obtain a number of results about the distribution of $N(R)$ including its asymptotic expectation and a bound on the probability that $N(R) = 0$. For $\varepsilon > 0$ we prove most Latin squares of order n have $N(R) \geqslant n^{3/2-\varepsilon}$. We also provide data from a computer enumeration of Latin rectangles for small k, n.

1. INTRODUCTION AND PRELIMINARY DEFINITIONS

For any positive integer c, let $I_c = \{1, 2, 3, ..., c\}$ and $\mu_c = \frac{1}{2}\binom{c}{2}$. A $k \times n$ *Latin rectangle* is a $k \times n$ array with entries from I_n with the property that no symbol is repeated within any row or column. Not surprisingly, a $n \times n$ Latin rectangle is a *Latin square*. An *intercalate* is a Latin 2×2 subsquare, or in other words a 2×2 submatrix containing only 2 distinct symbols. Note that the cells involved in an intercalate need not be contiguous.

Various papers have dealt with the construction of so-called N_2 *Latin squares*, or Latin squares containing no intercalates. In [5, 6, 9] such squares are shown to exist for all orders other than $n = 2$ and $n = 4$. Upper bounds on the number of intercalates have also been investigated in [3]. However, little work seems to have been published on the distribution between these two extremes or on the proportion of Latin squares which are N_2. We will show that such squares are very rare.

We begin with some notation. Let $L(k, n)$ denote the set of $k \times n$ Latin rectangles, which we think of as a probability space equipped with measure $P(\cdot)$ corresponding to the discrete uniform distribution. For any subset

* Current address: Department of Mathematics and Statistics, University of Melbourne, Parkville, Victoria 3052, Australia.

0097-3165/99 $30.00

$S \subset L(k, n)$ define $P_S(\cdot)$ to be the conditional probability $P(\cdot \mid R \in S)$. We consider $k = k(n)$ and $r = r(n)$ to be integer functions of n such that $2 \leqslant r \leqslant k \leqslant n$. This paper focuses on the number of intercalates within the first r rows of elements of $L(k, n)$, particularly the asymptotic behaviour as $n \to \infty$. Our results naturally apply to any other selection of r rows, since intercalates survive permutations of the rows (or of the columns or symbols) in a rectangle. To distinguish the intercalates we wish to count, we introduce the term *consequential intercalates*, meaning intercalates lying entirely within the first r rows of an element of $L(k, n)$.

Notation. For each rectangle $R \in L(k, n)$ we define the following counts of consequential intercalates within R.

(a) $N^a(R)$ is the number wholly contained within the first a rows of R,

(b) $N_a(R)$ is the number involving entries in the ath row,

(c) $N_{a,b}(R)$ is the number which involve the entry R_{ab} in row a, column b of R,

(d) $N_{a,b,c}(R)$ is the number which involve both the R_{ab} and R_{ac} entries.

We use both N and $N(R)$ as shorthands for $N^r(R)$, and sometimes think of them as random variables on the space $L(k, n)$. Also, let $\theta_i(R)$ be the number of entries in row i of R which are involved in consequential intercalates. The presentation of the main results of this paper begins in Section 7. A one line summary is that as n increases, $N(R)$ exhibits characteristics of a Poisson random variable with a mean of μ_r. In a way this is not surprising given the following simple observation:

PROPOSITION 1. *For $k = 2$, $\lim_{n \to \infty} P(N(R) = 0) = e^{-1/2}$ and the expected number of intercalates, $E(N) \to \frac{1}{2}$.*

Proof. The second row of a random $R \in L(2, n)$ is obtained by applying D_R to the first row, where D_R is a random derangement of I_n. It is easily seen that the number of intercalates in R is equal to the number, t, of 2-cycles in D_R. By the principle of inclusion-exclusion the probability of a random derangement having no 2-cycles is

$$\frac{\sum_{2t+f \leqslant n} (-1)^{t+f} \, n!/(2^t t! \, f!)}{\sum_{f \leqslant n} (-1)^f \, (n!/f!)} \to \sum_{t \geqslant 0} (-1)^t \frac{1}{2^t t!} = e^{-1/2}.$$

The calculation for the expected number of intercalates is similar, and we omit it. ∎

It should be noted that C. D. Godsil and the first author have already shown in [2] that $E(N) = \mu_k(1 + O(k^2/n^2))$ provided $k < \frac{1}{5}n$. Our method of proof will be similar to the switching argument used to prove this result.

2. SOME BASIC RESULTS ABOUT INTERCALATES

In this section we make some simple observations about intercalates which will be used implicitly throughout the remainder of the paper.

Two distinct intercalates can intersect in at most one entry (this is an application of a simple theorem that the intersection of Latin subsquares is itself a Latin subsquare). It follows that different intercalates occupying the same pair of rows (or the same pair of columns) must be disjoint and hence that no $r \times n$ Latin rectangle may contain more than $\mu_r n$ intercalates. We know from work of Heinrich and Wallis [3] that for $n > 3$ there is an $n \times n$ Latin square, SQ_n with at least $\frac{1}{45}n^3$ intercalates (the constant $\frac{1}{45}$ is not explicit in [3], nor is it important for our purposes). This implies that for all $n > 3$ and $r \leqslant n$ there is an $r \times n$ Latin rectangle with at least $\frac{1}{45}n^3\binom{n-2}{r-2}/\binom{n}{r} \geqslant \frac{4}{45}\mu_r n$ intercalates. Note this is within a constant factor of the theoretical bound given above.

As previously noted, it is known that for all $n > 4$ there exists an N_2 Latin square (that is, one without any intercalates), and it follows immediately that for any $r \leqslant n$ there exists an $r \times n$ Latin rectangle without any intercalates, provided $n > 4$.

Finally, note that an entry in a Latin rectangle can be part of at most $k-1$ intercalates, and any intercalate can intersect with at most $4(k-2)$ other intercalates.

3. EXTENSIONS

One foundation of our results will be the ability to infer information about Latin rectangles from information about their sub-rectangles. The basic ideas behind this section appear in [10, p. 136].

DEFINITION. For $R_1, R_2 \in L(k, n)$ we say that R_1 is *i-related* to R_2 (written $R_1 \sim_i R_2$) if the two rectangles are identical or differ only in the ith row. If $R_3 \in L(k', n)$ for $k \leqslant k' \leqslant n$ and the first k rows of R_3 exactly match R_1 then R_3 is an *extension* of R_1 (written $R_3 \sqsupseteq R_1$).

A measure of the number of rectangles *i*-related to a given rectangle (being the number of possible replacements for the ith row) can be found as follows.

PROPOSITION 2. *Suppose* $R \in L(k+1, n)$ *where* $1 \leqslant k < n$. *Then*

$$n!\left(1-\frac{k}{n}\right)^n \leqslant |\{R' \in L(k+1, n)\colon R' \sim_{k+1} R\}| \leqslant ((n-k)!)^{n/(n-k)}.$$

Proof. The number of $(k+1) \times n$ rectangles $(k+1)$-related to R is equal to $\operatorname{per}(A_R)$, the permanent of the $n \times n$ matrix defined by

$$(A_R)_{ij} = \begin{cases} 0, & \text{if } R_{xi} = j \text{ for some } x \in I_k; \\ 1, & \text{otherwise.} \end{cases}$$

Clearly A_R is a (0,1)-matrix with exactly $n-k$ positive entries in each row and column. Applying the Egorychev–Falikman theorem (formerly the van der Waerden conjecture) we find that $n!\,(1-(k/n))^n \leqslant \operatorname{per}(A_R)$. Also, by an upper bound on permanents due to Brègman we know that $\operatorname{per}(A_R) \leqslant ((n-k)!)^{n/(n-k)}$, which is what we need. For statements of the two general results just used, see [10, 12]. ∎

For properties of Latin rectangles which are inherited by extensions (such as the presence of intercalates) Proposition 2 will allow us to lift probabilistic information from $L(k, n)$ to $L(k+1, n)$, via the following result.

PROPOSITION 3. *For fixed* $A, B \in L(k, n)$ *where* $2 \leqslant k < n$ *and random* $R \in L(k+1, n)$, $P(R \sqsupseteq A)/P(R \sqsupseteq B) \leqslant (\sqrt{3n})^{k/(n-k)}$. *If* $4k \leqslant n$ *then* $P(R \sqsupseteq A)/P(R \sqsupseteq B) \leqslant k$.

Proof. Using Proposition 2 and Stirling's formula we see that

$$\frac{P(R \sqsupseteq A)}{P(R \sqsupseteq B)} \leqslant \frac{((n-k)!)^{n/(n-k)}}{n!\,(1-k/n)^n}$$
$$\leqslant \sqrt{\frac{(2\pi(n-k))^{n/(n-k)}}{2\pi n}} \exp\left(\frac{n}{12(n-k)^2}\right).$$

Now $(1/n)(n-k)^{n/(n-k)} < (n/e)^{k/(n-k)}$ so

$$\frac{P(R \sqsupseteq A)}{P(R \sqsupseteq B)} \leqslant \sqrt{\left(\frac{2\pi n}{e}\right)^{k/(n-k)}} \exp\left(\frac{k}{8(n-k)}\right) \leqslant (\sqrt{3n})^{k/(n-k)}.$$

If in addition, $2 \leqslant k \leqslant n/4$ then $(3n)^k < k^{2(n-k)}$ and $P(R \sqsupseteq A)/P(R \sqsupseteq B) \leqslant k$. ∎

Of course, we would like to be able to "lift" to extensions by more than one row...

PROPOSITION 4. *Suppose* $2 \leqslant k < k' \leqslant n$ *and fix* $A, B \in L(k, n)$. *Let* $R \in L(k', n)$ *be randomly chosen. Then* $P(R \sqsupseteq A)/P(R \sqsupseteq B) \leqslant (\sqrt{3n}\,)^{n \log(n-k)}$.

Proof. First note that every $(n-1) \times n$ Latin rectangle has a unique extension in $L(n, n)$, so we can assume without loss of generality that $k' < n$. By repeatedly applying Proposition 3 ($k' - k$ times) we have $P(R \sqsupseteq A)/P(R \sqsupseteq B) \leqslant (\sqrt{3n})^{\omega}$ where

$$\omega = \sum_{i=k}^{k'-1} \frac{i}{n-i} \leqslant \sum_{j=n-k'+1}^{n-k} \frac{n}{j} \leqslant n \log\left(\frac{n-k}{n-k'}\right). \quad \blacksquare$$

To conclude this section we note that Propositions 3 and 4 also hold (trivially) when $k = 1$. Symmetry dictates that, for arbitrary k', all $1 \times n$ rectangles have the same number of extensions in $L(k', n)$. In this paper, however, we are only interested in rectangles with at least two rows.

4. OBSTRUCTIONS ARE RARE

Certain configurations of intercalates are not able to be counted by the switching procedure outlined in the next section. In this section we show that such hindrances are so uncommon that omitting them makes negligible impact on the overall count. In the process of finding probability bounds we will need a number of somewhat arbitrary cutoffs. We give them definite values, and collect their definitions here, for the sake of clarity.

DEFINITION. Henceforth, unless explicitly stated otherwise, we assume that $k = o(n^u)$ for some fixed u satisfying $0 < u < 1$ and that n is large. The following constants will be used universally.

$$\beta = \left\lfloor \frac{3+2u}{1-u} \right\rfloor, \qquad \delta = \frac{5}{7},$$

$$\gamma = \delta^{-1/(2\beta)}, \qquad \alpha = \max\left\{42, \left\lceil \frac{3\gamma^2}{\gamma - 1} \right\rceil\right\}.$$

The intersection of multiple intercalates proves to be an impediment to our attempts to count them, so our first step is to bound the chance of it happening.

PROPOSITION 5. *Suppose that* $5k < n$ *and that* $m \geqslant 1$ *is a fixed integer. The probability* p_m *that a random* $R \in L(k, n)$ *has an entry which is part of* m *or more intercalates is bounded by* $k^{m+1} n^{1-m} \exp(O(k/n))$.

Proof. First note that if $m \geqslant k$ then $p_m = 0$. Otherwise by a theorem proved by Godsil and McKay [2, Theorem 4.7] we know that

$$P(R_{1j} = R_{j1} = j,\ R_{jj} = 1 \text{ for all } j \in I_{m+1}) = n^{-3m-1} \exp\left(O\left(\frac{k(3m+1)}{n-2k-m}\right)\right).$$

Since all instances of an entry being part of m intercalates can be reached from this one representative by permuting rows, columns and/or symbols we conclude that

$$p_m \leqslant \frac{1}{m!}\left(\frac{n!}{(n-m-1)!}\right)^2 \frac{k!}{(k-m-1)!} n^{-3m-1} \exp\left(O\left(\frac{k(3m+1)}{n-2k-m}\right)\right).$$

The result follows. ∎

In Proposition 5 we bounded the chance of having many intercalates overlap on a single entry. The other situation which turns out to be obstructive is a single row having most of its entries involved in intercalates. In order to study the probability of this occurring we need to introduce some terms and then prove a simple technical lemma.

DEFINITION. A *star* is a complete bipartite graph $K_{1,m}$ for some $m \geqslant 0$. We say the star is *trivial* if $m = 0$, otherwise it is *non-trivial*. We define a (k, n)*-constellation of magnitude* h as a graph C having n vertices of which exactly $n - h$ are of degree 0. We also require that every component of C is a star and that there is a component labelling λ which maps trivial components of C into I_n and non-trivial components into I_k.

PROPOSITION 6. *Any graph without isolated vertices contains a spanning forest of non-trivial stars.*

Proof. Let u_1 and u_2 be adjacent vertices in a graph G of minimum degree $\geqslant 1$. We distinguish two cases. Either there is some vertex u_3 whose sole neighbour is one of u_1 and u_2, or no such vertex exists. In the latter case make $T = \{u_1, u_2\}$. In the former, suppose without loss of generality that u_1 is adjacent to u_3 and define T to be the set consisting of u_1 and all the degree 1 neighbours of u_1. In both cases the subgraph H induced by T is a non-trivial star and $G \backslash H$ is a graph without isolated vertices. Proceed inductively. ∎

We are now ready to prove the rarity of the second obstruction.

PROPOSITION 7. *Fix* $R_1 \in L(k, n)$ *and* $h \in I_n$. *Selecting* $R \in L(k, n)$ *at random we have* $P(\theta_r(R) = h \mid R \sim_r R_1) \leqslant \delta^{h-14k}$.

Proof. Let $\Omega_h = \{R \in L(k, n): \theta_r(R) = h \text{ and } R \sim_r R_1\}$. We show that each $R \in \Omega_h$ is associated with a (not necessarily unique) constellation from which R is recoverable. It will follow that $|\Omega_h|$ is no greater than the number of (k, n)-constellations of magnitude h.

Suppose that $R \in \Omega_h$. Form a graph G_R on vertices $\{v_1, v_2, ..., v_n\}$ by making v_i and v_j adjacent if and only if $N_{r,i,j}(R) > 0$. Since $\theta_r(R) = h$ there are exactly h vertices in G_R of positive degree. Thus by Proposition 6 there is a spanning subgraph C_R of G_R which satisfies our definition of a constellation, once the component labelling λ is defined. For each isolated vertex v_i of C_R make $\lambda(v_i) = R_{ri}$. Next, suppose $v_i, v_{a_1}, v_{a_2}, ... v_{a_l}$ are the vertices of a non-trivial component $\mathscr{C}$ of C_R and v_i is adjacent to v_{a_j} for each $j \in I_l$. Then there are less than k possible values for R_{ri} since it must occur in each of columns $a_1, a_2, ... a_l$ of R_1. Furthermore, the value of R_{ri} determines R_{ra_j} for each $j \in I_l$ since $N_{r,i,a_j}(R) = 1$. Hence, by appropriate indexing, we can choose $\lambda(\mathscr{C}) \in I_k$ in such a manner that $R_{ri}, R_{ra_1}, R_{ra_2}, ... R_{ra_l}$ can be recovered from our choice. Once λ has been fully defined in this way we see that C_R is a constellation incorporating all the information in row r of R, and hence determining $R \sim_r R_1$ completely. As desired, this shows that we can bound $|\Omega_h|$ by counting (k, n)-constellations of magnitude h, a feat easily achieved by means of an exponential generating function.

$$\begin{aligned}|\Omega_h| &\leqslant [x^h]\, n! \sum_{j \geqslant 0} \frac{1}{j!} \left(\frac{1}{2} kx^2 + \sum_{i \geqslant 3} \frac{1}{(i-1)!} kx^i \right)^j \\ &= [x^h]\, n! \exp\left(kx \left(e^x - 1 - \frac{1}{2} x \right) \right),\end{aligned}$$

where $[x^h]\, f(x)$ denotes the coefficient of x^h in the power series expansion of $f(x)$. Since our generating function has all coefficients positive it follows that

$$[x^h] \exp(kx(e^x - 1 - \tfrac{1}{2}x)) \leqslant x^{-h} \exp(kx(e^x - 1 - \tfrac{1}{2}x))$$

holds for all $x > 0$. In particular, setting $x = 1/\delta$ yields

$$|\Omega_h| \leqslant [x^h]\, n! \exp(kx(e^x - 1 - \tfrac{1}{2}x)) \leqslant n!\, \delta^h e^{10k/3}.$$

Finally, we see that by applying Proposition 2

$$\begin{aligned}P(\theta_r(R) = h \mid R \sim_r R_1) &\leqslant \delta^h e^{10k/3} \left(1 - \frac{k}{n}\right)^{-n} \\ &\leqslant \delta^h e^{10k/3} \exp(kn/(n-k)) \\ &\leqslant \delta^h e^{14k/3}\end{aligned}$$

if $k < n/4$ (and otherwise $h \leqslant n \leqslant 4k$ so the result is trivially true). As $\delta < e^{-1/3}$ the proof is complete. ∎

PROPOSITION 8. *Choosing* $R \in L(k, n)$ *randomly gives* $P(\exists i \in I_k\colon \theta_i(R) \geqslant h) \leqslant \frac{7}{2} k\delta^{h-14k}$ *for every* $h \in I_n$.

Proof. Considering permutations of the rows of R gives,

$$P(\exists i \in I_k\colon \theta_i(R) \geqslant h) \leqslant k\, P(\theta_r(R) \geqslant h) = k \sum_{i=h}^{n} P(\theta_r(R) = i).$$

Our result comes via Proposition 7 which yields a geometric series bounding this sum. ∎

It is time to explicitly decide which rectangles we can work with and which we cannot.

DEFINITION. Let $K = \lceil \alpha \max\{k, \log n\} \rceil$. Define $\mathscr{L} \subseteq L(k, n)$ by $R \in \mathscr{L}$ if and only if $\theta_i(R) \leqslant K$ and $N_{i,j}(R) \leqslant \beta$ for every $i \in I_k$ and $j \in I_n$. The subsets S_c given by $S_c = \{R \in \mathscr{L}\colon N(R) = c\}$ are the principal objects of interest in our counting argument. For each $R \in L(k, n)$ we also define $\Psi(R) = \{i \in I_k\colon \theta_i(R) \leqslant K - 2\}$.

We wish to draw conclusions about $L(k, n)$ by studying $\mathscr{L}$, so it is important that most $k \times n$ Latin rectangles are in $\mathscr{L}$.

PROPOSITION 9. *If* $k = o(n^u)$ *for* $0 < u < 1$ *then* $|\mathscr{L}| = |L(k, n)| \times (1 - o(n^{-2-u}))$.

Proof. As a result of the definitions just made, Proposition 5, and Proposition 8,

$$\frac{|\mathscr{L}|}{|L(k, n)|} = 1 - O(k\delta^{K-14k}) - k^{\beta+2} n^{-\beta} \exp\left(O\left(\frac{k}{n}\right)\right) = 1 - o(n^{-2-u})$$

since $k^{\beta+2} n^{-\beta} \leqslant n^{\omega}$, where

$$\omega = u\left(\frac{3+2u}{1-u} + 1\right) + 1 - \frac{3+2u}{1-u} = -2 - u. \quad ∎$$

5. SWITCHING ENTRIES

A common technique for generating Latin rectangles is to switch entries within a row of an existing rectangle (for example, [2, 8]). In this paper we use such a method to create and destroy intercalates. It is easiest to explain

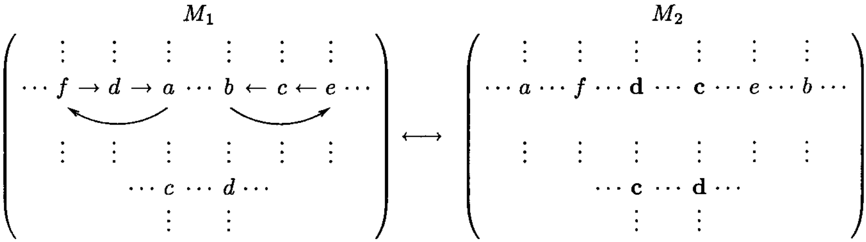

FIG. 1. In the left hand matrix (M_1) two triples of entries in the same row are cycled, to produce a new intercalate in the right hand matrix (M_2). Naturally, the switching process can be reversed in order to destroy the intercalate in M_2.

the procedure intuitively before providing a formal definition. Our switching process is displayed in Fig. 1. Methods for generating Latin squares using other sorts of local perturbation are discussed in [4, 11].

However, Fig. 1 is deceptively simple. Complications arise from two sources. The first is that switching entries in a Latin rectangle sometimes produces a matrix which is not a Latin rectangle because some symbol is duplicated within a column. The second is that each time an entry is moved there is potential for multiple intercalates to be destroyed, whilst some number of new intercalates may be created. To keep things simple we consider only switchings which fit the following guidelines (with reference to Fig. 1):

(a) Both M_1 and M_2 are Latin rectangles,

(b) M_1 has no consequential intercalates which are not present in M_2,

(c) M_2 may have a number of consequential intercalates which are not present in M_1,

(d) The entries labelled a, b, e, and f are not in consequential intercalates in M_2 (or M_1).

Note that a simpler switching technique which involves two direct swaps of pairs of entries does not have the flexibility we require. For example, consider $J \in L(3, 3m)$ formed by the juxtaposition of some number m of order 3 Latin squares, as shown in Fig. 2.

$$J = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \\ 2 & 3 & 1 & 5 & 6 & 4 & 8 & 9 & 7 & \cdots \\ 3 & 1 & 2 & 6 & 4 & 5 & 9 & 7 & 8 & \end{pmatrix}$$

FIG. 2. Juxtaposition of order 3 Latin squares.

$$J' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & \\ 2 & 3 & 1 & 5 & 6 & 4 & 8 & 9 & 7 & \cdots \\ 6 & 1 & 4 & 3 & 2 & 5 & 9 & 7 & 8 & \end{pmatrix}$$

FIG. 3. Two intercalates introduced by $J_{31} \leftrightarrow J_{34}$ and $J_{33} \leftrightarrow J_{35}$.

Suppose we were to try to create a single intercalate in J (which is intercalate-free) by means of two swaps of pairs of entries within one row, such as $J_{31} \leftrightarrow J_{34}$ and $J_{33} \leftrightarrow J_{35}$. Any such operation introduces two intercalates, with each relocated entry becoming part of one of them (Fig. 3).

We are ready to formalize the idea behind our switching process. For any $k \times n$ matrix A, denote by $sw^i_{xyz}(A)$ the $k \times n$ matrix, A', which is identical to A except that $A'_{ix} = A_{iz}$, $A'_{iy} = A_{ix}$ and $A'_{iz} = A_{iy}$. For any set S let 2^S denote the power set of S. Next, for $c \geqslant 0$, $a \geqslant -c$, $a \neq 0$ and $i \in I_r$ define $F^i_{c,a}\colon S_c \to 2^{S_{c+a}}$ as follows. For each $A \in S_c$, $A' \in S_{c+a}$ we have $A' \in F^i_{c,a}(A)$ if and only if there are distinct $x, y, z, x', y', z' \in I_n$ satisfying

(a) $A' = sw^i_{xyz}(sw^i_{x'y'z'}(A))$, (1)

(b) $N_{i,y}(M) = N_{i,z}(M) = N_{i,y'}(M) = N_{i,z'}(M) = 0$, for $M \in \{A, A'\}$, (2)

(c) $N_{i,x}(A) = N_{i,x'}(A) = 0$ and $N_{i,x}(A') + N_{i,x'}(A') = 1 + a$ if $a > 0$, (3)

(d) $N_{i,x}(A') = N_{i,x'}(A') = 0$ and $N_{i,x}(A) + N_{i,x'}(A) = 1 - a$ if $a < 0$. (4)

The following facts, holding for every $c \geqslant 0$ and $A \in S_c$, follow straight from the definitions of $F^i_{c,a}$ and $\mathscr{L}$. The proofs will be omitted.

- $A' \in F^i_{c,a}(A)$ if and only if $A \in F^i_{c+a,-a}(A')$,
- $F_{c,a}(A)$ is in $1:2$ correspondence with valid choices of $\{x, y, z, x', y', z'\}$,
- $F^i_{c,a}(A) \cap F^j_{c,b}(A) = \varnothing$ unless $i = j$ and $a = b$,
- $F^i_{c,a}(A) = \varnothing$ unless $|a| \leqslant 2\beta$.

Some other pertinent observations about switchings are noted now. There are k occurrences of each symbol in $R \in \mathscr{L}$, so there are $k-1$ columns to which a given entry may not be relocated. Likewise, an entry may not be replaced by any of the $k-1$ other symbols in its column. So long as these restrictions are obeyed the result of our switching process will be a Latin rectangle. Similarly, fewer than k different symbols can form an intercalate when placed in a given position in R, and for any symbol there are fewer than k positions in each row of R at which that symbol would form an intercalate. Finally, note that a switching from $R \in \mathscr{L}$ which creates no intercalates cannot create something in $L(k, n)\backslash\mathscr{L}$.

These simple observations are the basis for the next three propositions, which aim to gauge the size of $F^i_{c,a}$.

PROPOSITION 10. *The bounds* $|F^i_{c,1}(A)| \leqslant n^2\binom{n}{2}(r-1)$ *and* $|\bigcup_{a>1} F^i_{c,a}(A)| \leqslant kn^3(r-1)$ *hold for every* $A \in S_c$ *and* $i \in I_r$.

Proof. To satisfy (3) there must be $j \in I_r\backslash\{i\}$ for which $A_{jx} = A_{iz'}$ and $A_{jx'} = A_{iz}$. There are $\binom{n}{2}(r-1)$ ways to choose j, x and x' with $x < x'$; each determines z and z'. Both y and y' have fewer than n possibilities so $|F^i_{c,1}(A)| \leqslant |\bigcup_{a>0} F^i_{c,a}(A)| \leqslant n^2\binom{n}{2}(r-1)$. Now providing A' satisfies conditions (1) and (2) observe that $N(A') > N(A)+1$ cannot hold unless $N(sw^i_{xyz}(A)) > N(A)$ or $N(sw^i_{x'y'z'}(A)) > N(A)$. For each choice of j, x there are fewer than $2k$ choices of x' for which such extra intercalates are generated, given (2). Hence $|\bigcup_{a>1} F^i_{c,a}(A)| \leqslant kn^3(r-1)$ as required. ∎

PROPOSITION 11. $|F^i_{c,1}(A)| \geqslant \frac{1}{2}n^3(n-O(K))(|\Psi(A)|-1)$ *uniformly for* $A \in S_c$, $i \in \Psi(A)$.

Proof. Choose $j \in \Psi(A)\backslash\{i\}$ and distinct $x, y, z, x', y', z' \in I_n$ such that

(a) $A_{jx} = A_{iz'}$ and $A_{jx'} = A_{iz}$,

(b) $N_{i,s}(A) = 0$ for $s \in \{x, y, z, x', y', z'\}$,

(c) $N_{j,s}(A) = 0$ for $s \in \{x, x'\}$,

(d) $sw^i_{xyz}(A) \in \mathscr{L}$ and $sw^i_{x'y'z'}(A) \in \mathscr{L}$,

(e) $N_{i,s}(sw^i_{xyz}(sw_{x'y'z'}(A))) = 0$ for $s \in \{y, z, y', z'\}$.

The choices of x and x' determine z' and z by (a). Of the $n(n-1)$ choices for x, x' at most $n(4\theta_i(A) + 2\theta_j(A)) \leqslant 6nK$ breach (b) or (c) and at most a further $4nk$ cause problems for (d) or (e). Likewise, at most $4k$ choices of either y or y' breach (d) or (e). Crucially, we have enough freedom in choosing y and y' to ensure that (2) holds for the A' defined by (1), thus avoiding the difficulty discussed in Figs. 2 and 3. Also, our insistence on (c) and $i, j \in \Psi(A)$ ensures there are no "overflow" problems, so $A' \in \mathscr{L}$. Moreover A' has exactly the same consequential intercalates as A, with the addition of the single intercalate $A'_{ix} = A'_{jx'}$, $A'_{ix'} = A'_{jx}$. Counting half the choices for $\{x, y, z, x', y', z'\}$ gives $\frac{1}{2}(n-O(K))^4$ from which the result follows. ∎

PROPOSITION 12. $|\bigcup_{a<0} F^i_{c,a}(A)| = N_i(A)\, n^3(n-O(K))$ *uniformly for* $A \in S_c$, $i \in I_r$.

Proof. Condition (4) implies $N_{i,x,x'}(A)=1$ so there are $N_i(A)$ options with $x<x'$. By the same arguments as used in Proposition 11 there are then $n-O(K)$ ways to choose each of y, z, y', z' in order to satisfy (2) and (4), for some negative a. ▮

As a result of the observations made so far, we can count the number of switchings into and out from S_c, deriving the equation

$$\sum_{a=1}^{2\beta}\sum_{A\in S_{c-a}}\sum_{i\in I_r}|F^i_{c-a,a}(A)|=\sum_{A\in S_c}\sum_{i\in I_r}\left|\bigcup_{a<0}F^i_{c,a}(A)\right|$$
$$=\sum_{A\in S_c}\sum_{i\in I_r}N_i(A)\,n^3(n-O(K))$$
$$=2cn^3(n-O(K))\,|S_c|. \tag{3}$$

6. COMPARATIVE SIZES OF THE S_C

The aim of this section is to establish the relationship between the $|S_c|$ for varying c.

PROPOSITION 13. *When n is sufficiently large, $|S_{c-1}|\leqslant(\gamma\alpha/\alpha-3\gamma)\,|S_c|$ for all $c\leqslant\gamma\mu_r$.*

Proof. For $c\leqslant\gamma\mu_r$ and $A\in S_c$ a simple count in the worst case scenario yields,

$$|\Psi(A)|\geqslant r-\frac{4c}{K-1}\geqslant r-\frac{r(r-1)\gamma}{\alpha r-1}\geqslant r\left(1-\frac{\gamma}{\alpha}\right)$$

from which it follows that

$$\binom{|\Psi(A)|}{2}\geqslant\binom{r}{2}\left(1-3\frac{\gamma}{\alpha}+\frac{\gamma^2}{\alpha^2}\right). \tag{6}$$

Hence, by (5) and Proposition 11,

$$|S_c|\,2cn^4\leqslant\sum_{A\in S_{c-1}}\sum_{i\in\Psi(A)}|F^i_{c-1,1}(A)|$$
$$\geqslant\sum_{A\in S_{c-1}}\binom{|\Psi(A)|}{2}n^3(n-O(K))$$
$$\geqslant|S_{c-1}|\binom{r}{2}\left(1-\frac{3\gamma}{\alpha}\right)n^4$$

for sufficiently large n, by (6). The result then follows from $c\leqslant\gamma\mu_r$. ▮

We next prove a simple technical lemma before examining the chance of an entire row being clear of consequential intercalates.

PROPOSITION 14. *Let* $\{t_i\}_{i\geqslant 0}$ *be a positive sequence obeying the recurrence relation* $it_i = at_{i-1} + b\sum_{j=2}^{m} t_{i-j}$ *for all* $i \geqslant 1$, *where* a, b *and* m *are positive constants and* t_i *is interpreted as being zero if* $i<0$. *Then* $\sum_{i\geqslant 0} t_i < t_0 e^a m^b$ *and* $\sum_{i\geqslant 0} it_i < t_0(a+bm)\, e^a m^b$.

Proof. Define $f(x) = \sum_{i\geqslant 0} t_i x^i$. The given recurrence means $(df/dx) = (a + b\sum_{j=2}^{m} x^{j-1})\, f$. The solution to this differential equation is $f(x) = t_0 \exp(ax + b\sum_{j=2}^{m} (1/j)\, x^j)$. Evaluating $f(1)$ yields $\sum_{i\geqslant 0} t_i = t_0 \exp(a + b \times \sum_{j=2}^{m} (1/j)) < t_0 e^{a+b\log m} = t_0 e^a m^b$. Similarly, we find $\sum_{i\geqslant 0} it_i = (df/dx)(1) = t_0(a + b\sum_{j=2}^{m} 1) \exp(a + b\sum_{j=2}^{m} (1/j)) < t_0(a+bm)\, e^a m^b$. ▮

PROPOSITION 15. *For large* n, $P(N_r(R) = 0 \mid R \sim_r R_1) \geqslant \frac{1}{2} e^{1-r}$ *for every* $R_1 \in \mathscr{L}$.

Proof. Define $T_i = \{R \in \mathscr{L} : N_r(R) = i,\, R \sim_r R_1\}$ and let $c = N^{r-1}(R_1)$. Then by counting the rth row switchings out from and into T_i we see in turn that

$$\sum_{A\in T_i} \left| \bigcup_{a<0} F^r_{c+i,\,a}(A) \right| = \sum_{a>0} \sum_{A\in T_{i-a}} |F^r_{c+i-a,\,a}(A)|,$$

$$\sum_{A\in T_i} N_r(A)\, n^3(n - O(K)) = \sum_{A\in T_{i-1}} |F^r_{c+i-1,\,1}(A)| + \sum_{a=2}^{2\beta} \sum_{A\in T_{i-a}} |F^r_{c+i-a,\,a}(A)|,$$

$$|T_i|\, in^3(n-O(K)) \leqslant n^2 \binom{n}{2} (r-1)\, |T_{i-1}| + kn^3(r-1) \sum_{a=2}^{2\beta} |T_{i-a}|,$$

making use of Propositions 10 and 12. So for sufficiently large n,

$$i\, |T_i| \leqslant \frac{2}{3}(r-1)\, |T_{i-1}| + O\left(\frac{k}{n}\right)(r-1) \sum_{a=2}^{2\beta} |T_{i-a}|.$$

Now, applying Proposition 14 shows that for large enough n,

$$\sum_{i\geqslant 0} |T_i| < |T_0| \exp\left(\frac{2}{3}(r-1) + O\left(\frac{k}{n}\right)(r-1)\log(2\beta)\right) < |T_0|\, e^{r-1}.$$

Hence $|T_0| \geqslant e^{1-r}\, |\{R \in \mathscr{L} : R \sim_r R_1\}|$, which proves $P_{\mathscr{L}}(N_r(R) = 0 \mid R \sim_r R_1) \geqslant e^{1-r}$. To get the required result we need only observe that $P(R \in \mathscr{L} \mid R \sim_r R_1) \geqslant \frac{1}{2}$ for all $R_1 \in \mathscr{L}$ when n is large. This is not a bold claim! Given Proposition 7, the only hurdle is that an entry may be included in more than β intercalates. The chance of this happening can be bounded by yet another switching argument. ▮

The importance of the next result is that it shows that most rectangles in $\mathscr{L}$ do not contain a row that is in danger of "overflowing" if we create another intercalate (that is, the resulting rectangle will also be in $\mathscr{L}$). Note that we use P_c as shorthand for P_{S_c}.

PROPOSITION 16. *If* $c < \gamma\mu_{r-1}$ *then* $P_c(\Psi(R) \neq I_r) \leqslant \beta K^3 \delta^{K/2-17k}$ *for large n.*

Proof. Our strategy is to consider possible replacements for the rth row of R and thereby show that R is much more likely to have an intercalate-free row than a row which is too full to be counted in $\Psi(R)$. Since the result is trivial when $r=2$, we may assume $r \geqslant 3$.

Let L' be a system of distinct representatives of equivalence classes in $L(k,n)$ under $\sim_r$. Define $S'_i = \{R \in L' \colon N^{r-1}(R) = i\}$ for each i. The S'_i behave exactly as the S_i except with both r and k reduced by 1 (which is safe, since $r \geqslant 3$). In particular, for $i \leqslant \gamma\mu_{r-1}$ Proposition 13 gives $|S'_{i-1}| \leqslant |S'_i|\, \gamma\alpha/(\alpha - 3\gamma)$. Hence if $i \leqslant j \leqslant \gamma\mu_{r-1}$ and $4k < n$ then Proposition 3 yields

$$\frac{P(\exists A \in S'_i \colon R \sim_r A)}{P(\exists A \in S'_j \colon R \sim_r A)} \leqslant k \left(\frac{\gamma\alpha}{\alpha - 3\gamma}\right)^{j-i}. \tag{7}$$

Now, since any permutation of the first r rows of a random $R \in \mathscr{L}$ is equally likely,

$$\frac{1}{r} P_c(\Psi(R) \neq I_r) \leqslant P_c(\theta_r(R) \geqslant K-1)$$

$$\leqslant \sum_{i \in I_m} \frac{P_c(\theta_r(R) \geqslant K-1,\, N^{r-1}(R) = c-i)}{P_c(\theta_r(R) = 0,\, N^{r-1}(R) = c)},$$

where $m = \lfloor K\beta/2 \rfloor$, so every $A \in \mathscr{L}$ must satisfy $N_r(A) \leqslant m$. Note that

$$\frac{P_c(\theta_r(R) \geqslant K-1,\, N^{r-1}(R) = c-i)}{P_c(\theta_r(R) = 0,\, N^{r-1}(R) = c)} \leqslant \frac{\sum_{A \in S'_{c-i}} P(\theta_r(R) \geqslant K-1,\, R \sim_r A)}{\sum_{A \in S'_c} P(\theta_r(R) = 0,\, R \sim_r A)}$$

$$\leqslant 7\delta^{K-1-14k} e^{r-1} \frac{P(\exists A \in S'_{c-i} \colon R \sim_r A)}{P(\exists A \in S'_c \colon R \sim_r A)}$$

by using Propositions 7 and 15. Now compiling our results so far gives

$$P_c(\Psi(R) \neq I_r) \leqslant 7rme^{r-1} \delta^{K-14k-1} k \left(\frac{\gamma\alpha}{\alpha - 3\gamma}\right)^m$$

$$\leqslant \beta K^3 \delta^{K-17k} q^{K/2},$$

where $q = (\gamma\alpha/\alpha - 3\gamma)^{\beta} \leqslant \gamma^{2\beta} = 1/\delta$ since $\alpha \geqslant 3\gamma^2/(\gamma - 1)$. The required result follows. ∎

Next another technical lemma, this one dealing with the weight of the tails of the exponential power series.

PROPOSITION 17. *For each positive* $\lambda \neq 1$ *define* $d_\lambda = e^{\lambda - 1}\lambda^{-\lambda} < 1$. *Then for* $x \geqslant 1/\lambda$, $\sum_{i=0}^{\lambda x} (1/i!)\, x^i \geqslant e^x(1 - (\lambda - 1)^{-1}\, d_\lambda^x)$ *if* $\lambda > 1$ *and* $\sum_{i=0}^{\lambda x} (1/i!)\, x^i \leqslant e^x(x+1)\, d_\lambda^x$ *if* $\lambda < 1$.

Proof. First notice that $\lambda x \geqslant 1$ means

$$\frac{x^{\lambda x}}{(\lambda x)!} < \frac{x^{\lambda x}}{(\lambda x/e)^{\lambda x}} = e^x \left(\frac{e^{\lambda - 1}}{\lambda^\lambda}\right)^x = e^x\, d_\lambda^x.$$

Then observe that if $\lambda < 1$,

$$\sum_{i=0}^{\lambda x} \frac{x^i}{i!} \leqslant (\lambda x + 1)\, \frac{x^{\lambda x}}{(\lambda x)!} < e^x(x+1)\, d_\lambda^x$$

whereas if $\lambda > 1$ then

$$\sum_{i=0}^{\lambda x} \frac{x^i}{i!} = e^x - \sum_{i > \lambda x} \frac{x^i}{i!} \geqslant e^x - \frac{x^{\lambda x}}{(\lambda x)!} \sum_{i > 0} \left(\frac{x}{\lambda x}\right)^i > e^x \left(1 - \frac{d_\lambda^x}{\lambda - 1}\right). \quad ∎$$

Our next result concludes this section. It achieves our goal of measuring the relative sizes of the $|S_c|$.

PROPOSITION 18. *If* $\sqrt{\log(n/k)} \leqslant r \leqslant k = o(n)$ *then* $|\mathcal{L}| = \sum_{c \geqslant 0} |S_c| = e^{\mu_r(1 + O(K/n))}\, |S_0|$. *Also,* $|S_c| = c^{-1}\mu_r\, |S_{c-1}|\, (1 + O(K/n))$ *uniformly for* $c \leqslant \gamma\mu_{r-1}$.

Proof. Using the bounds from Proposition 10 in (5) gives

$$\begin{aligned} c\,|S_c| &\leqslant \frac{1}{2n^3(n - O(K))} \Bigg(\sum_{A \in S_{c-1}} \sum_{i \in I_r} n^2 \binom{n}{2} (r-1) \\ &\quad + \sum_{a=2}^{2\beta} \sum_{A \in S_{c-a}} \sum_{i \in I_r} kn^3(r-1) \Bigg) \\ &= \frac{1}{2n(n - O(K))} \Bigg(|S_{c-1}|\, |I_r|\, \tbinom{n}{2}(r-1) + \sum_{a=2}^{2\beta} |S_{c-a}|\, |I_r|\, kn(r-1) \Bigg) \\ &\leqslant \mu_r \left(1 + O\left(\frac{K}{n}\right)\right) |S_{c-1}| + \mu_r O\left(\frac{k}{n}\right) \sum_{a=2}^{2\beta} |S_{c-a}|. \end{aligned} \tag{8}$$

So by applying Proposition 14 to $\{|S_c|\}_{c\geqslant 0}$ we find

$$\sum_{c\geqslant 0} |S_c| \leqslant |S_0| \exp\left(\mu_r\left(1+O\left(\frac{K}{n}\right)+O\left(\frac{k}{n}\right)\log(2\beta)\right)\right) \leqslant e^{\mu_r(1+O(K/n))}\,|S_0|. \tag{9}$$

Now if $c\leqslant \gamma\mu_{r-1}$ then Proposition 13 and (8) imply

$$c\,|S_c| \leqslant \mu_r\left(1+O\left(\frac{K}{n}\right)\right)|S_{c-1}|, \tag{10}$$

but (5) together with Propositions 11 and 16 shows that

$$\begin{aligned} c\,|S_c| &\geqslant \frac{1}{2n^3(n-O(K))}\sum_{A\in S_{c-1}}\sum_{i\in\Psi(A)}|F^i_{c-1,1}(A)| \\ &\geqslant \frac{1}{2n^3(n-O(K))}\sum_{A\in S_{c-1}} n^3(n-O(K))\binom{|\Psi(A)|}{2} \\ &\geqslant |S_{c-1}|\,\mu_r\left(1+O\left(\frac{K}{n}\right)\right). \end{aligned} \tag{11}$$

From (10) it then follows that $|S_c| = c^{-1}\mu_r\,|S_{c-1}|\,(1+O(K/n))$ for $1\leqslant c\leqslant\gamma\mu_{r-1}$. Note that $r\geqslant\sqrt{\log(n/k)}\to\infty$ so $\mu_r/\mu_{r-1}\to 1$. Hence by Proposition 17 there is some constant $d\in(0,1)$ such that

$$\begin{aligned} \sum_{c\geqslant 0}|S_c| &\geqslant \sum_{c=0}^{\gamma\mu_{r-1}}\frac{1}{c!}\left(\mu_r\left(1+O\left(\frac{K}{n}\right)\right)\right)^c|S_0| \\ &\geqslant |S_0|\exp\left(\mu_r\left(1+O\left(\frac{K}{n}\right)\right)\right)(1-d^{\mu_r}). \end{aligned}$$

Finally note that $d^{\mu_r}=O(k/n)$ since $r^2\geqslant\log(n/k)$. ▮

7. PROVING THE MAIN RESULTS

The groundwork is now complete and we are ready to prove our theorems. In this section we deal with the case where $r\to\infty$, while in the next we treat constant r. Our first theorem deals with the probability of a sub-rectangle being free of intercalates. In particular it shows that N_2 Latin squares are rare.

THEOREM 1. *Let* n, $k=k(n)$ *and* $r=r(n)$ *satisfy* $n\geqslant k\geqslant r\to\infty$ *and suppose* $u\in(0,1)$ *is fixed. Let* $P(N^r(R)=0)$ *be the probability of a randomly chosen* $k\times n$ *Latin rectangle,* R, *having no intercalates contained wholly within the first* r *rows. Then as* $n\to\infty$

(a) $P(N^r(R)=0)=\exp(-\mu_r+o(1))$ *if* $k=o(n^u)$ *and* $r=O(\sqrt{\log(n/k)})$,

(b) $P(N^r(R)=0)=\exp(-\mu_r(1+O(n^{u-1})))$ *if* $k=o(n^u)$ *and* $r\geqslant\sqrt{\log(n/k)}$,

(c) $P(N^r(R)=0)\leqslant\exp(-\frac{1}{4}n^{2v}(1+o(1)))$ *if* $r\geqslant n^v$ *for some* $v\in(\frac{1}{2},1)$.

Proof. We prove (b) first. In this case $|\mathscr{L}|=|S_0|\exp(\mu_r(1+O(K/n)))$ by Proposition 18. Now the definition of $\mathscr{L}$ tells us $S_0=\{R\in\mathscr{L}: N^r(R)=0\}=\{R\in L(k,n): N^r(R)=0\}$. Also $|L(k,n)|=|\mathscr{L}|(1+O(K/n))$ by Proposition 9, so

$$P(N^r(R)=0)=\frac{|S_0|}{|L(k,n)|}=\exp\left(-\mu_r\left(1+O\left(\frac{K}{n}\right)\right)\right)$$
$$=\exp(-\mu_r(1+O(n^{u-1}))). \tag{12}$$

Scanning the derivation of (12) reveals that the only time we used $r\geqslant\sqrt{\log(n/k)}$ rather than just $r\to\infty$ was in the last line of Proposition 18. For $r=O(\sqrt{\log(n/k)})$ we use instead that $1-d^{\mu_r}=e^{o(1)}$ for $d\in(0,1)$ and $r\to\infty$. Also, $\mu_r O(k/n)=o(1)$ when $k=o(n^u)$ and $r^2=O(\log n)$. Hence, $|L(k,n)|=|S_0|\exp(\mu_r+o(1))$ which gives (a).

Now suppose $v\in(\frac{1}{2},1)$ and $\rho=\lceil n^v\rceil$. Applying (12) when $k=r=\rho$ gives that $P(N^\rho(R)=0)=\exp(-\frac{1}{4}n^{2v}(1+O(n^{v-1})))$. However, $n\log^2 n=o(n^{2v})$ so by using Proposition 4,

$$P(N^\rho(R)=0)=\exp(-\tfrac{1}{4}n^{2v}(1+o(1)))$$

holds even when $k>r=\rho$. Finally, $P(N^r(R)=0)\leqslant P(N^\rho(R)=0)$ for all $r>\rho$. ∎

COROLLARY 1. *For any* $\varepsilon>0$ *the probability of a randomly chosen Latin square of order* n *being* N_2 *is* $O(\exp(-n^{2-\varepsilon}))$ *as* $n\to\infty$.

Proof. Substitute $r=n$ and $v=1-\varepsilon/3$ in Theorem 1(c). ∎

COROLLARY 2. *Let* $\varepsilon>0$ *be an arbitrary constant. With probability approaching* 1 *as* $n\to\infty$, *a random Latin square of order* n *contains at least* $n^{3/2-\varepsilon}$ *intercalates.*

Proof. For each $R \in L(n,n)$ define a graph G_R which has n vertices corresponding to the rows of R. Two vertices v_i, v_j of G_R are adjacent if and only if R contains an intercalate which involves both the ith and jth rows of R. We choose to have $r = \lceil n^{(1+\varepsilon)/2} \rceil$ and $k = n$ so that $P(N^r(R) = 0) \leqslant \exp(-\frac{1}{4} n^{1+\varepsilon}(1+o(1)))$, by Theorem 1(c). Note that if $N^r(R) = 0$ then G_R contains an independent set of r vertices. In fact the probability of G_R containing any such independent set is at most $\binom{n}{r} P(N^r(R) = 0) = o(1)$. If G_R does not contain such a set then by Turán's theorem G_R contains at least $\frac{1}{2} n^{(3-\varepsilon)/2}$ edges. We conclude that the probability of R containing fewer than $n^{3/2-\varepsilon}$ intercalates is $o(1)$. ∎

Corollary 3. *With probability approaching* 1 *as* $n \to \infty$, *a randomly chosen row of a random Latin square of order n contains entries involved in intercalates.*

Proof. Let H_R be the set of rows in $R \in L(n,n)$ which do not contain entries involved in intercalates. Clearly G_R contains an independent set of size $|H_R|$ so with probability approaching one, $|H_R| < n^{1/2+\varepsilon}$ by the proof of the previous corollary. It follows that the probability of a randomly selected row being in H_R approaches 0. ∎

It is fair to point out that each of these corollaries is likely to be weaker than a best possible result. We make the following conjecture.

Conjecture. Let R represent a randomly chosen Latin square of order n and suppose $\varepsilon > 0$. The probability of each of the following events approaches 0 as $n \to \infty$:

(a) R contains a row in which no entry is involved in an intercalate,

(b) R contains fewer than $\mu_n(1-\varepsilon)$ intercalates,

(c) R contains more than $\mu_n(1+\varepsilon)$ intercalates.

It is of some interest to note that the squares which are included in event (a) were studied in [5] and used to construct sets of disjoint Steiner triple systems. Meanwhile, parts (b) and (c) of the conjecture are bolstered by our next theorem.

Theorem 2. *Let* n, $k = k(n)$ *and* $r = r(n)$ *satisfy* $n^u \geqslant k \geqslant r \to \infty$ *for some fixed* $u \in (0,1)$ *and suppose* $r^2 k = o(n)$ *as* $n \to \infty$. *The expected number of intercalates within the first r rows of a randomly chosen* $k \times n$ *Latin rectangle, R, is* $E(N^r(R)) = \mu_r(1+o(1))$. *Also, for each* $\varepsilon > 0$, $P(|N^r(R) - \mu_r| > r^{1+\varepsilon}) \to 0$ *as* $n \to \infty$.

Proof. We prove only the $r \geqslant \sqrt{\log(n/k)}$ case. The remaining case will follow by replacing $\mu_r O(K/n)$ with $o(1)$ throughout and recycling the proof of Theorem 1(a).

By (8) and Proposition 14 we know that

$$\sum_{c \geqslant 0} c\, |S_c| < |S_0|\, \mu_r \left(1 + O\left(\frac{K}{n}\right)\right) e^{\mu_r(1+O(K/n))}.$$

Also, by Proposition 18,

$$\begin{aligned}\sum_{c \geqslant 0} c\, |S_c| &> \sum_{c=0}^{\gamma\mu_r - 1} \frac{c}{c!}\, |S_0| \left(\mu_r \left(1 + O\left(\frac{K}{n}\right)\right)\right)^c \\ &> |S_0|\, \mu_r \left(1 + O\left(\frac{K}{n}\right)\right) \sum_{c=0}^{\gamma\mu_{r-1}-1} \frac{1}{c!} \left(\mu_r \left(1 + O\left(\frac{K}{n}\right)\right)\right)^c \\ &> |S_0|\, \mu_r \left(1 + O\left(\frac{K}{n}\right)\right) e^{\mu_r(1+O(K/n))} (1 - (1-\gamma)^{-1}\, d^{\mu_r})\end{aligned}$$

for some fixed $d \in (0, 1)$ by Proposition 17. Note that $d^{\mu_r} = O(k/n)$ since $r^2 \geqslant \log(n/k)$ so it must be that

$$\sum_{c \geqslant 0} c\, |S_c| = |S_0|\, \mu_r e^{\mu_r(1+O(K/n))}.$$

Now by Proposition 9, $|L(k, n)| = |\mathscr{L}|\, (1 + o(n^{-2-u}))$. But no $k \times n$ rectangle may have more than $\mu_k n = O(n^{2u+1})$ intercalates. It follows that

$$E(N) = o(n^{u-1}) + (1 - o(n^{-2-u})) \frac{1}{|\mathscr{L}|} \sum_{R \in \mathscr{L}} N(R) = \mu_r e^{\mu_r O(K/n)}$$

by Proposition 18 again. Since $Kr^2 = o(n)$ we have that $E(N) = \mu_r e^{o(1)} = \mu_r(1 + o(1))$ as predicted. To prove our other assertion let $\varsigma = \lfloor r^{1+\varepsilon} \rfloor$ where $\varepsilon > 0$ is a small constant. Then, making use of Proposition 17,

$$\begin{aligned}P_{\mathscr{L}}(N^r(R) \leqslant \mu_r + \varsigma) &= \sum_{c=0}^{\mu_r+\varsigma} \frac{|S_c|}{|\mathscr{L}|} \\ &\geqslant \exp\left(-\mu_r \left(1 + O\left(\frac{K}{n}\right)\right)\right) \sum_{c=0}^{\mu_r+\varsigma} \frac{1}{c!} \left(\mu_r \left(1 + O\left(\frac{K}{n}\right)\right)\right)^c \\ &\geqslant (1 - o(1))\, e^{\mu_r O(K/n)}.\end{aligned}$$

Similarly,

$$P_{\mathscr{L}}(N^r(R) < \mu_r - \varsigma) \leqslant \sum_{c=0}^{\mu_r - \varsigma} \frac{|S_c|}{|\mathscr{L}|} = o(1)\, e^{\mu_r O(K/n)}.$$

Hence

$$\begin{aligned} P(|N^r(R) - \mu_r| > \varsigma) &\leqslant P_{\mathscr{L}}(|N^r(R) - \mu_r| > \varsigma) + P(R \notin \mathscr{L}) \\ &\leqslant 1 - (1 - o(1))\, e^{\mu_r O(K/n)} + o(n^{-2-u}) \end{aligned}$$

which tends to zero as required because $\mu_r K = o(n)$. ▮

8. WHEN r IS INDEPENDENT OF n

It is worth considering the case of constant r separately, because its simplicity allows strong results. For this section only we change the definition of β so that $\beta = r$. That is, we simply use the natural restriction on the number of consequential intercalates overlapping on any entry.

THEOREM 3. *Let $2 \leqslant r \leqslant k(n) \leqslant n$ for $k = o(n)$ where r is constant. The distribution of $N^r(R)$, the number of intercalates contained within the first r rows of a randomly selected $R \in L(k, n)$, converges uniformly to a Poisson distribution with mean μ_r as $n \to \infty$.*

Proof. The proof follows Proposition 18 closely. Note that (8) and (9) are both valid here. Now, as $\beta = r$ we have $P(N_{i,j}(R) \geqslant \beta) = 0$ for all i and j. It follows that (11) holds for $c \leqslant K/2$, when $S_c = \{R \in L(k, n) \colon N(R) = c\}$ and $\Psi(A) = I_r$ for all $A \in S_c$. Thus,

$$\sum_{c \geqslant 0} |S_c| \geqslant \sum_{c=0}^{K/2} \frac{1}{c!} \left(\mu_r \left(1 + O\left(\frac{K}{n}\right) \right) \right)^c |S_0| \geqslant |S_0| \exp\left(\mu_r \left(1 + O\left(\frac{K}{n}\right) \right) \right)$$

by Proposition 17 because $d_{K/2} < d_{\log n} = o(1/n)$. We then have by (9) and Proposition 8,

$$|L(k, n)| = \sum_{c \geqslant 0} |S_c| + \left| \bigcup_i \{R \colon \theta_i(R) > K\} \right| = e^{\mu_r(1 + O(K/n))}\, |S_0|.$$

Now suppose $g \geqslant 0$ is a fixed integer, and $c \leqslant g \ll K$. We can use (11) instead of Proposition 13 to show (10) and hence $c\,|S_c| = \mu_r\,|S_{c-1}| \times (1 + O(K/n))$ by (11). Therefore

$$P(N(R) = g) = \frac{|S_g|}{|L(k, n)|} = \frac{(\mu_r(1 + O(K/n)))^g}{g!\, e^{\mu_r(1 + O(K/n))}} \to \frac{\mu_r^g}{g!\, e^{\mu_r}}.$$

This shows pointwise convergence. To get the required uniformity, note that for any $\varepsilon > 0$, Proposition 8 says that only finitely many g have $P(N(R) = g) > \varepsilon$ for some n. (In fact uniformity of convergence is automatic for discrete variables.) ▮

COROLLARY. *Let* $2 \leqslant r \leqslant k = o(n)$ *for constant r. The expected number of intercalates* $E(N) \to \mu_r$ *as* $n \to \infty$.

Proof. Proposition 8 is sufficient to show that the family of random variables $\{N(\cdot)\}_{n=2}^{\infty}$ is uniformly integrable and hence $E(N)$ tends to the expectation of the limit. ▮

9. RESULTS FOR SMALL SQUARES

The theoretical results given to date have dealt with the distribution of intercalates in large rectangles. As a counterpoint, in Table I we display the results of a computer enumeration of small order Latin squares and rectangles; and the number of intercalates contained therein. Since intercalates survive permutations of the rows and columns it suffices to enumerate $L_R(k, n)$, the set of reduced $k \times n$ Latin rectangles. A rectangle is *reduced* if its first row is in natural order, and its first column contains I_k, again in natural order. Note that $|L(k, n)| = n!\,(n-1)!\,|L_R(k, n)|/(n-k)!$. The values of $|L_R(k, n)|$ we present here appeared previously in [7].

We first look at the mean number $E(N)$ of intercalates across all $k \times n$ Latin rectangles. On the basis of Theorem 2 we might expect this value to be close to μ_k. The actual values (to 4 decimal places) can be found in Table I. We are also interested in the proportion of Latin rectangles which contain no intercalates, namely $|S_0|/|L(k, n)|$. Theorem 1 suggests that this ratio will be approximately $e^{-\mu_k}$, which is to say that the values in the final column of Table I should also be close to μ_k.

Since much attention has focused on N_2 squares we provide separate counts of them in Table II. We again give the count in terms of reduced squares, so the total number is $n!\,(n-1)!$ times the value cited. In addition we give the number of isotopy classes of N_2 squares of each order. An *isotopy class* is an equivalence class under permutations of the rows, columns and symbols.

A *conjugate* of a Latin square is obtained by permuting the roles of columns, rows and symbols (for example, by transposing the matrix). The closure of an isotopy class under conjugation is a *main class*. In [1] Denniston provided representatives of the only 3 main classes of N_2 squares of order 8. The first of Denniston's main classes is the union of 2

TABLE I

Data from Enumeration of Reduced $k \times n$ Rectangles

n	k	$\|L_R(k, n)\|$	$E(N)$	μ_k	$-\log\left(\frac{\|S_0\|}{\|L(k, n)\|}\right)$
2	2	1	1	0.5	—
3	2	1	0	0.5	0
3	3	1	0	1.5	0
4	2	3	0.6667	0.5	0.4055
4	3	4	3	1.5	—
4	4	4	6	3	—
5	2	11	0.4545	0.5	0.6061
5	3	46	1.3043	1.5	1.0561
5	4	56	2.1429	3	1.2528
5	5	56	3.5714	5	2.2336
6	2	53	0.5094	0.5	0.5046
6	3	1064	1.5226	1.5	1.5060
6	4	6552	3.0879	3	2.7099
6	5	9408	5.5102	5	3.3322
6	6	9408	8.2653	7.5	5.4604
7	2	309	0.4984	0.5	0.4863
7	3	35792	1.4973	1.5	1.3966
7	4	1293216	3.0032	3	2.6698
7	5	11270400	5.0096	5	4.2173
7	6	16942080	7.5204	7.5	6.1290
7	7	16942080	10.5286	10.5	6.8606
8	2	2119	0.5002	0.5	0.5010
8	3	1673792	1.5015	1.5	1.4705
8	4	420909504	3.0051	3	2.8818
8	5	27206658048	5.0119	5	4.7139
8	6	335390189568	?	7.5	6.9320
8	7	535281401856	10.5523	10.5	9.5044
8	8	535281401856	14.0697	14	12.4502
9	2	16687	0.5000	0.5	0.5013
9	3	103443808	1.5003	1.5	1.4774
9	4	207624560256	?	3	2.9061
9	5	112681643083776	?	5	4.7690
9	6	12952605404381184	?	7.5	7.0501
9	7	224382967916691456	?	10.5	9.7415
9	8	377597570964258816	14.0204	14	12.8239
9	9	377597570964258816	18.0262	18	16.3596
10	2	148329	0.5000	0.5	0.4998
10	3	8154999232	1.5003	1.5	1.4790
10	9	7580721483160132811489280	18.0240	18	?
10	10	7580721483160132811489280	22.5300	22.5	?

TABLE II

Number of Reduced N_2 Squares of Order $n \leqslant 9$

Order	Reduced N_2 squares	Isotopy classes	Main classes
2	0	0	0
3	1	1	1
4	0	0	0
5	6	1	1
6	40	1	1
7	17760	4	2
8	2096640	14	3
9	29659631400	9802	1707

isotopy classes, while the the remaining two contain 6 isotopy classes each. Hence we have some independent validation of Table II.

10. LARGER SUBSQUARES

An order n Latin square may have a subsquare of any order up to $n/2$. Until now we have only considered intercalates—order 2 subsquares. It is natural to ask all the same questions about larger subsquares. While we are unable to provide detailed answers at this stage, a few indications may prove enlightening.

Let $E_m(S)$ denote the expected number of order m subsquares in a randomly chosen $S \in L(n, n)$. Some computed values of $E_m(S)$ are given in Table III.

The method behind Proposition 5 shows that for fixed $m \leqslant k \ll n$ the chance of a $k \times n$ Latin rectangle having an order m subsquare is at most $k^m n^{m(2-m)} \exp(O(k/n))$. Thus it seems likely that most Latin squares will

TABLE III

$E_m(S)$ to 5 Significant Figures

	$m=2$	3	4	5
$n=4$	6	0	0	0
5	3.5714	0	0	0
6	8.2653	0.20408	0	0
7	10.529	0.066636	0	0
8	14.070	0.047606	3.4710×10^{-4}	0
9	18.026	0.053368	7.2909×10^{-6}	0
10	22.530	0.053620	6.2864×10^{-6}	1.0846×10^{-9}

not have a subsquare of order greater than 3. The case of order 3 subsquares stands to be the most difficult. It may be that $\lim_{n\to\infty} E_3(S)$ exists and is finite and positive.

Conjecture. $\lim_{n\to\infty} E_3(S) = \frac{1}{18}$.

The reasoning behind this conjecture is that there are $\binom{n}{3}^3$ ways to choose 3 columns, 3 rows and 3 symbols. For each choice of 3 symbols there are 12 possible Latin squares of order 3. We expect that any specific order 3 subsquare will have close to a $1/n^9$ chance of occurring. Note that $12\binom{n}{3}^3 n^{-9} \to \frac{1}{18}$ as $n \to \infty$.

Next we consider the opposite extreme to intercalates, namely subsquares of order $n/2$ in Latin squares of even order n. We will need the following corollary of Proposition 2, which was all but proved in [13],

$$|L(n, n)| = e^{-2n^2} n^{n^2} \exp(O(n \log^2 n)). \tag{13}$$

It is not hard to see that if $S \in L(n, n)$ contains a subsquare of order $n/2$ then in fact S decomposes into 4 separate subsquares of that order. Any given S may have a number of such decompositions, but certainly no more than $\binom{n}{n/2}$ of them. Note that $\binom{n}{n/2} = o(2^n)$ by Stirling's approximation. Thus to construct a suitable S there are $o(2^{3n})$ choices for how the decomposition splits the rows, columns and symbols between the 4 subsquares. Hence

$$E_{n/2}(S) = 2^{O(n)} \frac{|L(n/2, n/2)|^4}{|L(n, n)|} = 2^{-n^2} \exp(O(n \log^2 n))$$

by (13). Clearly the probability of finding a subsquare of order $n/2$ is of the same order.

Although there is still much to prove, it seems the subtext of the title of this paper should read, "... however, almost all of those subsquares are of order 2."

REFERENCES

1. R. H. F. Denniston, Remarks on Latin squares with no subsquares of order two, *Utilitas Math.* **13** (1978), 299–302.
2. C. D. Godsil and B. D. McKay, Asymptotic enumeration of Latin rectangles, *J. Combin. Theory Ser. B* **48** (1990), 19–44.
3. K. Heinrich and W. D. Wallis, The maximum number of intercalates in a Latin square, *in* "Combinatorial Mathematics, VIII, Geelong, 1980," Lecture Notes in Math., Vol. 884, pp. 221–233, Springer-Verlag, New York/Berlin, 1981.
4. M. T. Jacobson and P. Matthews, Generating uniformly distributed random Latin squares, *J. Combin. Des.* **4** (1996), 405–437.

5. A. Kotzig, C. C. Lindner, and A. Rosa, Latin squares with no subsquares of order two and disjoint Steiner triple systems, *Utilitas Math.* **7** (1975), 287–294.
6. A. Kotzig and J. Turgeon, On certain constructions for Latin squares with no Latin subsquares of order two, *Discrete Math.* **16** (1976), 263–270.
7. B. D. McKay and E. Rogoyski, *Electron. J. Combin.* **2** (1995), N3.
8. B. D. McKay and N. C. Wormald, Uniform generation of random Latin rectangles, *Combin. Math. Combin. Comput.* **9** (1991), 179–186.
9. M. McLeish, On the existence of Latin squares with no subsquares of order two, *Utilitas Math.* **8** (1975), 41–53.
10. H. Minc, "Permanents," Encyclopedia Math. Appl., Addison–Wesley, Reading, MA, 1978.
11. A. O. Pittenger, Mappings of Latin squares, *Linear Algebra Appl.* **261** (1997), 251–268.
12. A. Schrijver, Bounds on permanents and the number of 1-factors and 1-factorizations of bipartite graphs, *in* "London Math. Soc. Lecture Note Ser.," Vol. 82, pp. 107–134, Cambridge Univ. Press, Cambridge, UK, 1983.
13. J. H. van Lint and R. M. Wilson, "A Course in Combinatorics," Cambridge Univ. Press, Cambridge, UK, 1992.