Transversals in Latin Arrays with Many Distinct Symbols

Darcy Best, Kevin Hendrey, Ian M. Wanless, Tim E. Wilson, and David R. Wood
School of Mathematical Sciences, Monash University, Clayton Vic 3800, Australia,
E-mail: darcy.best@monash.edu, kevin.hendrey@monash.edu, ian.wanless@monash.edu, timothy.e.wilson@monash.edu, david.wood@monash.edu

Received December 30, 2016; revised April 01, 2017

Published online 17 July 2017 in Wiley Online Library (wileyonlinelibrary.com).
DOI 10.1002/jcd.21566

Abstract: An array is *row-Latin* if no symbol is repeated within any row. An array is *Latin* if it and its transpose are both row-Latin. A *transversal* in an $n \times n$ array is a selection of n different symbols from different rows and different columns. We prove that every $n \times n$ Latin array containing at least $(2-\sqrt{2})n^2$ distinct symbols has a transversal. Also, every $n \times n$ row-Latin array containing at least $\frac{1}{4}(5-\sqrt{5})n^2$ distinct symbols has a transversal. Finally, we show by computation that every Latin array of order 7 has a transversal, and we describe all smaller Latin arrays that have no transversal. © 2017 Wiley Periodicals, Inc. J. Combin. Designs 26: 84–96, 2018

Keywords: *transversal; Latin array; Latin square; row-Latin*

1. INTRODUCTION

This paper deals with square arrays of symbols. By an *entry* of such an array A, we mean a triple (i, j, A_{ij}) where A_{ij} is the symbol in cell (i, j) of A. A *partial transversal of length k* in an array is a selection of k entries no pair of which agree in any of their three coordinates. A *transversal* of an $n \times n$ array is a partial transversal of length n and a *near transversal* is a partial transversal of length $n-1$. An array is *Latin* if no symbol appears more than once in any row or column. Thus, an $n \times n$ Latin array may contain anywhere from n to n^2 distinct symbols. If it has just n distinct symbols, then it is a *Latin square*. Transversals of Latin squares were first studied to construct mutually orthogonal Latin squares. Since then they have garnered a lot of interest in their own right and led to several famous long-standing conjectures (see [14] for a survey).

For even orders n, there are at least $n^{n^{3/2}(1/2-o(1))}$ Latin squares that do not have transversals [5]. However, for $n \times n$ Latin arrays, as the number of distinct symbols increases,

© 2017 Wiley Periodicals, Inc.

there must come a point beyond which it becomes impossible to avoid transversals. This paper is motivated by the question of when this threshold occurs. Let $\ell(n)$ be the least positive integer such that $\ell(n) \geqslant n$ and every Latin array of order n with at least $\ell(n)$ distinct symbols contains a transversal. This function was introduced by Akbari and Alipour [1] who calculated $\ell(n)$ for $n \leqslant 4$ and showed that $\ell(5) \geqslant 7$ and $\ell(2^k - 2) > 2^k$ for every integer $k > 2$. Counterintuitively, every Latin square of order 5 contains a transversal, but there is a Latin array of order 5 with six symbols and no transversal. Hence, it is not always true that increasing the number of symbols increases the number of transversals. Nevertheless, $\ell(n)$ is well defined since an $n \times n$ Latin array with n^2 different symbols certainly has a transversal. Akbari and Alipour put forward the following conjectures:

Conjecture 1. *For every integer $n \geqslant 3$, we have $\ell(n) \leqslant n^2/2$.*

Conjecture 2. *For every integer c, there exists a positive integer n such that $\ell(n) > n + c$.*

Up until this point, it was unknown whether there is some constant $c < 1$ such that $\ell(n) \leqslant cn^2$ for every integer $n > 1$. In Sections 2 and 3, we provide two independent proofs of such a result. The proof in Section 3 gives a better bound, but the other is of independent interest since it demonstrates an entirely different (probabilistic) approach. In Section 4, we determine $\ell(n)$ exactly for $n \leqslant 7$.

On first glance, Conjecture 1 seems very generous and that maybe $\ell(n)$ even has a linear upper bound. However, the problem is deceptively hard, and the following observation gives some hint as to why.

Proposition 1. *Let k be a nonnegative integer. If $\ell(n) \leqslant 2kn + n - k^2 - k$ for all n, then every Latin square of order n has a partial transversal of length $n - k$.*

Proof. Let L be any Latin square of order n. Let M be a Latin array of order $n + k$, which has L as the top-left $n \times n$ subarray and all remaining entries are new distinct symbols. The number of symbols in M is $n + 2nk + k^2 \geqslant \ell(n + k)$, so there must be a transversal in M. This transversal hits at most $2k$ cells in the last k rows or columns of M, so it must intersect the copy of L in at least $n - k$ cells, each of which contains a different symbol. ■

Putting $k = 1$, we see that if $\ell(n) \leqslant 3n - 2$ for all n, then every Latin square has a near transversal. This would prove a famous conjecture attributed to Brualdi (see [14]). Indeed, any linear upper bound on $\ell(n)$ would imply the existence of a constant c such that every Latin square of order n has a partial transversal of length $n - c$. The best result to date [11] is that every Latin square has a partial transversal of length $n - O(\log^2 n)$.

There is a broader setting in which quadratically many symbols is known to be best possible, namely row-Latin arrays. An array is *row-Latin* if no symbol appears more than once in any row. For every positive integer n, let $\ell_r(n)$ be the least positive integer such that $\ell_r(n) \geqslant n$ and every $n \times n$ row-Latin array with at least $\ell_r(n)$ distinct symbols contains a transversal. Barát and Wanless [4] showed that $\ell_r(n) > \frac{1}{2}n^2 - O(n)$. In Section 3, we prove that $\ell_r(n) \leqslant \lceil \frac{1}{4}(5 - \sqrt{5})n^2 \rceil$ for every integer $n > 1$.

A related problem is when repeats are allowed within a row or column, but a restriction is placed on how many times a symbol can occur in the entire square. It has been shown in [9, 10] that a transversal must exist if no symbol occurs more than cn times in a square of order n, where c is a small constant. This means that if each symbol occurs roughly the same number of times then linearly many symbols are enough to ensure a transversal.

2. PROBABILISTIC APPROACH

In this section, we use probabilistic methods to prove a bound on $\ell(n)$. Similar probabilistic methods have been used in a number of other studies of transversals of arrays [2, 3, 9, 10].

Let $\mathcal{B} = \{B_1, \dots, B_t\}$ be a set of events in a probability space. Usually the events $\mathcal{B}$ are called the *bad events* because the aim is for them to not occur. Define $\overline{B_i}$ to be the complement of the event B_i. A graph G with vertex set $\mathcal{B}$ is a *lopsidependency graph* if for all $B_i \in \mathcal{B}$ and for every subset S of the complement of the closed neighborhood of B_i in G,

$$\mathbb{P}\Big(B_i \Big| \bigcap_{j \in S} \overline{B_j}\Big) \leqslant \mathbb{P}(B_i). \tag{1}$$

Lopsidependency graphs were introduced by Erdős and Spencer [9] and are useful because they have fewer edges than a naively defined dependency graph. Intuitively, a lopsidependency graph says that the probability of an event does not increase when conditioned on an arbitrary set of nonadjacent events not occurring.

The clique Lovász local lemma by Kolipaka et al. [12] gives a condition under which none of the bad events occur. Specializing their formulation, we get:

Lemma 1. *Let $\mathcal{B} = \{B_1, \dots, B_t\}$ be a set of events with lopsidependency graph G. Let $\{K_1, \dots, K_s\}$ be a set of cliques in G covering all the edges, and assume $\kappa \geqslant \max_i |K_i|$. Suppose that no event B_i is in more than μ of the cliques $K_1, \dots, K_s$. If there exists $x \in (0, 1/\kappa)$ such that $\mathbb{P}(B_i) \leqslant x(1 - \kappa x)^{\mu-1}$ for all $1 \leqslant i \leqslant t$, then*

$$\mathbb{P}\left(\bigcap_{i=1}^{t} \overline{B_i}\right) > 0.$$

We use this lemma to prove:

Theorem 1. *Let L be a Latin array of order n. If L has at least $(229n^2 + 27n)/256 \approx 0.8945n^2$ distinct symbols, then L has a transversal.*

Proof. Suppose L has at least $n^2 - cn^2 - dn$ distinct symbols. Let σ be a permutation picked uniformly at random from the symmetric group on $\{1, 2, \dots, n\}$. Think of σ as choosing the cells $(i, \sigma(i))$ for $1 \leqslant i \leqslant n$, which might correspond to a transversal. Define the bad events,

$$\mathcal{B} = \big\{(i, j, i', j') : 1 \leqslant i < i' \leqslant n, \sigma(i) = j, \sigma(i') = j', L_{ij} = L_{i'j'}\big\}.$$

These events correspond to σ choosing a pairs of cells in L that contain the same symbol. To prove that a transversal exists we just need to prove that, with positive probability, none of the bad events occur.

The next task is to define the lopsidependency graph that will be used in applying Lemma 1. Let G be a graph with vertex set $\mathcal{B}$. An edge $\{(a, b, x, y), (a', b', x', y')\}$ is in G if and only if at least two of the cells (a, b), (x, y), (a', b') and (x', y') share a row or column. This occurs only if at least one of $x = x'$, $x = a'$, $a = x'$, $a = a'$, $y = y'$, $y = b'$, $b = y'$ or $b = b'$. Erdős and Spencer [9] showed that G is a lopsidependency graph.

Let $\mathcal{K} = \{K_1, K_2, \ldots, K_{2n}\}$ be a set of cliques of G defined as follows. Each clique corresponds to a row or column of L. An event (a, b, x, y) is in a clique K_i if (a, b) or (x, y) is in the row or column corresponding to K_i. Note that $K_i \in \mathcal{K}$ is a clique because the events in K_i share a row or column (the one corresponding to K_i) and so they are adjacent in G. These cliques cover every edge of G because two events are adjacent only if they share a row or column.

Each event in $\mathcal{B}$ corresponds to two cells in distinct rows and columns, so each event is within exactly four cliques. Thus, we take $\mu = 4$. To find the bound κ, consider a clique $K \in \mathcal{K}$ which, without loss of generality, corresponds to the first row. Each event in K corresponds to two cells of L, one in the first row and one not in that row. Let D be the set of cells outside the first row that are included in some event in K. Each cell in D shares a symbol with exactly one cell in the first row. Hence, $|K| = |D|$ and the cells not in D contain as many distinct symbols as L does. Hence $n^2 - |K| \geqslant n^2 - cn^2 - dn$, which means that we may take $\kappa = cn^2 + dn$.

Taking $x = 1/(4\kappa)$, we find that to apply Lemma 1 we need

$$\frac{1}{n(n-1)} = \mathbb{P}(B_i) \leqslant x(1-\kappa x)^3 = \frac{27}{256\kappa},$$

which is satisfied when $c = 27/256$ and $d = -27/256$. ■

3. A BETTER BOUND

In this section, we prove a better bound on $\ell(n)$ using nonprobabilistic methods. We start by proving results about general square arrays, then later use these results to give bounds on the number of symbols in transversal-free row-Latin arrays and transversal-free Latin arrays.

We call a symbol in an array A a *singleton* if it occurs exactly once in A and a *clone* otherwise. We define $R_i(A)$ and $C_j(A)$ to be the set of symbols occurring in row i and column j of A, respectively. Let $A(i \mid j)$ denote the array formed from A by deleting row i and column j and let $\Psi_{ij}(A)$ be the set of symbols that appear in A and not in $A(i \mid j)$.

Lemma 2. *Let A be a transversal-free array of order n. If $A(n \mid n)$ has a transversal and if $|R_n(A) \cup C_n(A)| \geqslant (k+1)n - 1$, then A has at most*

$$\frac{1}{2}(k^2 - 2k + 2)n^2 + \frac{1}{2}(3k-2)n$$

distinct symbols.

Proof. Assume that T is a near transversal of A that does not meet the last row or column and minimizes the number of symbols that it has from $R_n(A) \cup C_n(A)$.

We call a symbol *large* if it appears in both T and $R_n(A) \cup C_n(A)$ and *small* otherwise. Let λ be the number of large symbols. Permute the first $n-1$ rows and columns of A so that T is located along the main diagonal and all of the large symbols of T appear in the top λ rows. For $1 \leqslant i < n$, note that A_{in} and A_{ni} cannot be two different small symbols. Otherwise, $(T \setminus \{(i, i, A_{ii})\}) \cup \{(i, n, A_{in}), (n, i, A_{ni})\}$ would be a transversal of A. So there are at most $n-1$ distinct small symbols in the last row and column. Thus,

$$\lambda \geqslant |R_n(A) \cup C_n(A)| - (n-1) \geqslant (k+1)n - 1 - (n-1) = kn. \quad (2)$$

We now define a subset Γ of the entries of A in which each symbol in A is represented exactly once. We populate Γ in three steps. First, $T \subseteq \Gamma$. Second, for every small symbol s that occurs in the last row or column, select one such entry containing s and add it to Γ. Note that s cannot appear in T, by the definition of "small." Finally, for every symbol s' in A that does not appear in T or in the last row or column, select one entry with the symbol s' and add it to Γ.

We claim that if (i, j) is in the top λ rows of A with $i < j < n$, then at most one of (i, j, A_{ij}) and $(j, i, A_{j,i})$ can be in Γ. Suppose otherwise, and consider

$$\big(T \setminus \big\{(i, i, A_{ii}), (j, j, A_{jj})\big\}\big) \cup \big\{(i, j, A_{ij}), (j, i, A_{ji})\big\}. \tag{3}$$

Note that the symbol A_{ii} is contained in the last row or column of A. By the definition of Γ, we know that (i, j) and (j, i) do not have the same symbol and neither one shares a symbol with any entry in T or in the last row or column. So (3) is a near transversal that contains fewer symbols in $R_n(A) \cup C_n(A)$ than T, contradicting the choice of T. This implies that within the first λ rows and columns of $A(n \mid n)$, there are at least

$$(n-2) + (n-3) + \cdots + (n - \lambda - 1) = \lambda n - \frac{\lambda(\lambda+3)}{2}$$

entries not contained in Γ. Within the last row and column of A, there are at most $n - 1$ entries in Γ (all containing small symbols), so at least n entries are not in Γ. Thus, the number of distinct symbols in A is

$$|\Gamma| \leqslant n^2 - \left(\lambda n - \frac{\lambda(\lambda+3)}{2}\right) - n = \frac{1}{2}\lambda^2 - \left(n - \frac{3}{2}\right)\lambda + n(n-1). \tag{4}$$

This quadratic in λ decreases weakly on the integer points in the interval $kn \leqslant \lambda \leqslant n - 1$. Given (2), we may substitute $\lambda = kn$ into (4) to get the desired result. ■

Lemma 3. *Let A be an $n \times n$ array with βn^2 distinct symbols. If there are $d \geqslant 1$ clones in row i, then there is some clone A_{ij} such that*

$$|R_i(A) \cup C_j(A)| \geqslant |R_i(A)| + \frac{\beta n^2 - (n-d)(n-1) - |R_i(A)|}{d}.$$

Proof. We will endeavor to find a column j such that $|C_j(A) \setminus R_i(A)|$ is large. Without loss of generality, assume that the rightmost d columns of row i contain clones. First, remove all occurrences of the symbols in $R_i(A)$ from the array. Now, arbitrarily select a representative entry for each of the remaining symbols in the array. Note that there are no representatives in row i and so there are at most $(n-d)(n-1)$ representatives in the first $n - d$ columns. Of the original βn^2 symbols, at least $\beta n^2 - (n-d)(n-1) - |R_i(A)|$ must have their representative in the last d columns. By the pigeon-hole principle, the desired clone A_{ij} occurs in one of the last d columns. ■

Let $\mathcal{A}$ be some class of square arrays of symbols, which has the following two properties: (i) if any row and column of an array in $\mathcal{A}$ is deleted, the resulting array is in $\mathcal{A}$ and (ii) if in one entry of the array, the symbol is changed to a new symbol that appears nowhere else in the array, then the resulting array is in $\mathcal{A}$. Note that $\mathcal{L}$, the set of all Latin arrays, and $\mathcal{R}$, the set of all row-Latin arrays, both satisfy the requirements listed.

Let $\frac{1}{2} \leqslant \alpha \leqslant 1$. Define $\mathcal{M}_\mathcal{A}(\alpha)$ to be the set of transversal-free arrays in $\mathcal{A}$ whose ratio of number of distinct symbols to cells is at least α. Suppose that $\mathcal{M}_\mathcal{A}(\alpha)$ is nonempty. Define $\mathcal{M}^*_\mathcal{A}(\alpha) \subseteq \mathcal{M}_\mathcal{A}(\alpha)$ by the rule that if $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$, then no array in $\mathcal{M}_\mathcal{A}(\alpha)$ has an order smaller than A and no array in $\mathcal{M}_\mathcal{A}(\alpha)$ of the same order as A contains more distinct symbols than A. For example, both $\mathcal{M}^*_\mathcal{L}(1/2)$ and $\mathcal{M}^*_\mathcal{R}(1/2)$ consist solely of the Latin squares of order 2. For the remainder of the section, we will bound the number of symbols in arrays by examining properties of the arrays in $\mathcal{M}^*_\mathcal{A}(\alpha)$.

Lemma 4. *Let $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$ be an array of order n. If A_{ij} is a singleton, then $|\Psi_{ij}(A)| > \alpha(2n-1)$ and $R_i(A)$ (resp., $C_j(A)$) contains more than $(2\alpha-1)n$ symbols that appear only in row i (resp., column j) of A.*

Proof. Any array of order 1 has a transversal, so $n \geqslant 2$. There is no transversal T of $A(i \mid j)$, or else $T \cup \{(i, j, A_{ij})\}$ would be a transversal of A. As $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$, we have that $A(i \mid j) \notin \mathcal{M}_\mathcal{A}(\alpha)$, so the number of distinct symbols in $A(i \mid j)$ is strictly less than $\alpha(n-1)^2$. Thus,

$$|\Psi_{ij}(A)| > \alpha n^2 - \alpha(n-1)^2 = \alpha(2n-1).$$

At most $n-1$ of the symbols in $\Psi_{ij}(A) \setminus \{A_{ij}\}$ appear in $C_j(A)$, so at least

$$|\Psi_{ij}(A)| - (n-1) > \alpha(2n-1) - (n-1) \geqslant (2\alpha - 1)n$$

symbols appear in row i and nowhere else in A. A similar argument applies to $C_j(A)$. ■

Lemma 5. *Let $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$ be an array of order n. If A_{ij} is a clone and $|R_i(A) \cup C_j(A)| \geqslant (k+1)n - 1$, then A has at most*

$$\frac{1}{2}(k^2 - 2k + 2)n^2 + \frac{1}{2}(3k-2)n$$

distinct symbols.

Proof. Without loss of generality, $i = j = n$. Create A' by changing the symbol in the (n, n) cell of A to a symbol that did not previously appear in A. Since A_{ij} is a clone in A, we know that A' contains strictly more symbols than A. Since $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$, we conclude that A' has a transversal, although A does not. Hence there is a near transversal of A that does not meet row n or column n. By applying Lemma 2, the result follows. ■

In the best case, Lemma 5 falls just short of proving Conjecture 1.

Corollary 1. *Let $A \in \mathcal{M}^*_\mathcal{A}(\alpha)$ be an array of order n. If A_{ij} is a clone and $|R_i(A) \cup C_j(A)| = 2n - 1$, then A has at most $(n^2 + n)/2$ distinct symbols.*

Lemmas 3–5 form the main framework needed to bound the number of symbols. We will utilize Lemmas 3 and 4 in different ways to find an entry (i, j, k), where k is a clone and row i and column j contain many different symbols. We then apply Lemma 5 to bound the number of symbols overall. The following subsections concentrate on specific classes for $\mathcal{A}$.

A. Row-Latin Arrays

In this subsection, we consider $\mathcal{A} = \mathcal{R}$, the set of row-Latin arrays.

Lemma 6. *Let $M \in \mathcal{M}^*_{\mathcal{R}}(\alpha)$ be a row-Latin array of order n. There exists a clone M_{ij} for which $|R_i(M) \cup C_j(M)| \geqslant 2\alpha n - 1$.*

Proof. First suppose that there is a clone M_{ij} that appears in the same column as a singleton. By Lemma 4, $C_j(M)$ contains at least $(2\alpha - 1)n$ symbols that appear only in $C_j(M)$. One of these symbols may be M_{ij}, but

$$|R_i(M) \cup C_j(M)| = |R_i(M)| + |C_j(M) \setminus R_i(M)| \geqslant n + (2\alpha - 1)n - 1 = 2\alpha n - 1,$$

as required.

Hence, we may assume that no column contains a singleton and a clone. Let d be the number of columns that contain clones.

If $d \leqslant n/2$, then we can find a transversal in the following way. Let R be the $n \times d$ subarray of M that contains the clones of M. A result of Drisko [6] implies that M has a partial transversal of length d that is wholly inside R. Since this partial transversal covers all columns that contain clones, it can trivially be extended to a transversal using singletons.

So we may assume that $d > n/2$. Since each row contains d clones, we may use Lemma 3 with $\beta \geqslant \alpha$ to find some clone M_{ij} such that

$$|R_i(M) \cup C_j(M)| \geqslant \frac{\alpha - 1}{d} n^2 + 2n - 1 > 2(\alpha - 1)n + 2n - 1 = 2\alpha n - 1. \qquad \blacksquare$$

We now show one of our main results, that row-Latin arrays with many symbols must have a transversal.

Theorem 2. *Let L be a row-Latin array of order n. If L has at least $\frac{1}{4}(5 - \sqrt{5})n^2 \approx 0.6910n^2$ distinct symbols, then L has a transversal.*

Proof. Aiming for a contradiction, suppose that $L \in \mathcal{M}_{\mathcal{R}}(\alpha)$ for $\alpha = (5 - \sqrt{5})/4$. Then there exists $M \in \mathcal{M}^*_{\mathcal{R}}(\alpha)$. Let M have order m. By Lemma 6, there is a clone M_{ij} such that $|R_i(M) \cup C_j(M)| \geqslant 2\alpha m - 1$. By Lemma 5, the number of distinct symbols in M is at most

$$\frac{1}{2}\left((2\alpha - 1)^2 - 2(2\alpha - 1) + 2\right) m^2 + \frac{1}{2}\left(3(2\alpha - 1) - 2\right) m = \alpha m^2 - \frac{1}{4}(3\sqrt{5} - 5)m.$$

This contradicts the fact that M has at least αm^2 distinct symbols, and we are done. $\blacksquare$

B. Latin Arrays

In this subsection, we consider $\mathcal{A} = \mathcal{L}$, the set of Latin arrays.

We call a Latin array L of order *n focused* if every singleton in L occurs in a row or a column that contains only singletons and $|\Psi_{ij}(L)| = 2n - 1$ for some (i, j) (i.e., row i and column j contain only singletons). We deal with focused and unfocused arrays separately.

For focused arrays we use the following simple adaptation of a result of Woolbright [15]. The original proof was for Latin squares, but it works without change for Latin arrays (in fact for row-Latin arrays, but we do not need that).

Theorem 3. *Let L be an $n \times n$ Latin array and $0 \leqslant t < n$. If $(n-t)^2 > t$, then L has a partial transversal of length $t+1$.*

In the following result, recall that we assume $\alpha \geqslant 1/2$.

Lemma 7. *Let $M \in \mathcal{M}^*_{\mathcal{L}}(\alpha)$ be a Latin array of order n. If M is focused, then M contains at most $\frac{1}{8}(6-\sqrt{2})n^2 \approx 0.5732n^2$ distinct symbols.*

Proof. Let $\delta = \lceil (2\alpha - 1)n \rceil$. Suppose that M has r rows and c columns that contain singletons. Permute the rows and columns of M so that these singletons occur in the top r rows and leftmost c columns. Since M is focused, $\min(r, c) \geqslant 1$ and the bottom-right $(n-r) \times (n-c)$ subarray does not contain any singletons. Thus, if we consider any singleton in the last row or last column, we get $\min(r, c) \geqslant \delta$ by Lemma 4.

If $\alpha \geqslant 3/4$, then $\min(r, c) \geqslant n/2$ and so $\{(i, n-i+1) : 1 \leqslant i \leqslant n\}$ is a set of cells containing only singletons, contradicting the fact that M has no transversal. So $\alpha < 3/4$.

Let M' be the subarray formed by the last $n-\delta$ rows and columns of M. Suppose that M has a partial transversal of length $n-2\delta$ wholly inside M'. Then this partial transversal can easily be extended to a transversal by selecting singletons in the first δ rows and δ columns of M. By assumption M has no transversal, so applying Theorem 3 to M' we find that $(\delta+1)^2 \leqslant n - 2\delta - 1$. Hence

$$0 \geqslant \delta^2 + 4\delta + 2 - n \geqslant (2\alpha-1)^2 n^2 + (8\alpha - 5)n + 2. \tag{5}$$

From the discriminant of this quadratic we learn that $32\alpha^2 - 48\alpha + 17 \geqslant 0$. Since $\alpha < 3/4$ we have $\alpha \leqslant (6-\sqrt{2})/8$. ∎

For any $\alpha > 1/2$, it is noteworthy that (5) fails for all large n. So we get an asymptotic version of Conjecture 1 holding for focused Latin arrays. We are not able to reach such a strong conclusion for the unfocused case.

Lemma 8. *Let $M \in \mathcal{M}^*_{\mathcal{L}}(\alpha)$ be a Latin array of order n. If M is unfocused, then there exists some clone M_{ij} such that $|R_i(M) \cup C_j(M)| \geqslant (\alpha+1)n - 1$.*

Proof. First, we consider the case that M has some row or column that contains only clones. Without loss of generality, row i contains only clones. By Lemma 3, there is some clone M_{ij} such that $|R_i(M) \cup C_j(M)| \geqslant n + (\alpha n^2 - n)/n = (\alpha+1)n - 1$.

Second, we consider the case that every row and column of M contains a singleton. Since M is unfocused, there is some singleton M_{ik} such that there is a clone in both row i and column k. By Lemma 4, we have $|\Psi_{ik}(M)| > \alpha(2n-1)$. Each symbol in $\Psi_{ik}(M)$ appears in either $R_i(M)$ or $C_k(M)$. Also, M_{ik} appears in both $R_i(M)$ and $C_k(M)$, so without loss of generality, $R_i(M)$ contains at least $(|\Psi_{ik}(M)|+1)/2 > \alpha(n-1/2) + 1/2 \geqslant \alpha n$ symbols that are in $\Psi_{ik}(M)$. Let M_{ij} be a clone in the same row as M_{ik}. Except possibly for M_{ij}, none of the n symbols in $C_j(M)$ are in $\Psi_{ik}(M)$. Hence, $|R_i(M) \cup C_j(M)| \geqslant \alpha n + n - 1$ as required. ∎

We now show a stronger result than Theorem 2 holds for Latin arrays.

Theorem 4. *Let L be a Latin array of order n. If L has at least $(2-\sqrt{2})n^2 \approx 0.5858n^2$ distinct symbols, then L has a transversal.*

Proof. Aiming for a contradiction, suppose that $L \in \mathcal{M}_{\mathcal{L}}(\alpha)$ for $\alpha = 2-\sqrt{2}$. Then there exists $M \in \mathcal{M}^*_{\mathcal{R}}(\alpha)$. Let M have order m. Note that M cannot be focused, by

Lemma 7. So, there is a clone M_{ij} such that $|R_i(M) \cup C_j(M)| \geqslant (\alpha + 1)m - 1$, by Lemma 8. By Lemma 5, the number of distinct symbols in M is at most

$$\frac{1}{2}\left(\alpha^2 - 2\alpha + 2\right) m^2 + \frac{1}{2}(3\alpha - 2)\, m = \alpha m^2 - \frac{1}{2}(3\sqrt{2} - 4)m.$$

This contradicts the fact that M has at least αm^2 distinct symbols, and we are done. ■

4. SMALL VALUES

We now shift our attention to small values of n where we can compute $\ell(n)$ exactly. Akbari and Alipour [1] determined $\ell(n)$ for $n \leqslant 4$. We extend this search to $n \leqslant 7$ and catalog all Latin arrays of small orders with no transversals. For $n \geqslant 8$, computing $\ell(n)$ seems challenging. We will mention a couple of unsuccessful attempts to find examples that would provide some insight.

Following [8], we say that two Latin arrays are *trisotopic* if one can be changed into the other by permuting rows, permuting columns, permuting symbols, and/or transposing. The set of all Latin arrays trisotopic to a given array is a *trisotopy class*. The number of transversals is a trisotopy class invariant, so to find all transversal-free Latin arrays of a given order it suffices to consider trisotopy class representatives. However, for orders $n > 5$ it becomes difficult to construct a representative of every trisotopy class. The following method allows us to push our results a couple of orders further.

Let L be a transversal-free Latin array. In the first two rows of L, select two entries that do not share a column or symbol (this can always be done for $n \geqslant 3$). Without loss of generality, we may assume that these two entries are $(1, 1, x)$ and $(2, 2, y)$. Let L' be the bottom-right $(n - 2) \times (n - 2)$ subarray of L where all occurrences of x and y are replaced with a hole (i.e., a cell with no symbol; we forbid holes from being chosen in a transversal or partial transversal). There cannot be a partial transversal of length $n - 2$ in L', otherwise the corresponding entries in L, together with $(1, 1, x)$ and $(2, 2, y)$, would form a transversal of L.

Thus, to search for transversal-free Latin arrays of order n, we first build a catalog $\mathcal{C}_{n-2}$ of trisotopy class representatives of transversal-free partial Latin arrays of order $n - 2$ with at most two holes in each row and each column. Starting with this catalog, we can reverse the argument above. At least one representative of each trisotopy class of transversal-free Latin array of order n can be obtained by taking an element of $\mathcal{C}_{n-2}$, filling its holes with x and y, then extending it to a Latin array of order n.

By the above technique, we are able to give a complete catalog of the transversal-free trisotopy classes for orders $n \leqslant 7$. Table I gives the value of $\ell(n)$ and the number of trisotopy classes with a specific number of symbols.

Representatives of the trisotopy classes of transversal-free Latin arrays of orders 4 and 5 are:

$$\begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} \\ \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{a} \\ \mathtt{c} & \mathtt{d} & \mathtt{a} & \mathtt{b} \\ \mathtt{d} & \mathtt{a} & \mathtt{b} & \mathtt{c} \end{pmatrix}, \quad \begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} \\ \mathtt{b} & \mathtt{c} & \mathtt{a} & \mathtt{e} \\ \mathtt{c} & \mathtt{a} & \mathtt{d} & \mathtt{b} \\ \mathtt{e} & \mathtt{d} & \mathtt{b} & \mathtt{a} \end{pmatrix}, \quad \begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} \\ \mathtt{b} & \mathtt{c} & \mathtt{a} & \mathtt{e} & \mathtt{f} \\ \mathtt{c} & \mathtt{a} & \mathtt{b} & \mathtt{f} & \mathtt{d} \\ \mathtt{e} & \mathtt{d} & \mathtt{f} & \mathtt{c} & \mathtt{a} \\ \mathtt{d} & \mathtt{f} & \mathtt{e} & \mathtt{a} & \mathtt{b} \end{pmatrix}, \quad \begin{pmatrix} \mathtt{f} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} \\ \mathtt{b} & \mathtt{c} & \mathtt{a} & \mathtt{e} & \mathtt{f} \\ \mathtt{c} & \mathtt{a} & \mathtt{b} & \mathtt{f} & \mathtt{d} \\ \mathtt{e} & \mathtt{d} & \mathtt{f} & \mathtt{c} & \mathtt{a} \\ \mathtt{d} & \mathtt{f} & \mathtt{e} & \mathtt{a} & \mathtt{b} \end{pmatrix}.$$

TABLE I. Values of $\ell(n)$ and the number of trisotopy classes of transversal-free Latin arrays.

		Trisotopy classes			
n	$\ell(n)$	n Symbols	$n+1$ Symbols	$n+2$ Symbols	Total
2	3	1	–	–	1
3	3	–	–	–	0
4	6	1	1	–	2
5	7	–	2	–	2
6	9	8	19	1	28
7	7	–	–	–	0

Note that our two representatives of order 5 differ only in their first entry. Both can be completed to Latin squares of order 6; in the first case this Latin square has no transversals, but in the second case it has eight transversals.

Many of the transversal-free Latin arrays for order 6 also turn out to be quite similar to one another. There are exactly 28 trisotopy classes for $n = 6$. Previously, nine of these classes were known: eight Latin squares and the array constructed by Akbari and Alipour [1] by removing two rows and columns from the elementary abelian Cayley table of order 8. We will now describe the 19 transversal-free trisotopy classes of order 6 with seven symbols. We will denote their representative arrays by $L_1, L_2, \ldots, L_{19}$. Let

$$L_1 = \begin{pmatrix} \boxed{\mathtt{a}} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} & \mathtt{f} \\ \mathtt{b} & \boxed{\mathtt{c}} & \mathtt{a} & \mathtt{f} & \mathtt{d} & \mathtt{e} \\ \mathtt{c} & \mathtt{a} & \boxed{\mathtt{b}} & \mathtt{e} & \mathtt{f} & \mathtt{d} \\ \mathtt{d} & \mathtt{e} & \mathtt{f} & \mathtt{g} & \mathtt{b} & \mathtt{c} \\ \mathtt{f} & \mathtt{d} & \mathtt{e} & \mathtt{b} & \mathtt{g} & \mathtt{a} \\ \mathtt{e} & \mathtt{f} & \mathtt{d} & \mathtt{c} & \mathtt{a} & \mathtt{g} \end{pmatrix} \quad \text{and} \quad L' = \begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} & \boxed{\mathtt{f}} \\ \mathtt{c} & \mathtt{f} & \mathtt{b} & \boxed{\mathtt{e}} & \mathtt{d} & \mathtt{a} \\ \mathtt{b} & \mathtt{c} & \mathtt{e} & \mathtt{f} & \boxed{\mathtt{a}} & \mathtt{d} \\ \mathtt{d} & \mathtt{e} & \boxed{\mathtt{f}} & \mathtt{a} & \mathtt{b} & \mathtt{c} \\ \boxed{\mathtt{e}} & \mathtt{d} & \mathtt{a} & \mathtt{c} & \mathtt{f} & \mathtt{b} \\ \mathtt{f} & \boxed{\mathtt{a}} & \mathtt{d} & \mathtt{b} & \mathtt{c} & \mathtt{e} \end{pmatrix}.$$

From L', we define $L_2, \ldots, L_8$ by changing some entries on the main diagonal to a new symbol, g, in the following way. Let

$$R' \in \big\{\{1, 2, 3, 4, 5, 6\}, \{1, 2, 4, 5, 6\}, \{1, 3, 4, 5\}, \{1, 3, 6\}, \{1, 4\}, \{2, 3, 5, 6\}, \{3, 4, 5, 6\}\big\}.$$

For all $r \in R'$, change the symbol on the main diagonal in row r of L' to g. It turns out that changing the shaded entries in L_1 to g results in an array that is trisotopic to L_2. Next, L_9 is obtained by changing the symbol of the shaded entries in L' to a new symbol, g. Let

$$L_{10} = \begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} & \mathtt{f} \\ \mathtt{b} & \mathtt{c} & \mathtt{g} & \mathtt{a} & \mathtt{f} & \mathtt{e} \\ \mathtt{c} & \mathtt{f} & \mathtt{d} & \mathtt{g} & \mathtt{a} & \mathtt{b} \\ \mathtt{d} & \mathtt{a} & \mathtt{f} & \mathtt{e} & \mathtt{g} & \mathtt{c} \\ \mathtt{e} & \mathtt{g} & \mathtt{a} & \mathtt{f} & \mathtt{c} & \mathtt{d} \\ \mathtt{g} & \mathtt{e} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{a} \end{pmatrix} \quad \text{and} \quad L'' = \begin{pmatrix} \mathtt{a} & \mathtt{b} & \mathtt{c} & \mathtt{d} & \mathtt{e} & \mathtt{f} \\ \mathtt{b} & \mathtt{c} & \mathtt{a} & \mathtt{e} & \mathtt{f} & \mathtt{d} \\ \mathtt{c} & \mathtt{a} & \mathtt{b} & \mathtt{f} & \mathtt{d} & \mathtt{e} \\ \mathtt{d} & \mathtt{e} & \mathtt{f} & \mathtt{a} & \mathtt{c} & \mathtt{b} \\ \mathtt{e} & \mathtt{f} & \mathtt{d} & \mathtt{c} & \mathtt{b} & \mathtt{a} \\ \mathtt{f} & \mathtt{d} & \mathtt{e} & \mathtt{b} & \mathtt{a} & \mathtt{c} \end{pmatrix}.$$

From L_{10}, we can either change the symbol in the (3, 3) cell to $\mathtt{e}$, giving L_{11}, or change the symbol in the (4, 4) cell to $\mathtt{b}$, giving L_{12}.

The array L_{13} is obtained by changing the $\mathtt{d}$ in rows 2 and 3 of L'' to $\mathtt{g}$, as well as changing the $\mathtt{f}$ in row 2 to $\mathtt{d}$. Next, L_{14} is obtained by changing the $\mathtt{e}$ in row 3 of L_{13} to $\mathtt{d}$. From the Latin square L'', any subset of entries that contain $\mathtt{d}$ may be changed to a new symbol, $\mathtt{g}$. This gives rise to five trisotopy classes. In particular, we define $L_{15}, \ldots, L_{19}$ by changing some occurrences of $\mathtt{d}$ to a new symbol, $\mathtt{g}$, in the following way. Let

$$R'' \in \big\{\{1\}, \{1, 2\}, \{1, 2, 3\}, \{1, 3, 5\}, \{1, 4\}\big\}.$$

For all $r \in R''$, change the $\mathtt{d}$ in row r of L'' to $\mathtt{g}$.

One can check that $L_{15}, \ldots, L_{19}$ are transversal-free by exhaustive computation, but next we give a reason why they have no transversals. The argument is in the style of the highly successful Δ-lemma (see [14]). Let L be any Latin array obtained by replacing any subset of the occurrences of $\mathtt{d}$ in L'' by $\mathtt{g}$. Define functions ρ, ν to $\mathbb{Z}_3$ by

$$\begin{gathered}\rho(1) = \rho(4) = 0, \quad \rho(2) = \rho(5) = 1, \quad \rho(3) = \rho(6) = 2,\\ \nu(\mathtt{a}) = \nu(\mathtt{d}) = \nu(\mathtt{g}) = 0, \quad \nu(\mathtt{b}) = \nu(\mathtt{e}) = 1, \quad \nu(\mathtt{c}) = \nu(\mathtt{f}) = 2.\end{gathered}$$

Define a function Δ from the entries of L to $\mathbb{Z}_3$ by $\Delta(r, c, s) = \rho(r) + \rho(c) - \nu(s)$. Let D denote the bottom-right 3×3 subsquare of L. Suppose that T is a transversal of L and that $\bar{s}$ is the only symbol in $\{\mathtt{a}, \mathtt{b}, \ldots, \mathtt{g}\}$ that does not appear in T. Then

$$\sum_{(r,c,s)\in T} \Delta(r, c, s) = 2\sum_{i=1}^{6} \rho(i) - \sum_{(r,c,s)\in T} \nu(s) = \nu(\bar{s}). \tag{6}$$

Also, if T includes x entries in D then overall it has $2x$ entries with symbols in $\{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$, which means that $x = 1$ and $\bar{s} \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}$. However, $\Delta(r, c, s) = 0$ for all entries of L, except those in D, where $\Delta(r, c, s) = \nu(s)$. Hence to satisfy (6), the symbol in the only entry of T in D has to be $\bar{s}$, contradicting the fact that this symbol does not appear in T.

The argument we have just presented is specific to order $n = 6$ and does not seem to easily generalize to arrays of larger orders. When performing the search for transversal-free Latin arrays of order $n = 7$, we found 15,611,437 trisotopy classes of transversal-free partial Latin arrays of order 5 with at most two holes in each row and column. Table II provides counts of the trisotopy classes based on number of holes and number of symbols. Since none of these arrays can extend to a Latin array of order 7 with no transversals, we have the following result.

Theorem 5. *Every Latin array of order 7 has a transversal.*

The approach that we used to prove Theorem 5 is infeasible for $n \geqslant 8$, although we did examine certain interesting sets of Latin arrays of order 8. There are 68 different transversal-free Latin squares of order 8, up to trisotopy. We also considered all Latin arrays that are obtained by removing one row and one column from a Latin square of order 9. We could immediately eliminate any square of order 9 that contains a transversal through every entry. Latin squares that do not contain a transversal through every entry are called *confirmed bachelor squares*. The confirmed bachelor squares of order 9 were generated for [7], providing us with a set of trisotopy class representatives. None of these squares has an order 8 transversal-free subarray. Lastly, we searched all Latin arrays of

TABLE II. Counts of trisotopy classes of transversal-free 5 × 5 partial Latin arrays, categorized by number of symbols and number of holes.

	Number of symbols										
Number of holes	3	4	5	6	7	8	9	10	11	12	13
0	–	–	–	2	–	–	–	–	–	–	–
1	–	–	1	17	–	–	–	–	–	–	–
2	–	–	9	271	13	–	–	–	–	–	–
3	–	–	137	4,893	1,179	61	5	–	–	–	–
4	–	–	1,484	54,911	31,342	5,539	1,906	462	62	4	–
5	–	3	10,686	341,251	319,750	58,257	9,823	1,175	86	4	–
6	–	19	48,436	1,155,690	1,420,192	299,951	33,366	1,953	56	–	–
7	–	151	124,275	2,045,859	2,754,143	670,137	63,480	2,676	30	–	–
8	–	632	159,295	1,720,463	2,198,260	549,316	43,912	1,710	78	8	1
9	–	916	80,609	557,285	603,320	134,056	7,120	148	7	1	–
10	3	320	9,420	40,418	34,218	6,014	159	1	–	–	–

order 8 with exactly 9 symbols where one of the symbols appears at most 4 times. None of these were transversal-free. The arrays that we have checked are a tiny subset of all Latin arrays of order 8. Without theoretical insight, it seems hopeless to check them all. So all that we can conclude at this stage is that $\ell(8) \geqslant 9$.

It is known that all Latin squares of order 9 have transversals (see, e.g., [7]). We tried, unsuccessfully, to build a transversal-free Latin array of order 9. We did this by removing a row and column from Latin squares of order 10. The squares that we used were representatives of all trisotopy classes for which the autoparatopy group has order 3 or higher, as generated for [13].

The results of our investigations lead us to be skeptical that Conjecture 2 is true. However, proving that it is false is likely to be extremely hard, for the reasons explained after Proposition 1. Yet, it also seems hard to prove a subquadratic bound on $\ell(n)$, or even to prove Conjecture 1. For $\ell_r(n)$ we know more. Thanks to [4] and Theorem 2, we know that $\frac{1}{2}n^2 - O(n) < \ell_r(n) \leqslant \lceil \frac{1}{4}(5 - \sqrt{5})n^2 \rceil$.

ACKNOWLEDGMENTS

The authors are grateful to Gweneäl Joret and János Barát for interesting discussions. After submission of our manuscript we learnt of [3], which obtains our Theorem 1 by a similar method and also a weaker version of our Theorem 4. D. Best's research was supported by Endeavour Postgraduate Scholarship and the NSERC PGS-D. I. M. Wanless' research was supported by ARC grant DP150100506. D. R. Wood's research was supported by ARC grant FT1310464.

REFERENCES

[1] S. Akbari and A. Alipour, Transversals and multicolored matchings, J Combin Des 12(5) (2004), 325–332.

[2] N. Alon, J. Spencer, and P. Tetali, Covering with Latin transversals, Discrete Appl Math 57(1) (1995), 1–10.

[3] J. Barát and Z. L. Nagy, Transversals in generalized Latin squares, arXiv:1701.08220 [math.CO].

[4] J. Barát and I. M. Wanless, Rainbow matchings and transversals, Australas J Combin 59 (2014), 211–217.

[5] N. Cavenagh and I. M. Wanless, Latin squares with no transversals, The Electronic Journal of Combinatorics, 24(2), #P2.45, 2017.

[6] A. A. Drisko, Transversals in row-Latin rectangles, J Combin Theory Ser A 84(2) (1998), 181–195.

[7] J. Egan and I. M. Wanless, Latin squares with restricted transversals, J Combin Des 20(7) (2012), 344–361.

[8] J. Egan and I. M. Wanless, Enumeration of MOLS of small order, Math Comp 85(298) (2016), 799–824.

[9] P. Erdős and J. Spencer, Lopsided Lovász local lemma and Latin transversals, Discrete Appl Math 30(2–3) (1991), 151–154.

[10] H.-L. Fu and C.-C. Lee, Transversals in $m \times n$ arrays, J Stat Theory Pract 6(1) (2012), 139–146.

[11] P. Hatami and P. W. Shor, A lower bound for the length of a partial transversal in a Latin square, J Combin Theory Ser A 115(7) (2008), 1103–1113.

[12] K. Kolipaka, M. Szegedy, and Y. Xu, "A sharper local lemma with improved applications," in *Approximation, Randomization, and Combinatorial Optimization*, Gupta A., Jansen K., Rolim J., Servedio R., vol. 7408, *Lecture Notes in Computer Science*, Springer, Heidelberg, 2012, pp. 603–614.

[13] B. D. McKay, A. Meynert, and W. Myrvold, Small Latin squares, quasigroups, and loops, J Combin Des 15(2) (2007), 98–119.

[14] I. M. Wanless, "Transversals in Latin squares: A survey," in *Surveys in Combinatorics 2011*, Robin Chapman. vol. 392, *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 2011, pp. 403–437.

[15] D. E. Woolbright, An $n \times n$ Latin square has a transversal with at least $n - \sqrt{n}$ distinct symbols, J Combin Theory Ser A 24(2) (1978), 235–237.