A partial latin squares problem posed by Blackburn

Ian M. Wanless
Department of Computer Science
Australian National University
ACT 0200 Australia
e-mail: imw@cs.anu.edu.au

Blackburn asked for the largest possible density of filled cells in a partial latin square with the property that whenever two distinct cells P_{ab} and P_{cd} are occupied by the same symbol the 'opposite corners' P_{ad} and P_{bc} are blank. We show that, as the order n of the partial latin square increases, a density of at least $\exp\left(-c(\log n)^{1/2}\right)$ is possible using a diagonally cyclic construction, where c is a positive constant. The question of whether a constant density is achievable remains, but we show that a density exceeding $\frac{1}{5}(\sqrt{11}-1)(1+4/n)$ is not possible.

We say that a partial latin square P has the Blackburn property if whenever two distinct cells P_{ab} and P_{cd} are occupied by the same symbol the 'opposite corners' P_{ad} and P_{bc} are blank. The problem of filling as many cells without violating this property was posed by Simon Blackburn [2]. His motivation was an application in perfect hash families and the problem was originally posed in those terms. Examples of order 6 and 8 of partial latin squares with the Blackburn property are shown in (1). The example of order 8 is a re-arrangement of an example given by Blackburn [2].

$$\begin{pmatrix} 1 & - & - & - & 5 & 4 \\ - & 1 & - & 5 & - & 3 \\ - & - & 1 & 6 & 3 & - \\ 3 & 4 & - & 2 & - & - \\ 6 & - & 4 & - & 2 & - \\ - & 6 & 5 & - & - & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & - & - & - & 5 & 7 & 4 & - \\ - & 1 & - & - & 8 & 6 & - & 4 \\ - & - & 1 & - & 3 & - & 6 & 7 \\ - & - & - & 1 & - & 3 & 8 & 5 \\ 6 & 7 & 4 & - & 2 & - & - & - \\ 8 & 5 & - & 4 & - & 2 & - & - \\ 3 & - & 5 & 7 & - & - & 2 & - \\ - & 3 & 8 & 6 & - & - & - & 2 \end{pmatrix} \tag{1}$$

Bulletin of the ICA, Volume 42 (2004), 76-80

By the density of a partial latin square of order n we mean the number of filled cells divided by n^2. The density of both examples in (1) is $\frac{1}{2}$. In this note we investigate asymptotics for $\delta(n)$, the maximum density of any partial latin square of order n with the Blackburn property. The principal question, which shall remain open, is whether $\delta(n) = o(1)$ or whether a constant density is achievable.

Let k_σ denote the number of occurrences of a symbol σ. By relabelling if necessary, we may assume that

$$k_1 \geq k_2 \geq k_3 \geq \cdots \geq k_n \geq 0. \tag{2}$$

We say that a cell (i,j) is *wasted* by a symbol if that symbol occurs in row i and also in column j (but not in the cell (i,j) itself) and hence the cell (i,j) must be vacant to obey the Blackburn property. By $w(\sigma)$ we denote the number of cells wasted by a symbol σ and by $w(\sigma \cap \tau)$ we denote the number of cells wasted by σ and also wasted by τ. It should be clear that $w(\sigma) = k_\sigma(k_\sigma - 1)$ for every σ. We also have:

Lemma 1 $w(\sigma \cap \tau) \leq \lfloor \frac{1}{2} k_\sigma \rfloor \lceil \frac{1}{2} k_\sigma \rceil \leq \frac{1}{4} k_\sigma^2$ *for any symbols* σ *and* τ.

Proof: We partition the occurrences of σ into (a) those which lie in the same row as an occurrence of τ, (b) those which lie in the same column as an occurrence of τ, and (c) those which have neither property (a) nor (b).

We remark that by the Blackburn property, (a) and (b) are disjoint sets. If a, b, c denote the respective numbers of occurrences of σ of types (a), (b) and (c) then $w(\sigma \cap \tau) = ab$. Thus, subject to the restriction that $a + b + c = k_\sigma$ we see that $w(\sigma \cap \tau)$ is maximised by taking $c = 0$ and $\{a, b\} = \{\lfloor \frac{1}{2} k_\sigma \rfloor, \lceil \frac{1}{2} k_\sigma \rceil\}$. ⊙

Blackburn [2] observed that from Corollary 2 in his joint paper [3] with Wild it follows that $\delta(n) \leq \frac{1}{2} n + o(1)$. With the above lemma we can improve the constant $\frac{1}{2}$ to approximately 0.463 as follows:

Theorem 1 $\delta(n) < \frac{1}{5}(\sqrt{11} - 1)(1 + \frac{4}{n})$ *for all* n.

Proof: For $n \leq 3$ the result follows trivially from $\delta(n) \leq 1$, so we assume in the remainder of the proof that $n \geq 4$.

Using Lemma 1 we know that the number W of cells which are wasted by at least one of the symbols satisfies

$$\begin{aligned} W &\geq w(1) + w(2) + w(3) + w(4) \\ &\quad - w(1 \cap 2) - w(1 \cap 3) - w(2 \cap 3) - w(1 \cap 4) - w(2 \cap 4) - w(3 \cap 4) \\ &\geq k_1(k_1 - 1) + k_2(k_2 - 1) + k_3(k_3 - 1) + k_4(k_4 - 1) \\ &\quad - \frac{1}{4}k_2^2 - \frac{1}{4}k_3^2 - \frac{1}{4}k_3^2 - \frac{1}{4}k_4^2 - \frac{1}{4}k_4^2 - \frac{1}{4}k_4^2 \\ &= k_1^2 + \frac{3}{4}k_2^2 + \frac{1}{2}k_3^2 + \frac{1}{4}k_4^2 - k_1 - k_2 - k_3 - k_4. \end{aligned} \tag{3}$$

Suppose that $t = k_1+k_2+k_3+k_4$ is fixed. The minimum of (3) subject to (2) is then achieved by taking $k_1 = k_2 = k_3 = k_4 = \frac{1}{4}t$. However, we know that $k_i \le k_4 \le \frac{1}{4}t$ for all $i \ge 4$ so that $\sum k_i \le t + \frac{1}{4}t(n-4) = nt/4$. Hence, since $\delta(n) = \sum k_i/n^2 \le t/(4n)$ the theorem must be true if $t/(4n) \le (\sqrt{11}-1)/5$. So we suppose that $t > 4n(\sqrt{11}-1)/5$, in which case (3) shows that

$$W \ge \frac{5}{2}\left(\frac{t}{4}\right)^2 - t > \frac{5}{2}\left(\frac{1}{5}(\sqrt{11}-1)n\right)^2 - \frac{4}{5}n(\sqrt{11}-1).$$

Therefore $\delta(n) \le 1 - W/n^2 < \frac{1}{5}(\sqrt{11}-1)(1+4/n)$ as required. ⊙

As stated, Theorem 1 shows that density of $\frac{1}{2}$ is not achievable for $n > 50$. In fact, from (3) we find that density of $\frac{1}{2}$ is not achievable when $n > 16$, because in that case $\frac{5}{2}(t/4)^2 - t > \frac{1}{2}n^2$ whenever $t/4 \ge \frac{1}{2}n$.

We saw in (1) that density of $\frac{1}{2}$ is attainable, though it seems plausible that the order 8 example given there is the largest possible. Of course, the partial latin squares

$$\begin{pmatrix} 1 & 3 & - \\ - & 2 & 1 \\ 2 & - & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & - & - & 2 \\ - & 1 & - & 3 \\ - & - & 1 & 4 \\ 3 & 4 & 2 & - \end{pmatrix} \qquad (4)$$

show that density exceeding $\frac{1}{2}$ is achievable for smaller n. By ad hoc use of the techniques of Theorem 1 and the stronger form of Lemma 1 it is not difficult to show that the examples in (1) and (4) have the highest possible density for their respective orders.

We next develop a constructive lower bound for $\delta(n)$.

Let n be an odd integer. We say that a subset S of $\mathbb{Z}_n$ satisfies the *law of the excluded middle* (LEM) if whenever x and y are distinct elements of S, the element $\frac{1}{2}(x+y)$ of $\mathbb{Z}_n$ is not an element of S. Let $\varepsilon(n)$ denote the cardinality of the largest subset of $\mathbb{Z}_n$ which satisfies LEM. We have:

Theorem 2 *$\delta(n) \ge \varepsilon(n)/n$ for every odd integer n.*

Proof: Suppose that S is a subset of $\mathbb{Z}_n$ satisfying LEM. We form a partial latin square P as follows, where all calculations are modulo n. If i and j are such that $j - i \in S$ then we put $P_{ij} = \frac{1}{2}(i+j)$, otherwise we leave P_{ij} blank. It is clear that the P so formed is a partial latin square. There is no duplication within a row because $\frac{1}{2}(i+j_1) \equiv \frac{1}{2}(i+j_2) \bmod n$ if and only if $j_1 \equiv j_2 \bmod n$ since n is odd. For a similar reason there is no duplication within columns.

The structure of P is best understood by considering its diagonals. We define a diagonal to be the set of entries of P occupying cells (i,j) where

$j - i \equiv d \bmod n$ for some fixed d. The diagonals of P are either entirely blank or have every entry filled. On the filled diagonals the entries of $\mathbb{Z}_n$ occur in cyclic order since $P_{i+1,j+1} \equiv P_{i,j} + 1 \bmod n$. We say that P is *diagonally cyclic*.

We now argue that P has the Blackburn property. Let i_1, j_1, i_2, j_2 be such that cells (i_1, j_1) and (i_2, j_2) are occupied by the same symbol in P but $i_1 \neq i_2$ and $j_1 \neq j_2$. By definition of P this means that $\frac{1}{2}(i_1 + j_1) \equiv \frac{1}{2}(i_2 + j_2) \bmod n$ and hence $j_1 - i_2 \equiv j_2 - i_1 \equiv \frac{1}{2}(j_1 - i_1 + j_2 - i_2) \bmod n$. Since our two cells are filled we know that $j_1 - i_1 \in S$ and $j_2 - i_2 \in S$. Crucially, P is diagonally cyclic so no symbol occurs twice on the same diagonal, which means that these are distinct elements of S. The definition of S then says that $\frac{1}{2}(j_1 - i_1 + j_2 - i_2) \notin S$ so that cells (i_1, j_2) and (i_2, j_1) must be blank. Thus the Blackburn property is achieved.

If we choose S to be as large as possible then there are $\varepsilon(n)$ filled cells in each row of P, so the density is $\varepsilon(n)/n$. ⊙

A recent survey of applications for diagonally cyclic latin squares may be found in [7]. The problem of which diagonally cyclic partial latin squares can be completed to diagonally cyclic latin squares has been studied by Grüttmüller [5, 6]. In our case, of course, P can trivially be completed to the diagonally cyclic square defined by $L_{ij} \equiv \frac{1}{2}(i + j)$ for all $i, j \in \mathbb{Z}_n$.

Although it seems a very natural question, the author is unaware of any work towards finding large subsets of $\mathbb{Z}_n$ which obey LEM. However, a related problem in $\mathbb{Z}$ has been well studied. We say that a set S of non-negative integers has the three term arithmetic progression (3-TAP) property if it contains no three terms which are in arithmetic progression. For a given positive integer n let $m(n)$ denote the cardinality of the largest subset of $\{1, 2, \ldots, n\}$ with the 3-TAP property. Behrend [1] showed that $m(n) > n \exp\left(-c(\log n)^{1/2}\right)$ for some constant $c > 0$.

Observe that $m(n) \geq \varepsilon(n)$ since any set S satisfying LEM automatically has the 3-TAP property. To see this note that if S contained a 3 term arithmetic progression a, $a+d$ and $a+2d$ then putting $x = a$ and $y = a+2d$ violates LEM since $\frac{1}{2}(x + y) = a + d$.

Also $m(n) \leq \varepsilon(2n + 1)$ as we now argue. Suppose S is a subset of $\{1, 2, \ldots, n\}$ with the 3-TAP property. We embed S in $\mathbb{Z}_{2n+1}$ and look at pairs $x, y \in S$ where $x < y$. If $y - x$ is even, say $y - x = 2k$ then $\frac{1}{2}(x + y) = x + k \notin S$ since otherwise the triple $(x, x + k, x + 2k = y)$ would violate the 3-TAP condition. So suppose that $y - x$ is odd, say $y - x = 2k + 1$. Then in $\mathbb{Z}_{2n+1}$ we have $\frac{1}{2}(x + y) \equiv x + k + n + 1$. But $1 \leq x < x + k + 1 \leq y \leq n$ so that $n + 1 \leq x + k + n + 1 \leq 2n$. This means that $\frac{1}{2}(x + y) \notin S$, so S satisfies LEM in $\mathbb{Z}_{2n+1}$.

Putting together the last two results we see that $\varepsilon(n)$ and $m(n)$ agree to within a constant factor. Hence we can couple Behrend's result with Theorem 2 to furnish the following lower bound.

Theorem 3 *There is a constant $c > 0$ such that $\delta(n) > \exp\left(-c(\log n)^{1/2}\right)$ for all n.*

Note that our construction as described above only worked for odd n. However, for even n we can take the construction for $n - 1$ and extend it with an empty row and column. This only changes the density by a factor of $(1 - 1/n)^2 = 1 + o(1)$.

In closing, we remark that Bourgain [4] proved that

$$m(n) = O\left(n(\log\log n/\log n)^{1/2}\right)$$

so that the highest density achievable by our method is $O\left((\log\log n/\log n)^{1/2}\right)$. Thus the question of whether a constant density is achievable remains open.

References

[1] F. A. Behrend, On sets of integers which contain no three terms in arithmetical progression, *Proc. Nat. Acad. Sci. U.S.A.* **32**, (1946), 331–332.

[2] S. R. Blackburn, Perfect hash families: probabilistic methods and explicit constructions, *J. Comb. Theory Series A*, **92** (2000), 54–60.

[3] S. R. Blackburn and P. R. Wild, Optimal linear perfect hash families, *J. Comb. Theory Series A*, **83** (1998), 233–250.

[4] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.* **9** (1999), 968–984.

[5] M. Grüttmüller, Completing partial latin squares with two prescribed diagonals, Preprints aus dem Fachbereich Mathematik, Universität Rostock.

[6] M. Grüttmüller, Completing partial latin squares with prescribed diagonals, Preprints aus dem Fachbereich Mathematik, Universität Rostock.

[7] I. M. Wanless, Diagonally cyclic latin squares, submitted for publication.