

# On the size of the minimum critical set of a Latin square

M. Ghandehari<sup>a</sup>, H. Hatami<sup>b</sup> and E. S. Mahmoodian<sup>a</sup>

<sup>a</sup>*Department of Mathematical Sciences*

<sup>b</sup>*Department of Computer Engineering*

Sharif University of Technology

P.O. Box 11365–9415, Tehran, I.R. Iran

## Abstract

A critical set in an  $n \times n$  array is a set  $C$  of given entries, such that there exists a unique extension of  $C$  to an  $n \times n$  Latin square and no proper subset of  $C$  has this property. For a Latin square  $L$ ,  $\text{scs}(L)$  denotes the size of the smallest critical set of  $L$ , and  $\text{scs}(n)$  is the minimum of  $\text{scs}(L)$  over all Latin squares  $L$  of order  $n$ . We find an upper bound for the number of partial Latin squares of size  $k$  and prove that

$$n^2 - (e + o(1))n^{10/6} \leq \max \text{scs}(L) \leq n^2 - \frac{\sqrt{\pi}}{2}n^{9/6}.$$

This improves a result of N. Cavenagh (Ph.D. thesis, The University of Queensland, 2003) and disproves one of his conjectures. Also it improves the previously known lower bound for the size of the largest critical set of any Latin square of order  $n$ .

KEYWORDS: Critical sets; Latin squares; Partial Latin Squares.

## 1 Introduction

A Latin square of order  $n$  is an  $n \times n$  array of integers, chosen from the set  $X = \{1, 2, \dots, n\}$  such that each element of  $X$  occurs exactly once in each row and exactly once in each column. A Latin square can also be written as a set of ordered triples  $\{(i, j; k) \mid \text{symbol } k \text{ occurs in cell } (i, j) \text{ of the array}\}$ . A partial Latin square  $P$  of order  $n$  is an  $n \times n$  array with entries chosen from the set  $X = \{1, 2, \dots, n\}$ , such that each element of  $X$  occurs at most once in each row and at most once in each column.

Hence there are cells in the array that may be empty, but the cells that are filled have been filled so as to conform with the Latin property of the array. Note that a partial Latin square of order  $n$  is not necessarily completable to a Latin square of order  $n$ . Let  $P$  be a partial Latin square of order  $n$ , then  $|P|$  is said to be the **size** of the partial Latin square and the set of positions  $\mathcal{S}_P = \{(i, j) \mid (i, j; k) \in P\}$  is said to determine the **shape** of  $P$ .

A partial Latin square  $C$  contained in a Latin square  $L$  is said to be **uniquely completable** if  $L$  is the only Latin square of order  $n$  with  $k$  in the cell  $(i, j)$  for every  $(i, j; k) \in C$ . A **critical set**  $C$  contained in a Latin square  $L$  is a partial Latin square that is uniquely completable, with no proper subset of  $C$  satisfying this requirement. We say a partial Latin square  $P$  **forces** an entry  $e = (i, j; k)$  into  $P$ , if  $P \cup \{(i, j; k')\}$  is not a partial Latin square, for every  $k' \neq k$ . The name “critical set” and the concept were invented by statistician John Nelder, about 1977, and his ideas were first published in a note [4]. For a Latin square  $L$ ,  $\text{lcs}(L)$  and  $\text{scs}(L)$  respectively, denote the size of the **largest critical sets** and **smallest critical sets** of  $L$ . Let  $\text{lcs}(n)$  be the maximum of  $\text{lcs}(L)$  over all Latin squares  $L$  of size  $n$ , and  $\text{scs}(n)$  be the minimum of  $\text{scs}(L)$  over all Latin squares  $L$  of size  $n$ . Determining  $\text{lcs}(n)$  and  $\text{scs}(n)$  are open questions, see for example [3]. We introduce some new bounds for  $\text{lcs}(n)$ , and for  $\max \text{scs}(L)$ .

In Section 2 we show that every Latin square has a critical set of size at most  $n^2 - \frac{\sqrt{\pi}}{2}n^{3/2}$ , and in Section 3 we give an upper bound for the number of partial Latin squares of order  $n$  and size  $k$ . By using this upper bound, we prove in Section 4 that there exist Latin squares which do not contain any critical set of size less than  $n^2 - (e + o(1))n^{5/3}$ . This result improves the previously known lower bound given in [2]:

$$\text{lcs}(n) \geq n^2 \left(1 - \frac{2 + \ln 2}{\ln n}\right) + n \left(1 + \frac{2 \ln 2 + \ln(2\pi)}{\ln n}\right) - \frac{\ln 2}{\ln n}.$$

Note that the two bounds given in Sections 2 and 4 show that:

$$n^2 - (e + o(1))n^{10/6} \leq \max \text{scs}(L) \leq n^2 - \frac{\sqrt{\pi}}{2}n^{9/6}.$$

Most of our proofs are involved with calculations, and the following well-known inequalities will be used frequently.

$$\binom{a}{b} = \binom{a}{a-b} \leq \left(\frac{ea}{b}\right)^b, \tag{1}$$

where  $a$  and  $b$  are natural numbers.

$$\sqrt{2\pi n} \left(\frac{n}{e}\right)^n \leq n! \leq \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}. \tag{2}$$

## 2 The upper bound

In this section we use the probabilistic method to obtain an upper bound for the size of the smallest critical set of any arbitrary Latin square of order  $n$ .

**Theorem 1** *Every Latin square  $L$  of order  $n$  contains a critical set of size less than  $n^2 - n\frac{\sqrt{n\pi}}{2}$ .*

**Proof.** Assign to each entry  $e = (i, j; k)$  of  $L$  a “birth time”  $x_e$ . These  $x_e$  are independent real variables, each with a uniform distribution in  $[0, 1]$ . Next, order the entries according to increasing birth time, giving the ordering  $e_1, e_2, \dots, e_{n^2}$ . So we have  $x_{e_i} < x_{e_{i+1}}$ . Now begin from the empty set  $C$ , and for every  $1 \leq i \leq n^2$ , if the partial Latin square  $\{e_1, e_2, \dots, e_{i-1}\}$  does not force  $e_i$ , add  $e_i$  to  $C$ . It is trivial that the constructed set  $C$  is a uniquely completable set. We want to calculate the expected size of  $C$ . Consider an entry  $e = (i, j; k)$  with the birth time  $x_e$ . For every element  $k'$  in  $\{1, 2, \dots, n\} \setminus \{k\}$ , there exists a cell in the  $i$ -th row of  $L$  and a cell in the  $j$ -th column with value  $k'$ . Since birth times have uniform distribution in  $[0, 1]$ , the probability that at least one of these two entries has birth time less than  $x_e$  is  $1 - (1 - x_e)^2$ . Thus the entry  $e$  is forced by the previous entries with probability  $(1 - (1 - x_e)^2)^{n-1}$ . So we have

$$E(|C|) = n^2 \left(1 - \int_0^1 (1 - (1 - x)^2)^{n-1} dx\right),$$

and if  $\sin(\alpha) = 1 - x$ , then

$$E(|C|) = n^2 \left(1 - \int_0^{\frac{\pi}{2}} \cos^{2n-1} \alpha \, d\alpha\right) = n^2 \left(1 - \frac{(2n-2)(2n-4)\dots 2}{(2n-1)(2n-3)\dots 1}\right).$$

We know that  $\frac{\pi}{2} \leq \frac{2}{1} \frac{2}{3} \frac{4}{3} \frac{4}{5} \dots \frac{2n-2}{2n-1} \frac{2n}{2n-1}$  (see for example [5], page 188). So  $E(|C|) \leq n^2 - n\frac{\sqrt{n\pi}}{2}$ . This implies that there exists a uniquely completable set of size less than  $n^2 - n\frac{\sqrt{n\pi}}{2}$ . ■

## 3 Number of partial Latin squares

In this section we give an upper bound for the number of partial Latin squares of order  $n$  and of size  $k$ . This result will be used in Section 4. The following lemma is a corollary of Brégman’s well-known inequality (see for example [6], page 83). Note that for an  $m \times n$  matrix,  $A = [a_{ij}]$ ,  $m \leq n$ , the permanent is defined as

$$\text{per}(A) = \sum_{\sigma} \prod_{i=1}^m a_{i\sigma(i)},$$

where  $\sigma$  is a one to one function from  $\{1, \dots, m\}$  to  $\{1, \dots, n\}$ .

**Lemma 1** Let  $A$  be a  $(0, 1)$ -matrix of order  $m \times n$ , which has  $r_i$  ones in row  $i$ ,  $1 \leq i \leq m$ . Then

$$\text{per}(A) \leq \frac{n!^{\frac{n-m}{n}}}{(n-m)!} \prod_{i=1}^m (r_i!)^{1/r_i}. \quad \blacksquare$$

**Theorem 2** Let  $S$  be a set of the cells in an  $n \times n$  array that has exactly  $r_i$  cells in the  $i$ -th row and  $c_j$  cells in the  $j$ -th column. The number of partial Latin squares of shape  $S$  is less than or equal to

$$\left( \prod_{i=1}^n n!^{\frac{n-r_i}{n}} \frac{1}{(n-r_i)!} \right) \left( \prod_{i=1}^n \prod_{j=0}^{c_i-1} (n-j)!^{\frac{1}{n-j}} \right).$$

**Proof.** Suppose that all cells in the first  $t-1$  rows of  $S$  are filled, and denote the constructed partial Latin square by  $P_{t-1}$ . Then for the  $t$ -th row construct an  $r_t$  by  $n$   $(0, 1)$ -matrix  $A_t$  as in the following. For every cell  $(t, y_i)$  in  $S$  ( $1 \leq i \leq r_t$ ), let the cell  $(i, k)$  of  $A_t$  be 1 if  $k$  does not occur in the column  $y_i$  of  $P_{t-1}$  and 0 otherwise. Note that the  $t$ -th row of  $S$  can be filled in exactly  $\text{per}(A_t)$  ways. Now if  $l_{y_i}$  is the number of cells in the first  $t-1$  rows of the  $y_i$ -th column of  $S$ , then by Lemma 1

$$\text{per}(A_t) \leq \frac{n!^{\frac{n-r_t}{n}}}{(n-r_t)!} \prod_{i=1}^{r_t} (n-l_{y_i})!^{\frac{1}{n-l_{y_i}}}.$$

So by multiplying right sides together for  $t = 1, \dots, n$ , we achieve an upper bound for the number of ways that  $S$  can be filled. It is easy to see that the product is

$$\left( \prod_{i=1}^n n!^{\frac{n-r_i}{n}} \frac{1}{(n-r_i)!} \right) \left( \prod_{i=1}^n \prod_{j=0}^{c_i-1} (n-j)!^{\frac{1}{n-j}} \right). \quad \blacksquare$$

**Theorem 3** The number of partial Latin squares of order  $n$  and of size  $k$  is bounded above, by:

$$\binom{n^2}{k} \frac{n!^{2n-\frac{k}{n}} e^{n(3+\frac{\ln(2\pi n)^2}{4})}}{(n-\frac{k}{n})!^{2n} e^k}.$$

**Proof.** Let  $S_P$  be a shape which has  $r_i$  cells in the  $i$ -th row and  $c_j$  cells in the  $j$ -th column. Then  $r_1 + r_2 + \dots + r_n = c_1 + c_2 + \dots + c_n = k$ . First we show that  $\prod_{i=1}^n n!^{\frac{n-r_i}{n}} \frac{1}{(n-r_i)!}$  achieves its maximum value when  $r_i = k/n$ , for all  $1 \leq i \leq n$ . Recall that when  $x$  is any real number,  $x!$  is defined as  $x! = \Gamma(x+1)$ . We have

$$\prod_{i=1}^n n!^{\frac{n-r_i}{n}} \frac{1}{(n-r_i)!} = n!^{(n-\frac{k}{n})} \prod_{i=1}^n \frac{1}{(n-r_i)!}.$$

Since  $\ln \Gamma(n)$  is convex, the expression is maximized when  $r_i$  are all equal. Hence

$$\prod_{i=1}^n n!^{\frac{n-r_i}{n}} \frac{1}{(n-r_i)!} \leq \frac{n!(n-\frac{k}{n})}{(n-\frac{k}{n})!^n} \quad (3)$$

Next consider  $\prod_{i=1}^n \prod_{j=0}^{c_i-1} (n-j)!^{\frac{1}{n-j}}$ . By Inequality (2) we have  $(n-j)! \leq (\frac{n-j}{e})^{n-j} \sqrt{2\pi(n-j)} e^{\frac{1}{12(n-j)}}$ , so that

$$\begin{aligned} \prod_{i=1}^n \prod_{j=0}^{c_i-1} (n-j)!^{\frac{1}{n-j}} &\leq \prod_{i=1}^n \prod_{j=0}^{c_i-1} \frac{n-j}{e} (2\pi(n-j))^{\frac{1}{2(n-j)}} e^{\frac{1}{12(n-j)^2}} \\ &\leq \left( \prod_{i=1}^n \prod_{j=0}^{c_i-1} \frac{n-j}{e} \right) \left( \prod_{i=1}^n (2\pi i)^{\frac{n}{2i}} e^{\frac{n}{12i^2}} \right). \end{aligned}$$

Now knowing that  $\frac{\pi^2}{6} \geq \sum_{i=1}^{\infty} \frac{1}{i^2}$ , we have

$$\ln \left( \prod_{i=1}^n (2\pi i)^{\frac{n}{2i}} e^{\frac{n}{12i^2}} \right) = n \sum_{i=1}^n \left( \frac{\ln(2\pi i)}{2i} + \frac{1}{12i^2} \right) \leq n \left( \frac{\pi^2}{72} + \frac{\ln(2\pi)}{2} + \int_1^n \left( \frac{\ln(2\pi x)}{2x} \right) dx \right).$$

The integral is equal to  $\frac{\ln(2\pi n)^2}{4}$ . So we have

$$\ln \left( \prod_{i=1}^n (2\pi i)^{\frac{n}{2i}} e^{\frac{n}{12i^2}} \right) \leq n \left( \frac{\pi^2}{72} + \frac{\ln(2\pi)}{2} + \frac{\ln(2\pi n)^2}{4} \right) \leq n \left( 3 + \frac{\ln(2\pi n)^2}{4} \right).$$

For  $\prod_{i=1}^n \prod_{j=0}^{c_i-1} \frac{n-j}{e}$ , we have

$$\prod_{i=1}^n \prod_{j=0}^{c_i-1} \frac{n-j}{e} = \prod_{i=1}^n \frac{n!}{(n-c_i)! e^{c_i}} = \frac{n!^n}{e^k} \prod_{i=1}^n \frac{1}{(n-c_i)!}.$$

And again since  $\ln \Gamma(n)$  is convex, this expression achieves its maximum value, when  $c_i$  are all equal, i.e.

$$\left( \prod_{i=1}^n \prod_{j=0}^{c_i-1} (n-j)!^{\frac{1}{n-j}} \right) \leq \frac{n!^n e^{n(3 + \frac{\ln(2\pi n)^2}{4})}}{(n-\frac{k}{n})!^n e^k}. \quad (4)$$

Note that we can choose the shape of these partial Latin squares in  $\binom{n^2}{k}$  ways. This fact and Inequalities (3) and (4) lead to the result of the theorem.  $\blacksquare$

## 4 The lower bound

**Theorem 4** *There exists a Latin square  $L$  such that  $\text{scs}(L) \geq n^2 - (e + o(1))n^{5/3}$ .*

**Proof.** As a result of van der Waerden conjecture, we have

$$L(n) \geq \frac{(n!)^{2n}}{n^{n^2}},$$

(see for example Theorem 17.2 in [6]) where  $L(n)$  is the number of Latin squares of order  $n$ .

If every Latin square has a critical set of size at most  $k$ , then obviously the number of critical sets of size at most  $k$  is greater than or equal to  $L(n)$ , and as a result, the number of uniquely completable partial Latin squares of size  $k$  is greater than or equal to  $L(n)$ . By Theorem 3, we know that the number of partial Latin squares of size  $k$  is at most

$$\binom{n^2}{k} \frac{n!^{2n - \frac{k}{n}} e^{n(3 + \frac{\ln(2\pi n)^2}{4})}}{(n - \frac{k}{n})! 2^n e^k}.$$

So

$$\frac{(n!)^{2n}}{n^{n^2}} \leq \binom{n^2}{k} \frac{n!^{2n - \frac{k}{n}} e^{n(3 + \frac{\ln(2\pi n)^2}{4})}}{(n - \frac{k}{n})! 2^n e^k}$$

Let  $c = 1 - \frac{k}{n^2}$ . Then by Inequality (1),  $\binom{n^2}{k} = \binom{n^2}{n^2 - k} \leq (\frac{e}{c})^{cn^2}$ . Hence

$$\frac{n!^{n - cn}}{n^{n^2}} \leq \frac{e^{cn^2} e^{n \ln(2\pi n)^2}}{c^{cn^2} (cn)! 2^n e^{n^2 - cn^2}},$$

or

$$\frac{n^{n^2 - cn^2}}{e^{n^2 - cn^2} n^{n^2}} \leq \frac{e^{3cn^2} e^{n \ln(2\pi n)^2}}{c^{cn^2} (cn)^{2cn^2} e^{n^2 - cn^2}},$$

or

$$c^{3c} n^c \leq e^{3c} e^{\frac{\ln(2\pi n)^2}{n}}. \quad (5)$$

Let  $n$  be a fixed and sufficiently large number. If  $c \geq \frac{1}{n^{1/3}}$ , then  $c^{3c} n^c e^{-3c}$  increases as  $c$  increases, and if  $c = \frac{e^{1 + \frac{1}{\sqrt{n}}}}{n^{1/3}}$ , then  $c^{3c} n^c > e^{3c} e^{\frac{\ln(2\pi n)^2}{n}}$ . So Inequality (5) implies that  $c \leq \frac{e + o(1)}{n^{1/3}}$ , or  $k \geq n^2 - (e + o(1))n^{5/3}$ . ■

Cavenagh ([1], Corollary 10.8, page 147) proved that  $\max \text{scs}(L) < n^2 - O(n^{4/3})$ . Theorem 1 improves this result. Also Theorem 4 shows that a conjecture of Cavenagh ([1], Conjecture 10.9, page 147) which states that  $\max \text{scs}(L) \leq \frac{n^2}{2}$ , is not true for sufficiently large  $n$ .

## Acknowledgements

We thank R. Tusserkani for his valuable discussions leading towards the result of this paper.

## References

- [1] N. J. Cavenagh. *Latin trades and critical sets in Latin squares*. PhD thesis, The University of Queensland, January 2003.
- [2] H. Hatami and E. S. Mahmoodian. A lower bound for the size of the largest critical sets in latin squares. *Bull. Inst. Combin. Appl.*, 38:19–22, 2003.
- [3] A. D. Keedwell. Critical sets in latin squares and related matters: an update. *Utilitas Math*, to appear.
- [4] J. Nelder. Critical sets in latin squares. *CSIRO (Commonwealth Scientific and Industrial Research Organisation), Division of Math. and Stats, Newsletter*, 38, 1977.
- [5] M. R. Spiegel. *Mathematical Handbook of Formulas and Tables*. McGraw Hill, 1968.
- [6] J. H. van Lint and R. M. Wilson. *A Course in Combinatorics*. Cambridge University Press, 1992.