

Zero-sum squares in bounded discrepancy $\{-1, 1\}$ -matrices

Alma R. Arévalo

Instituto de Matemáticas
Universidad Nacional Autónoma de México
Mexico City, Mexico

arevalo@ciencias.unam.mx

Amanda Montejano*

Facultad de Ciencias
Universidad Nacional Autónoma de México
Juriquilla, Mexico

amandamontejano@ciencias.unam.mx

Edgardo Roldán-Pensado[†]

Centro de Ciencias Matemáticas
Universidad Nacional Autónoma de México
Morelia, Mexico

e.roldan@im.unam.mx

Submitted: Jun 2, 2020; Accepted: Oct 1, 2021; Published: Oct 22, 2021

© The authors. Released under the CC BY-ND license (International 4.0).

Abstract

For $n \geq 5$, we prove that every $n \times n$ matrix $\mathcal{M} = (a_{i,j})$ with entries in $\{-1, 1\}$ and absolute discrepancy $|\text{disc}(\mathcal{M})| = |\sum a_{i,j}| \leq n$ contains a zero-sum square except for the split matrices (up to symmetries). Here, a square is a 2×2 submatrix of \mathcal{M} with entries $a_{i,j}, a_{i+s,s}, a_{i,j+s}, a_{i+s,j+s}$ for some $s \geq 1$, and a split matrix is a matrix with all entries above the diagonal equal to -1 and all remaining entries equal to 1 . In particular, we show that for $n \geq 5$ every zero-sum $n \times n$ matrix with entries in $\{-1, 1\}$ contains a zero-sum square.

Mathematics Subject Classifications: 05D10, 05B20

1 Introduction

An *Erickson matrix* is a square binary matrix that contains no *squares* (defined below) with constant entries. In [7], Erickson asked for the maximum value of n for which there exists an $n \times n$ Erickson matrix. In [1] Axenovich and Manske gave an upper bound of

*Supported by CONACyT project 282280 and PAPIIT project IN116519.

[†]Supported by CONACyT project 282280.

around $2^{2^{40}}$. This gargantuan bound was later improved by Bacher and Eliahou in [2] using computational means to the optimal value of 15.

This paper is devoted to studying a zero-sum analogue of Erickson matrices considering matrices with entries in $\{-1, 1\}$. For this purpose, of course, we need to take into account the discrepancy or deviation of the matrix.

Discrepancy theory is an important branch in combinatorics with deep connections to many other areas in mathematics (see [5] for a good general reference on this topic). In particular, one important result is Tao's recent proof of the Erdős discrepancy conjecture, [10], which states that any sequence of the form $f : \mathbb{N} \rightarrow \{-1, 1\}$ satisfies $\sup_{n,d} \left| \sum_{j=1}^n f(jd) \right| = \infty$.

In recent years, we have witnessed the study of zero-sum structures becoming increasingly popular. Some examples related to our work are the following. Caro et al. proved in [6] that for any finite sequence $f : [1, n] \rightarrow \{-1, 1\}$ satisfying that $|\sum_{i=1}^n f(i)|$ is small, there is a set of consecutive numbers $B \subset [1, n]$ for which $|\sum_{i \in B} f(i)|$ is also small (in particular, small can mean zero-sum). Another interesting work is [4], where Buttkewitz and Elsholtz proved the existence of zero-sum arithmetic progressions with four terms in certain sequences $f : \mathbb{N} \rightarrow \{-1, 1\}$. Balister et al. studied matrices where, for some fixed integer p , the sum of each row and each column is a multiple of p [3]; they showed that these matrices appear in any large enough integer square matrix.

Throughout this paper a matrix with entries in $\{-1, 1\}$ will be called a *binary matrix*. Given an $n \times m$ binary matrix $\mathcal{M} = (a_{i,j})$, the *discrepancy* of \mathcal{M} is the sum of all its entries, that is

$$\text{disc}(\mathcal{M}) = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} a_{i,j}. \quad (1)$$

Note that if a^+ is the number of entries in \mathcal{M} equal to 1 and a^- is the number of entries in \mathcal{M} equal to -1 then

$$\text{disc}(\mathcal{M}) = a^+ - a^- = 2a^+ - nm = nm - 2a^-. \quad (2)$$

We define a *zero-sum matrix* \mathcal{M} as a binary matrix with $\text{disc}(\mathcal{M}) = 0$.

A *square* S in $\mathcal{M} = (a_{i,j})$ is a 2×2 sub-matrix of \mathcal{M} of the form

$$S = \begin{pmatrix} a_{i,j} & a_{i,j+s} \\ a_{i+s,j} & a_{i+s,j+s} \end{pmatrix}$$

for some positive integer s . A *zero-sum square* is a square S with $\text{disc}(S) = 0$. Note that a square in \mathcal{M} is not zero-sum if and only if it has at least 3 equal entries.

We are interested in studying matrices \mathcal{M} which do not contain zero-sum squares, we call these matrices *zero-sum-square-free*.

Note that this may also be seen as a 2-coloring of an $n \times m$ rectangular grid. In this case, zero-sum is the same as balanced.

An $n \times m$ binary matrix $\mathcal{M} = (a_{i,j})$ is called *t-split* if for some $0 \leq t < n + m$,

$$a_{i,j} = \begin{cases} -1 & \text{if } i + j \leq t + 1, \\ 1 & \text{otherwise.} \end{cases}$$

If either \mathcal{M} , its negative, or its horizontal or vertical reflections are t -split for some t , we say that \mathcal{M} is *split*. This is relevant since split matrices are always zero-sum-square-free. We are also interested in the possible discrepancies they can have.

Observation 1. For an $n \times m$ t -split matrix \mathcal{M} with $n \leq m$,

$$\text{disc}(\mathcal{M}) = \begin{cases} nm - t(t+1) & \text{if } t \leq n, \\ nm + n(n-1) - 2nt & \text{if } n < t \leq m, \\ (n+m-t-1)(n+m-t) - nm & \text{if } m < t. \end{cases}$$

From this we may conclude the following.

Corollary 2. Let \mathcal{M} be a t -split binary matrix such that $|\text{disc}(\mathcal{M})| \leq n$. If \mathcal{M} is of size $n \times n$, then $t \in \{n-1, n\}$ and $|\text{disc}(\mathcal{M})| = n$. If \mathcal{M} is of size $n \times (n+1)$, then $t = n$ and $\text{disc}(\mathcal{M}) = 0$.

In particular, the discrepancy of a square split matrix never vanishes.

Now we are ready to state our main theorem.

Theorem 3. Let $n \geq 5$. Every $n \times n$ non-split binary matrix \mathcal{M} with $|\text{disc}(\mathcal{M})| \leq n$ contains a zero-sum square. In particular, every $n \times n$ zero-sum matrix \mathcal{M} contains a zero-sum square.

Theorem 3 and Corollary 2 immediately yield the following.

Corollary 4. Let \mathcal{M} be an $n \times n$ binary matrix. If $n \geq 5$ and $|\text{disc}(\mathcal{M})| \leq n-1$, then \mathcal{M} contains a zero-sum square.

Our proof method suggests that a stronger result may hold.

Conjecture 5. For every $C > 0$, there is an integer N with the following property: For all $n \geq N$, every $n \times n$ non-split binary matrix \mathcal{M} with $|\text{disc}(\mathcal{M})| \leq Cn$ contains a zero-sum square.

There is a more general question. Let $f : \mathbb{N} \rightarrow \mathbb{N}$ be the function associating to each $n \in \mathbb{N}$ the largest possible integer $f(n)$ such that every $n \times n$ non-split binary matrix \mathcal{M} satisfying $|\text{disc}(\mathcal{M})| \leq f(n)$ contains a zero-sum square. Obviously $f(n) < n^2$. In fact, we have that $f(n) \leq \frac{n^2}{2} + o(n^2)$, as shown by the $n \times n$ matrix $\mathcal{M} = (a_{i,j})$ defined by

$$a_{i,j} = \begin{cases} -1 & \text{if } i, j \text{ are both even,} \\ 1 & \text{otherwise.} \end{cases}$$

This is a zero-sum-square-free matrix and its discrepancy is about $\frac{n^2}{2}$. Theorem 3 implies $f(n) \geq n$ if $n \geq 5$. It would be very interesting to determine whether $f(n)$ is linear or quadratic in n .

This paper is organized as follows. Section 2 is devoted to particular cases, which were analyzed by computer. In Section 3 we give a stronger version of Theorem 3 and its proof. Finally, Section 4 contains our conclusions and some open questions.

2 Small cases

Since the proof of Theorem 3 uses induction, we must analyze some of the smaller cases to obtain our induction basis. It is possible to do this by hand but the amount of work is quite large, so we aid ourselves with a computer program.

Our program takes three positive parameters as input: n , m and d , which should satisfy $n \leq m$ and $d \leq nm$. The output is a list of all $n \times m$ binary matrices which are zero-sum-square-free and satisfy $\text{disc}(\mathcal{M}) = d$. To do this we use a standard backtracking algorithm that explores all binary matrices with the desired properties. The code is written in C++ and is available at

<https://github.com/edyrol/ZeroSumSquares>.

We are mainly interested in two types of zero-sum-square-free matrices: square matrices (with $m = n$) and almost-square matrices (with $m = n + 1$). These are the sizes of matrices we need to understand in order to prove Theorem 3.

Recall that split matrices are zero-sum-square-free, so we always find these examples.

Lemma 6. *Let \mathcal{M} be a zero-sum-square-free binary matrix with $|\text{disc}(\mathcal{M})| \leq 2n$. If \mathcal{M} is of size $n \times (n + 1)$ and $4 \leq n \leq 11$, then \mathcal{M} is either a split matrix or it is one of 28 exceptional 4×5 matrices. If \mathcal{M} is of size $n \times n$ and $5 \leq n \leq 11$, then \mathcal{M} is either a split matrix or it is one of 32 exceptional 5×5 matrices.*

Although Lemma 6 mentions 60 exceptional matrices, there are essentially only 11. The rest can be obtained by taking the symmetries of these 11 (generated by reflections and rotations, and their negatives). These 11 matrices are shown in Figure 1.

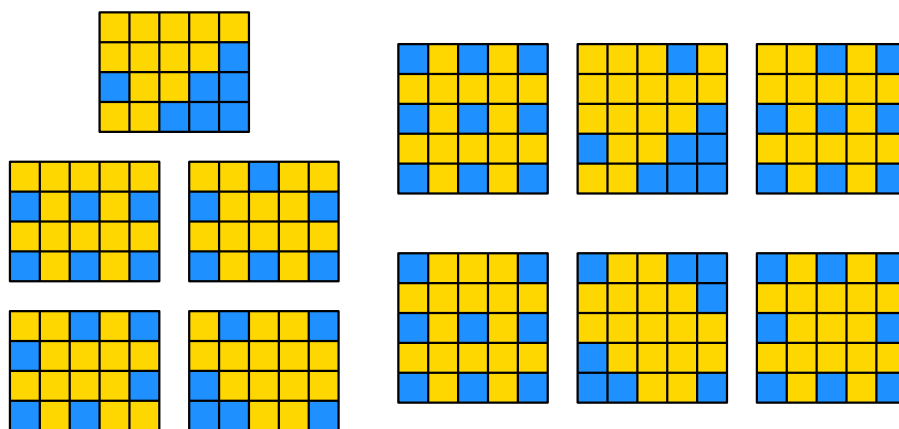


Figure 1: Non-split zero-sum-square-free binary matrices.

The computer program does not take too long to run. Using a home computer with an i7-3770 3.40GHz processor and compiling the program with GCC 8.1.0, it takes less than a second to analyze a 9×9 matrix with fixed discrepancy. For the larger matrices, it can take a couple of minutes. For example, depending on the discrepancy, it takes between 30 and 50 seconds to analyze an 11×11 binary matrix and between 1.5 and 3 minutes for an 11×12 binary matrix.

3 Proof

Our proof of Theorem 3 uses an induction argument. The main idea in the induction step is to split a large zero-sum-square-free matrix \mathcal{M} into four square (with equal side-lengths) or almost-square (with side-lengths differing by 1) sub-matrices. Since it is not always possible to only use squares, we are forced to understand the behavior of both square and almost-square zero-sum-square-free matrices. For the induction to work, we prove the following stronger statement.

Theorem 7. *Let $n \geq 5$ and $m \in \{n, n + 1\}$. Every $n \times m$ non-split binary matrix \mathcal{M} with $|\text{disc}(\mathcal{M})| \leq n$ contains a zero-sum square.*

The basis of the induction is given by the computer analysis described in Section 2. It is not indispensable to use a computer to prove Lemma 6, although doing it by hand would require either substantial case analysis or a clever argument that has eluded us.

For the rest of the proof we proceed as follows: assuming that the discrepancy of \mathcal{M} is not too large, we find a relatively large sub-matrix \mathcal{N} of \mathcal{M} with small discrepancy. By the induction hypothesis, if we assume that \mathcal{M} is a zero-sum-square-free matrix, we conclude that \mathcal{N} must be split. It turns out that having a relatively large split sub-matrix \mathcal{N} determines the value of many other entries of \mathcal{M} . From those values we find that, either \mathcal{M} is itself split as desired, or we can estimate $\text{disc}(\mathcal{M})$ and find that it is larger than n which contradicts the hypothesis of Theorem 7.

For integers h, j, k, l satisfying $1 \leq h < k \leq n$ and $1 \leq j < l \leq m$, we define a *block of \mathcal{M}* as the $(k - h + 1) \times (l - j + 1)$ sub-matrix

$$\mathcal{M}[h, k; j, l] = \begin{pmatrix} a_{h,j} & a_{h,j+1} & \cdots & a_{h,l-1} & a_{h,l} \\ a_{h+1,j} & a_{h+1,j+1} & \cdots & a_{h+1,l-1} & a_{h+1,l} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k-1,j} & a_{k-1,j+1} & \cdots & a_{k-1,l-1} & a_{k-1,l} \\ a_{k,j} & a_{k,j+1} & \cdots & a_{k,l-1} & a_{k,l} \end{pmatrix}.$$

The next lemma shows that a block \mathcal{M}' in a zero-sum-square-free matrix \mathcal{M} is split if a certain sub-block of \mathcal{M}' is also split. It is divided into four instances. Parts (a) and (b) refer to blocks obtained by removing the first row and the first column of \mathcal{M}' , respectively. Parts (c) and (d) refer to blocks obtained by removing the last row and the last column of \mathcal{M}' , respectively.

Lemma 8. *Let \mathcal{M} be a zero-sum-square-free $n \times m$ matrix with $n \geq 5$. Let $1 \leq h < k \leq n$ and $1 \leq j < l \leq m$ be integers such that $k - h = l - j = b \geq 2$. Consider the block $\mathcal{M}' = \mathcal{M}[h, k; j, l]$ of size $(b + 1) \times (b + 1)$.*

- (a) *If $\mathcal{M}[h + 1, k; j, l]$ is b -split then \mathcal{M}' is $(b + 1)$ -split.*
- (b) *If $\mathcal{M}[h, k; j + 1, l]$ is b -split then \mathcal{M}' is $(b + 1)$ -split.*
- (c) *If $\mathcal{M}[h, k - 1; j, l]$ is b -split then \mathcal{M}' is b -split.*
- (d) *If $\mathcal{M}[h, k; j, l - 1]$ is b -split then \mathcal{M}' is b -split.*

Proof. If $\mathcal{M}[h + 1, k; j, l]$ (respectively $\mathcal{M}[h, k; j + 1, l]$) is b -split, we need to prove that all entries in the top row (respectively in the leftmost column) of \mathcal{M}' are equal to -1 . If $\mathcal{M}[h, k; j, l - 1]$ (respectively $\mathcal{M}[h, k - 1; j, l]$) is b -split, we need to prove that all entries in the rightmost column (respectively bottom row) of \mathcal{M}' are equal to 1 . Since the arguments are analogous for each case, we only show the first one. Assume that $\mathcal{M}[h + 1, k; j, l]$ is b -split then, for every $1 \leq i \leq b$,

$$a_{h+i, l-i} = -1 \text{ and } a_{h+i, l} = 1. \tag{3}$$

Consider now the square

$$S = \begin{pmatrix} a_{h, l-i} & a_{h, l} \\ a_{h+i, l-i} & a_{h+i, l} \end{pmatrix}$$

and recall that, since \mathcal{M}' is a zero-sum-square-free matrix, any square S in \mathcal{M}' has at least 3 equal entries. Thus, (3) implies that $a_{h, l-i} = a_{h, l}$ for every $1 \leq i \leq b$. Therefore, the elements in the first row of \mathcal{M}' , $a_{h, j}, \dots, a_{h, l}$, are all equal. Finally, since $a_{h, j} = a_{h, j+1}$ and $a_{h+1, j} = a_{h+1, j+1} = -1$, the same argument for the square

$$S = \begin{pmatrix} a_{h, j} & a_{h, j+1} \\ a_{h+1, j} & a_{h+1, j+1} \end{pmatrix}$$

implies that $a_{h, j} = a_{h, j+1} = -1$, so all entries in the top row of \mathcal{M}' are equal to -1 . This shows that \mathcal{M}' is indeed $(b + 1)$ -split. \square

Once we have a t -split block \mathcal{M}' , we can also deduce the values of other entries which are not necessarily adjacent to \mathcal{M}' .

Lemma 9. *Let \mathcal{M} be a zero-sum-square-free $n \times m$ matrix, where $\mathcal{M}' = \mathcal{M}[1, k; 1, l]$ is t -split with $t < k < n$ and $t < l < m$.*

If $l < r \leq \min(t + l - 1, n)$, then the entries $a_{r, i}$ have the same value for

$$i \in [1, \lfloor (t + l - r + 1)/2 \rfloor] \cup [r - t + 1, l].$$

Analogously, if $k < c \leq \min(t + k - 1, m)$, then the entries $a_{i, c}$ have the same value for

$$i \in [1, \lfloor (t + k - c + 1)/2 \rfloor] \cup [c - t + 1, k].$$

Proof. Assume $1 \leq i \leq (t + l - r + 1)/2$ and consider the square

$$S = \begin{pmatrix} a_{r-l+i, i} & a_{r-l+i, l} \\ a_{r, i} & a_{r, l} \end{pmatrix}.$$

Note that, since \mathcal{M}' is t -split and $r - l + 2i \leq t + 1$, it follows that $a_{r-l+i, i} = -1$ and $a_{r-l+i, l} = 1$. So two entries of S have opposite values and therefore $a_{r, i} = a_{r, l}$.

If $r - t + 1 \leq i \leq l$, consider the square

$$S = \begin{pmatrix} a_{r+1-i, 1} & a_{r+1-i, i} \\ a_{r, 1} & a_{r, i} \end{pmatrix}.$$

Since \mathcal{M}' is t -split and $r + 1 - i \leq t$, it follows that $a_{r+1-i,1} = -1$. Furthermore, since $i \leq l$, we have that $a_{r+1-i,i} = 1$. So two entries of S have opposite values and therefore $a_{r,1} = a_{r,i}$.

In conclusion, since $a_{r,i} = a_{r,l}$ for any $i \in [1, (t + l - r + 1)/2]$, we have that $a_{r,1} = a_{r,l}$. If $i \in [r - t + 1, l]$ then $a_{r,i} = a_{r,1}$. Therefore, all of these values are equal. The proof for columns is analogous. \square

Proof of Theorem 7. By Lemma 6 we know the theorem holds for any $n \leq 11$ and $m \in \{n, n + 1\}$. Let $\mathcal{M} = (a_{i,j})$ be a $n \times m$ binary matrix with $n \geq 12$, $m \in \{n, n + 1\}$ and $|\text{disc}(\mathcal{M})| \leq n$. We need to prove that either \mathcal{M} is split or it contains a zero-sum square, so we assume henceforth that \mathcal{M} is zero-sum-square-free.

As stated before, we use induction on n , so we may assume that the theorem holds true for all square and almost-square binary matrices with smaller dimensions than those of \mathcal{M} .

We consider the four blocks of \mathcal{M} formed by splitting \mathcal{M} vertically and horizontally as evenly as possible. To be precise, let

$$\begin{aligned} \mathcal{M}_1 &= \mathcal{M} \left[1, \left\lfloor \frac{n}{2} \right\rfloor; \quad 1, \left\lfloor \frac{m}{2} \right\rfloor \right], \\ \mathcal{M}_2 &= \mathcal{M} \left[\left\lfloor \frac{n}{2} \right\rfloor + 1, n; \quad 1, \left\lfloor \frac{m}{2} \right\rfloor \right], \\ \mathcal{M}_3 &= \mathcal{M} \left[1, \left\lfloor \frac{n}{2} \right\rfloor; \quad \left\lfloor \frac{m}{2} \right\rfloor + 1, m \right] \text{ and} \\ \mathcal{M}_4 &= \mathcal{M} \left[\left\lfloor \frac{n}{2} \right\rfloor + 1, n; \quad \left\lfloor \frac{m}{2} \right\rfloor + 1, m \right]. \end{aligned}$$

Note that, for $1 \leq i \leq 4$, each block \mathcal{M}_i is either a square or an almost-square matrix. Also, the smallest side of any \mathcal{M}_i is $\lfloor \frac{n}{2} \rfloor$ and the largest is $m - \lfloor \frac{m}{2} \rfloor = \lceil \frac{m}{2} \rceil \leq \lceil \frac{n+1}{2} \rceil \leq \lfloor \frac{n}{2} \rfloor + 1$. Therefore, the side-lengths of each \mathcal{M}_i are in the set $\{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor + 1\}$.

Claim 10. *Either one of the four matrices \mathcal{M}_i satisfies $|\text{disc}(\mathcal{M}_i)| < \lfloor \frac{n}{2} \rfloor$ or two of these four matrices have discrepancies with opposite signs.*

Proof. If this is not the case and that the four matrices satisfy $\text{disc}(\mathcal{M}_i) \geq \lfloor \frac{n}{2} \rfloor$, then $n \geq \text{disc}(\mathcal{M}) = \sum \text{disc}(\mathcal{M}_i) \geq 4 \lfloor \frac{n}{2} \rfloor$ which is a contradiction. If the four matrices satisfy $\text{disc}(\mathcal{M}_i) \leq -\lfloor \frac{n}{2} \rfloor$ we obtain a contradiction in the same way. Therefore two of the \mathcal{M}_i have discrepancies with opposite signs. \square

What we actually wish to find is a relatively large block of \mathcal{M} with small discrepancy. By an interpolation argument this is easily achievable.

Claim 11. *By exchanging 1 and -1 if necessary, we may assume that there is an almost-square $\lfloor \frac{n}{2} \rfloor$ -split block \mathcal{N} with side-lengths in the set $\{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor + 1\}$ such that $|\text{disc}(\mathcal{N})| < \lfloor \frac{n}{2} \rfloor$.*

Proof. Claim 10 either provides the block we want or it gives us two blocks \mathcal{N}_+ and \mathcal{N}_- from the set $\{\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \mathcal{M}_4\}$ with $\text{disc}(\mathcal{N}_+) > 0$ and $\text{disc}(\mathcal{N}_-) < 0$.

We can construct a sequence $\mathcal{N}_- = \mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_k = \mathcal{N}_+$ of blocks of \mathcal{M} with the following properties:

- The side-lengths of every \mathcal{N}_i are in $\{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor + 1\}$.
- For each $1 \leq i < k$, one of \mathcal{N}_i and \mathcal{N}_{i+1} can be obtained from the other by removing one row or one column.

In other words, we start with $\mathcal{N}_1 = \mathcal{N}_-$ and start moving towards \mathcal{N}_+ . In each step we add or remove a row or column to \mathcal{N}_i taking care to always leave \mathcal{N}_{i+1} with side-lengths in the set $\{\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor + 1\}$. Note that in each step we switch from square to almost-square and vice-versa.

At some point the discrepancy changes from negative to positive, so assume that $\text{disc}(\mathcal{N}_i) < 0$ and $\text{disc}(\mathcal{N}_{i+1}) > 0$ for some $1 \leq i < k$. Since, at each step the discrepancy changes by at most $\lfloor \frac{n}{2} \rfloor + 1$, we conclude that either \mathcal{N}_i or \mathcal{N}_{i+1} must have absolute discrepancy at most $(\lfloor \frac{n}{2} \rfloor + 1)/2 < \lfloor \frac{n}{2} \rfloor$. Let \mathcal{N} be this block.

Now we can use our induction hypothesis on \mathcal{N} . Since $|\text{disc}(\mathcal{N})| < \lfloor \frac{n}{2} \rfloor$, either \mathcal{N} contains a zero-sum square, or it must necessarily be split. Furthermore, by Corollary 2, if \mathcal{N} is zero-sum-square-free, then it must be an almost-square block and have discrepancy exactly 0. \square

In the following Claim we prove that several entries of \mathcal{M} are forced. Note that, if $\mathcal{N} = \mathcal{M}[p, r; q, s]$, the $\lfloor \frac{n}{2} \rfloor$ -th diagonal of \mathcal{N} is contained in the $(p + q + \lfloor \frac{n}{2} \rfloor - 2)$ -th diagonal of \mathcal{M} . So, to simplify things, we define

$$t = p + q + \lfloor \frac{n}{2} \rfloor - 2 \geq \lfloor \frac{n}{2} \rfloor. \quad (4)$$

Claim 12. *By relabeling the entries of \mathcal{M} and exchanging 1 and -1 if necessary, we may assume that the block $\mathcal{M}_0 = \mathcal{M}[1, t + 1; 1, t + 1]$ is t -split.*

Proof. We start with the block $\mathcal{N} = \mathcal{M}[p, r; q, s]$ described in Claim 11. We repeatedly apply Lemma 8 to obtain a sequence of split matrices $\mathcal{N} = \mathcal{N}_1, \dots, \mathcal{N}_k$ in the following way. Assume that $\mathcal{N}_i = \mathcal{M}[h, k; j, l]$ is a b -split block. The block \mathcal{N}_i is either square or almost-square and b differs from the side-lengths of \mathcal{N}_i by at most 1. There are four possibilities.

- If $b = k - h + 1 = j - l$, then it follows from parts (a) and (c) of Lemma 8 that $\mathcal{M}[h - 1, k; j, l]$ is $(b + 1)$ -split (if $h > 1$) and $\mathcal{M}[h, k + 1; j, l]$ is b -split (if $k < n$).
- If $b = k - h = j - l + 1$, then parts (b) and (d) of Lemma 8 imply that $\mathcal{M}[h, k; j - 1, l]$ is $(b + 1)$ -split (if $j > 1$) and $\mathcal{M}[h, k; j, l + 1]$ is b -split (if $l < m$).
- If $b = k - h = j - l$ then we can remove the last row or column from \mathcal{N} and apply parts (c) and (d) of Lemma 8 to show that $\mathcal{M}[h - 1, k; j, l]$ (if $h > 1$) and $\mathcal{M}[h, k; j - 1, l]$ (if $j > 1$) are b -split.
- If $b = k - h + 1 = j - l + 1$ then we can remove the first row or column from \mathcal{N} and apply parts (a) and (b) of Lemma 8 to show that $\mathcal{M}[h, k + 1; j, l]$ (if $k < n$) and $\mathcal{M}[h, k; j, l + 1]$ (if $l < m$) are b -split.

In any case, let \mathcal{N}_{i+1} be any of the larger split blocks described above, whenever possible.

The process can only stop at $\mathcal{N}_k = \mathcal{M}[h, k; j, l]$ if either $\mathcal{N}_k = \mathcal{M}$ or \mathcal{N}_k is square with $(h, j) = (1, 1)$ or $(k, l) = (n, m)$. If $(h, j) = (1, 1)$, we are done. If $(k, l) = (n, m)$ then we may relabel the entries of \mathcal{M} , exchanging $(1, 1)$ and (n, m) and exchange 1 and -1 to obtain the desired result. \square

Claim 13. *The block*

$$\mathcal{M}_1 = \mathcal{M} \left[1, \min \left(n, \left\lfloor \frac{3t}{2} \right\rfloor \right); 1, \min \left(m, \left\lfloor \frac{3t}{2} \right\rfloor \right) \right]$$

is t -split.

Proof. We start with the block $\mathcal{M}_0 = \mathcal{M}[1, t+1; 1, t+1]$ from Claim 12 and repeatedly apply Lemma 9 in the following way.

If $\mathcal{M}' = \mathcal{M}[1, k; 1, k]$ is t -split with $t+1 \leq k < m$, apply Lemma 9 for columns with $l = k$ and $c = k+1$. The values $a_{i,c}$ are all equal for $i \in [1, k]$ whenever

$$(c - t + 1) - 1 \leq (t + k - c + 1)/2,$$

which is equivalent to $k+1 = c \leq 3t/2$. If this is the case, consider the square

$$\begin{pmatrix} a_{k-1,k} & a_{k-1,k+1} \\ a_{k,k} & a_{k,k+1} \end{pmatrix}.$$

Since $a_{k-1,k} = a_{k,k} = 1$ and $a_{k-1,k+1} = a_{k,k+1}$, then $a_{k,k+1}$ and therefore every $a_{i,k+1}$ with $i \in [1, k]$ is 1. Thus, the block $\mathcal{M}[1, k; 1, k+1]$ is t -split as long as $k+1 \leq \lfloor 3t/2 \rfloor$.

Now, starting with $\mathcal{M}[1, k; 1, k+1]$, apply Lemma 9 for rows with $l = k+1$ and $r = k+1$. The values $a_{r,i}$ are all equal for $i \in [1, l]$ whenever

$$(r - t + 1) - 1 \leq (t + l - r + 1)/2,$$

which is equivalent to $k+1 = r \leq 3t/2 + 1/2$. In the same way as before, we may conclude that $a_{i,k+1} = 1$ if $i \in [1, k+1]$. Thus, the block $\mathcal{M}[1, k+1; 1, k+1]$ is t -split whenever $k+1 \leq \lfloor 3t/2 \rfloor$.

This process stops when either r or c exceeds $\lfloor 3t/2 \rfloor$ or the corresponding dimension of \mathcal{M} . \square

In view of the previous Claim we may assume that $m \geq \lfloor \frac{3t}{2} \rfloor + 1$, otherwise $\mathcal{M} = \mathcal{M}_1$ is a split matrix. Since $m \leq n+1$, this implies that

$$t \leq \frac{2n+1}{3}, \tag{5}$$

which will be relevant later. In the case in which \mathcal{M}_1 does not cover \mathcal{M} , we may infer the values of additional entries of \mathcal{M} . This is done in a similar way to Claim 13, although we are no longer able to obtain a t -split matrix. Instead, we obtain five regions outside of \mathcal{M}_1 for which $a_{i,j} = 1$. These are illustrated in Figure 2.

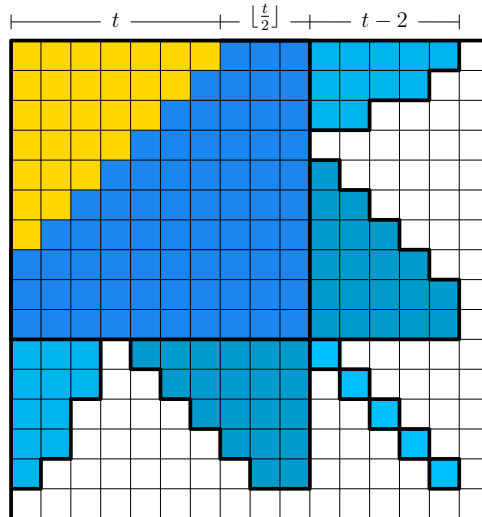


Figure 2: The regions described in Claims 12, 13 and 14. The colors yellow and blue represent values of -1 and 1 , respectively. The matrix \mathcal{M} is not actually large enough to contain all of the marked squares.

The first region has a triangular shape bounded by the first column outside of \mathcal{M}_1 , the first row of \mathcal{M} and a certain line of slope $\frac{1}{2}$. The second region is bounded by the first column outside of \mathcal{M}_1 , the last row of \mathcal{M}_1 and a line of slope -1 . Of course, there are corresponding regions to these below \mathcal{M}_1 . Lastly, the entries of the diagonal of \mathcal{M} which are outside of \mathcal{M}_1 must also have value 1 . This is formalized in the following claim.

Claim 14. Let $T = \lfloor \frac{3t}{2} \rfloor$, then $a_{i,j} = 1$ and $a_{j,i} = 1$ whenever $T < j$ and any of the following hold:

- (a) $i \leq \lfloor \frac{T+t+1-j}{2} \rfloor$,
- (b) $j - t < i \leq T$, or
- (c) $i = j$.

Proof. We start with the t -split block $\mathcal{M}_1 = \mathcal{M}[1, T; 1, T]$ from Claim 13.

We inductively deduce the values in column j starting with $j = T + 1$ and increasing j one by one. Take $k = l = T$ and $c = j$ in Lemma 9 for columns.

Note that (4) implies that $j \leq T + t - 2$, so two things happen; the values $a_{i,j}$ are all equal for

$$i \in [1, \lfloor (t + T - j + 1)/2 \rfloor] \cup [j - t + 1, T]$$

and the interval $[j - t + 1, T]$ contains at least two elements. By considering the square

$$\begin{pmatrix} a_{T-1,j-1} & a_{T-1,j} \\ a_{T,j-1} & a_{T,j} \end{pmatrix}$$

and using the fact that the elements $a_{T-1,j-1}$ and $a_{T,j-1}$ from the previous column have value 1, we conclude that all the $a_{i,j}$ described above are equal to 1.

Analogously, using Lemma 9 for rows, we can say the same for $a_{j,i}$.

Condition $i \in [1, \lfloor (t+T-j+1)/2 \rfloor]$ is equivalent to $i \leq \lfloor (T+t+1-j)/2 \rfloor$ which proves part (a) of the claim, while $i \in [j-t+1, T]$ is equivalent to $j-t+1 \leq i \leq T$ which proves part (b).

To prove part (c), for $T < i = j \leq T+t-2$, consider the square

$$\begin{pmatrix} a_{1,1} & a_{1,i} \\ a_{i,1} & a_{i,i} \end{pmatrix}.$$

Since $a_{1,1} = -1$ and $a_{i,1} = a_{1,i} = 1$, we must have that $a_{i,i} = 1$. □

Now we can bound the discrepancy of \mathcal{M} . Recall from (2) that it is enough to know the number of positive entries a^+ of \mathcal{M} in order to compute $\text{disc}(\mathcal{M})$. Since $2a^+ - nm = \text{disc}(\mathcal{M}) \leq n$, we have that

$$a^+ \leq \frac{n + nm}{2}. \tag{6}$$

If this inequality is violated, it means that \mathcal{M} is not larger than \mathcal{M}_1 . So, all that remains is to bound from below the number of positive entries a^+ of \mathcal{M} .

Let $R = n - T$ and define

$$\begin{aligned} a_0 &= \frac{t(t-1)}{2} + \left\lfloor \frac{t}{2} \right\rfloor^2 + 2t \left\lfloor \frac{t}{2} \right\rfloor, \\ a_1 &= 2 \sum_{j=T+1}^n \left\lfloor \frac{T+t+1-j}{2} \right\rfloor \\ &= 2 \sum_{k=1}^R \left\lfloor \frac{t+1-k}{2} \right\rfloor, \\ a_2 &= 2 \sum_{j=T+1}^n (T-j+t) \\ &= 2 \sum_{k=1}^R (t-k) \text{ and} \\ a_3 &= R. \end{aligned}$$

A simple calculation gives the following claim.

Claim 15.

$$a^+ \geq a_0 + a_1 + a_2 + a_3.$$

Proof. The number of positive entries in \mathcal{M}_1 , described in Claim 13, is a_0 .

If $m = n$, then a_1 , a_2 and a_3 equal the number of positive entries described in parts (a), (b) and (c) of Claim 14, respectively.

If $m = n + 1$, by ignoring the positive entries in the last column of \mathcal{M} , we obtain that a_1 , a_2 and a_3 are lower bounds for the number of positive entries described in parts (a), (b) and (c) of Claim 14, respectively. \square

Recall that we are currently dealing with $n \geq 12$ and, from (4) and (5),

$$\left\lfloor \frac{n}{2} \right\rfloor \leq t \leq \left\lfloor \frac{2n+1}{3} \right\rfloor. \quad (7)$$

Before simplifying this lower bound, we can check that (6) cannot be satisfied for small values of n . The following claim can be easily verified with aid from a computer.

Claim 16. For $12 \leq n \leq 15$ and $\left\lfloor \frac{n}{2} \right\rfloor \leq t \leq \left\lfloor \frac{2n+1}{3} \right\rfloor$, we have that

$$a_0 + a_1 + a_2 + a_3 > \frac{(n^2 + 2n)}{2} \geq \frac{(n + nm)}{2}.$$

Therefore, we may assume that $n \geq 16$. What follows is a series of algebraic manipulations to obtain a simpler lower bound for a^+ which can be analyzed analytically.

Claim 17. For $n \geq 16$ and $\left\lfloor \frac{n}{2} \right\rfloor \leq t \leq \left\lfloor \frac{2n+1}{3} \right\rfloor$, we have that

$$a_0 + a_1 + a_2 + a_3 \geq \frac{23n^2 - 70n - 77}{32}.$$

Proof. We can remove the integer parts in $a_0 + a_1 + a_2 + a_3$ by using that, for any integer x , the inequalities $\frac{x-1}{2} \leq \left\lfloor \frac{x}{2} \right\rfloor \leq \frac{x}{2}$ are satisfied. It is convenient to do this in two parts, first we apply these inequalities but leave the variable r as it is. This gives

$$\begin{aligned} a_0 + a_1 + a_2 + a_3 &= \frac{t(t-1)}{2} + \left\lfloor \frac{t}{2} \right\rfloor^2 + 2t \left\lfloor \frac{t}{2} \right\rfloor + 2 \sum_{k=1}^R \left(\left\lfloor \frac{t+1-k}{2} \right\rfloor \right) \\ &\quad + 2 \sum_{k=1}^R (t-k) + R \\ &\geq \frac{t(t-1)}{2} + \left(\frac{t-1}{2} \right)^2 + t(t-1) + 2 \sum_{k=1}^R \left(\frac{t-k}{2} \right) \\ &\quad + 2 \sum_{k=1}^R (t-k) + R \\ &= \frac{7t^2}{4} - 2t + \frac{1}{4} + 3Rt - \frac{3}{2}R^2 - \frac{1}{2}R. \end{aligned}$$

Since $R = n - \left\lfloor \frac{3t}{2} \right\rfloor$ we have that $n - \frac{3t}{2} \leq R \leq n - \frac{3t-1}{2}$, using this on the last expression

we obtain

$$\begin{aligned} a_0 + a_1 + a_2 + a_3 &\geq \frac{7t^2}{4} - 2t + \frac{1}{4} + 3 \left(n - \frac{3t}{2} \right) t \\ &\quad - \frac{3}{2} \left(n - \frac{3t-1}{2} \right)^2 - \frac{1}{2} \left(n - \frac{3t-1}{2} \right) \\ &= -\frac{49t^2}{8} + \frac{15nt}{2} + t - \frac{3n^2}{2} - 2n - \frac{3}{8}. \end{aligned}$$

To minimize this last expression think of n as fixed and consider it as a function of t . Then this is an upside-down parabola and, from (7), the relevant values for t are contained in the interval $\left[\frac{n-1}{2}, \frac{2n+1}{3}\right]$. Therefore the parabola is bounded from below by the minimum between the values at $t = \frac{n-1}{2}$ and $t = \frac{2n+1}{3}$. These are, respectively,

$$\frac{23n^2 - 70n - 77}{32} \quad \text{and} \quad \frac{14n^4 - 28n - 13}{18}.$$

The former gives the smallest value. □

To conclude the proof, notice that the parabolas $\frac{1}{32}(23n^2 - 70n - 77)$ and $\frac{1}{2}(n^2 + 2n)$ intersect twice, once in the interval $(-1, 0)$ and a second time in the interval $(15, 16)$. Since $n \geq 16$, we have

$$a^+ \geq a_0 + a_1 + a_2 + a_3 \geq \frac{23n^2 - 70n - 77}{32} > \frac{n^2 + 2n}{2}.$$

This contradicts (6), so $\mathcal{M} = \mathcal{M}_1$ and therefore \mathcal{M} is a split matrix. □

4 Conclusions and further work

We were able to give an elemental proof of Theorem 3, but we are sure that there is a deeper result in the direction of Conjecture 5. It is also likely that something can be said for non-square matrices. The fact that the final bound given for a^+ is significantly smaller than n^2 suggests that a much stronger theorem should hold. It is possible to strengthen our proof to obtain a stronger version of Theorem 7 with something like $|\text{disc}(\mathcal{M})| \leq 2n$ instead of $|\text{disc}(\mathcal{M})| \leq n$, however significantly more work is required to establish this and it is probably not worth the effort.

In [9] Erickson matrices were generalized to *3-squares*. A k -square in a matrix \mathcal{M} is a $k \times k$ sub-matrix of \mathcal{M} contained in k rows of \mathcal{M} of the form $i, i + s, \dots, i + (k - 1)s$ and k columns of \mathcal{M} of the form $j, j + s, \dots, j + (k - 1)s$. We could ask about zero-sum- k -square-free binary matrices but this does not make sense when k is odd. However, the case when k is even seems interesting. For odd k we can ask about binary matrices which do not have k -squares of sum ± 1 .

Lastly, we should point out that with the aid of Claims 12 and 14, or with stronger versions of this claim, zero-sum-square-free matrices of much larger sizes may be analyzed

by a computer. This might be useful for generalizing our results. However, a different type of computer search might likely be much more useful. SAT-solvers have been used for finding lower bounds in Ramsey-like problems (see e.g. [8]) but it is not obvious how to include the discrepancy condition here. Perhaps linear integer programming could work. Since we did not need to analyze anything larger than an 11×12 matrix, we did not work much on making our program efficient.

Acknowledgements

The authors would like to thank the anonymous referee for his comments which improved the paper greatly. They are also thankful for the facilities provided by the Banff International Research Station “Casa Matemática Oaxaca” during the “Zero-Sum Ramsey Theory: Graphs, Sequences and More” workshop (19w5132).

References

- [1] M. Axenovich and J. Manske, *On monochromatic subsets of a rectangular grid*, *Integers* **8** (2008), A21.
- [2] R. Bacher and S. Eliahou, *Extremal binary matrices without constant 2-squares*, *J. Comb.* **1** (2010), no. 1, 77–100.
- [3] P. Balister, Y. Caro, C. Rousseau and R. Yuster, *Zero-sum square matrices*, *Eur. J. Combin.*, **23** (2002), no. 5, 489–497.
- [4] Y. Buttkewitz and C. Elsholtz, *Patterns and complexity of multiplicative functions*, *J. London Math. Soc.* (2) **84** (2011), no. 3, 578–594.
- [5] B. Chazelle, *The discrepancy method: randomness and complexity*, Cambridge University Press, 2001.
- [6] Y. Caro, A. Hansberg, and A. Montejano, *Zero-sum subsequences in bounded-sum $\{-1, 1\}$ -sequences*, *J. Combin. Theory Ser. A* **161** (2019), 387–419.
- [7] M. J. Erickson, *Introduction to combinatorics*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., New York, 1996, A Wiley-Interscience Publication.
- [8] P. R. Herwig, M. J. H. Heule, P. M. van Lambalgen, and H. van Maaren, *A new method to construct lower bounds for van der Waerden numbers*, *Electron. J. Combin.* **14** (2007), no. 1, #R6.
- [9] D. Robilliard, A. Boumaza, and V. Marion-Poty, *Meta-heuristic search and square Erickson matrices*, *IEEE Congress on Evolutionary Computation, IEEE*, 2010, pp. 1–8.
- [10] T. Tao, *The Erdős discrepancy problem*, *Discrete Anal.* (2016), Paper No. 1, 29.