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

Tom Johnston

Mathematical Institute, University of Oxford, United Kingdom.
thomas.johnston@maths.ox.ac.uk

Abstract

Given a matrix $M = (a_{i,j})$ a *square* is a 2×2 submatrix with entries $a_{i,j}$, $a_{i,j+s}$, $a_{i+s,j}$, $a_{i+s,j+s}$ for some $s \geq 0$, and a *zero-sum square* is a square where the entries sum to 0. Recently, Arévalo, Montejano and Roldán-Pensado [1] proved that all large $n \times n$ $\{-1, 1\}$-matrices M with discrepancy $|\sum a_{i,j}| \leq n$ contain a zero-sum square unless they are diagonal. We improve this bound by showing that all large $n \times n$ $\{-1, 1\}$-matrices M with discrepancy at most $n^2/4$ are either diagonal or contain a zero-sum square.

1 Introduction

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

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

In 1996 Erickson [11] asked for the largest n such that there exists an $n \times n$ binary matrix M with no squares which have constant entries. An upper bound was first given by Axenovich and Manske [2], before the answer 14 was determined by Bacjer and Eliahou in [3].

Recently, Arévalo, Montejano and Roldán-Pensado [1] initiated the study of a zero-sum variant of Erickson's problem. Here we wish to avoid *zero-sum squares*, squares with entries that sum to 0.

Zero-sum problems have been well-studied since the classic Erdős-Ginsburg-Ziv Theorem in 1961 [10]. Much of the research has been on zero-sum problems in finite abelian groups (see the survey [12] for details), but problems have also been studied in other settings such as on graphs (see e.g. [5,6,7,9]). Of particular relevance is the result of Balister, Caro, Rousseau and Yuster in [4] on submatrices of integer valued matrices where the rows and columns sum to $0 \mod p$, and the result of Caro, Hansberg and Montejano on zero-sum subsequences in bounded sum $\{-1,1\}$-sequences [8].

Given an $n \times m$ matrix $M = (a_{i,j})$ define the *discrepancy* of M as the sum of the entries, that is

$$\operatorname{disc}(M) = \sum_{\substack{1 \le i \le n \\ 1 \le j \le m}} a_{i,j}.$$

We say a square S is a *zero-sum square* if $\operatorname{disc}(S) = 0$, or equivalently,

$$a_{i,j} + a_{i,j+s} + a_{i+s,j} + a_{i+s,j+s} = 0.$$

We will be interested in $\{-1,1\}$-matrices M which do not contain any zero-sum squares, and we shall call such matrices *zero-sum square free*. Clearly matrices with at most one -1 are zero-sum square free and, in general, there are many such matrices when the number of -1s is low. Instead, we will be interested in matrices which have a similar number of 1s and -1s or, equivalently, matrices with small discrepancy (in absolute value).

An $n \times m$ $\{-1,1\}$-matrix $M = (a_{i,j})$ is said to be *t-diagonal* for some $0 \le t \le n + m - 1$ if

$$a_{i,j} = \begin{cases} 1 & i + j \le t + 1, \\ -1 & i + j \ge t + 2. \end{cases}$$

We say a matrix M is *diagonal* if there is some t such that a t-diagonal matrix N can be obtained from M by applying vertical and horizontal reflections. Diagonal matrices are of particular interest since they can have low discrepancy, yet they never contain a zero-sum square.

Arévalo, Montejano and Roldán-Pensado [1] proved that, except when $n \le 4$, every $n \times n$ non-diagonal $\{-1,1\}$-matrix M with $|\operatorname{disc}(M)| \le n$ has a zero-sum square. They remark that it should be possible to extend their proof to give a bound of $2n$, and they conjecture that the bound Cn should hold for any $C > 0$ when n is large enough relative to C.

Conjecture 1 (Conjecture 3 in [1])**.** *For every $C > 0$ there is a integer N such that whenever $n \geq N$ the following holds: every $n \times n$ non-diagonal $\{-1, 1\}$-matrix M with $|\operatorname{disc}(M)| \leq Cn$ contains a zero-sum square.*

We prove this conjecture in a strong sense with the following theorem.

Theorem 2. *Let $n \geq 5$. Every $n \times n$ non-diagonal $\{-1, 1\}$-matrix M with $|\operatorname{disc}(M)| \leq n^2/4$ contains a zero-sum square.*

The best known construction for a non-diagonal zero-sum square free matrix has discrepancy close to $n^2/2$, and our computer experiments suggest that this construction is in fact optimal. Closing the gap between the upper and lower bounds remains a very interesting problem and we discuss it further in Section 3.

2 Proof

For $p \leq r$ and $q \leq s$ define the *consecutive submatrix* $M[p : r, q : s]$ by

$$M[p : r, q : s] = \begin{pmatrix} a_{p,q} & a_{p,q+1} & \cdots & a_{p,s} \\ a_{p+1,q} & a_{p+1,q+1} & \cdots & a_{p+1,s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r,q} & a_{r+1,q} & \cdots & a_{r,s} \end{pmatrix}.$$

Throughout the rest of this paper, we will assume that all submatrices except squares are consecutive submatrices.

We start by stating the following lemma from [1] which, starting from a small t'-diagonal submatrix M', determines many entries of the matrix M. An example application is shown in Figure 1.

Lemma 3 (Claim 3 in [1])**.** *Let M be an $n \times n$ $\{-1, 1\}$-matrix with no zero-sum squares, and suppose that there is a submatrix $M' = M[p : p+s, q : q+s]$ which is t'-diagonal for some $2 \leq t' \leq 2s - 3$. Let $t = t + p + q - 2$ and suppose $t \leq n$.*

1. *The submatrix*

$$N = M[1 : \min(t + \lfloor t/2 \rfloor, n), 1 : \min(t + \lfloor t/2 \rfloor, n)]$$

is t-diagonal.

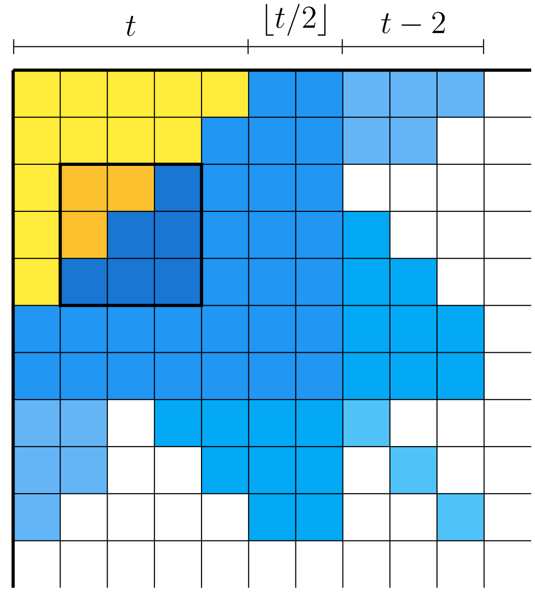


Figure 1: The entries known from applying Lemma 3. The yellow squares represent -1s and the blue squares represent 1s. The submatrix M' is show in a darker shade.

Furthermore, both $a_{i,j} = 1$ *and* $a_{j,i} = 1$ *whenever* $t + \lfloor t/2 \rfloor < j \leq t + \lfloor t/2 \rfloor + t - 2$ *and one of the following holds:*

2. $j - t \leq i \leq t + \left\lfloor \frac{t}{2} \right\rfloor$*;*

3. $i \leq \left\lfloor \frac{t}{2} \right\rfloor - \left\lfloor \frac{j-t-\lfloor t/2 \rfloor -1}{2} \right\rfloor$*;*

4. $i = j$*.*

Note that we can apply this lemma even when it is a reflection of M' which is t-diagonal; we just need to suitably reflect M and potentially multiply by -1, and then undo these operations at the end. The matrix N will always contains at least one of $a_{1,1}$, $a_{1,n}$, $a_{n,1}$ and $a_{n,n}$, and if N contains two, then M is diagonal.

We will also make use of the following observation. This will be used in conjunction with the above lemma to guarantee the existence of some additional 1s, which allows us to show a particular submatrix has positive discrepancy.

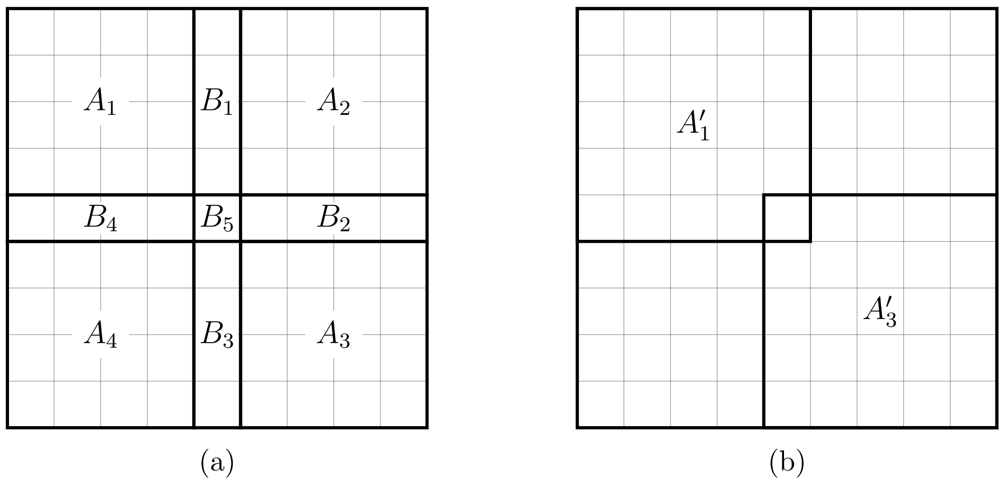


Figure 2: A subset of the regions used in the proof of Lemma 5.

Observation 4. *Let M be an $n \times n$ $\{-1, 1\}$-matrix with no zero-sum squares, and suppose that $a_{i,i} = 1$ for every $i \in [n]$. Then at least one of $a_{i,j}$ and $a_{j,i}$ is 1. In particular, $a_{i,j} + a_{j,i} \geq 0$ for all $1 \leq i, j \leq n$.*

The final lemma we will need to prove Theorem 2 is a variation on Claims 1 and 2 from [1]. The main difference between Lemma 5 and the result used by Arévalo, Montejano and Roldán-Pensado is that we will always find a square submatrix. This simplifies the proof of Theorem 2.

Lemma 5. *For $n \geq 8$, every $n \times n$ $\{-1, 1\}$-matrix M with $|\operatorname{disc}(M)| \leq n^2/4$ has an $n' \times n'$ submatrix M' with $|\operatorname{disc}(M')| \leq (n')^2/4$ for some $(n-1)/2 \leq n' \leq (n+1)/2$.*

Proof. We only prove this in the case n is odd as the case n is even is similar, although simpler. Partition the matrix M into 9 regions as follows. Let the four $(n-1)/2 \times (n-1)/2$ submatrices containing $a_{1,1}$, $a_{1,n}$, $a_{n,1}$ and $a_{n,n}$ be $A_1, \ldots, A_4$ respectively. Let the $(n-1)/2 \times 1$ submatrix between A_1 and A_2 be B_1 and define B_2, B_3 and B_4 similarly. Finally, let the central entry be B_5. The partition is shown in Figure 2a.

As these partition the matrix M, we have

$$\operatorname{disc}(M) = \operatorname{disc}(A_1) + \cdots + \operatorname{disc}(A_4) + \operatorname{disc}(B_1) + \cdots + \operatorname{disc}(B_5). \qquad (1)$$

Let the overlapping $(n+1)/2 \times (n+1)/2$ submatrices containing $a_{1,1}$, $a_{1,n}$, $a_{n,1}$ and $a_{n,n}$ be $A'_1, \ldots, A'_4$, as indicated in Figure 2b. The submatrices $B_1, \ldots, B_4$ each appear twice in the A'_i and B_5 appears four times and, by subtracting these overlapping regions, we obtain a second equation for $\operatorname{disc}(M)$:

$$\operatorname{disc}(M) = \operatorname{disc}(A'_1) + \cdots + \operatorname{disc}(A'_4) \\ - \operatorname{disc}(B_1) - \cdots - \operatorname{disc}(B_4) - 3\operatorname{disc}(B_5). \quad (2)$$

If any of the A_i or A'_i have $|\operatorname{disc}(A_i)| \leq (n-1)^2/16$ or $|\operatorname{disc}(A'_i)| \leq (n+1)^2/16$ respectively, we are done, so we may assume that this is not the case. First, suppose that $\operatorname{disc}(A_i) > (n-1)^2/16$ and $\operatorname{disc}(A'_i) > (n+1)^2/16$ for all $i = 1, 2, 3, 4$. Since $n-1$ is even and $\operatorname{disc}(A_i) \in \mathbb{Z}$, we must have $\operatorname{disc}(A_i) \geq (n-1)^2/16 + 1/4$, and similarly, $\operatorname{disc}(A'_i) \geq (n+1)^2/16 + 1/4$. Adding the equations (1) and (2) we get the bound

$$n^2/2 \geq 2\operatorname{disc}(M) \geq (n+1)^2/4 + (n-1)^2/4 + 2 - 2\operatorname{disc}(B_5),$$

which reduces to $\operatorname{disc}(B_5) \geq 5/4$. This gives a contradiction since B_5 is a single square. Similarly we get a contradiction if, for every i, both $\operatorname{disc}(A_i) < -(n-1)^2/16$ and $\operatorname{disc}(A'_i) < -(n+1)^2/16$.

This only leaves the case where two of the 8 submatrices have different signs. If $A'_i > (n+1)^2/16$, then, for $n \geq 8$,

$$A_i > (n+1)^2/16 - n > -(n-1)^2/16,$$

and either $|\operatorname{disc}(A_i)| \leq (n-1)^2/16$, a contradiction, or $\operatorname{disc}(A_i) > 0$. By repeating the argument when $\operatorname{disc}(A'_i)$ is negative, it follows that A_i and A'_i have the same sign for every i. In particular, two of the A_i must have different signs, and we can apply an interpolation argument as in [1].

Without loss of generality we can assume that $\operatorname{disc}(A_1) > (n-1)^2/16$ and $\operatorname{disc}(A_2) < -(n-1)^2/16$. Consider the sequence of matrices $N_0, \ldots, N_{(n+1)/2}$ where

$$N_i = M[1 : (n-1)/2, 1 + i : i + (n-1)/2].$$

We claim that there is a j such that $|\operatorname{disc}(N_j)| \leq (n-1)^2/16$, which would complete the proof of the lemma. By definition, $N_0 = A_1$ and $N_{(n+1)/2} = A_2$ so there must be some j such that $\operatorname{disc}(N_{j-1}) > 0$ and $\operatorname{disc}(N_j) \leq 0$. Since the submatrices N_{j-1} and N_j share most of their entries $|\operatorname{disc}(N_{j-1}) - \operatorname{disc}(N_j)| \leq (n-1)$, and as $(n-1)^2/8 > (n-1)$, it cannot be the case that $\operatorname{disc}(N_{j-1}) > (n-1)^2/16$ and $\operatorname{disc}(N_j) < -(n-1)^2/16$. This means there must be some j such that $|\operatorname{disc}(N_j)| \leq (n-1)^2/16$, as required. □

Armed with the above results, we are now ready to prove our main result, but let us first give a sketch of the proof which avoids the calculations in the main proof.

Sketch proof of Theorem 2. Assume we have an $n \times n$ $\{-1,1\}$-matrix M with $|\operatorname{disc}(M)| \leq n^2/4$ which is zero-sum square free. We will prove the result by induction, so we assume that the result is true for $5 \leq n' < n$.

Applying Lemma 5 gives a submatrix M' with low discrepancy. Since M' must also be zero-sum square free, we know that it is diagonal by the induction hypothesis. Applying Lemma 3 then gives us a lot of entries M and, in particular, a submatrix N with high discrepancy. Since we are assuming that M has low discrepancy, the remainder $M \setminus N$ of M not in N must either have low discrepancy or negative discrepancy. In both cases we will find B, a submatrix of M with low discrepancy. When the discrepancy of $M \setminus N$ is low, we use an argument similar to the proof of Lemma 5, and when the discrepancy of $M \setminus N$ is negative, we find a positive submatrix using Observation 4 and use an interpolation argument.

By the induction hypothesis, B must also be diagonal and we can apply Lemma 3 to find many entries of M. By looking at specific $a_{i,j}$, we will show that the two applications of Lemma 3 contradict each other. $\square$

We now give the full proof of Theorem 2, complete with all the calculations. To start the induction, we must check the cases $n < 30$ which is done using a computer. The problem is encoded as a SAT problem using PySAT [13] and checked for satisfiability with the CaDiCaL solver. The code to do this is attached to the arXiv submission.

Proof of Theorem 2. We will use induction on n. A computer search gives the result for all $n < 30$, so we can assume that $n \geq 30$ and that the result holds for all $5 \leq n' < n$.

Suppose, towards a contradiction, that M is an $n \times n$ matrix with no zero-sum squares and $|\operatorname{disc}(M)| \leq n^2/4$. By Lemma 5, we can find an $n' \times n'$ submatrix $M' = M[p : p+s, q : q+s]$ with $(n-1)/2 \leq n' \leq (n+1)/2$ and $|\operatorname{disc}(M')| \leq (n')^2/4$. By the induction hypothesis and our assumption that M doesn't contain a zero-sum square, the matrix M' must be diagonal. By reflecting M and switching -1 and 1 as necessary, we can assume that the submatrix M' is t'-diagonal for some t', and that $t := t' + p + q - 2 \leq n$.

We will want to apply Lemma 3, for which we need to check $2 \le t' \le 2s-3$. If $t' \le 1$ or $t' \ge 2s-2$, then the discrepancy of M' is

$$|\operatorname{disc}(M')| \ge (n')^2 - 1 > (n')^2/4,$$

which contradicts our choice of M'. In fact, since $\operatorname{disc}(M') \le (n')^2/4$ and $\operatorname{disc}(M') \le (n')^2 - t'(t'+1)$ we find

$$t \ge t' \ge \frac{1}{2}\left(\sqrt{3(n')^2+1} - 1\right) \approx 0.433n. \tag{3}$$

If $t + \lfloor t/2 \rfloor \ge n$, the matrix M is t-diagonal and we are done, so we can assume that this is not the case, and that $t \le 2n/3$. We will also need the following bound on $2t + \lfloor t/2 \rfloor - 2$, which follows almost immediately from (3).

Claim 1. *We have*

$$2t + \lfloor t/2 \rfloor - 2 \ge n - 1.$$

Proof. Substituting $n' \ge (n-1)/2$ into (3) gives the following bound on t.

$$t \ge \frac{1}{4}\left(\sqrt{3n^2 - 6n + 7} - 2\right)$$

We now lower bound $\lfloor t/2 \rfloor$ by $(t-1)/2$ to find

$$\begin{aligned} 2t + \lfloor t/2 \rfloor - 2 &\ge 2t + \frac{t-5}{2} \\ &\ge \frac{5}{8}\sqrt{3n^2 - 6n + 7} - \frac{15}{4} \end{aligned}$$

The right hand side grows like $\frac{\sqrt{75}}{8}n$ asymptotically, which is faster than n, so the claim is certainly true for large enough n. In fact, the equation $\frac{5}{8}\sqrt{3n^2 - 6n + 7} - \frac{15}{4} \ge n - 1$ can be solved explicitly to obtain the following the bound on n:

$$n \ge \frac{1}{11}\left(251 + 20\sqrt{166}\right) \approx 46.2.$$

This still leaves the values $30 \le n \le 46$ for which the bounds above are not sufficient. These cases can be checked using a computer. $\square$

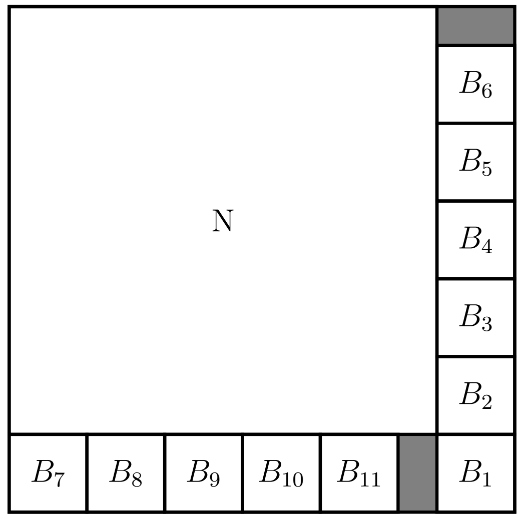


Figure 3: The matrix M with the submatrices N and $B_1, \ldots, B_{11}$. The entries of M which are not in any of the submatrices are shown in grey.

Let $k = \lceil 5n/6 \rceil$ and let N be the $k \times k$ sub-matrix in the top left corner which contains $a_{1,1}$ i.e. $N = M[1:k, 1:k]$. We will apply Lemma 3 and Observation 4 to guarantee lots of 1s in N, and therefore ensure N has large discrepancy. This will mean that the rest of M which is not in N must have low discrepancy, and we can find another diagonal submatrix, B.

Claim 2. *There is an $(n-k) \times (n-k)$ submatrix B which is disjoint from N and with $|\operatorname{disc}(B)| \leq (n-k)^2/4$.*

Proof. Consider the 11 $(n-k) \times (n-k)$ disjoint submatrices of M $B_1, \ldots, B_{11}$ given by

$$B_i = \begin{cases} M[k:n, n-ik:n-(i-1)k] & \text{if } i \leq 6 \\ M[(i-7)(n-k):(i-6)(n-k), k:n] & \text{if } i > 6, \end{cases}$$

and shown in Figure 3. The submatrix B_1 contains $a_{n,n}$ and sits in the bottom right of M, while the others lie along the bottom and right-hand edges of M.

If one of the B_i satisfies $|\operatorname{disc}(B_i)| \leq (n-k)^2/4$, we are done, so suppose this is not the case.

We start by using Observation 4 to show that $\operatorname{disc}(B_1) > 0$. Let the entries of B be $b_{i,j}$ where $1 \leq i, j \leq n-k$. By Claim 1, $2t + \lfloor t/2 \rfloor - 2 \geq n-1$

and, applying Lemma 1, $b_{i,i} = 1$ for all $i \leq n-k-1$. Further, by Observation 4, we have $b_{i,j} + b_{j,i} \geq 0$ for all $1 \leq i, j \leq n-k-1$. This means

$$\text{disc}(B_1) \geq (n-k-1) - (2(n-k)-1) = -(n-k)$$

For $(n-k) \geq 5$, $(n-k) < (n-k)^2/4$ so we must have $\text{disc}(B_1) > (n-k)^2/4$.

As $\text{disc}(B_1) > 0$, if $\text{disc}(B_i) < 0$ for any $i \neq 1$, we can use an interpolation argument as in Lemma 5 to find the claimed matrix. The argument only requires

$$2(n-k) < \frac{(n-k)^2}{2}$$

which is true for $(n-k) > 4$.

We must now be in the case where $\text{disc}(B_i) > (n-k)^2/4$ for every i. The bulk of the work in this case will be bounding the discrepancy of the matrix N, and then the discrepancy of M. There are $2nk - 12(n-k)^2 \leq 10(n-k)$ entries of M in the gaps between the B_i i.e. there are at most $10(n-k)$ entries $a_{i,j}$ which are not contained in either N or one of the B_i. In particular, we have

$$\begin{aligned} \text{disc}(M) &\geq \text{disc}(N) + \text{disc}(B_1) + \cdots + \text{disc}(B_{11}) - 10(n-k) \\ &> \text{disc}(N) + 11(n-k)^2/4 - 10(n-k) \end{aligned} \tag{4}$$

Let $s = \min\{k, t + \lfloor t/2 \rfloor\}$ so that $M[1:s, 1:s]$ is t diagonal, and let $r = k - s$ be the number of remaining rows. Let $a_1, \ldots, a_4$ be the number of 1s in N guaranteed by Lemma 3, and let a_5 be the number of additional 1s guaranteed by also applying Observation 4. This guarantees that at least one of $a_{i,j}$ and $a_{j,i}$ is 1 for all $(t+2)/2 \leq i, j \leq r$, and $a_5 \geq r(r-1)$.

We have the following bounds.

$$\begin{aligned} a_1 &= s^2 - \frac{t(t+1)}{2}, \\ a_2 &= 2\sum_{i=1}^{r}(t-i), \\ a_3 &= 2\sum_{i=1}^{r}\left(\left\lfloor \frac{t}{2} \right\rfloor - \left\lfloor \frac{i-1}{2} \right\rfloor\right), \\ a_4 &= r, \\ a_5 &\geq r(r-1). \end{aligned}$$

Let us first consider the case where $s = k$, so that N is t-diagonal. In this case $a_2 = \cdots = a_5 = 0$, and we can easily write down the discrepancy of N as $k^2 - t(t+1)$. Since $k \geq 5n/6$, we get the bound

$$\mathrm{disc}(N) \geq \frac{25n^2}{36} - t(t+1).$$

Substituting this into (4) and using the bounds $(n-5)/6 \leq n-k \leq n/6$ we get

$$\begin{aligned}\mathrm{disc}(N) &> \frac{25n^2}{36} - t(t+1) + \frac{11}{4}\left(\frac{n-5}{6}\right)^2 - \frac{10n}{6} \\ &= \frac{1}{144}\left(111n^2 - 350n - 144t^2 - 144t + 275\right).\end{aligned}$$

For $n \geq 4$, the righthand side is greater than $n^2/4$ whenever

$$t < \frac{1}{12}\left(\sqrt{75n^2 - 350n + 311} - 6\right) \approx 0.721n + o(n).$$

Since we have assumed $t \leq 2n/3$, we get a contradiction for all sufficiently large n. In fact, we get a contradiction for all $n \geq 39$. The remaining cases need to be checked using exact values for the floor and ceiling functions which we do with the help of a computer.

Now we consider the case where $s = t + \lfloor t/2 \rfloor$ which is very similar, although more complicated. To be in this case, we must have $t + \lfloor t/2 \rfloor \leq k$ which implies

$$t + \frac{t-1}{2} \leq \frac{5(n+1)}{6},$$

and $t \leq (5n+8)/9 \approx 0.556n$.

Start by using the bounds $(t-1)/2 \le \lfloor t/2 \rfloor$ and $\lfloor (i-1)/2 \rfloor \le (i-1)/2$ to get

$$\begin{aligned} a_1 + \cdots + a_5 &\ge \left(t + \frac{t-1}{2}\right)^2 - \frac{t(t+1)}{2} + r(2t-r-1) + r(t-1) \\ &\quad - \frac{r(r-1)}{2} + r + r(r-1) \\ &= \frac{7t^2}{4} - 2t - \frac{r^2}{2} + 3rt - \frac{5r}{2} + \frac{1}{4}. \end{aligned}$$

By definition, $r = k - t - \lfloor t/2 \rfloor$, so we get the bounds $5n/6 - t - t/2 \le r \le 5(n+1)/6 - t - (t-1)/2$, and substituing these in gives

$$\begin{aligned} a_1 + \cdots + a_5 &\ge \frac{7}{4}t^2 - 2t + \frac{1}{4} - \frac{1}{2}\left(\frac{5(n+1)}{6} - t - \frac{t-1}{2}\right)^2 + 3t\left(\frac{5n}{6} - t - \frac{t}{2}\right) \\ &\quad - \frac{5}{2}\left(\frac{5(n+1)}{6} - t - \frac{t-1}{2}\right) \\ &= \frac{1}{72}\left(-25n^2 + 270nt - 230n - 279t^2 + 270t - 286\right) \end{aligned}$$

Plugging this into (4) and using the bounds $5n/6 \le k \le 5(n+1)/6$ we get

$$\begin{aligned} \operatorname{disc}(M) &> 2(a_1 + \cdots a_5) - \left(\frac{5(n+1)}{6}\right)^2 + \frac{11}{4}\left(\frac{n-5}{6}\right)^2 - \frac{10n}{6} \\ &\ge \frac{1}{48}\left(-63n^2 + 360nt - 490n - 372t^2 + 360t - 323\right). \end{aligned}$$

When $n \ge 27$, this is greater than $n^2/4$ whenever

$$\begin{aligned} &\frac{1}{186}\left(90n + 90 - \sqrt{1125n^2 - 29370n - 21939}\right) < \\ &\qquad t < \frac{1}{186}\left(90n + 90 + \sqrt{1125n^2 - 29370n - 21939}\right), \end{aligned}$$

or approximately,

$$0.304n < t < 0.664n.$$

We have the bounds

$$\frac{1}{4}\left(\sqrt{3n^2 - 6n + 7} - 2\right) \le t \le \frac{5n+8}{9},$$

and so, for $n \geq 36$, $\mathrm{disc}(M) > n^2/4$.

This again leaves a few cases which we check with the help of a computer (although they could feasibly be checked by hand). $\square$

Given a submatrix B as in the above claim we apply the induction hypothesis, noting that $(n-k) \geq 5$ since $n \geq 30$, to find that B is diagonal. Let C be the diagonal submatrix obtained from applying Lemma 4 to B, and let C be ℓ-diagonal up to rotation. Note that $\ell \geq 3$ as $(n-k) \geq 5$, and we can assume $\ell \leq 2n/3$ as M is not diagonal.

Hence, C contains exactly one of $a_{1,1}$, $a_{1,n}$, $a_{n,1}$ and $a_{n,n}$, and we will split into cases based on which one C contains. We will also sometimes need to consider cases for whether the entry is 1 or -1, but in all cases we will find a contradiction.

From Lemma 3 applied to M' and Claim 1, we already know some of the entries and we highlight some important entries in the following claim.

Claim 3. *We have*

1. $a_{j,1} = a_{1,j} = \begin{cases} 1 & t+1 \leq j \leq n-1, \\ -1 & 1 \leq j \leq t, \end{cases}$

2. $a_{2,t} = a_{t,2} = 1,$

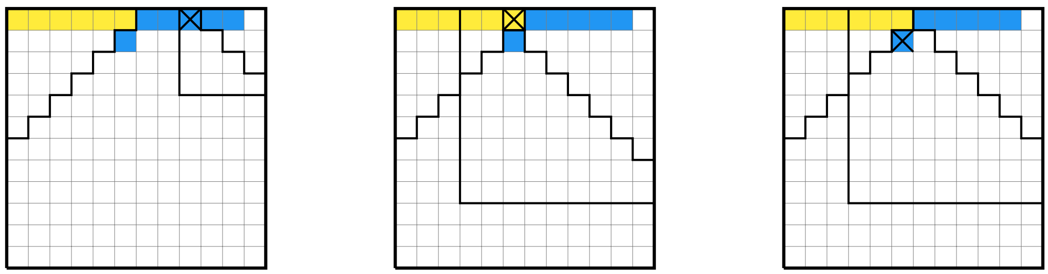
3. $a_{i,i} = 1$ *for all* $(t+2)/2 \leq i \leq n-1.$

Suppose the submatrix C contains $a_{1,1}$ so sits in the top-left corner. Since $M[1 : t + \lfloor t/2 \rfloor, 1 : t + \lfloor t/2 \rfloor]$ is t-diagonal, C must also be t-diagonal. As C was found by applying Lemma 3 to B, it must contain a -1 from B. Hence, $t \geq 5n/6$ which is a contradiction as we assumed that $t \leq 2n/3$.

Suppose instead that C contains $a_{1,n}$ so sits in the top-right corner. Since $\ell \geq 3$, if the corner entry is -1, so is the entry $a_{1,n-1}$, but this contradicts Claim 3. Suppose instead the corner entry is 1. Since C is ℓ-diagonal up to rotation we have, for all $1 \leq i, (n-j+1) \leq \ell + \lfloor \ell/2 \rfloor$,

$$a_{i,j} = \begin{cases} -1 & i + (n-j+1) \geq \ell + 2, \\ 1 & \text{otherwise}. \end{cases} \tag{5}$$

If $n - \ell > t$, then $a_{1,n-\ell} = -1$ by (5) and $a_{1,n-\ell} = 1$ by Claim 3. Suppose $n - \ell < t$. Then $a_{1,t} = 1$ by (5) and $a_{1,t} = -1$ as $M[1 : t, 1 : t]$ is t-diagonal.



(a) $n - \ell > t$ (b) $n - \ell < t$ (c) $n - \ell = t$

Figure 4: The three cases when C contains $a_{1,n}$ and $a_{1,n} = 1$. The yellow squares represent some of the $a_{i,j}$ which are known to be -1 from Claim 3 and the blue squares those which are 1. The square which gives the contradiction is marked with a cross.

Finally, when $n - \ell = t$, we have $a_{2,t} = -1$ by (5) and $a_{2,t} = 1$ from Claim 3. Some illustrative examples of these three cases are shown in Figure 4.

The case where C contains $a_{n,1}$ is done in the same way with the rows and columns swapped.

This leaves the case where C contains $a_{n,n}$. Since $\ell \geq 3$, if the entry $a_{n,n}$ equals -1, so does the entry $a_{n-1,n-1}$, and this contradicts Claim 3. If instead $a_{n,n} = 1$, we consider the entry $a_{i,i}$ where $i = n + 1 - \lceil (l+2)/2 \rceil$, which must be -1. However, since $\ell \leq 2n/3$,

$$n + 1 - \lceil (l+2)/2 \rceil \geq n + -\frac{n}{3} - \frac{1}{2} > \frac{n}{3} + 1 \geq \frac{t+2}{2},$$

and $a_{i,i} = 1$ by Claim 3. This final contradiction is shown in Figure 5. □

We remark that it should be possible to improve the bound $n^2/4$ using a similar proof provided one can check a large enough base case. Indeed, we believe that all the steps in the above proof hold when the bound is increased to $n^2/3$, but only when n is large enough. For example, Claim 1 fails for $n = 127$ and our proof of Claim 2 fails for $n = 86$. Checking base cases this large is far beyond the reach of our computer check, and some new ideas would be needed here.

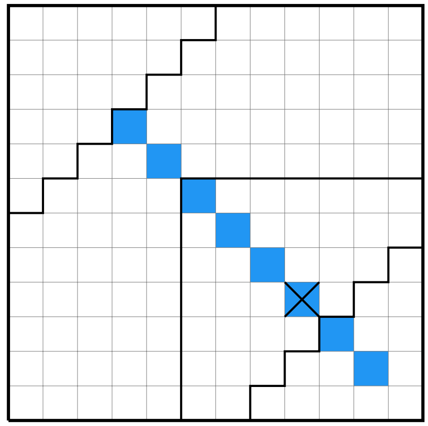


Figure 5: The case where C contains $a_{n,n}$ and $a_{n,n} = 1$. The square marked with a cross gives a contradiction.

3 Open problems

The main open problem is to determine the correct lower bound for the discrepancy of a non-diagonal $\{-1,1\}$-matrix with no zero-sum squares. We have improved the lower bound to $n^2/4$, but this does not appear to be optimal.

The best known construction is the following example by Arévalo, Montejano and Roldán-Pensado [1]. Let $M = (a_{i,j})$ be given by

$$a_{i,j} = \begin{cases} -1 & i \text{ and } j \text{ are odd,} \\ 1 & \text{otherwise.} \end{cases}$$

This has discrepancy $n^2/2$ when n is even and $(n-1)^2/2 - 1$ when n is odd. With the help of a computer we have verified that this construction is best possible when $9 \leq n \leq 32$, and we conjecture that this holds true for all $n \geq 9$. In fact, our computer search shows that the above example is the unique zero-sum square free non-diagonal matrix with minimum (in magnitude) discrepancy, up to reflections and multiplying by -1.

We note that the condition $n \geq 9$ is necessary, as shown by the 8×8 zero-sum square free $\{-1,1\}$-matrix with discrepancy 30 given in Figure 6.

Conjecture 6. *Let $n \geq 9$. Every $n \times n$ non-diagonal $\{-1,1\}$-matrix M with*

$$|\operatorname{disc}(M)| \leq \begin{cases} \frac{n^2}{2} - 1 & n \textit{ is even} \\ \frac{(n-1)^2}{2} - 2 & n \textit{ is odd} \end{cases}$$

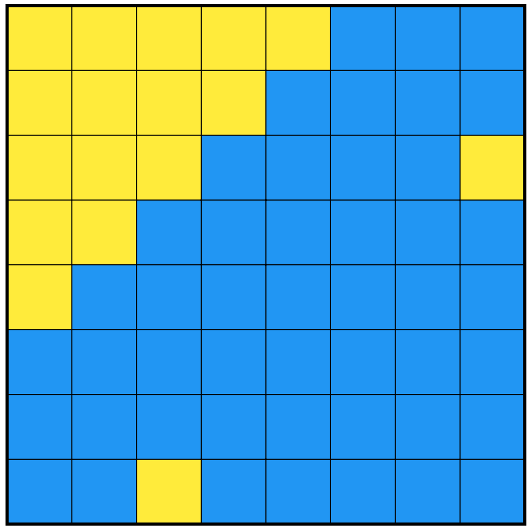


Figure 6: An 8×8 $\{-1, 1\}$-matrix with no zero-sum squares and discrepancy 30. The yellow squares represent a -1 and the blue squares represent a 1.

contains a zero-sum square.

Arévalo, Montejano and Roldán-Pensado prove their result for both $n \times n$ and $n \times (n+1)$ matrices, and computational experiments suggest that Theorem 2 holds for $n \times (n+1)$ matrices as well. More generally, what is the best lower bound for a general $n \times m$ matrix when n and m are large?

Problem 1. *Let $f(n,m)$ be the minimum $d \in \mathbb{N}$ such that there exists an $n \times m$ non-diagonal $\{-1,1\}$ matrix M with $|\operatorname{disc}(M)| \leq d$. What are the asymptotics of $f(n,m)$?*

References

[1] A. R. Arévalo, A. Montejano and E. Roldán-Pensado. Zero-sum squares in bounded discrepancy {-1, 1}-matrices. *arXiv preprint arXiv:2005.07813*, 2020.

[2] M. Axenovich and J. Manske. On monochromatic subsets of a rectangular grid. *Integers*, **8**(1):A21, 2008.

[3] R. Bacher and S. Eliahou. Extremal binary matrices without constant 2-squares. *Journal of Combinatorics*, **1**(1):77–100, 2010.

[4] P. Balister, Y. Caro, C. Rousseau and R. Yuster. Zero-sum square matrices. *European Journal of Combinatorics*, **23**(5):489–497, 2002.

[5] A. Bialostocki. Zero sum trees: a survey of results and open problems. In *Finite and infinite combinatorics in sets and logic*, pages 19–29. Springer, 1993.

[6] Y. Caro, A. Hansberg, J. Lauri and C. Zarb. On zero-sum spanning trees and zero-sum connectivity. *arXiv preprint arXiv:2007.08240*, 2020.

[7] Y. Caro, A. Hansberg and A. Montejano. Zero-sum K_m over $\mathbb{Z}$ and the story of K_4. *Graphs and Combinatorics*, **35**(4):855–865, 2019.

[8] Y. Caro, A. Hansberg and A. Montejano. Zero-sum subsequences in bounded-sum $\{-1, 1\}$-sequences. *Journal of Combinatorial Theory, Series A*, **161**:387–419, 2019.

[9] Y. Caro and R. Yuster. On zero-sum and almost zero-sum subgraphs over $\mathbb{Z}$. *Graphs and Combinatorics*, **32**(1):49–63, 2016.

[10] P. Erdos, A. Ginzburg and A. Ziv. Theorem in the additive number theory. *Bull. Res. Council Israel F*, **10**:41–43, 1961.

[11] M. Erickson. *Introduction to Combinatorics*. Wiley Series in Discrete Mathematics and Optimization. Wiley, 1996.

[12] W. Gao and A. Geroldinger. Zero-sum problems in finite abelian groups: a survey. *Expositiones Mathematicae*, **24**(4):337–369, 2006.

[13] A. Ignatiev, A. Morgado and J. Marques-Silva. PySAT: A Python toolkit for prototyping with SAT oracles. In *SAT*, pages 428–437, 2018.