INVERSE LITTLEWOOD-OFFORD PROBLEMS AND THE SINGULARITY OF RANDOM SYMMETRIC MATRICES

HOI H. NGUYEN

Abstract. Let M_n denote a random symmetric n by n matrix, whose upper diagonal entries are iid Bernoulli random variables (which take value -1 and 1 with probability 1/2). Improving the earlier result by Costello, Tao and Vu [4], we show that M_n is non-singular with probability $1 - O(n^{-C})$ for any positive constant C. The proof uses an inverse Littlewood-Offord result for quadratic forms, which is of interest of its own.

1. Introduction

Let A_n denote a random n by n matrix, whose entries are iid Bernoulli random variables which take values ± 1 with probability 1/2. Let p_n be the probability that A_n is singular. A classical result of Komlós [1, 13] shows

$$p_n = O(n^{-1/2}). \tag{1}$$

By considering the event that two rows or two columns of A_n are equal (up to a sign), it is clear that

$$p_n \geq (1 + o(1))n^2 2^{1-n}.$$

It has been conjectured by many researchers that in fact this bound is best possible.

Conjecture 1.1.

$$p_n = (\frac{1}{2} + o(1))^n.$$

In a breakthrough paper, Kahn, Komlós and Szemerédi [9] proved that

$$p_n = O(.999^n).$$

Another significant improvement is due to Tao and Vu [24], who used inverse theory from additive combinatorics to show that $p_n = O((3/4)^n)$. The most recent record is due to Bourgain, Vu and Wood [2], who improved it to $p_n = O((1/\sqrt{2})^n)$.

2000 *Mathematics Subject Classification.* 11B25.

Another popular model of random matrices is that of random symmetric matrices; this is one of the simplest models that has non-trivial correlations between the matrix entries. Let M_n denote a random symmetric n by n matrix, whose upper diagonal entries are iid Bernoulli random variables.

Let q_n be the probability that M_n is singular. Despite its obvious similarity to p_n, less is known concerning the bound for q_n. A significant new difficulty is that the symmetry ensures that the determinant $\det(M_n)$ is a quadratic function of each row, as opposed to $\det(A_n)$ which is a linear function of each row.

As far as we can trace, the question to determine whether q_n tends to zero together with n was first posed by Weiss in the early nineties. This simple looking question had been open until a recent breakthrough paper by Costello, Tao and Vu [4], who showed

$$q_n = n^{-1/8+o(1)}.$$

To prove this result, Costello, Tao and Vu introduced and studied a *quadratic* variant of the classical Erdős-Littlewood-Offord inequality concerning the concentration of random variables. Note that this classical inequality plays a key role in the work of Komlós to establish (1).

Although the bound $q_n = n^{-1/8+o(1)}$ can be improved further by applying the more recent inequalities from [3], it seems that the approach developed by Costello, Tao and Vu cannot give any bound better than $n^{-1/2+o(1)}$.

In this paper we show that q_n decays faster than any polynomial in n.

Theorem 1.2 (Main theorem)**.** *We have*

$$q_n = O(n^{-C})$$

for any positive constant C, where the implied constant depends on C.

One may hope to combine our approach and the "replacement technique" from [9] and [24] to improve the bound further to exponential decay. However, we have not been able to do so. It is commonly believed that (see [26])

Conjecture 1.3.

$$q_n = (\frac{1}{2} + o(1))^n.$$

Notation. Here and later, asymptotic notations such as O, Ω, Θ, and so for, are used under the assumption that $n \to \infty$. A notation such as $O_C(.)$ emphasizes that the hidden constant in O depends on C. If $a = \Omega(b)$, we write $b \ll a$ or $a \gg b$.

For a matrix A we use the notations $\mathbf{r}_i(A)$ and $\mathbf{c}_j(A)$ to denote its i-th row and j-th column respectively; we use the notation $A(ij)$ to denote its ij entry.

2. The approach

Let $x = (x_1, \dots, x_n)$ be the first row of M_n, and $a_{ij}, 2 \le i, j \le n$, be the cofactors of M_{n-1} obtained by removing x and x^T from M_n. We have

$$\det(M_n) = x_1^2 \det(M_{n-1}) + \sum_{2 \le i,j \le n} a_{ij} x_i x_j.$$

Roughly speaking, the main approach of [4] is to show that with high probability (with respect to M_{n-1}) most of the a_{ij} are nonzero. It then follows that, via the so called quadratic Littlewood-Offord inequality (Theorem 5.1),

$$\mathbf{P}_x(\det(M_n) = 0) = n^{-1/8+o(1)}.$$

In this paper we adapt the reversed approach, which consists of two main steps outlined below.

(1) If $\mathbf{P}_x(\det(M_n) = 0 | M_{n-1}) \ge n^{-O(1)}$, then there is a strong additive structure among the cofactors a_{ij}.

(2) With respect to M_{n-1}, a strong additive structure among the a_{ij} occurs with negligible probability.

The first step, which is at the heart of our paper, concentrates on the study of inverse Littlewood-Offord problem for linear forms and quadratic forms. We will provide an almost complete answer to this problem throughout Section 3, 4, and 5.

For the rest of this section we sketch the proof of Theorem 1.2.

We first show that it is enough to consider the case of M_n having rank $n-1$, thanks to the following result.

Lemma 2.1. *For any* $1 \le k \le n-2$,

$$\mathbf{P}(\mathrm{rank}(M_n) = k \le n-2) \le 0.1 \times \mathbf{P}(\mathrm{rank}(M_{2n-k-1}) = 2n-k-2).$$

We deduce Lemma 2.1 from a useful observation by Odlyzko.

Lemma 2.2 (Odlyzko's lemma,[17])**.** *Let H be a linear subspace in $\mathbf{R}^n$ of dimension at most $k \le n$. Then it contains at most 2^k vectors from $\{-1,1\}^n$.*

Proof. (of Lemma 2.1) Because M_n has rank k, the subspace spanned by its rows intersects $\{-1,1\}^n$ in a set H of no more than 2^k vectors. Thus the probability that the subvector formed by the last n components of the first row of M_{n+1} does not belong to H is at least $1 - 2^{-n+k}$. Hence,

$$\mathbf{P}(\mathrm{rank}(M_{n+1}) = k+2 | \mathrm{rank}(M_n) = k) \ge 1 - 2^{-n+k}.$$

In general, for $1 \le t \le n-k$ we have

$$\mathbf{P}(\mathrm{rank}(M_{n+t}) = k+2t | \mathrm{rank}(M_{n+t-1}) = k+2(t-1)) \ge 1 - 2^{-n+t+k-1}.$$

Because the rows (and columns) added to M_{n+t-1} each step (to create M_{n+t}) are independent, we have

$$\begin{aligned} &\mathbf{P}(\mathrm{rank}(M_{2n-k-1}) = 2n-k-2 | \mathrm{rank}(M_n) = k) \ge \\ &\ge \prod_{t=1}^{n-k-1} \mathbf{P}\big(\mathrm{rank}(M_{n+t}) = k+2t | \mathrm{rank}(M_{n+t-1}) = k+2(t-1)\big) \\ &\ge (1-2^{-n+k})(1-2^{-n+k+1})\dots(1-2^{-1}) \ge 0.1. \end{aligned}$$

□

Next we show that in the case of M_n having rank $n-1$, it suffices to assume that $\mathrm{rank}(M_{n-1}) \ge n-2$, thanks to the following simple observation.

Lemma 2.3. *Assume that M_n has rank $n-1$. Then there exists $1 \le i \le n$ such that the removal of the i-th row and the i-column of M_n results in a symmetric matrix M_{n-1} of rank at least $n-2$.*

Proof. (of Lemma 2.3) With out loss of generality, assume that the last $n-1$ rows of M_n span a subspace of dimension $n-1$. Then the matrix obtained from M_n by removing the first row and the first column has rank at least $n-2$. □

To prove Theorem 1.2, it thus suffices to prove

Theorem 2.4.

$$\mathbf{P}(\det(M_n) = 0, \mathrm{rank}(M_{n-1}) = n-1) = O(n^{-C}).$$

Theorem 2.5.

$$\mathbf{P}(\det(M_n) = 0, \mathrm{rank}(M_{n-1}) = n-2) = O(n^{-C}).$$

We will prove Theorem 2.4 by relying on a structural lemma stated below, which follows from our study of the inverse Littlewood-Offord problem for linear forms in Step 1.

Lemma 2.6 (Structural theorem, degenerate case)**.** *Let $\epsilon < 1$ and C be positive constants. Assume that M_{n-1} has rank $n-2$ and that*

$$\mathbf{P}_x(\sum_{i,j} a_{ij} x_i x_j = 0 | M_{n-1}) \ge n^{-C}.$$

Then there is a nonzero vector $u = (u_1, \dots, u_{n-1})$ with the following properties.

- *All but n^ϵ elements of u_i belong to a symmetric proper generalized arithmetic progression of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.*
- $u_i \in \{p/q : p, q \in \mathbf{Z}, |p|, |q| = n^{O_{C,\epsilon}(n^\epsilon)}\}$ *for all i.*
- *u is orthogonal to $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1}.*

We refer the reader to Section 3 for a definition of generalized arithmetic progression. Theorem 2.5 follows from a similar structural lemma, which can be deduced from our study of the inverse Littlewood-Offord problem for quadratic forms in Step 1.

Lemma 2.7 (Structural theorem, non-degenerate case)**.** *Let $\epsilon < 1$ and C be positive constants. Assume that M_{n-1} has rank $n-1$ and that*

$$\mathbf{P}_x(\sum_{i,j} a_{ij} x_i x_j = 0 | M_{n-1}) \geq n^{-C}.$$

Then there exists a nonzero vector $u = (u_1, \dots, u_{n-1})$ with the following properties.

- *All but n^ϵ elements of u_i belong to a proper symmetric generalized arithmetic progression of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.*
- $u_i \in \{p/q : p, q \in \mathbf{Z}, |p|, |q| = n^{O_{C,\epsilon}(n^\epsilon)}\}$ *for all i.*
- *u is orthogonal to $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1}.*

The rest of the paper is organized as follows. In Sections 3-5, we discuss the inverse Littlewood-Offord problem in details. As applications, we prove Lemma 2.6 and Lemma 2.7 in Section 9 and Section 10 respectively. We conclude by proving Theorem 2.4 and Theorem 2.5 in Section 11.

3. The inverse Littlewood-Offord problem for linear forms

Let $x_i, i = 1, \dots, n$ be iid Bernoulli random variables, taking values ± 1 with probability $\frac{1}{2}$. Given a multiset A of n real number $a_1, \dots, a_n$, we define the random walk S with steps in A to be the random variable $S := \sum_{i=1}^n a_i x_i$. The *concentration probability* is defined to be

$$\rho(A) := \sup_a \mathbf{P}(S = a).$$

Motivated by their study of random polynomials, in the 1940s Littlewood and Offord [15] raised the question of bounding $\rho(A)$. (We call this the *forward* Littlewood-Offord problem, in contrast with the *inverse* Littlewood-Offord problem discussed later.) They showed that if the a_i are nonzero then $\rho(A) = O(n^{-1/2} \log n)$. Shortly after the Littlewood-Offord paper, Erdős [5] gave a beautiful combinatorial proof of the refinement

$$\rho(A) \le \frac{\binom{n}{n/2}}{2^n} = O(n^{-1/2}). \tag{2}$$

The results of Littlewood-Offord and Erdős are classics in combinatorics and have generated an impressive wave of research, particularly from the early 1960s to the late 1980s.

One direction of research was to generalize Erdős' result to other groups. For example, in 1966 and 1970 [12], Kleitman extended Erdős' result to complex numbers and normed vectors, respectively. Several results in this direction can be found in [11].

Another direction was motivated by the observation that (2) can be improved significantly by making additional assumptions about V. The first such result was discovered by Erdős and Moser [6], who showed that if a_i are distinct, then $\rho(A) = O(n^{-3/2}\log n)$. This bound was then sharpened to $\rho(A) = O(n^{-3/2})$ by Sárkőzy and Szemerédi [20]. Another famous result regarding this result of Erdős and Moser is that of Stanley [21], who shows that if a_i are distinct then $\rho(A) \le \rho(A_0)$, where $A_0 := \{-\lfloor n/2\rfloor, \dots, \lfloor n/2\rfloor\}$.

In [8] (see also in [25]), Halász proved very general theorems that imply the Sárkőzy-Szemerédi theorem and many others. One of his results can be formulated as follows.

Theorem 3.1. *Let l be a fixed integer and R_l be the number of solutions of the equation $a_{i_1} + \dots + a_{i_l} = a_{j_1} + \dots + a_{j_l}$. Then*

$$\rho(A) = O(n^{-2l-\frac{1}{2}} R_l).$$

We remark that the Erdős-Littlewood-Offord inequality (2) and Theorem 3.1 of Halász can be extended to the continuous setting. This type of concentration has been vastly investigated in the literature, we refer the reader to [7, 8, 14, 18] for further reading. We mention here an asymptotic result of Kanter [10], which generalizes (2) and is closely related to our discussion

Theorem 3.2. *Let Φ be a symmetric convex measurable set in a vector space V, and $a_i \in V$. Assume that there are $\Theta(n)$ indices i such that $a_i \notin \Phi$. Then we have*

$$\sup_a \mathbf{P}(S \in a + \Phi) = O(n^{-1/2}).$$

Let us now turn to the main goal of this section.

Motivated by inverse theorems from additive combinatorics (see [25, Chapter 5]) and a variant for random sums in [24, Theorem 5.2], Tao and Vu [23] brought a different view to the problem. Instead of trying to improve the bound further by imposing new assumptions as done in the forward problems, they tried to provide the complete picture by finding the underlying reason as to why the concentration probability is large (say, polynomial in n).

Note that the (multi)-set A has 2^n subsums, and $\rho(A) \geq n^{-C}$ means that at least $\frac{2^n}{n^C}$ among these take the same value. This observation suggests that the set should have a very strong additive structure. To determine this structure, let us introduce an important concept in additive combinatorics, *generalized arithmetic progressions* (GAPs).

A set Q is a *GAP of rank* r if it can be expressed as in the form

$$Q = \{g_0 + m_1 g_1 + \cdots + m_r g_r | N_i \leq m_i \leq N_i' \text{ for all } 1 \leq i \leq r\}$$

for some $g_0, \ldots, g_r, N_1, \ldots, N_r, N_1', \ldots, N_r'$.

It is convenient to think of Q as the image of an integer box $B := \{(m_1, \ldots, m_r) \in \mathbf{Z}^r | M_i \leq m_i \leq M_i'\}$ under the linear map

$$\Phi : (m_1, \ldots, m_r) \mapsto g_0 + m_1 g_1 + \cdots + m_r g_r.$$

The numbers g_i are the *generators* of P, the numbers N_i', N_i are the *dimensions* of P, and $\mathrm{Vol}(Q) := |B|$ is the *volume* of B. We say that Q is *proper* if this map is one to one, or equivalently if $|Q| = \mathrm{Vol}(Q)$. For non-proper GAPs, we of course have $|Q| < \mathrm{Vol}(Q)$. If $-N_i = N_i'$ for all $i \geq 1$ and $g_0 = 0$, we say that Q is *symmetric*.

We next consider an example of A where $\rho(A)$ is large. For a positive integer l we denote the set $\{a_1 + \cdots + a_l | a_i \in A\}$ by lA.

Example 3.3 (Structure implies large concentration probability). *Let Q be a proper symmetric GAP of rank r and volume N. Let $a_1, \ldots, a_n$ be (not necessarily distinct) elements of P. The random variable $S = \sum_{i=1}^n a_i x_i$ takes values in the GAP nP. Because $|nP| \leq \mathrm{Vol}(nB) = n^r N$, the pigeonhole principle implies that $\rho(V) \geq \Omega(\frac{1}{n^r N})$. In fact, by using the second moment method, one can improve the bound to $\Omega(\frac{1}{n^{r/2} N})$. If we set $N = n^{C - r/2}$ for some constant $C \geq r/2$, then*

$$\rho(V) = \Omega(\frac{1}{n^C}). \tag{3}$$

The example above shows that, if the elements of A belong to a symmetric proper GAP with a small rank and small cardinality, then $\rho(V)$ is large. A few years ago, Tao and Vu [22, 23] proved several versions showing that this is essentially the only reason. We present here an optimal version due to Vu and the current author.

Theorem 3.4 (Optimal inverse Littlewood-Offord theorem for linear forms). [16, Theorem 2.5] *Let $\epsilon < 1$ and C be positive constants. Assume that*

$$\rho(A) \geq n^{-C}.$$

Then, for any $n^\epsilon \leq n' \leq n$, there exists a proper symmetric GAP Q of rank $r = O_{C,\varepsilon}(1)$ that contains all but at most n' elements of A (counting multiplicity), where

$$|Q| = O_{C,\varepsilon}(\rho(A)^{-1} n'^{-\frac{r}{2}}).$$

Our method can be extended to more general distributions. We just cite one below for our later applications.

Let $0 < \mu \le 1$ be a positive parameter. Let η^μ be a random variable such that $\eta^\mu = 1$ or -1 with probability $\mu/2$, and $\eta^\mu = 0$ with probability $1 - \mu$.

Theorem 3.5. *The conclusion of Theorem 3.4 also holds if the x_i are iid copies of η^μ.*

Remark 3.6. In their work to obtain the bound $p_n = O((3/4)^n)$, Tao and Vu studied a similar inverse problem.

Let $0 < \mu < 1/4$ be a parameter, and let $\epsilon < 1$ be a positive constant.

Define

$$\rho^{(\mu)}(A) := \sup_{a \in \mathbf{R}} \mathbf{P}(\sum_{i=1}^{n} a_i \eta_i^\mu = a).$$

It can be shown that $\rho(A) \le \rho^{(\mu)}(A)$. In [24], Tao and Vu characterized those A where $\rho(A)$ is comparable to $\rho_\mu(A)$,

$$\rho(A) \ge \epsilon \rho^{(\mu)}(A).$$

4. The inverse Littlewood-Offord problem for bilinear forms

Let x_i, y_j be iid Bernoulli random variables, let $A = (a_{ij})$ be an $n \times n$ matrix of real entries. We define the *bilinear concentration probability* of A by

$$\rho_b(A) := \sup_{a \in \mathbf{R}} \mathbf{P}(\sum_{i,j} a_{ij} x_i y_j = a).$$

More generally , if x_i, y_j are iid copies of η^μ, then the *weighted bilinear concentration probability* of A is defined by

$$\rho_b^{(\mu)}(A) = \sup_{a \in \mathbf{R}} \mathbf{P}(\sum_{i,j} a_{ij} x_i y_j = a).$$

As an application of the Littlewood-Offord-Erdős inequality (2), it has been shown in [3] (also in [4] with a weaker bound) that

Theorem 4.1 (Bilinear Littlewood-Offord inequality)**.** *Suppose that there are $\Theta(n)$ indices i such that for each i there are $\Theta(n)$ indices j such that $a_{ij} \neq 0$. Then*

$$\rho_b(A) = O(n^{-1/2}).$$

The bound $O(n^{-1/2})$ is sharp, as the bilinear form $\sum_{i,j} x_i y_j$ shows.

The bilinear Littlewood-Offord inequality for the continuous setting was also studied in the literature. For instance, as an application of Kanter's inequality (Theorem 3.2), it follows from a result of Rosiński and Samorodnitsky [19] that

Theorem 4.2. *Let Φ be a symmetric convex measurable set in a vector space V, and $a_i \in V$. Assume that there are $\Theta(n)$ indices i such that for each i there are $\Theta(n)$ indices j such that $a_{ij} \notin \Phi$. Then we have*

$$\sup_{a\in V} \mathbf{P}(\sum_{i,j} a_{ij}x_iy_j \in a+\Phi) = O(n^{-1/16}).$$

Rosiński and Samorodnitsky also studied concentration inequalities for more general multilinear forms. We refer the reader to [19] for further reading.

Motivated by the inverse Littlewood-Offord results for linear forms, our goal is to find the reason as to why $\rho_b(A)$ is large.

Question 4.3. *Is it true that if $\rho_b(A)$ is large then there must be a "structural" relation among the entries of A?*

To answer this question, we first consider a few examples of A.

Example 4.4 (Additive structure implies large concentration probability)**.** *Let Q be a proper symmetric GAP of rank $r = O(1)$ and of size $n^{O(1)}$. Assume that $a_{ij} \in Q$, for all a_{ij}. Then for any $x_i, y_j \in \{\pm 1\}$,*

$$\sum_{i,j} a_{ij}x_iy_j \in n^2 Q.$$

Thus, by the pigeon-hole principle, we have

$$\rho_b(A) \geq n^{-2r}|Q|^{-1} = n^{-O(1)}.$$

Our next example shows that if the a_{ij} are "separable", then $\rho_b(A)$ is also large.

Example 4.5 (Algebraic structure implies large concentration probability)**.** *Assume that*

$$a_{ij} = k_ib_j + l_jb'_i,$$

where b_j, b'_i are arbitrary real numbers and $k_i, l_j \in \mathbf{Z}, |k_i|, |l_j| = n^{O(1)}$, such that

$$\mathbf{P}_x(\sum_i k_ix_i = 0) = n^{-O(1)}$$

and

$$\mathbf{P}_y(\sum_j l_j y_j = 0) = n^{-O(1)}.$$

Then we have

$$\mathbf{P}_{x,y}(\sum_{i,j} a_{ij}x_i y_j = 0) = \mathbf{P}\left(\sum_i k_i x_i \sum_j b_j y_j + \sum_i b_i' x_i \sum_j l_j y_j = 0\right) = n^{-O(1)}.$$

Remark 4.6. In the above example, the assumption that k_i, l_j are integers seems unnecessary. However, because $\mathbf{P}_x(\sum_i k_i x_i = 0) = n^{-O(1)}$ and $\mathbf{P}_y(\sum_j l_j y_j = 0) = n^{-O(1)}$, Theorem 3.4 implies that most of the k_i and l_j belong to a GAP of bounded size. Thus, without loss of generality, we may assume that k_i, l_j are bounded integers.

Our last example shows that a combination of additive structure and algebraic structure also implies high bilinear concentration probability.

Example 4.7 (Structure implies large concentration probability)**.** *Assume that $a_{ij} = a_{ij}' + a_{ij}''$, where $a_{ij}' \in Q$, a proper symmetric GAP of rank $O(1)$ and size $n^{O(1)}$, and*

$$a_{ij}'' = k_{i1}b_{1j} + \dots k_{ir}b_{rj} + l_{1j}b_{i1}' + \dots + l_{rj}b_{ir}',$$

where $b_{1j}, \dots, b_{rj}, b_{i1}', \dots, b_{ir}'$ are arbitrary and $k_{i1}, \dots, k_{ir}, l_{1j}, \dots, l_{rj}$ are integers bounded by $n^{O(1)}$, and $r = O(1)$ such that

$$\mathbf{P}_x\left(\sum_i k_{i1}x_i = 0, \dots, \sum_i k_{ir}x_i = 0\right) = n^{-O(1)}$$

and

$$\mathbf{P}_y\left(\sum_j l_{1j}y_j = 0, \dots, \sum_j l_{rj}y_j = 0\right) = n^{-O(1)}.$$

Then we have

$$\sum_{i,j} a_{ij}x_i y_j = \sum_{i,j} a_{i,j}' x_i y_j + \sum_i k_{i1}x_i \sum_j b_{1j}y_j + \dots + \sum_i k_{ir}x_i \sum_j b_{rj}y_j$$
$$+ \sum_i b_{i1}' x_i \sum_j l_{1j}y_j + \dots + \sum_i b_{ir}' x_i \sum_j l_{rj}y_j.$$

Thus,

$$\mathbf{P}_{x,y}\left(\sum_{i,j} a_{ij}x_iy_j \in n^2Q\right) = n^{-O(1)}.$$

It then follows, by the pegion-hole principle, that $\rho_b(A) = n^{-O(1)}$.

The above examples demonstrate that if the a_{ij} can be decomposed into additive and algebraic structural parts, then $\rho_b(A)$ is large. Our inverse result asserts that these are essentially the only ones that have large bilinear concentration probability.

Theorem 4.8 (Inverse Littlewood-Offord theorem for bilinear forms)**.** *Let $\epsilon < 1, C$ be positive constants. Assume that*

$$\rho_b(A) \geq n^{-C}.$$

Then there exist index sets I_0, J_0, both of size $O_{C,\epsilon}(1)$, and index sets I, J, both of size $n - O_C(n^\epsilon)$, with $I \cap I_0 = \emptyset, J \cap J_0 = \emptyset$, and there exist integers $k, l, k_{ii_0}, l_{jj_0}, i_0 \in I_0, j_0 \in J_0, i \in I, j \in J$, all of size bounded by $n^{O_{C,\epsilon}(1)}$, such that the following hold for all $i \in I$:

- *for any $j \in J$,*

$$a_{ij} = \frac{a'_{ij}}{kl} - \frac{\sum_{i_0\in I_0} k_{ii_0}a_{i_0j}}{k} - \frac{\sum_{j_0\in J_0} l_{j_0j}a_{ij_0}}{l};$$

- *all but $O_C(n^\epsilon)$ entries a'_{ij} belong to a proper symmetric GAP Q_i depending on i, which has rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.*

Although Theorem 4.8 is enough for our later application, it does not yet reflect the examples given, namely the additive structures Q_i corresponding to each row can be totally different. In the next theorem we show that these GAPs can be unified into a structure similar to a GAP.

Theorem 4.9 (Inverse Littlewood-Offord theorem for bilinear forms, common structure)**.** *Let $\epsilon < 1, C$ be positive constants. Assume that*

$$\rho_b(A) \geq n^{-C}.$$

Then there exist index sets I_0, J_0, both of size $O_{C,\epsilon}(1)$, and index sets I, J, both of size $n - O_C(n^\epsilon)$, with $I \cap I_0 = \emptyset, J \cap J_0 = \emptyset$, and there exist integers $k, l, k_{ii_0}, l_{jj_0}, i_0 \in I_0, j_0 \in J_0, i \in I, j \in J$, all of size bounded by $n^{O_{C,\epsilon}(1)}$, such that for all $i \in I$ the following hold:

- *for any $j \in J$,*

$$a_{ij} = \frac{a'_{ij}}{kl} - \frac{\sum_{i_0 \in I_0} k_{ii_0} a_{i_0 j}}{k} - \frac{\sum_{j_0 \in J_0} l_{j_0 j} a_{ij_0}}{l};$$

- *all but $O_C(n^\epsilon)$ entries a'_{ij} belong to a set Q (independent of i) of the form*

$$Q = \{ \sum_{h=1}^{O_{C,\epsilon}(1)} (p_h/q_h) \cdot g_h; p_h, q_h \in \mathbf{Z}, |p_h|, |q_h| = n^{O_{C,\epsilon}(1)} \}.$$

Our proof of Theorem 4.8 and 4.9 can be extended (rather automatically) to other Bernoulli distributions.

Theorem 4.10. *Let $0 < \mu \le 1$ be a constant. Then the conclusions of Theorem 4.8 and Theorem 4.9 also hold if we assume that $\rho_b^{(\mu)}(A) \ge n^{-C}$.*

Remark 4.11. The inverse Littlewood-Offord problem for bilinear forms was also studied in [3], but only for the case $\rho_b(A) \ge n^{-1+o(1)}$.

5. The inverse Littlewood-Offord problem for quadratic forms

Let x_i be iid Bernoulli random variables, let $A = (a_{ij})$ be an $n \times n$ symmetric matrix of real entries. We define the *quadratic concentration probability* of A by

$$\rho_q(A) := \sup_{a \in \mathbf{R}} \mathbf{P}(\sum_{i,j} a_{ij} x_i x_j = a).$$

More general, if x_i are iid copies of η^μ, then the *weighted quadratic concentration probability* of A is defined by

$$\rho_q^{(\mu)}(A) := \sup_{a \in \mathbf{R}} \mathbf{P}(\sum_{i,j} a_{ij} x_i x_j = a).$$

It was shown in [3, 4], as an application of Theorem 4.1, that

Theorem 5.1 (Quadratic Littlewood-Offord inequality)**.** *Suppose that there are $\Theta(n)$ indices i such that for each i there are $\Theta(n)$ indices j such that $a_{ij} \neq 0$. Then*

$$\rho_q(A) \le n^{-1/2+o(1)}.$$

The bound $n^{-1/2+o(1)}$ is almost best possible, as demonstrated by the quadratic form $\sum_{ij} x_i x_j$.

A more general version of Theorem 5.1 also appeared in the mentioned paper of Rosiński and Samorodnitsky.

Theorem 5.2. [19, Theorem 3.1] *Let Φ be a symmetric convex measurable set in a vector space V, and $a_i \in V$. Assume that there are $\Theta(n)$ indices i such that for each i there are $\Theta(n)$ indices j such that $a_{ij} \notin \Phi$. Then we have*

$$\sup_{a\in V} \mathbf{P}(\sum_{i,j} a_{ij}x_ix_j \in a+\Phi) = O(n^{-1/16}).$$

Motivated by the inverse Littlewood-Offord results for linear forms and bilinear forms, we would like to characterize those A which have large quadratic concentration probability.

We first consider a few examples of A when $\rho_q(A)$ is large, based on the examples given in the previous sections.

Example 5.3 (Additive structure implies large concentration probability)**.** *Let Q be a proper symmetric GAP of rank $r = O(1)$ and of size $n^{O(1)}$. Assume that $a_{ij} \in Q$, then for any $x_i \in \{\pm 1\}$*

$$\sum_{i,j} a_{ij}x_ix_j \in n^2Q.$$

Thus, by the pigeon-hole principle,

$$\rho_q(A) \ge n^{-2r}|Q|^{-1} = n^{-O(1)}.$$

Similar to Example 4.5, our next example shows that if the a_{ij} are separable, then $\rho_q(A)$ is large.

Example 5.4 (Algebraic structure implies large concentration probability)**.** *Assume that*

$$a_{ij} = k_ib_j + k_jb_i$$

where $k_i \in \mathbf{Z}, |k_i| = n^{O(1)}$ and such that $\mathbf{P}_x(\sum_i k_ix_i = 0) = n^{-O(1)}$.

Then we have

$$\mathbf{P}(\sum_{i,j} a_{ij}x_ix_j = 0) = \mathbf{P}(\sum_i k_ix_i \sum_j b_jx_j = 0) = n^{-O(1)}.$$

In our last example, we show that a combination of both structures also implies high quadratic concentration probability.

Example 5.5 (Structure implies large concentration probability)**.** *Assume that $a_{ij} = a_{ij}' + a_{ij}''$, where $a_{ij}' \in Q$, a proper symmetric GAP of rank $O(1)$ and size $n^{O(1)}$, and*

$$a''_{ij} = k_{i1}b_{1j} + k_{j1}b_{1i} + \cdots + k_{ir}b_{rj} + k_{jr}b_{ri},$$

where $b_{1i}, \dots, b_{ri}$ *are arbitrary and* $k_{i1}, \dots, k_{ir}$ *are integers bounded by* $n^{O(1)}$*, and* $r = O(1)$ *such that*

$$\mathbf{P}_x\left(\sum_i k_{i1}x_i = 0, \dots, \sum_i k_{ir}x_i = 0\right) = n^{-O(1)}.$$

Then we have

$$\sum_{i,j} a_{ij}x_ix_j = \sum_{i,j} a'_{i,j}x_ix_j + (\sum_i k_{i1}x_i)(\sum_j b_{1j}x_j) + \cdots + (\sum_i k_{ir}x_i)(\sum_i b_{rj}x_j).$$

Thus,

$$\mathbf{P}_x(\sum_{i,j} a_{ij}x_ix_j \in n^2Q) = n^{-O(1)}.$$

It then follows, by the pigeon-hole principle, that $\rho_q(A) = n^{-O(1)}$*.*

Next we state our main result which asserts that the examples above are essentially the only ones that have high quadratic concentration probability.

Theorem 5.6 (Inverse Littlewood-Offord theorem for quadratic forms)**.** *Let* $\epsilon < 1, C$ *be positive constants. Assume that*

$$\rho_q(A) \geq n^{-C}.$$

Then there exist index sets I_0 *and* I *of size* $O_{C,\epsilon}(1)$ *and* $n - O_C(n^\epsilon)$ *respectively, and* $I \cap I_0 = \emptyset$*, and there exist integers* $k, k_{ii_0} \in \mathbf{Z}, i_0 \in I_0, i \in I$*, all bounded by* $n^{O_{C,\epsilon}(1)}$*, such that the following hold for all* $i \in I$*:*

- *for any* $j \in I$*,*
$$a_{ij} = a'_{ij}/k^2 - k\sum_{i_0 \in I_0} k_{ii_0}a_{i_0j}/k^2 - k\sum_{i_0 \in I_0} k_{ji_0}a_{i_0i}/k^2;$$
- *all but* $O_C(n^\epsilon)$ *entries* a'_{ij} *belong to a proper symmetric GAP* Q_i *depending on* i*, which has rank* $O_{C,\epsilon}(1)$ *and size* $n^{O_{C,\epsilon}(1)}$*.*

Similar to Theorem 4.9, we show that the structures Q_i from Theorem 5.6 can be unified into a structure similar to a GAP.

Theorem 5.7 (Inverse Littlewood-Offord theorem for quadratic forms, common structure)**.** *Let $\epsilon < 1, C$ be positive constants. Assume that*

$$\rho_q(A) \geq n^{-C}.$$

Then there exist index sets I_0, I of size $O_{C,\epsilon}(1)$ and $n - O_C(n^\epsilon)$ respectively, with $I \cap I_0 = \emptyset$, and there exist integers $k, k_{ii_0}, i_0 \in I_0, i \in I$, all of size bounded by $n^{O_{C,\epsilon}(1)}$, such that for all $i \in I$ the following hold:

- *for any $j \in I$,*

$$a_{ij} = a'_{ij}/k^2 - k \sum_{i_0 \in I_0} k_{ii_0} a_{i_0 j}/k^2 - k \sum_{i_0 \in I_0} k_{ji_0} a_{i_0 i}/k^2;$$

- *all but $O_C(n^\epsilon)$ entries a'_{ij} belong to a set Q (independent of i) of the form*

$$Q = \{ \sum_{h=1}^{O_C(1)} (p_h/q_h) \cdot g_h; p_h, q_h \in \mathbf{Z}, |p_h|, |q_h| = n^{O_{C,\epsilon}(1)} \}.$$

Remark 5.8. The conclusions of Theorem 5.6 and Theorem 5.7 also hold if we assume that

$$\rho_q^{(\mu)}(A) \geq n^{-C}.$$

We invite the reader to prove this result using the approach presented in Section 8.

Remark 5.9. The inverse Littlewood-Offord problem for quadratic forms was also studied in [3], but only in the case $\rho_q(A) \geq n^{-1/2+o(1)}$.

6. A rank reduction argument and the full rank assumption

This section, which can be read independently of the rest of this paper, provides a technical lemma we will need for later sections. Informally, it says that if we can find a proper symmetric GAP that contains a given set (in the spirit of Sections 3, 4 and 5), then we can assume this containment is non-degenerate. More details follow.

Assume that $P = \{m_1 g_1 + \cdots + m_r g_r | -M_i \leq m_i \leq M_i\}$ is a proper symmetric GAP, which contains a set $U = \{u_1, \ldots u_n\}$.

We consider P together with the map $\Phi : P \to \mathbf{R}^r$ which maps $m_1 g_1 + \cdots + m_r g_r$ to $(m_1, \ldots, m_r)$. Because P is proper, this map is bijective.

We know that P contains U, but we do not know yet that U is non-degenerate in P in the sense that the set $\Phi(U)$ has full rank in $\mathbf{R}^r$. In the later case, we say U *spans* P.

Theorem 6.1. *Assume that U is a subset of a proper symmetric GAP P of size r, then there exists a proper symmetric GAP Q that contains U such that the following hold.*

- $\mathrm{rank}(Q) \le r$ *and* $|Q| \le O_r(1)|P|$;
- U *spans* Q, *that is,* $\phi(U)$ *has full rank in* $\mathbf{R}^{\mathrm{rank}(Q)}$.

To prove Theorem 6.1, we will rely on the following lemma.

Lemma 6.2 (Progressions lie inside proper progressions)**.** [25, Chapter 3.] *There is an absolute constant* C *such that the following holds. Let* P *be a GAP of rank* r *in* $\mathbf{R}$. *Then there is a symmetric proper GAP* Q *of rank at most* r *containng* P *and*

$$|Q| \le r^{Cr^3}|P|.$$

Proof. (of Theorem 6.1) We shall mainly follow [24, Section 8].

Suppose that $\Phi(U)$ does not have full rank, then it is contained in a hyperplane of $\mathbf{R}^r$. In other words, there exist integers $\alpha_1, \dots, \alpha_r$ whose common divisor is one and $\alpha_1 m_1 + \dots + \alpha_r m_r = 0$ for all $(m_1, \dots, m_r) \in \Phi(U)$.

Without loss of generality, we assume that $\alpha_r \neq 0$. We select w so that $g_r = \alpha_r w$, and consider P' be the GAP generated by $g_i' := g_i - \alpha_i w$ for $1 \le i \le r-1$. The new symmetric GAP P' will continue to contain U, because we have

$$\begin{aligned} m_1 g_1' + \dots + m_{r-1} g_{r-1}' &= m_1 g_1 + \dots + m_r g_r - w(\alpha_1 m_1 + \dots + \alpha_r g_r) \\ &= m_1 g_1 + \dots + m_r g_r \end{aligned}$$

for all $(m_1, \dots, m_r) \in \Phi(U)$.

Also, note that the volume of P' is $2^{r-1} M_1 \dots M_{r-1}$, which is less than the volume of P.

We next use Lemma 6.2 to guarantee that P' is symmetric and proper without increasing the rank.

Iterate the process if needed. Because we obtain a new proper symmetric GAP whose rank strictly decreases each step, the process must terminate after at most r steps.

□

7. Proof of Theorem 4.8 , Theorem 4.9, and Theorem 4.10

We begin by applying Theorem 3.5.

Lemma 7.1. *Let* $\epsilon < 1$, $0 < \mu \le 1$, *and* C *be positive constants. Assume that* $\rho_b^{(\mu)}(A) \ge n^{-C}$. *Then the following holds with probability at least* $\frac{3}{4} n^{-C}$ *with respect to* $y = (y_1, \dots, y_n)$. *There exist a proper symmetric GAP* Q_y *of rank* $O_{C,\epsilon,\mu}(1)$ *and size* $O_{C,\epsilon,\mu}(1/\rho_b^{(\mu)})$ *and a set* I_y *of* $n - n^\epsilon$ *indices such that for each* $i \in I_y$ *we have*

$$\langle \mathbf{r}_i, y\rangle \in Q_y.$$

Proof. (of Lemma 7.1) For short we write

$$\sum_{i,j} a_{ij}x_iy_j = \sum_{i=1}^n x_i\langle \mathbf{r}_i, y\rangle.$$

We say that a vector $y = (y_1, \ldots, y_n)$ is *good* if

$$\mathbf{P}_x(\sum_{i=1}^n x_i\langle \mathbf{r}_i, y\rangle = a) \geq \rho_b^{(\mu)}/4.$$

We call y *bad* otherwise.

First, we estimate the probability p of a randomly chosen vector $y = (y_1, \ldots, y_n)$ being bad by an averaging method.

$$\mathbf{P}_y\mathbf{P}_x \sum_{i=1}^n \langle \mathbf{r}_i, y\rangle = \rho_b^{(\mu)}$$

$$p\rho_b^{(\mu)}/4 + 1 - p \geq \rho_b^{(\mu)}.$$

$$(1 - \rho_b^{(\mu)})/(1 - \rho_b^{(\mu)}/4) \geq p.$$

Thus, the probability of a randomly chosen vector being good is at least

$$1 - p \geq (3\rho_b^{(\mu)}/4)/(1 - \rho_b^{(\mu)}/4) \geq 3\rho_b^{(\mu)}/4.$$

Next, we consider a good vector $y \in G$. By definition, we have

$$\mathbf{P}_x(\sum_{i=1}^n x_i\langle \mathbf{r}_i, y\rangle = a) \geq \rho_b^{(\mu)}/4.$$

A direct application of Theorem 3.5 to the sequence $\langle \mathbf{r}_i, y\rangle$, $i = 1, \ldots, n$ yields the desired result. □

By Theorem 6.1, we may assume that the $\langle \mathbf{r}_i, y\rangle$ span Q_y. From now on we fix such a Q_y for each y.

Let G be the collection of good vectors. Thus,

$$\mathbf{P}_y(y \in G) \geq 3\rho_b^{(\mu)}/4. \tag{4}$$

Next, for each $y \in G$, we choose from I_y s indices $i_{y_1}, \dots, i_{y_s}$ such that $\langle \mathbf{r}_{i_{y_j}}, y\rangle$ span Q_y, where s is the rank of Q_y. We note that $s = O_{C,\epsilon,\mu}(1)$ for all s.

Consider the tuples $(i_{y_1}, \dots, i_{y_s})$ for all $y \in G$. Because there are $\sum_s O_{C,\epsilon,\mu}(n^s) = n^{O_{C,\epsilon,\mu}(1)}$ possibilities these tuples can take, there exists a tuple, say $(1, \dots, r)$ (by rearranging the rows of A if needed, we may assume so), such that $(i_{y_1}, \dots, i_{y_s}) = (1, \dots, r)$ for all $y \in G'$, a subset of G satisfying

$$\mathbf{P}_y(y \in G') \geq \mathbf{P}_y(y \in G)/n^{O_{C,\epsilon,\mu}(1)} = \rho_b^{(\mu)}(A)/n^{O_{C,\epsilon,\mu}(1)}. \tag{5}$$

For each $1 \leq i \leq r$, we express $\langle \mathbf{r}_i, y\rangle$ in terms of the generators of Q_y for each $y \in G'$,

$$\langle \mathbf{r}_i, y\rangle = c_{i1}(y)g_1(y) + \dots + c_{ir}(y)g_r(y),$$

where $c_{i1}(y), \dots c_{ir}(y)$ are integers bounded by $n^{O_{C,\epsilon,\mu}(1)}$, and $g_i(y)$ are the generators of Q_y.

We will show that there are many y that correspond to the same coefficients c_{ij}.

Consider the collection of the coefficient-tuples $\Big(\big(c_{11}(y), \dots, c_{1r}(y)\big); \dots; \big(c_{r1}(y), \dots c_{rr}(y)\big)\Big)$ for all $y \in G'$. Because the number of possibilities these tuples can take is at most

$$(n^{O_{C,\epsilon,\mu}(1)})^{r^2} = n^{O_{C,\epsilon,\mu}(1)}.$$

There exists a coefficient-tuple, say $\Big((c_{11}, \dots, c_{1r}), \dots, (c_{r1}, \dots c_{rr})\Big)$, such that

$$\Big(\big(c_{11}(y), \dots, c_{1r}(y)\big); \dots; \big(c_{r1}(y), \dots c_{rr}(y)\big)\Big) = \Big((c_{11}, \dots, c_{1r}), \dots, (c_{r1}, \dots c_{rr})\Big)$$

for all $y \in G''$, a subset of G' satisfying

$$\mathbf{P}_y(y \in G'') \geq \mathbf{P}_y(y \in G')/n^{O_{C,\epsilon,\mu}(1)} \geq \rho_b^{\mu}(A)/n^{O_{C,\epsilon,\mu}(1)}. \tag{6}$$

In summary, there exist r tuples $(c_{11}, \dots, c_{1r}), \dots, (c_{r1}, \dots c_{rr})$, whose components are integers bounded by $n^{O_{C,\epsilon,\mu}(1)}$, such that the following hold for all $y \in G''$.

- $\langle \mathbf{r}_i, y\rangle = c_{i1}g_1(y) + \dots + c_{jr}g_r(y)$, for $i = 1, \dots, r$.
- The vectors $(c_{11}, \dots, c_{1r}), \dots, (c_{r1}, \dots c_{rr})$ span $\mathbf{Z}^{\mathrm{rank}(Q_y)}$.

Next, because $|I_y| \geq n - n^\epsilon$ for each $y \in G''$, there is a set I of size $n - 3n^\epsilon$ such that $I \cap \{1, \dots, r\} = \emptyset$ and for each $i \in I$ we have

$$\mathbf{P}_y(i \in I_y, y \in G'') \geq \mathbf{P}_y(y \in G'')/2. \tag{7}$$

Indeed, let I' be the set of i satisfying (7). Then, as

$$\sum_i \sum_{y \in G'', i \in I_y} 1 = \sum_{y \in G''} \sum_{i \in I_y} 1 \geq (n - n^\epsilon)|G''|,$$

we have $\sum_{i \in I'} |G''| + \sum_{i \notin I'} |G''|/2 \geq (n - n^\epsilon)|G''|$. Hence,

$$|I'||G''| + (n - |I'|)|G''|/2 \geq (n - n^\epsilon)|G''|,$$

from which we deduce that $|I'| \geq n - 2n^\epsilon$. To obtain I we just remove the elements of $\{1, \dots, r\}$ from I'.

Now fix an arbitrary row $\mathbf{r}$ of index from I. We concentrate on those $y \in G''$ where the index of $\mathbf{r}$ belongs to I_y.

Because $\langle \mathbf{r}, y \rangle \in Q_y$, we can write

$$\langle \mathbf{r}, y \rangle = c_1(y) g_1(y) + \dots c_r(y) g_r(y)$$

where $c_i(y)$ are integers bounded by $n^{O_{C,\epsilon,\mu}(1)}$.

For short, we denote the vector $(c_{i1}, \dots, c_{ir})$ by $\mathbf{v}_i$ for each i. We will also denote the vector $(c_1(y), \dots c_r(y))$ by $\mathbf{v}_{\mathbf{r},y}$.

Because Q_i is spanned by $\langle \mathbf{r}_1, y \rangle, \dots, \langle \mathbf{r}_r, y \rangle$, we have $k = \det(\mathbf{v}_1, \dots \mathbf{v}_r) \neq 0$, and that

$$k\langle \mathbf{r}, y \rangle + \det(\mathbf{v}_{\mathbf{r},y}, \mathbf{v}_2, \dots, \mathbf{v}_r)\langle \mathbf{r}_1, y \rangle + \dots + \det(\mathbf{v}_{\mathbf{r},y}, \mathbf{v}_1, \dots, \mathbf{v}_{r-1})\langle \mathbf{r}_r, y \rangle = 0.$$

Next, because each coefficient of the identity above is bounded by $n^{O_{C,\epsilon,\mu}(1)}$, there exists a subset $G''_{\mathbf{r}}$ of G'' such that all $y \in G''_{\mathbf{r}}$ correspond to the same identity, and

$$\mathbf{P}_y(y \in G''_{\mathbf{r}}) \geq (\mathbf{P}_y(y \in G'')/2)/(n^{O_{C,\epsilon,\mu}(1)})^r = \rho_b^{(\mu)}/n^{O_{C,\epsilon,\mu}(1)}. \tag{8}$$

In other words, there exist integers $k_1, \dots, k_r$, all bounded by $n^{O_{C,\epsilon,\mu}(1)}$, such that

$$k\langle \mathbf{r}, y \rangle + k_1\langle \mathbf{r}_1, y \rangle + \dots + k_r\langle \mathbf{r}_r, y \rangle = 0$$

for all $y \in G''_{\mathbf{r}}$.

Note that k is independent of $\mathbf{r}$ and y. We thus conclude below.

Lemma 7.2 (The rows are mutually orthogonal to many $\{-1,0,1\}$ vectors)**.** *Let i be any index of I. Then there are numbers $k_{i1},\dots,k_{ir} \in \mathbf{Z}$, all bounded by $n^{O_{C,\epsilon,\mu}(1)}$, such that*

$$\mathbf{P}_y\big(k\langle \mathbf{r}_i, y\rangle + \sum_{j=1}^{r} k_{ij}\langle \mathbf{r}_j, y\rangle = 0\big) = \rho_b^{(\mu)}/n^{O_{C,\epsilon,\mu}(1)}.$$

Putting Lemma 7.1, Lemma 7.2, and Theorem 3.5 together, we obtain the following result.

Theorem 7.3 (Refined row relation)**.** *Let $0 < \epsilon \le 1$, $0 < \mu \le 1$, and C be positive constants. Assume that $\rho_b^{(\mu)}(A) \ge n^{-C}$. Then there exist a set I_0 of size $O_{C,\epsilon,\mu}(1)$, a set I of size $n - 3n^\epsilon$ with $I \cap I_0 = \emptyset$, and there exists a nonzero integer k of size $n^{O_{C,\epsilon,\mu}(1)}$ such that the following holds for all $i \in I$: there exists a proper symmetric GAP Q_i of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$, an index set J_i of size $n - n^\epsilon$, and integers $k_{ii_0}, i_0 \in I_0$, all bounded by $n^{O_{C,\epsilon,\mu}(1)}$, such that the following holds for all $j \in J_i$*

$$\sum_{i_0 \in I_0} k_{ii_0} a_{i_0 j} + k a_{ij} \in Q_i.$$

Because the role of rows and columns of A can be swapped, we obtain a similar conclusion for the columns of A.

Theorem 7.4 (Refined column relation)**.** *Let $0 < \epsilon \le 1, 0 < \mu \le 1$, and C be positive constants. Assume that $\rho_b^{(\mu)}(A) \ge n^{-C}$. Then there exist a set J_0 of size $O_{C,\epsilon,\mu}(1)$, a set J of size $n - 3n^\epsilon$ with $J \cap J_0 = \emptyset$, and there exists a nonzero integer l of size $n^{O_{C,\epsilon,\mu}(1)}$ such that the following holds for all $j \in J$: there exists a proper symmetric GAP P_j of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$, an index set I_j of size $n - n^\epsilon$, and integers $l_{j_0 j}, j_0 \in J_0$, all bounded by $n^{O_{C,\epsilon,\mu}(1)}$, such that the following holds for all $i \in I_j$*

$$\sum_{j_0 \in J_0} l_{j_0 j} a_{ij_0} + l a_{ij} \in P_j.$$

Next we introduce the following two matrices.

Definition 7.5 (Row matrix)**.** L is an n by n matrix, whose i-th row, where $i \in I$, is defined by

$$\mathbf{r}_i(L)(j) := \begin{cases} k_{ij}, & \text{if } j \in I_0; \\ k, & \text{if } j = i; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The other entries of L are zero, except the diagonal terms which are set to be 1.

Definition 7.6 (Column matrix)**.** R is an n by n matrix, whose j-th column, where $j \in J$, is defined by

$$\mathbf{c}_j(R)(i) := \begin{cases} l_{ij}, & \text{if } i \in J_0; \\ l, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases} \qquad (10)$$

The other entries of R are zero, except the diagonal terms which are set to be 1.

Remark 7.7. For each $i \in I$, the non-singular matrix L acts on the left of A by rescaling $\mathbf{r}_i(A)$ by a factor of k, modulo $\sum_{i_0 \in I_0} k_{ii_0}\mathbf{r}_{i_0}$. For each $j \in J$, the non-singular matrix R acts on the right of A by rescaling $\mathbf{c}_j(A)$ by a factor of l, modulo $\sum_{j_0 \in J_0} l_{j_0 j}\mathbf{c}_{j_0}$.

Define

$$A' := LAR.$$

First, consider the matrix AR. By definition, $(AR)_{ij} \in P_j$ for all $i \in I_j$, where $j \in J$. By adding a constant number of generators to P_j we may assume that $(AR)_{ij} \in P_j$, where $i \in I_0$.

Next, consider the matrix $A' = LAR$. Suppose that $j \in J$, then we have

$$(LAR)_{ij} = k(AR)_{ij} + \sum_{i_0 \in I_0} k_{ii_0}(AR)_{i_0 j}.$$

Because $k, k_{ii_0} = n^{O_{C,\epsilon,\mu}(1)}$, it thus follows that $(LAR)_{ij} \in n^{O_{C,\epsilon,\mu}(1)} \cdot P_j$ whenever $i \in I_j \cap I$. To avoid notational complication, we keep the same notation P_j for this new proper symmetric GAP (which is still of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$, with possibly worse constants).

We have just shown that for each $j \in J$ there exists a proper symmetric GAP P_j of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that all but n^ϵ coordinates of the j-th column of A' belong to P_j.

Similarly, by viewing LAR as $(LA)R$, we infer that for each $i \in I$, there exists a proper symmetric GAP Q_i of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that all but n^ϵ coordinates of the i-th row of A' belong to Q_i.

Putting everything together, we obtain the following result.

Theorem 7.8 (Matrix relation)**.** *Let $0 < \epsilon \le 1$, $0 < \mu \le 1$, and C be positive constants. Assume that $\rho_b^{(\mu)}(A) \ge n^{-C}$. Then there exist index sets I_0, J_0, both of size $O_{C,\epsilon}(1)$, and index sets I, J, both of size $n - 3n^\epsilon$, with $I \cap I_0 = \emptyset, J \cap J_0 = \emptyset$, such that the following holds. There exist two matrices L, R defined by* (9) *and* (10) *respectively such that the matrix $A' = LAR$ possess the following properties.*

- *For each $i \in I$, there exist a subset $\mathbf{r}_i' \subset \mathbf{r}_i(A')$ of size $n - n^\epsilon$ and a proper symmetric GAP Q_i of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that $\mathbf{r}_i' \subset Q_i$.*
- *For each $j \in J$, there exist a subset $\mathbf{c}_j' \subset \mathbf{c}_j(A')$ of size $n - n^\epsilon$ and a proper symmetric GAP P_j of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that $\mathbf{c}_j' \subset P_j$.*

We now deduce Theorem 4.8. Assume that $i \in I$ and $j \in J$. We then have

$$\begin{aligned} a_{ij}' =& kla_{ij} + \sum_{i_0 \in I_0, j \in J_0} k_{ii_0} a_{i_0 j_0} l_{j_0 j} \\ &+ l \sum_{i_0 \in I_0} k_{ii_0} a_{i_0 j} + k \sum_{j_0 \in J_0} l_{j_0 j} a_{ij_0}. \end{aligned}$$

This identity implies

$$\begin{aligned} a_{ij} &= \frac{a_{ij}'}{kl} - \sum_{i_0 \in I_0, j_0 \in J_0} \frac{k_{ii_0} l_{j_0 j} a_{i_0, j_0}}{kl} \\ &- \sum_{i_0 \in I_0} \frac{k_{ii_0} a_{i_0 j}}{k} - \sum_{j_0 \in J_0} \frac{l_{j_0 j} a_{ij_0}}{l}. \end{aligned} \tag{11}$$

To complete the proof of Theorem 4.8 we just need to add $a_{i_0 j_0}$ to the set of the generators of Q_i.

To finish the proof of Theorem 4.9, it is enough to show that the proper symmetric GAPs from Theorem 7.8 can be unified.

Lemma 7.9. *Assume that for each $i \in I$, there exist a subset $\mathbf{r}_i' \subset \mathbf{r}_i$ of size $n - n^\epsilon$ and a proper symmetric GAP Q_i of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that $\mathbf{r}_i' \subset Q_i$, and for each $j \in J$, there exist a subset $\mathbf{c}_j' \subset \mathbf{c}_j$ of size $n - n^\epsilon$ and a proper symmetric GAP P_j of rank $O_{C,\epsilon,\mu}(1)$ and size $n^{O_{C,\epsilon,\mu}(1)}$ such that $\mathbf{c}_j' \subset P_j$. Then there exist a bounded number of generators $g_1, \dots, g_s$, where $s = O_{C,\epsilon,\mu}(1)$, such that the set $\{\sum_{h=1}^s (p_h/q_h) g_h, |p_h|, |q_h| = n^{O_{C,\epsilon,\mu}(1)}\}$ contains all but at most ϵn entries of all but at most ϵn rows of A.*

It is clear that Theorem 4.9 follows from Lemma 7.9. It thus remains to verify this lemma.

Proof. (of Lemma 7.9) Throughout the proof, if not specified, all the rows and columns will have index in I and J respectively. We assume that all the proper GAPs has rank at most $r = O_{C,\epsilon}(1)$.

By throwing away at most $\epsilon n/2$ rows, we may assume that for each row $\mathbf{r}_i$ all but at most $n^\epsilon/2\epsilon$ indices j satisfy $\mathbf{r}_i(j) \in \mathbf{c}_j' \subset P_j$. Let $\mathbf{r}_i'$ be the collection of these $\mathbf{r}_i(j)$ for each i.

Set

$$\delta = \epsilon/2r.$$

Consider an arbitrary $\mathbf{r}_i'$. It's components are combinations of the generators of Q_i. Thus we may view these elements of $\mathbf{r}_i'$ as vectors over $\mathbf{Z}^{\mathrm{rank}(Q_i)}$ (see Section 6). We say that the elements of $\mathbf{r}_i'$ are *independent* if their defining vectors are independent.

Next we will choose a subset $\mathbf{r}_i''$ of $\mathbf{r}_i'$ with the following properties.

(1) $|\mathbf{r}_i''| \geq (1-\epsilon)n$.

(2) Let H_i be the subspace generated by the defining vectors of the components of $\mathbf{r}_i''$. Then any hyperplane of H_i contains no more than $(1-\delta)|\mathbf{r}_i''|$ such defining vectors.

We show that there must exist such $\mathbf{r}_i''$.

Assume that $\mathbf{r}_i'$ does not have the above property. By definition of $\mathbf{r}_i'$, this means that (2) is not satisfied. We next pass to consider the set of at least $(1-\delta)|\mathbf{r}_i'|$ components that belong to a proper subspace. Assume that this set does not have the above properties either, we then keep iterating the process. Because the dimensions of the subspaces strictly decrease after each step, the process must terminate after at most r steps. By definition, the subset $\mathbf{r}_i''$ obtained at the time of termination has the desired properties.

Also,

$$|\mathbf{r}_i''| \geq |\mathbf{r}_i'| - r\delta|\mathbf{r}_i'| = (1-\epsilon/2)(n - n^{\epsilon}/2\epsilon) \geq (1-\epsilon)n.$$

Now we will group some generators from the P_j's to create a new set S.

We start with the first column $\mathbf{c}_{j_1}$ and put the generators of P_{j_1} into S. Assume that we already gathered the generators of $P_{j_1}, \dots, P_{j_k}$ after k steps.

To choose a P_j for the next step, we consider the defining vectors of $\mathbf{r}_i''(j_1), \dots \mathbf{r}_i''(j_k)$ for each i. Let $\dim(\mathbf{r}_i''(j_1), \dots \mathbf{r}_i''(j_k))$ denote the dimension of the subspace generated by these vectors.

By the definition of $\mathbf{r}_i''$, if $\mathbf{r}_i''(j_1), \dots \mathbf{r}_i''(j_k)$ do not generate H_i (in which case we say that $\mathbf{r}_i''$ is not *complete*), then there are at least $\delta(1-\epsilon)n \geq \delta n/2$ ways to choose P_j so that $\dim(\mathbf{r}_i''(j_1), \dots \mathbf{r}_i''(j_k), \mathbf{r}_i''(j)) = \dim(\mathbf{r}_i''(j_1), \dots \mathbf{r}_i''(j_k)) + 1$. In this case we say that there is an *increase in dimension* in $\mathbf{r}_i''$.

Hence after some k steps, if there are αn rows that are not complete, then, by the pigeonhole principle, there is a choice for P_j which results in an increase in dimension in at least $\alpha\delta n/2$ rows $\mathbf{r}_i''$.

Because the total of the dimensions is bounded by rn, there must be at least $(1-\epsilon)n$ rows that are complete after at most $2r/(\epsilon\delta) = 2r^2\epsilon^2$ steps. Let S be the collection of all the generators of P_j considered until this step. The size s of S is then at most $2r^3/\epsilon^2$.

Consider a row $\mathbf{r}_i''$ that is complete. Assume that its elements are generated by $\mathbf{r}_i''(j_1), \dots, \mathbf{r}_i''(j_r)$, where $\mathbf{r}_i''(j_k) \in P_{j_k}$, a GAP whose generators belong S. Let a be any element of $\mathbf{r}_i''$, and let $\mathbf{a}$ be its defining vector in Q_i, we then have

$$a = \det\big(\mathbf{a}, \mathbf{r}_i''(j_2), \dots, \mathbf{r}_i''(r)\big) \det\big(\mathbf{r}_i''(j_1), \dots, \mathbf{r}_i''(j_r)\big)^{-1} \cdot \mathbf{r}_i''(j_1) + \dots$$

$$+ \det\big(\mathbf{r}_i''(j_1), \dots, \mathbf{r}_i''(j_{r-1}), \mathbf{a}\big) \det\big(\mathbf{r}_i''(j_1), \dots, \mathbf{r}_i''(j_r)\big)^{-1} \cdot \mathbf{r}_i''(j_r).$$

Thus a can be written in the form $\sum_{h=1}^s (p_h/q_h) \cdot g_h$, where $|p_h|, |q_h| = n^{O_{C,\epsilon}(1)}$.

□

8. Proof of Theorem 5.6 and Theorem 5.7

In this section we will use the results from Section 7 to prove Theorem 5.6 and Theorem 5.7.

Let U be a random subset of $\{1, \dots, n\}$, where $\mathbf{P}(i \in U) = 1/2$ for each i. Let A_U be a submatrix of A defined by

$$A_U(ij) = \begin{cases} a_{ij} & \text{if either } i \in U, j \notin U \text{ or } i \notin U, j \in U, \\ 0 & \text{otherwise.} \end{cases}$$

We first apply the following lemma.

Lemma 8.1 (Concentration for bilinear forms controls concentration for quadratic forms)**.**

$$\rho_q(A)^8 \le \mathbf{P}_{v,w}(\sum_{i,j} A_U(ij) v_i w_j = 0),$$

where v_i, w_j are iid copies of $\eta^{1/2}$.

Proof. (of Lemma 8.1) We first write

$$\mathbf{P}_x(\sum_{i,j} a_{ij} x_i x_j = a) = \mathbf{E}_x \int_0^1 \exp\big(2\pi\sqrt{-1}(\sum_{i,j} a_{ij} x_i x_j - a)t\big) dt.$$

Hence,

$$\mathbf{P}_x(\sum_{i,j} a_{ij} x_i x_j = a) \le \int_0^1 \big|\mathbf{E}_x \exp(2\pi\sqrt{-1}(\sum_{i,j} a_{ij} x_i x_j)t) dt\big|.$$

Next we consider x as $(x_U, x_{\bar U})$, where $x_U, x_{\bar U}$ are the vectors corresponding to $i \in U$ and $i \notin U$ respectively. By the Cauchy-Schwarz inequality

$$\left(\int_0^1 \left|\mathbf{E}_x \exp(2\pi\sqrt{-1}(\sum_{i,j} a_{ij}x_ix_j)t)\right| dt\right)^4 \le \left(\int_0^1 \left|\mathbf{E}_x \exp(2\pi\sqrt{-1}(\sum_{i,j} a_{ij}x_ix_j)t)\right|^2 dt\right)^2$$

$$\le \left(\int_0^1 \mathbf{E}_{x_U}\left|\mathbf{E}_{x_{\bar U}} \exp(2\pi\sqrt{-1}(\sum_{i,j} a_{ij}x_ix_j)t)\right|^2 dt\right)^2$$

$$= \left(\int_0^1 \mathbf{E}_{x_U}\mathbf{E}_{x_{\bar U},x'_{\bar U}} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}x_i(x_j-x'_j) + \sum_{i\in\bar U, j\in\bar U} a_{ij}(x_ix_j - x'_ix'_j))t\Big) dt\right)^2$$

$$\le \int_0^1 \mathbf{E}_{x_{\bar U},x'_{\bar U}}\left|\mathbf{E}_{x_U} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}x_i(x_j-x'_j) + \sum_{i\in\bar U, j\in\bar U} a_{ij}(x_ix_j - x'_ix'_j))t\Big)\right|^2 dt$$

$$= \int_0^1 \mathbf{E}_{x_U,x'_U,x_{\bar U},x'_{\bar U}} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}(x_i-x'_i)(x_j-x'_j))t\Big) dt.$$

$$= \int_0^1 \mathbf{E}_{y_U,z_{\bar U}} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in\bar U, j\in U} a_{ij}y_iz_j)t\Big) dt,$$

where $y_U = x_U - x'_U$ and $z_{\bar U} = x_{\bar U} - x'_{\bar U}$, whose entries are iid copies of $\eta^{1/2}$.

Thus we have

$$\left(\int_0^1 \left|\mathbf{E}_x \exp(2\pi\sqrt{-1}(\sum_{i,j} a_{ij}x_ix_j)t)\right| dt\right)^8 \le \left(\int_0^1 \mathbf{E}_{y_U,z_{\bar U}} \exp\big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}y_iz_j)t\big) dt\right)^2$$

$$\le \int_0^1 \mathbf{E}_{y_U,z_{\bar U},y'_U,z'_{\bar U}} \exp\big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}y_iz_j - \sum_{i\in U, j\in\bar U} a_{ij}y'_iz'_j)t\big) dt.$$

Because $a_{ij} = a_{ji}$, we can write the last term as

$$\int_0^1 \mathbf{E}_{y_U,z'_{\bar U},y'_U,z_{\bar U}} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}y_iz_j + \sum_{j\in\bar U, i\in U} a_{ji}(-z'_j)y'_i)t\Big) dt$$

$$= \int_0^1 \mathbf{E}_{v,w} \exp\big(2\pi\sqrt{-1}(\sum_{i\in U, j\in\bar U} a_{ij}v_iw_j + \sum_{i\in\bar U, j\in U} a_{ij}v_iw_j)t\big) dt,$$

where $v := (y_U, -z'_{\bar U})$ and $w := (y'_U, z_{\bar U})$.

To conclude the proof we observe that the entries of v and w are iid copies of $\eta^{1/2}$, and

$$\int_0^1 \mathbf{E}_{v,w} \exp\Big(2\pi\sqrt{-1}(\sum_{i\in U j\in \bar{U}} a_{ij}v_iw_j + \sum_{i\in \bar{U}, j\in U} a_{ij}v_iw_j)t\Big)dt = \mathbf{P}_{v,w}(\sum_{i,j} A_U(ij)v_iw_j = 0).$$

□

Next, it follows from Lemma 8.1 that

$$\mathbf{P}_{v,w}(\sum_{i,j} A_U(ij)v_iw_j = 0) \ge n^{-8C}.$$

This inequality means that $\rho_q^{(1/2)}(A_U) \ge n^{-8C}$. We now apply Lemma 7.2.

Lemma 8.2. *There exist a set $I_0(U)$ of size $O_{C,\epsilon}(1)$ and a set $I(U)$ of size at least $n-n^\epsilon$ such that for any $i\in I$, there are integers $0\neq k(U)$ and $k_{ii_0}(U), i_0\in I_0(U)$, all bounded by $n^{O_{C,\epsilon}(1)}$, such that*

$$\mathbf{P}_y\big(\langle k(U)\mathbf{r}_{A_U}(i), y\rangle + \langle \sum_{i_0\in I_0} k_{ii_0}(U)\mathbf{r}_{A_U}(i_0), y\rangle = 0\big) = n^{-O_{C,\epsilon}(1)}.$$

Note that Lemma 8.2 holds for all U. We will try to obtain a similar conclusion for A.

As $I_0(U)\subset [n]^{O_{C,\epsilon}(1)}$ and $k(U)\le n$, there are only $n^{O_{C,\epsilon}(1)}$ possibilities that $(I_0(U), k(U))$ can take. Thus there exists a tuple (I_0, k) such that $I_0(U) = I_0$ and $k(U) = k$ for $2^n/n^{O_{C,\epsilon}(1)}$ different U. Let us denote this set of U by $\mathcal{U}$. Thus

$$|\mathcal{U}| \ge 2^n/n^{O_{C,\epsilon}(1)}.$$

Next, let I be the collection of i which belong to at least $|\mathcal{U}|/2$ index sets I_U. Then we have

$$|I||\mathcal{U}| + (n-|I|)|\mathcal{U}|/2 \ge (n-n^\epsilon)|\mathcal{U}|$$
$$|I| \ge n - 2n^\epsilon.$$

Fix an $i\in I$. Consider the tuples $(k_{ii_0}(U), i_0\in I_0)$ where $i\in I_U$. Because there are only $n^{O_{C,\epsilon}(1)}$ possibilities such tuples can take, there must be a tuple, say $(k_{ii_0}, i_0\in I_0)$, such that $(k_{ii_0}(U), i_0\in I_0) = (k_{ii_0}, i_0\in I_0)$ for at least $|\mathcal{U}|/2n^{O_{C,\epsilon}(1)} = 2^n/n^{O_{C,\epsilon}(1)}$ sets U.

Because $|I_0| = O_{C,\epsilon}(1)$, it is easy to see that there is a way to partition I_0 into $I_0'\cup I_0''$ such that there are $2^n/n^{O_{C,\epsilon}(1)}$ sets U above satisfying that $I_0''\subset U$ and $U\cap I_0' = \emptyset$. Let $\mathcal{U}_{I_0', I_0''}$ denote the collection of these U.

By passing to consider a subset of $\mathcal{U}_{I_0',I_0''}$ if needed, we may assume that either $i \notin U$ or $i \in U$ for all $U \in \mathcal{U}_{I_0',I_0''}$. Without loss of generality, we assume the first case that $i \notin U$. (The other case can be treated similarly).

Let $U \in \mathcal{U}_{I_0',I_0''}$ and $u = (u_1, \dots, u_n)$ be its characteristic vector, that is $u_j = 1$ if $j \in U$, and $u_j = 0$ otherwise. Then, by the definition of A_U, and because $I_0'' \subset U$ and $I_0' \cap U = \emptyset$, for $i_0' \in I_0'$ and $i_0'' \in I_0''$ we can respectively write

$$\langle \mathbf{r}_{i_0'}(A_U), y\rangle = \sum_{j=1}^n a_{i_0'j} u_j y_j, \text{ and } \langle \mathbf{r}_{i_0''}(A_U), y\rangle = \sum_{j=1}^n a_{i_0''j}(1-u_j) y_j.$$

Also, because $i \notin U$, we have

$$\langle \mathbf{r}_i(A_U), y\rangle = \sum_{j=1}^n a_{ij} u_j y_j.$$

Thus,

$$\begin{aligned}
&\langle k\mathbf{r}_i(A_U), y\rangle + \sum_{i_0 \in I_0} \langle k_{ii_0}\mathbf{r}_{i_0}(A_U), y\rangle \\
&= \langle k\mathbf{r}_i(A_U), y\rangle + \langle \sum_{i_0' \in I_0'} k_{ii_0'}\mathbf{r}_{i_0'}(A_U), y\rangle + \langle \sum_{i_0'' \in I_0''} k_{ii_0''}\mathbf{r}_{i_0''}(A_U), y\rangle \\
&= \sum_{j=1}^n k a_{ij} u_j y_j + \sum_{j=1}^n \sum_{i_0' \in I_0'} k_{ii_0'} a_{i_0'j} u_j y_j + \sum_{j=1}^n \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0''j}(1-u_j) y_j \\
&= \sum_{j=1}^n (k a_{ij} + \sum_{i_0' \in I_0'} k_{ii_0'} a_{i_0'j} - \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0''j}) u_j y_j + \sum_{j=1}^n \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0''j} y_j
\end{aligned}$$

Next, by Lemma 8.2, for each $U \in \mathcal{U}_{I_0',I_0''}$ we have

$$\mathbf{P}_y\big(\langle k\mathbf{r}_i(A_U), y\rangle + \sum_{i_0 \in I_0} \langle k_{ii_0}\mathbf{r}_{i_0}(A_U), y\rangle = 0\big) = n^{-O_{C,\epsilon}(1)}.$$

Also, note that

$$|\mathcal{U}_{I_0',I_0''}| = 2^n / n^{O_{C,\epsilon}(1)}.$$

Hence,

$$\mathbf{E}_y\mathbf{E}_U\big(k\langle \mathbf{r}_i(A_U),y\rangle+\sum_{i_0\in I_0}\langle k_{ii_0}\mathbf{r}_{i_0}(A_U),y\rangle=0\big)\ge n^{-O_{C,\epsilon}(1)}.$$

By applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}
n^{-O_{C,\epsilon}(1)} &\le \left(\mathbf{E}_y\mathbf{E}_U(k\langle \mathbf{r}_i(A_U),y\rangle+\sum_{i_0\in I_0}\langle k_{ii_0}\mathbf{r}_{i_0}(A_U),y\rangle=0)\right)^2\\
&\le \mathbf{E}_y\left(\mathbf{E}_U(k\langle \mathbf{r}_i(A_U),y\rangle+\sum_{i_0\in I_0}\langle k_{ii_0}\mathbf{r}_{i_0}(A_U),y\rangle=0\right)^2\\
&= \mathbf{E}_y\left(\mathbf{E}_u(\sum_{j=1}^n(ka_{ij}+\sum_{i_0'\in I_0'}k_{ii_0'}a_{i_0'j}-\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j})u_jy_j+\sum_{j=1}^n\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j}y_j=0)\right)^2\\
&\le \mathbf{E}_y\mathbf{E}_{u,u'}\big(\sum_{j=1}^n(ka_{ij}+\sum_{i_0'\in I_0'}k_{ii_0'}a_{i_0'j}-\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j})(u_j-u_j')y_j=0\big)\\
&= \mathbf{E}_z\big(\sum_{j=1}^n(ka_{ij}+\sum_{i_0'\in I_0'}k_{ii_0'}a_{i_0'j}-\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j})z_j=0\big)
\end{aligned}$$

where $z_j:=(u_j-u_j')y_j$, and in the last inequality we used the simple observation that $\mathbf{E}_{u,u'}(f(u)=0,f(u')=0)\le\mathbf{E}_{u,u'}(f(u)-f(u')=0)$.

Note that u_j-u_j' and y_j are iid copies of $\eta^{1/2}$. Hence z_j are iid copies of $\eta^{1/4}$.

Finally, by Theorem 3.5, the bound

$$n^{-O_{C,\epsilon}(1)}\le\mathbf{E}_z\big(\sum_{j=1}^n(ka_{ij}+\sum_{i_0'\in I_0'}k_{ii_0'}a_{i_0'j}-\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j})z_j=0\big)$$

implies that there exists a proper symmetric GAP Q_i of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ such that the following holds for all but at most n' elements of j

$$ka_{ij}+\sum_{i_0'\in I_0'}k_{ii_0'}a_{i_0'j}-\sum_{i_0''\in I_0''}k_{ii_0''}a_{i_0''j}\in Q_i.$$

We summarize below.

Theorem 8.3 (Refined row relation)**.** *Let $\epsilon<1$ and C be positive constants. Assume that $\rho_q(A)\ge n^{-C}$. Then there exist a set I_0 of size $O_{C,\epsilon}(1)$, a set I of size at least $n-2n^\epsilon$, a number $0\neq k=n^{O_{C,\epsilon}(1)}$ such that for any $i\in I$ there are integers $k_{ii_0},i_0\in I_0$, all bounded* *by $n^{O_{C,\epsilon}(1)}$, an index set J_i of size $n-n^\epsilon$, and a proper symmetric GAP Q_i of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ such that the following holds for all $j \in J_i$*

$$ka_{ij} + \sum_{i_0 \in I_0} k_{ii_0} a_{i_0 j} \in Q_i.$$

Clearly, we may assume that $I \cap I_0 = \emptyset$ by throwing away those i from I that also belong to I_0.

Let R be the matrix defined below.

Definition 8.4 (row matrix)**.** R is an n by n matrix, whose i-th row, where $i \in I$, is defined by

$$\mathbf{r}_i(R)(j) := \begin{cases} k_{ij}, & \text{if } j \in I_0; \\ k, & \text{if } j = i; \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The other entries of R are zero except the diagonal terms which are set to be 1.

We restate Theorem 8.3 in a more convenient way below.

Theorem 8.5 (Refined row relation, again)**.** *Let $\epsilon \le 1$ and C be positive constants. Assume that $\rho_q(A) \ge n^{-C}$. Then there exist a set I_0 of size $O_{C,\epsilon}(1)$, a set I of size at least $n - 2n^\epsilon$ satisfying $I \cap I_0 = \emptyset$, integers $0 \ne k, k_{ii_0}, i_0 \in I_0, i \in I$, all bounded by $n^{O_{C,\epsilon}(1)}$, and a matrix R defined by (12) such that the matrix $A' = RA$ possess the following properties: for each $i \in I$, there exist a subset $\mathbf{r}_i' \subset \mathbf{r}_i(A')$ of size $n - n^\epsilon$ and a proper symmetric GAP Q_i of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ such that $\mathbf{r}_i' \subset Q_i$.*

Next, because A is symmetric, we obtain a similar relation between the columns of A. Hence, we obtain the following key result.

Theorem 8.6 (Matrix relations)**.** *Let $\epsilon \le 1$ and C be positive constants. Assume that $\rho_q(A) \ge n^{-C}$. Then there exist a set I_0 of size $O_{C,\epsilon}(1)$, a set I of size at least $n - 2n^\epsilon$ satisfying $I \cap I_0 = \emptyset$, integers $0 \ne k, k_{ii_0}, i_0 \in I_0, i \in I$, all bounded by $n^{O_{C,\epsilon}(1)}$, and a matrix R defined by (12) such that the matrix $A' = RAR^T$ possess the following properties.*

- *For each $i \in I$, there exist a subset $\mathbf{r}_i' \subset \mathbf{r}_i(A')$ of size $n-n^\epsilon$ and a proper symmetric GAP Q_i of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ such that $\mathbf{r}_i' \subset Q_i$.*
- *For each $j \in I$, there exist a subset $\mathbf{c}_j' \subset \mathbf{c}_j(A')$ of size $n-n^\epsilon$ and a proper symmetric GAP P_j of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ such that $\mathbf{c}_j' \subset P_j$.*

We complete the proof of Theorem 5.6 by using (11), and Theorem 5.7 by using Lemma 7.9, noting that $k_{ij} = k_{ji}$ and $a_{ii_0} = a_{i_0 i}$.

Remark 8.7. In later application we will not need the whole strength of Theorem 8.6. It will suffice to apply Theorem 8.5.

9. PROOF OF LEMMA 2.6

We now prove Lemma 2.6 by using our inverse Littlewood-Offord result for linear forms presented in Section 3.

First of all, because $\mathrm{rank}(M_{n-1}) = n-2$, the cofactor matrix (a_{ij}) of M_{n-1} has rank 1. Because this matrix is symmetric, each entry a_{ij} must have the form $a_i a_j$, where not all the a_i are zeros.

We will show that the vector $u = (a_1, \dots, a_{n-1})$ satisfies the conclusions of Lemma 2.6.

Observe that

$$\det(M_n) = \sum_{1\le i,j\le n-1} a_{ij}x_i x_j = (\sum_{i=1}^{n-1} a_i x_i)^2.$$

Thus the assumption $\mathbf{P}(\det(M_n) = 0|M_{n-1}) \ge n^{-C}$ implies that

$$\mathbf{P}(\sum_{i=1}^{n-1} a_i x_i = 0|M_{n-1}) \ge n^{-C}.$$

By Theorem 3.4, all but n^ϵ elements of a_i belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$. Also, by the definition of the a_i, $u = (a_1, \dots, a_{n-1})$ is orthogonal to $n-2$ linearly independent rows of M_{n-1}. We finish the proof of Lemma 2.6 by using the following lemma.

Lemma 9.1 (Rational commensurability)**.** *Let $v = (v_1, \dots, v_{n-1})$ be a vector such that all but n^ϵ components v_i belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$, and that v is a normal vector of a hyperplane spanned by vectors of integral components bounded by $n^{O_{C,\epsilon}(1)}$. Then $\{v_1, \dots, v_{n-1}\} \subset \{(p/q)v_{i_0}, |p|, |q| = n^{O_{C,\epsilon}(n^\epsilon)}\}$ for some i_0.*

Proof. (of Lemma 9.1) Without loss of generality, we assume that $(v_{n-n^\epsilon}, \dots, v_{n-1})$ are the exceptional elements that may not belong to the GAP.

For each v_i, where $i < n - n^\epsilon$, there exist numbers v_{ij}, all bounded by $n^{O_{C,\epsilon}(1)}$, such that

$$v_i = v_{i1}g_1 + \dots v_{ir}g_r,$$

where $g_1, \dots, g_r$ are the generators of the GAP.

Note that by Theorem 6.1, one may assume that the vectors $(v_{i1}, \dots, v_{ir})$, where $i < n-n^\epsilon$, generate the whole space $\mathbf{R}^r$.

Consider the $n-1$ by $r+n^\epsilon$ matrix M_v whose i-th row is the vector $(v_{i1}, \dots, v_{ir}, 0, \dots, 0)$ if $i < n-n^\epsilon$, and $(0, \dots, 0, 1, 0, \dots, 0)$ if $n-n^\epsilon \le i$. Note that M_v has rank $r+n^\epsilon$.

We thus have

$$v^T = M_v \cdot u^T,$$

where $u = (g_1, \dots, g_r, v_{n-n^\epsilon}, \dots, v_{n-1})$.

Next, let $w_1, \dots, w_{n-2}$ be the vectors of integral entries bounded by $n^{O_{C,\epsilon}(1)}$ which are orthogonal to v. We form an $n-1$ by $n-1$ matrix M_w whose i-th row is w_i for $i \le n-2$, and the $n-1$-th row is e_{i_0}, a unit vector among the standard basis $\{e_1, \dots, e_{n-1}\}$ that is linearly independent to $w_1, \dots .w_{n-2}$.

By definition, we have $M_w v^T = (0, \dots, 0, v_{i_0})^T$, and hence

$$(M_w M_v) u^T = (0, \dots, 0, v_{i_0})^T.$$

The identity above implies that

$$(M_w M_v)(\frac{1}{v_{i_0}} u)^T = (0, \dots, 0, 1)^T. \tag{13}$$

Next we choose a submatrix M of size $r + n^\epsilon$ by $r + n^\epsilon$ of $M_w M_v$ thas has full rank. Then

$$M(\frac{1}{v_{i_0}} u)^T = x \tag{14}$$

for some x which is a subvector of $(0, \dots, 0, 1)$ from (13).

Observe that the entries of M are integers bounded by $n^{O_{C,\epsilon}(1)}$. Hence, the entries of M^{-1} are fractions whose numerators and denominators are integers bounded by $(n^{O_{C,\epsilon}(1)})^{r+n^\epsilon} = n^{O_{C,\epsilon}(n^\epsilon)}$.

Solving for g_i/v_{i_0} and v_j/v_{i_0} from (14), we conclude that each of these components can be written in the form p/q, where $|p|, |q| = n^{O_{C,\epsilon}(n^\epsilon)}$.

□

Remark 9.2. In principle, Lemma 9.1 is similar to Theorem 5.2 of [24].

10. Proof of Lemma 2.7

In this section we will apply the results from Section 5 and Section 8 to prove Lemma 2.7.

First, assume that

$$\mathbf{P}_x(\sum_{ij} a_{ij}x_ix_j = 0|M_{n-1}) \geq n^{-C}.$$

Let A be the matrix (a_{ij}). Then by Theorem 5.6 (or more explicitly, Theorem 8.5), there exists a non-singular matrix R (see Definition 8.4) such that most of the entries of each row of A' belong to proper symmetric GAPs of small ranks and small sizes, where $A' = RA$.

Set

$$M := M_{n-1}R^{-1}.$$

Because $M_{n-1}A = \det(M_{n-1}) \cdot I_{n-1} \neq 0$, we have

$$MA' = M_{n-1}A = \det(M_{n-1}) \cdot I_{n-1} \neq \mathbf{O}. \tag{15}$$

Next, it follows from the definition of R that

$$R^{-1}(ij) = \begin{cases} 1/k & \text{if } i \in I \text{ and } j = i; \\ -k_{ij}/k & \text{if } i \in I \text{ and } j \in I_0; \\ 1 & \text{if } i \notin I \text{ and } j = i; \\ 0 & \text{if } i \notin I \text{ and } j \neq i. \end{cases}$$

We thus have, by $M(ij) = \sum_{j'} M_{n-1}(ij')(R^{-1})(j'j)$, that

$$\begin{aligned} M(ij_0) &= \sum_{j' \in I} M_{n-1}(ij')(-k_{j'j_0}/k) + M_{n-1}(ij_0), \text{ if } j_0 \in I_0; \\ M(ij) &= M_{n-1}(ij)/k, \text{ if } j \in I; \\ M(ij) &= M_{n-1}(ij), \text{ if } j \notin I_0 \cup I. \end{aligned} \tag{16}$$

Because the entries of M_{n-1} are ± 1, and $k_{ii_0} = n^{O_{C,\epsilon}(1)}$, the entries of M are rational numbers of the form m/k, where $m \in \mathbf{Z}$ and $m = n^{O_{C,\epsilon}(1)}$. Furthermore, note that M also has full rank.

Let v be any column of A' whose all but n^ϵ components belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.

Because $MA' = \det(M_{n-1}) \cdot I_n \neq \mathbf{0}$, v is not a zero vector which is orthogonal to $n-2$ rows of M. Hence, it follows from Lemma 9.1 that $\{v_1, \dots, v_{n-1}\} \subset \{(p/q)v_{i_0}, |p|, |q| = n^{O_{C,\epsilon}(n^\epsilon)}\}$ for some i_0.

Next, consider a row $\mathbf{r}_i(M)$ that is orthogonal to v, where $i \in I$. Note that there are at least $|I| - 1 \geq n - 2n^{\epsilon} - 1$ such indices i. We have

$$\begin{aligned}
& \sum_j M(ij)v_j = \sum_{j_0 \in I_0} M(ij_0)v_{j_0} + \sum_{j \in I} M(ij)v_j + \sum_{j \notin I_0 \cup I} M(ij)v_j \\
= \; & -\sum_{j_0 \in I_0} \sum_{j' \in I} M_{n-1}(ij')k_{j'j_0}v_{j_0}/k + \sum_{j_0 \in I_0} M_{n-1}(ij_0)v_{j_0} + \sum_{j \in I} M_{n-1}(ij)v_j/k + \sum_{j \notin I_0 \cup I} M_{n-1}(ij)v_j \\
= \; & \sum_{j' \in I} M_{n-1}(ij')(-\sum_{j_0 \in I_0} k_{j'j_0}v_{j_0}/k) + \sum_{j_0 \in I_0} M_{n-1}(ij_0)v_{j_0} + \sum_{j \in I} M_{n-1}(ij)v_j/k + \sum_{j_0 \notin I_0 \cup I} M_{n-1}(ij)v_j \\
= \; & \sum_{j \in I} M_{n-1}(ij)(v_j/k - \sum_{j_0 \in I_0} k_{jj_0}v_{j_0}/k) + \sum_{j_0 \in I_0} M_{n-1}(ij_0)v_{j_0} + \sum_{j \notin I_0 \cup I} M_{n-1}(ij)v_j \\
= \; & 0. \qquad (17)
\end{aligned}$$

Define

$$u_j := \begin{cases} v_j & \text{if } j \notin I; \\ v_j/k - \sum_{j_0 \in I_0} k_{jj_0}v_{j_0}/k & \text{if } j \in I. \end{cases}$$

It then follows from (17) that

$$\sum_j M_{n-1}(ij)u_j = 0.$$

Thus, the vector $u = (u_1, \dots, u_{n-1})$ is orthogonal to $\mathbf{r}_i(M_{n-1})$. This holds for at least $n - 2n^{\epsilon} - 1$ rows of M_{n-1}.

Additionally, by the definition of u and v, all but n^{ϵ} coordinates of u belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$ (with probably worse parameters), and $\{u_1, \dots, u_{n-1}\} \subset \{(p/q)u_{j_0}, |p|, |q| = n^{O_{C,\epsilon}(n^{\epsilon})}\}$ for some j_0.

We conclude the proof by noting that, because v is not a zero vector, u is not either.

11. proof of Theorem 2.4 and Theorem 2.5

Assume that M_{n-1} has rank $n-2$ or $n-1$, and $\mathbf{P}(\det(M_n) = 0 | M_{n-1}) \geq n^{-C}$. We apply Lemma 2.6 and 2.7 to obtain a vector $u = (u_1, \dots, u_{n-1})$ of the following properties.

(1) All but n^{ϵ} elements of u_i belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.

(2) $u_i \in \{p/q : |p|, |q| = n^{O_{C,\epsilon}(n^{\epsilon})}\}$ for all i.

(3) u is orthogonal to $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1}.

Let $\mathcal{P}$ denote the collection of all u having property (1) and (2). For each $u \in \mathcal{P}$, let $\mathbf{P}_u$ be the probability, with respect to M_{n-1}, that u is orthogonal to $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1}. We shall prove the following key result.

Theorem 11.1. *We have*

$$\sum_{u\in\mathcal{P}} \mathbf{P}_u = O_{C,\epsilon}((1/2)^{(1-o(1))n}).$$

It is clear that Theorem 2.4 and Theorem 2.5 follow from Theorem 11.1.

In the sequel we will choose $0 < \delta$ to be small enough so that $\delta \cdot O_{C,\epsilon}(1) \le \epsilon/4$ for all constants $O_{C,\epsilon}(1)$ appearing in Lemma 2.6 and Lemma 2.7.

Let n_u denote the number on nonzero components of u. To prove Theorem 11.1 we decompose the sum $\sum_{u\in\mathcal{P}} \mathbf{P}_u$ into two parts depending on the magnitude of n_u.

Theorem 11.2. *The probability of a random symmetric matrix M_{n-1} having $n - O_{C,\epsilon}(n^\epsilon)$ rows being orthogonal to a vector $u \in \mathcal{P}$ having $n_u \le n^{1-\delta}$ is bounded by*

$$\sum_{u\in\mathcal{P}, n_u \le n^{1-\delta}} \mathbf{P}_u = O\left((1/2)^{(1-o(1))n}\right),$$

where the implied constants depend on C, ϵ and δ.

Theorem 11.3. *The probability of a random symmetric matrix M_{n-1} having $n - O_{C,\epsilon}(n^\epsilon)$ rows being orthogonal to a vector $u \in \mathcal{P}$ having $n_u \ge n^{1-\delta}$ is bounded by*

$$\sum_{u\in\mathcal{P}, n_u \ge n^{1-\delta}} \mathbf{P}_u = O(n^{-n^{1-\delta}/32}),$$

where the implied constants depend on C, ϵ and δ.

Proof. (of Theorem 11.2) By paying a factor $\binom{n-1}{O_{C,\epsilon}(n^\epsilon)} = O(n^{O_{C,\epsilon}(n^\epsilon)})$ in probability, we may assume that the first $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1} are orthogonal to u.

Also, by paying a factor $\binom{n}{n_u}$ in probability, we may assume that the first n_u components of u are nonzero. Thus we have

$$\sum_{i=1}^{n_u} u_i \mathbf{r}_i(M_{n-1}) = 0.$$

Which in turn implies that $\mathbf{r}_{n_u}(M_{n-1})$ lies in the subspace spanned by $\mathbf{r}_1(M_{n-1}), \dots, \mathbf{r}_{n_u-1}(M_{n-1})$.

Next, due to symmetry, $\mathbf{r}_{n_u}(M_{n-1})$ has $n_u - 1$ components that were already exposed in the first $n_u - 1$ rows (if we work with the general case that the rows in consideration are not necessarily the first n_u rows of M_{n-1}, then there are less dependencies: at most $n_u - 1$ components already exposed in the previous $n_u - 1$ rows.)

Let $\mathbf{r}'_{n_u}$ be the subvector obtained from $\mathbf{r}_{n_u}$ by removing the exposed components, and for each $1 \le i \le n_u - 1$ we let $\mathbf{r}'_i$ be the subvectors of $\mathbf{r}_i(M_{n-1})$ corresponding to the columns restricted by $\mathbf{r}'_{n_u}$.

By definition, each $\mathbf{r}'_i$ has $n - n_u$ components, and because $\mathbf{r}_{n_u}(M_{n-1})$ lies in the subspace spanned by $\mathbf{r}_1(M_{n-1}), \dots, \mathbf{r}_{n_u-1}(M_{n-1})$, so does $\mathbf{r}'_{n_u}$ in the subspace spanned by $\mathbf{r}'_1, \dots, \mathbf{r}'_{n_u-1}$. The probability for this event, by Lemma 2.2, is at most

$$2^{n_u-1-(n-n_u)} = 2^{2n_u-n-1}.$$

Thus we have

$$\sum_{u \in \mathcal{P}, n_u \le n^{1-\delta}} \mathbf{P}_u \le \sum_{n_u=1}^{n^{1-\delta}} n^{O_{C,\epsilon}(n^\epsilon)} \binom{n}{n_u} 2^{2n_u-n} = O\left((1/2)^{(1-o(1))n}\right),$$

where the implied constants depend on C, ϵ and δ.

□

Remark 11.4. In the proof of Theorem 11.2, because the assumption that u has many zero components is strong, we do not need the additional additive structure on the remaining components of u.

We next focus on the estimate for the minor term.

Proof. (of Theorem 11.3) By paying a factor of $\binom{n-1}{n_u}\binom{n_u}{n^\epsilon}$ in probability and without loss of generality, we may assume that u has the following properties:

- the first n_u components of u are nonzero;
- the first $n_0 := n_u - n^\epsilon$ components of u are non-exceptional (that is they all belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$.)

Because u is orthogonal to $n - O_{C,\epsilon}(n^\epsilon)$ rows of M_{n-1}, it is orthogonal to $n_1 := n_0 - O_{C,\epsilon}(n^\epsilon)$ rows among the first n_0 rows of M_{n-1}. By paying a factor of $\binom{n_0}{O_{C,\epsilon}(n^\epsilon)} = O(n_u^{O_{C,\epsilon}(n^\epsilon)})$ in probability, we may assume that these are the first n_1 rows of M_{n-1}. (One proceed similarly in the general case, occasionally with better bounds due to more independence among the entries.)

We will expose the first n_1 rows of M_{n-1} one by one. Let i be an index between 1 and n_1. Condition on the first $i-1$ rows of M_{n-1}, the probability that $\mathbf{r}_i(M_{n-1})$ is orthogonal to u is controlled by

$$\mathbf{P}_{x_i,\dots,x_{n_u}\in\{-1,1\}}(\sum_{j=i}^{n_u} x_j u_j = -\sum_{j=1}^{i-1} x_j u_j) \leq$$

$$\leq \rho_i(u) := \sup_{a\in\mathbf{R}} \mathbf{P}_{x_i,\dots,x_{n_0}\in\{-1,1\}}(\sum_{j=i}^{n_0} x_j u_j = a).$$

Observe that $\rho_1(u) \leq \cdots \leq \rho_{n_1}(u)$. With room to spare, we concentrate on $\rho_i(u)$ where $i \leq (1-\delta)n_0$ only.

Note that $(1-\delta)n_0 < n_1$, thus the probability that the first n_1 rows of M_{n-1} are orthogonal to u is bounded by

$$\mathbf{P}\left(\mathbf{r}_i(M_{n-1}), 1 \leq i \leq (1-\delta)n_0, \text{ are orthogonal to } u\right) \leq \prod_{i=1}^{(1-\delta)n_0} \rho_i. \quad (18)$$

Note that the nonzero $u_j, j = 1,\dots,n_0$, all belong to a proper symmetric GAP of rank $O_{C,\epsilon}(1)$ and size $n^{O_{C,\epsilon}(1)}$. It thus follows from the Erdős-Littlewood-Offord inequality (2) and Example 3.3) that for any $1 \leq i \leq (1-\delta)n_0$

$$n^{-O_{C,\epsilon}(1)} \leq \rho_i(u) = O((\delta n_0)^{-1/2}) = O((\delta n_u)^{-1/2}). \quad (19)$$

Next we fix a sequence $b_0, b_1, \dots, b_K$, where $b_0 = n^{-O_{C,\epsilon}(1)}$ is the left bound of (19) and $b_{i+1} := n^\delta b_i$, as well as K is the smallest index such that b_K exceeds the right bound of (19) (thus $K \leq O_{C,\epsilon}(1)\delta^{-1}$).

By the definition of the sequence b_i, for any $1 \leq i \leq (1-\delta)n_0$ we have

$$b_0 \leq \rho_i(u) \leq b_K.$$

In the next step, we classify u depending on how fast the concentration probabilities $\rho_i(u)$ grow.

Definition 11.5 (concentration sequence)**.** We say that a $u \in \mathcal{P}$ satisfying $n_u \geq n^{1-\delta}$ has concentration sequence $(m_1,\dots,m_K)$, where $m_1+\cdots+m_K = (1-\delta)n_0$, if there are exactly m_j terms $\rho_i(u)$ belonging to $[b_{j-1}, b_j)$.

Observe that the smaller δ we choose, the more detail we know about the distribution of $\rho_i(u)$.

Basing on concentration sequences, we say that $u \in \mathcal{P}$ belongs to $\mathcal{P}_{(m_1,\dots,m_K)}$ if its concentration sequence is $(m_1,\dots,m_K)$.

Our next lemma is to show that there is a collection of structures that contains all the elements of $\mathcal{P}_{(m_1,\dots,m_K)}$. This result will then enable us to compute $\mathbf{P}_u$ in a convenient way.

Theorem 11.6. *Assume that $u \in \mathcal{P}_{(m_1,\dots,m_K)}$. Then there exists a sequence of proper symmetric GAPs $Q_0, Q_1, \dots, Q_K$ such that*

(1) *$u_i \in Q_0$ for all $1 \le i \le n_0$;*

(2) *$u_j \in Q_i$ for all but n^ϵ/K indices j with $m_1 + \dots + m_{i-1} \le j < m_1 + \dots + m_i$;*

(3) *$|Q_i| \le c b_i^{-1} n^\delta/(n^\epsilon)^{1/2}$, where c is a constant depending only on C, ϵ and δ;*

(4) *all the generators of Q_i belong to the set $\{p/q, |p|, |q| \le n^{O_{C,\epsilon,\delta}(n^\epsilon)}\}$.*

Theorem 11.6 can be shown by applying Theorem 3.4 several times. To begin with, we set Q_0 to be the proper symmetric GAP that contains all the non-exceptional u.

Next, as

$$\rho_{m_1+\dots+m_{i-1}} \ge b_{i-1} = O(n^{-O(1)})$$

Theorem 3.4 implies that all but at most n^ϵ/K components u_j, where $m_1 + \dots + m_{i-1} \le j \le (1-\delta)n_0$, belong to a proper symmetric GAP Q_i of size

$$\begin{aligned} O_{C,\epsilon,\delta}(\rho^{-1}_{m_1+\dots+m_{i-1}}/(n^\epsilon)^{1/2}) &= O_{C,\epsilon,\delta}(b^{-1}_{i-1} n^\delta/(n^\epsilon)^{1/2}) \\ &= O_{C,\epsilon,\delta}(b_i^{-1} n^\delta/(n^\epsilon)^{1/2}). \end{aligned}$$

We keep this information only for those u_j where $m_1 + \dots + m_{i-1} \le j < m_1 + \dots + m_i$, and release other u_j for the next application of Theorem 3.4. By Theorem 6.1, we may assume that these u_j span Q_i, and thus (4) holds because $u_j \in \{p/q, |p|, |q| = n^{O_{C,\epsilon,\delta}(n^\epsilon)}\}$.

Now for each $u \in \mathcal{P}_{(m_1,\dots,m_K)}$, we reconsider the probability that the first n_1 rows of M_{n-1} are orthogonal to u. As shown in (18), this probability is bounded by $\prod_i \rho_i$. By definition of concentration sequence, we have

$$\prod_{i=1}^{(1-\delta)n_0} \rho_i \le \prod_{i=1}^{K} b_i^{m_i}. \tag{20}$$

In the next sequel we want to sum this bound over $u \in \mathcal{P}_{(m_1,\dots,m_K)}$.

Because each Q_i is determined by its $O_{C,\epsilon,\delta}(1)$ generators from the set $\{p/q, |p|, |q| \le n^{O_{C,\epsilon,\delta}(n^\epsilon)}\}$, and its dimensions from the integers bounded by $n^{O_{C,\epsilon,\delta}(1)}$, there are $n^{O_{C,\epsilon,\delta}(n^\epsilon)}$ ways to choose each Q_i. So the total number of ways to choose $Q_0, \dots, Q_K$ is

$$(n^{O_{C,\epsilon,\delta}(n^\epsilon)})^K = n^{O_{C,\epsilon,\delta}(n^\epsilon)}. \tag{21}$$

Next, after locating Q_i, the total number of ways to choose is

$$\prod_{i=1}^K \binom{m_i}{n^\epsilon/K} |Q_i|^{m_i - n^\epsilon/K} \le 2^{m_1+\dots+m_K} \prod_{i=1}^K |Q_i|^{m_i} = 2^{(1-\delta)n_0} \prod_{i=1}^K |Q_i|^{m_i},$$

where $\binom{m_i}{n^\epsilon/k}|Q_i|^{m_i-n^\epsilon/K}$ is the number of ways to choose u_j from each Q_i, following (2) of Theorem 11.6.

We then continue to estimate

$$\begin{aligned} 2^{(1-\delta)n_0} \prod_{i=1}^K |Q_i|^{m_i} &\le (2c)^{(1-\delta)n_0} \prod_{i=1}^K (b_i^{-1} n^\delta/(n^\epsilon)^{1/2})^{m_i} \\ &= (2c)^{(1-\delta)n_0} \prod_{i=1}^K b_i^{-m_i} n^{\delta(1-\delta)n_0} n^{-\epsilon(1-\delta)n_0/2} \\ &= O(\prod_{i=1}^K b_i^{-m_i} n^{-\epsilon n_0/4}), \end{aligned} \tag{22}$$

where in the last estimate we use the fact that $\delta \le \epsilon/16$.

For the remaining non-exceptional u_i, where $(1-\delta)n_0 \le i \le n_0$ or $u_j \notin Q_i$ from (2) of Theorem 11.6, we choose them from Q_0, which results in the bound

$$b_0^{\delta n_0 + n^\epsilon} = n^{\delta O_{C,\epsilon}(1) n_0} \le n^{\epsilon n_0/8}, \tag{23}$$

where we use the fact that δ is chosen so that $\delta \cdot O_{C,\epsilon}(1) \le \epsilon/16$, and $n^\epsilon = o(n^{1-\delta}) = o(n_0)$.

Regarding the exceptional elements u_i, where $n_0 < i \le n_u$, we may choose them from $\{p/q, |p|, |q| \le n^{O_{C,\epsilon}(n^\epsilon)}\}$, which results in the bound

$$(n^{O_{C,\epsilon}(n^\epsilon)})^{2n^\epsilon} = n^{O_{C,\epsilon}(n^{2\epsilon})}. \tag{24}$$

Putting the estimates (21), (22), (23) and (24) together we obtain the bound for total number of ways to choose u

$$n^{O_{C,\epsilon,\delta}(n^{2\epsilon})} n^{-\epsilon n_0/8} \prod_{i=1}^{K} b_i^{-m_i} \leq O(n^{-\epsilon n_0/16}) \prod_{i=1}^{K} b_i^{-m_i},$$

where we use the fact that $n^{2\epsilon} = o(n^{1-\delta}) = o(n_0)$.

Thus, according to (20) we have

$$\sum_{u \in \mathcal{P}_{(m_1,\dots,m_k)}} \mathbf{P}\left(\mathbf{r}_i(M_{n-1}), 1 \leq i \leq (1-\delta)n_0, \text{ are orthogonal to } u\right) = O(n^{-\epsilon n_u/16}).$$

Summing over the number of concentration sequences $(m_1, \dots, m_k)$ (which can be bounded cheaply by $n^K = n^{O_{C,\epsilon}(1)\delta^{-1}}$), over the positions of n_u nonzero components and n_0 non-exceptional components of u (which is bounded by $O(\binom{n-1}{n_u}\binom{n_u}{n^\epsilon})$), and over the position of n_1 rows of M_{n-1} that are orthogonal to u (which is bounded by $O(n_u^{O_{C,\epsilon}(n^\epsilon)})$), we hence obtain

$$\sum_{u \in \mathcal{P}, n_u \geq n^{1-\delta}} \mathbf{P}_u = O(n^{-\epsilon n^{1-\delta}/32}),$$

where the implied constant depends on C, ϵ and δ, completing the proof.

□

Acknowledgements. The author would like to thank R. Pemantle and V. Vu for enthusiastic encouragement. He is grateful to the referees for careful reading the paper and providing very helpful remarks.

References

[1] B. Bollobás, *Random Graphs*, Academic Press, New York.

[2] J. Bourgain, V. Vu and P. M. Wood, *On the singularity probability of discrete random matrices*, Journal of Functional Analysis 258 (2010), no.2, 559-603.

[3] K. Costello, *Bilinear and quadratic variants on the Littlewood-Offord problem, submitted.*

[4] K. Costello, T. Tao and V. Vu, *Random symmetric matrices are almost surely non-singular*, Duke Math. J. 135 (2006), 395-413.

[5] P. Erdős, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. 51 (1945), 898-902.

[6] P. Erdős and L. Moser, *Elementary Problems and Solutions*: Solutions: E736. Amer. Math. Monthly, 54 (1947), no. 4, 229-230.

[7] C. G. Esséen, *On the Kolmogorov-Rogozin inequality for the concentration function*, Z. Wahrsch. Verw. Gebiete 5 (1966), 210-216.

[8] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Period. Math. Hungar. 8 (1977), no. 3-4, 197-211.

[9] J. Kahn, J. Komlós and E. Szemerédi, *On the probability that a random* ± 1 *matrix is singular*, J. Amer. Math. Soc. 8 (1995), 223-240.

[10] M. Kanter, *Probability inequalities for convex sets*, J. Multivariate Anal, 6 (1976), 222-236.

[11] G. Katona, *On a conjecture of Erdős and a stronger form of Sperner's theorem.* Studia Sci. Math. Hungar 1 (1966), 59-63.
[12] D. Kleitman, *On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors*, Advances in Math. 5 (1970), 155-157.
[13] J. Komlós, *On the determinant of* $(0,1)$ *matrices*, Studia Sci. Math. Hungar. 2 (1967), 7-22.
[14] A. Kolmogorov, *Two uniform limit theorems for sums of independent random variables*, Theor. Probab. Appl. 1 (1956), 384-394.
[15] J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation.* III. Rec. Math. Mat. Sbornik N.S. 12 , (1943). 277–286.
[16] H. Nguyen and V. Vu, *Optimal Littlewood-Offord theorems*, to appear in Advances in Mathematics.
[17] A. Odlyzko, *On subspaces spanned by random selections of* ± 1 *vectors*, J. Combin. Theory Ser. A 47 (1988), no. 1, 124-133.
[18] B. A. Rogozin, *An estimate for concentration functions*, Theor. Probab. App1. 6 (1961), 94-97.
[19] J. Rosiński and G. Samorodnitsky, *Symmetrization and concentration inequality for multilinear forms with applications to zero-one laws for Lévy chaos*, The Annals of Probability, Vol. 24 1 (1996), 422-437.
[20] A. Sárközy and E. Szemerédi, *Über ein Problem von Erdős und Moser*, Acta Arithmetica 11 (1965), 205-208.
[21] R. Stanley, *Weyl groups, the hard Lefschetz theorem, and the Sperner property*, SIAM J. Algebraic Discrete Methods 1 (1980), no. 2, 168-184.
[22] T. Tao and V. Vu, *A sharp inverse Littlewood-Offord theorem*, to appear in Random Structures and Algorithms.
[23] T. Tao and V. Vu, *Inverse Littlewood-Offord theorems and the condition number of random matrices*, Annals of Mathematics (2) 169 (2009), no 2, 595-632
[24] T. Tao and V. Vu, *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. 20 (2007), 603628.
[25] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.
[26] V. Vu, *Discrete random matrices*, `http://arxiv.org/abs/math/0611321`.

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA

E-mail address: `hoing@math.upenn.edu`