A CONTINUOUS VARIANT OF THE INVERSE LITTLEWOOD-OFFORD PROBLEM FOR QUADRATIC FORMS

HOI H. NGUYEN

ABSTRACT. Motivated by the inverse Littlewood-Offord problem for linear forms, we study the concentration of quadratic forms. We show that if this form concentrates on a small ball with high probability, then the coefficients can be approximated by a sum of additive and algebraic structures.

1. Introduction

1.1. The Littlewood-Offord problem for linear forms.

Let ξ be a real random variable, and let $A = \{a_1, \ldots, a_n\}$ be a multiset in $\mathbf{R}^d$. For any $\beta > 0$, we define the *small ball probability* as

$$\rho_{\beta,\xi}(A) := \sup_{a \in \mathbf{R}^d} \mathbf{P_x}\big(a_1 x_1 + \cdots + a_n x_n \in B(a, \beta)\big),$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ and x_i are iid copies of ξ, and $B(x, \beta)$ denotes the closed disk of radius β centered at x in $\mathbf{R}^d$.

A classical result of Erdős [3] and Littlewood-Offord [7] asserts that if ξ has Bernoulli distribution and a_i are real numbers of magnitude $|a_i| \geq \beta$, then

$$\rho_{\beta,\xi}(A) = O(n^{-1/2}).$$

This remarkable inequality has generated an impressive way of research, particularly from the early 1960s to the late 1980s. We refer the reader to [4, 5, 6] and the references therein.

Motivated by inverse theorems from additive combinatorics (see [13, Chapter 5]), Tao and Vu brought a new view to the problem: find the underlying reason as to why the small ball probability is large (say, polynomial in n).

Typical examples of A, where $\rho_{\beta,\xi}$ is large, involve *generalized arithmetic progressions* (GAPs), an important concept from additive combinatorics.

A set $Q \subset \mathbf{R}^d$ is a *GAP of rank* r if it can be expressed in the form

$$Q = \{g_0 + k_1 g_1 + \cdots + k_r g_r \mid K_i \leq k_i \leq K_i' \text{ for all } 1 \leq i \leq r\}$$

for some $g_0, \ldots, g_r \in \mathbf{R}^d$, and some integers $K_1, \ldots, K_r, K_1', \ldots, K_r'$. It is convenient to think of Q as the image of an integer box $B := \{(x_1, \ldots, x_r) \in \mathbf{Z}^r \mid K_i \leq k_i \leq K_i'\}$ under the linear map

$$\Phi : (x_1, \ldots, x_r) \mapsto g_0 + x_1 g_1 + \cdots + x_r g_r.$$

The vectors g_i are the *generators* of Q, the numbers K_i' and K_i are the *dimensions* of Q, and $\mathrm{Vol}(Q) := |B|$ is the *volume* of Q. 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 $g_0 = 0$ and $-K_i = K_i'$ for all $i \geq 1$, we say that Q is *symmetric*.

Example 1.1. *Let $Q = \{\sum_{i=1}^r k_i g_i \mid -K_i \leq k_i \leq K_i\}$ be a proper symmetric GAP of rank $r = O(1)$ and size $N = n^{O(1)}$. Assume that ξ has Bernoulli distribution, and for each a_i there exists $q_i \in Q$ such that $\|a_i - q\|_2 \leq \delta$. Then, because the random sum $\sum_i q_i x_i$ takes value in the GAP $nQ := \{\sum_{i=1}^r k_i g_i \mid -nK_i \leq k_i \leq nK_i\}$, and because $|nQ| \leq n^r N = n^{O(1)}$, the pigeon-hole principle implies that $\sum_i q_i x_i$ takes some value in nQ with probability $n^{-O(1)}$. Thus we have*

$$(1) \qquad\qquad \rho_{n\delta,\xi}(A) = n^{-O(1)}.$$

The above example shows that if ξ has Bernoulli distribution and if a_i are *close* to a *GAP* of rank $O(1)$ and size $n^{O(1)}$, then A has large small ball probability.

It was shown (rather implicitly) by Tao and Vu in [16, 17, 15, 18] that these are essentially the only examples which have large small ball probability. An explicit version was given by Vu and the current author under the following condition.

Condition 1 (Anti-concentration). *There exist positive constants $0 < c_1 < c_2$ and c_3 such that*

$$\mathbf{P}(c_1 \leq |\xi - \xi'| \leq c_2) \geq c_3,$$

where ξ' is an independent copy of ξ.

We note that Bernoulli random variables $\eta^{(\mu)}$ (which equal ± 1 with probability $\mu/2$ and 0 with probability $1 - \mu$), where the parameters μ are bounded away from 0, are clearly of this type.

We say that a vector a is δ-close to a set Q if there exists $q \in Q$ such that $\|a - q\|_2 \leq \delta$.

Theorem 1.2 (Inverse Littlewood-Offord theorem for linear forms, [11]). *Let $0 < \epsilon < 1$ and $B > 0$. Let $\beta > 0$ be a parameter that may depend on n. Suppose that $\sum_i \|a_i\|_2^2 = 1$ and*

$$\rho := \rho_{\beta,\xi}(A) \geq n^{-B},$$

where x_i are iid copies of a random variable ξ satisfying Condition 1. Then, for any number n' between n^ϵ and n, there exists a proper symmetric GAP $Q = \{\sum_{i=1}^r k_i g_i \mid |k_i| \le K_i\}$ such that

- *At least $n - n'$ elements of a_i are β-close to Q.*
- *Q has small rank, $r = O_{B,\epsilon}(1)$, and small size*

$$|Q| \le \max\left\{O_{B,\epsilon}\left(\frac{\rho^{-1}}{\sqrt{n'}}\right), 1\right\}.$$

- *There is a non-zero integer $p = O_{B,\epsilon}(\sqrt{n'})$ such that all steps g_i of Q have the form $g_i = (g_{i1}, \dots, g_{id})$, where $g_{ij} = \beta p_{ij}/p$ with $p_{ij} \in \mathbf{Z}$ and $|p_{ij}| = O_{B,\epsilon}(\beta^{-1}\sqrt{n'})$.*

In this and all subsequent theorems, the hidden constants could also depend on d and c_1, c_2, c_3 of Condition 1. We could have written $O_{d,c_1,c_2,c_3}(.)$ everywhere, but these notations are somewhat cumbersome, and this dependence is not our focus, so we omit them. Theorem 1.2 was proven in [11] with $c_1 = 1, c_2 = 2$ and $c_3 = 1/2$, but the proof there extends to the general case rather automatically.

Notation. Let $x_1, \dots, x_n$ be real numbers, and let $a_1, \dots, a_n$ be vectors in $\mathbf{R}^d$. To simplify our presentation, we will denote the sum vector $\sum_i a_i x_i$ by $\mathbf{a} \cdot \mathbf{x}$, or $\mathbf{x} \cdot \mathbf{a}$, where $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{a} = (a_1, \dots, a_n)$. For instance, the small ball probability can be expressed as

$$\rho_{\beta,\xi}(A) = \sup_a \mathbf{P}_\mathbf{x}\big(\mathbf{x} \cdot \mathbf{a} \in B(a, \beta)\big).$$

1.2. **The Littlewood-Offord problem for quadratic forms.** Let ξ be a real random variable, and let $A = (a_{ij})$ be an $n \times n$ *symmetric* matrix whose entries are vectors of $\mathbf{R}^d$. For any $\beta > 0$, we define the *quadratic small ball probability* as

$$\rho_{\beta,\xi}(A) := \sup_{a,b_1,\dots,b_n \in \mathbf{R}^d} \mathbf{P}\left(\sum_{i,j} a_{ij} x_i x_j + \sum_i b_i x_i \in B(a, \beta)\right).$$

where $x_1, \dots, x_n$ are iid copies of ξ.

It follows from [12, Theorem 3.1] and [2, Corollary 4.4] that if ξ has Bernoulli distribution and if there are $\Theta(n)$ indices i for each of which there are $\Theta(n)$ indices j such that $\|a_{ij}\|_2 \ge \beta$, then the following holds for some explicit constant $c > 0$:

$$(2) \qquad\qquad \rho_{\beta,\xi}(A) = O(n^{-c}).$$

By using a recent result of Costello [1], one can improve the right hand side to $O(n^{-1/2+o(1)})$, which is asymptotically tight. It seems that one can improve the bound further by imposing new assumptions on a_{ij}. However, this is not our goal here. Motivated by the inverse Littewood-Offord problem for linear forms, we would like to find the underlying reason as to why the quadratic small ball probability is large (say, polynomial in n).

In the following examples, ξ has Bernoulli distribution, and for each a_{ij} there exists q_{ij} such that
$$\|a_{ij} - q_{ij}\|_2 \le \delta.$$

Example 1.3. *Let Q be a proper symmetric GAP of rank $r = O(1)$ and size $n^{O(1)}$. Assume that the approximated values q_{ij} belong to Q.*

Then, because the random sum $\sum_{i,j} q_{ij} x_i x_j$ takes value in the GAP $n^2 Q$, and because the size of $n^2 Q$ is $n^{O(1)}$, the pigeon-hole principle implies that $\sum_{i,j} q_{ij} x_i x_j$ takes some value in $n^2 Q$ with probability $n^{-O(1)}$. Passing back to a_{ij}, we obtain
$$\rho_{n^2\delta,\xi}(A) = n^{-O(1)}.$$

One observes that this example is similar to Example 1.1, in which case q_{ij} have additive structure. However, unlike what we have in the linear case, there are examples of different nature where the quadratic small ball probability can be large.

Example 1.4. *Assume that q_{ij} can be written as $q_{ij} = k_i b_j + k_j b_i$, where b_i are arbitrary in $\mathbf{R}^d$ and k_i are integers bounded by $n^{O(1)}$ such that*
$$\mathbf{P_x}\left(\sum_i k_i x_i = 0\right) = n^{-O(1)}.$$

Then, we have
$$\mathbf{P}\left(\sum_{i,j} q_{ij} x_i x_j = \mathbf{0}\right) = \mathbf{P}\left(\sum_i k_i x_i \sum_j b_j x_j = \mathbf{0}\right) = n^{-O(1)}.$$

Passing back to a_{ij}, we obtain
$$\rho_{n^2\delta,\xi}(A) = n^{-O(1)}.$$

Motivated by 1.3 and 1.4, we now consider a more complicated example.

Example 1.5. *Assume that $q_{ij} = q'_{ij} + q''_{ij}$, where $q'_{ij} \in Q$, a proper symmetric GAP of rank $O(1)$ and size $n^{O(1)}$, and $q''_{ij} = k_{i1}b_{1j} + k_{j1}b_{1i} + \cdots + k_{ir}b_{rj} + k_{jr}b_{ri}$, where $r = O(1)$, and $b_{1i}, \ldots, b_{ri}$ are arbitrary in $\mathbf{R}^d$, and $k_{i1}, \ldots, k_{ir}$ are integers bounded by $n^{O(1)}$ such that*
$$\mathbf{P_x}\left(\sum_i k_{i1} x_i = 0, \ldots, \sum_i k_{ir} x_i = 0\right) = n^{-O(1)}.$$

Observe that
$$\sum_{i,j} q_{ij} x_i x_j = \sum_{i,j} q'_{i,j} x_i x_j + \left(\sum_i k_{i1} x_i\right)\left(\sum_j b_{1j} x_j\right) + \cdots$$
$$\cdots + \left(\sum_i k_{ir} x_i\right)\left(\sum_j b_{rj} x_j\right).$$

Thus,

$$\sup_{q \in n^2 Q} \mathbf{P_x} \left(\sum_{i,j} q_{ij} x_i x_j = q \right) = n^{-O(1)}.$$

Passing to a_{ij}, we obtain

$$\rho_{n^2 \delta, \xi}(A) = n^{-O(1)}.$$

In this example, the matrix (q_{ij}) is a sum of two unrelated submatrices (q'_{ij}) and (q''_{ij}): one has entries belonging to a GAP of rank $O(1)$ and size $n^{O(1)}$, and one has rank $O(1)$.

Our main theorem partially demonstrates that if $\rho_{\beta, \xi}(A)$ is large, then a_{ij} are close to some q_{ij} taking the form of Example 1.5. We denote by $\mathbf{r}_i(A)$ the row $(a_{i1}, \ldots, a_{in})$ of A.

Theorem 1.6 (Inverse Littlewood-Offord theorem for quadratic forms). *Let $0 < \epsilon < 1$ and $B > 0$. Let $\beta > 0$ be a parameter that may depend on n. Assume that $a_{ij} = a_{ji}$, and*

$$\rho := \rho_{\beta, \xi}(A) \geq n^{-B}.$$

Then, there exist an integer $k \neq 0, |k| = n^{O_{B,\epsilon}(1)}$, a set of $r = O(1)$ rows $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_r}$ of A, and set I of size at least $n - 2n^\epsilon$ such that for each $i \in I$, there exist integers $k_{ii_1}, \ldots, k_{ii_r}$, all bounded by $n^{O_{B,\epsilon}(1)}$, such that the following holds.

$$(3) \qquad \mathbf{P_z} \left(\left\| \mathbf{z} \cdot (k\mathbf{r}_i(A) + \sum_j k_{ii_j} \mathbf{r}_{i_j}(A)) \right\|_2 \leq \beta n^{O_{B,\epsilon}(1)} \right) \geq n^{-O_{B,\epsilon}(1)},$$

where $\mathbf{z} = (z_1, \ldots, z_n)$ and z_i are iid copies of $\eta^{(1/2)}(\xi - \xi')$, where $\eta^{(1/2)}$ is a Bernoulli random variable of parameter $1/2$ which is independent of ξ and ξ'.

It follows from (3) and from Theorem 1.2 that for each $i \in I$, most of the entries of $k\mathbf{r}_i(A) + \sum_j k_{ii_j} \mathbf{r}_{i_j}(A)$ are $\beta n^{O_{B,\epsilon}(1)}$-close to a symmetric GAP of rank $O(1)$ and size $n^{O(1)}$. In other words, Theorem 1.6 asserts that, modulo some special linear combinations of $\mathbf{r}_{i_1}(A), \ldots, \mathbf{r}_{i_r}(A)$ (where the coefficients are integers bounded by $n^{O(1)}$), most of the components of $\mathbf{r}_i(A)$ are $\beta n^{O(1)}$-close to a symmetric GAP of rank $O(1)$ and size $n^{O(1)}$.

Theorem 1.6 seems to be useful. It plays a crucial role in our work [10] of establishing polynomial bounds for the singular value of random symmetric matrices. We remark that a discrete version of Theorem 1.6 was discussed in an earlier paper [8].

2. A RANK REDUCTION ARGUMENT AND THE FULL RANK ASSUMPTION

This section 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, then we can assume this containment is non-degenerate.

Assume that $P = \{k_1 g_1 + \cdots + k_r g_r \mid -K_i \leq k_i \leq K_i\}$ 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 $k_1 g_1 + \cdots + k_r g_r$ to $(k_1, \ldots, k_r)$. 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 2.1. *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 followings hold.*

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

To prove Theorem 2.1, we will rely on the following lemma.

Lemma 2.2 (Progressions lie inside proper progressions, [13])**.** *There is an absolute constant C depending in d such that the following holds. Let P be a GAP of rank r in $\mathbf{R}^d$. Then there is a symmetric proper GAP Q of rank at most r containing P and*

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

Proof. (of Theorem 2.1) We shall mainly follow [14, 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 $\alpha_1, \ldots, \alpha_r$ whose common divisor is 1 and $\alpha_1 k_1 + \cdots + \alpha_r k_r = 0$ for all $(k_1, \ldots, k_r) \in \Phi(U)$. 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 $1 \leq i \leq r - 1$. The new symmetric GAP P' will continue to contain U, because we have

$$k_1 g_1' + \cdots + k_{r-1} g_{r-1}' = k_1 g_1 + \cdots + k_r g_r - w(\alpha_1 k_1 + \cdots + \alpha_r g_r)$$
$$= k_1 g_1 + \cdots + k_r g_r$$

for all $(k_1, \ldots, k_r) \in \Phi(U)$. Also, note that the volume of P' is $2^{r-1} K_1 \cdots K_{r-1}$, which is less than the volume of P.

We next use Lemma 2.2 to guarantee that P' is symmetric and proper without increasing the rank. Iterate the process if needed. Because the rank of the newly obtained proper symmetric GAP decreases strictly after each step, the process must terminate after at most r steps. $\square$

3. A DECOUPLING LEMMA AND INVERSE PROBLEM FOR BILINEAR FORMS

As the first step to establish Theorem 1.6, we pass to bilinear forms by using a decoupling technique. Let U be a subset of $\{1, \ldots, n\}$. Let A_U be a

symmetric matrix of size n by n defined as

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

where we denoted by $A_U(ij)$ the ij entry of A_U.

Lemma 3.1 (Decoupling lemma). *Assume that*

$$\rho = \sup_{a, b_1, \ldots, b_n} \mathbf{P_x}\left(\left\| \sum_{i,j} a_{ij} x_i x_j + \sum_i b_i x_i - a \right\|_2 \leq \beta \right) \geq n^{-B}.$$

Then,

$$(4) \qquad \mathbf{P_{v,w}}\left(\left\| \sum_{i,j} A_U(ij) v_i w_j \right\|_2 = O_B\left(\beta \sqrt{\log n} \right) \right) = \Theta(\rho^8),$$

where $\mathbf{v} = (v_1, \ldots, v_n)$, $\mathbf{w} = (w_1, \ldots, w_n)$, *and* v_i, w_j *are iid copies of* $\xi - \xi'$.

We refer the reader to Appendix A for a proof of this lemma.

Lemma 3.1 asserts that if $\rho_{\beta,\xi}(A)$ is large then $\sum_{i,j} A_U(ij) v_i w_j$ has small norm with high probability. This fact allows us to deduce useful information for A_U (for all U) by combining with the following inverse-type result.

Theorem 3.2 (Inverse Littlewood-Offord theorem for bilinear forms). *Let* $0 < \epsilon < 1$ *and* $B > 0$. *Let* $\beta > 0$ *be a parameter that may depend on* n. *Assume that*

$$\sup_a \mathbf{P_{x,y}}\left(\left\| \sum_{i,j \leq n} a_{ij} x_i y_j - a \right\|_2 \leq \beta \right) \geq n^{-B},$$

where $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{y} = (y_1, \ldots, y_n)$, *and* x_i *and* y_i *are iid copies of a random variable* ξ *satisfying Condition 1. Then, there exist an integer* $k \neq 0, |k| = n^{O_{B,\epsilon}(1)}$, *a set of* $r = O(1)$ *rows* $\mathbf{r}_{i_1}, \ldots, \mathbf{r}_{i_r}$ *of* A, *and set* I *of size at least* $n - 2n^\epsilon$ *such that for each* $i \in I$, *there exist integers* $k_{ii_1}, \ldots, k_{ii_r}$, *all bounded by* $n^{O_{B,\epsilon}(1)}$, *such that the following holds.*

$$(5) \qquad \mathbf{P_y}\left(\left\| \mathbf{y} \cdot \left(k\mathbf{r}_i(A) + \sum_j k_{ii_j} \mathbf{r}_{i_j}(A) \right) \right\|_2 \leq \beta n^{O_{B,\epsilon}(1)} \right) \geq n^{-O_{B,\epsilon}(1)}.$$

For the rest of this section, we prove Theorem 3.2.

First of all, for minor technical reasons, it is convenient to assume ξ to have discrete distribution. The continuous case can be recovered by approximating the continuous distribution by a discrete one while holding n fixed. For short, we denote the vector $(a_{i1}, \ldots, a_{in})$ by $\mathbf{a}_i$. We begin by applying Theorem 1.2.

Lemma 3.3. *Let* $\epsilon < 1$, *and* B *be positive constants. Assume that*

$$\rho = \sup_a \mathbf{P_{x,y}}\left(\left| \sum_{i,j} a_{ij} x_i y_j - a \right| \leq \beta \right) \geq n^{-B}.$$

Then, the following holds with probability at least $3\rho/4$ with respect to $\mathbf{y} = (y_1, \ldots, y_n)$. There exist a proper symmetric GAP $Q_\mathbf{y} \subset \mathbf{R}^d$ of rank $O_{B,\epsilon}(1)$ and size $\max\{O_{B,\epsilon}(\rho^{-1}/n^{\epsilon/2}), 1\}$, and an index set $I_\mathbf{y}$ of size $n - n^\epsilon$ such that $\mathbf{a}_i \cdot \mathbf{y}$ is β-close to $Q_\mathbf{y}$ for all $i \in I_\mathbf{y}$.

Proof. (of Lemma 3.3) Write

$$\sum_{i,j} a_{ij} x_i y_j = \sum_{i=1}^n x_i(\mathbf{a}_i \cdot \mathbf{y}).$$

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

$$\mathbf{P}_\mathbf{x}\left(\left|\sum_{i=1}^n x_i(\mathbf{a}_i \cdot \mathbf{y}) - a\right| \leq \beta\right) \geq \frac{\rho}{4}.$$

We call $\mathbf{y}$ *bad* otherwise.

Let G denote the collection of good vectors. We are going to estimate the probability p of a randomly chosen vector $\mathbf{y} = (y_1, \ldots, y_n)$ being bad by an averaging method.

$$\mathbf{P}_\mathbf{y}\mathbf{P}_\mathbf{x}\left(\left|\sum_{i=1}^n x_i(\mathbf{a}_i \cdot \mathbf{y}) - a\right| \leq \beta\right) = \rho$$

$$p\frac{\rho}{4} + 1 - p \geq \rho$$

$$\frac{1-\rho}{1-\rho/4} \geq p.$$

Thus, the probability of a randomly chosen $\mathbf{y}$ belonging to G is at least

$$1 - p \geq \frac{3\rho/4}{1-\rho/4} \geq 3\frac{\rho}{4}.$$

Consider a good vector $\mathbf{y} \in G$. By definition, we have

$$\mathbf{P}_\mathbf{x}\left(\left|\sum_{i=1}^n x_i(\mathbf{a}_i \cdot \mathbf{y}) - a\right| \leq \beta\right) \geq \frac{\rho}{4}.$$

Next, if $\mathbf{a}_i \cdot \mathbf{y} = \mathbf{0}$ for all i, then the conclusion of the lemma holds trivially for $Q_\mathbf{y} := \mathbf{0}$. Otherwise, we apply Theorem 1.2 to the sequence $\{\mathbf{a}_i \cdot \mathbf{y}, i = 1, \ldots, n\}$ (after a rescaling). As a consequence, we obtain an index set $I_\mathbf{y}$ of size $n - n^\epsilon$ and a proper symmetric GAP $Q_\mathbf{y}$ of rank $O_{B,\epsilon}(1)$ and size $\max\{O_{B,\epsilon}(\rho^{-1}/n^{\epsilon/2}), 1\}$, together with its elements $q_i(\mathbf{y})$, such that $\|\mathbf{a}_i \cdot \mathbf{y} - q_i(\mathbf{y})\|_2 \leq \beta$ for all $i \in I_\mathbf{y}$. $\square$

We now work with $q_i(\mathbf{y})$, where $\mathbf{y} \in G$.

Common generating indices. By Theorem 2.1, we may assume that the $q_i(\mathbf{y})$ span $Q_\mathbf{y}$. We choose from $I_\mathbf{y}$ s indices $i_{y_1}, \ldots, i_{y_s}$ such that $q_{i_{y_j}}(\mathbf{y})$ span $Q_\mathbf{y}$, where s is the rank of $Q_\mathbf{y}$. Note that $s = O_{B,\epsilon}(1)$ for all $\mathbf{y} \in G$.

Consider the tuples $(i_{y_1}, \ldots, i_{y_s})$ for all $\mathbf{y} \in G$. Because there are $\sum_s O_{B,\epsilon}(n^s) = n^{O_{B,\epsilon}(1)}$ possibilities these tuples can take, there exists a

tuple, say $(1, \ldots, r)$ (by rearranging the rows of A if needed), such that $(i_{y_1}, \ldots, i_{y_s}) = (1, \ldots, r)$ for all $\mathbf{y} \in G'$, a subset G' of G which satisfies

$$(6) \qquad \mathbf{P_y}(\mathbf{y} \in G') \geq \frac{\mathbf{P_y}(\mathbf{y} \in G)}{n^{O_{C,\epsilon}(1)}} = \frac{\rho}{n^{O_{B,\epsilon}(1)}}.$$

Common coefficient tuple. For each $1 \leq i \leq r$, we express $q_i(\mathbf{y})$ in terms of the generators of $Q_\mathbf{y}$ for each $\mathbf{y} \in G'$,

$$q_i(\mathbf{y}) = c_{i1}(\mathbf{y})g_1(\mathbf{y}) + \cdots + c_{ir}(\mathbf{y})g_r(\mathbf{y}),$$

where $c_{i1}(\mathbf{y}), \ldots c_{ir}(\mathbf{y})$ are integers bounded by $n^{O_{B,\epsilon}(1)}$, and $g_i(\mathbf{y})$ are the generators of $Q_\mathbf{y}$. We will show that there are many $\mathbf{y}$ that correspond to the same coefficients c_{ij}.

Consider the collection of the coefficient-tuples

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

for all $\mathbf{y} \in G'$. Because the number of possibilities these tuples can take is at most

$$(n^{O_{B,\epsilon}(1)})^{r^2} = n^{O_{B,\epsilon}(1)},$$

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

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

for all $\mathbf{y} \in G''$, a subset of G' which satisfies

$$(7) \qquad \mathbf{P_y}(\mathbf{y} \in G'') \geq \frac{\mathbf{P_y}(\mathbf{y} \in G')}{n^{O_{B,\epsilon}(1)}} \geq \frac{\rho}{n^{O_{B,\epsilon}(1)}}.$$

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

- $q_i(\mathbf{y}) = c_{i1}g_1(\mathbf{y}) + \cdots + c_{jr}g_r(\mathbf{y})$, for $i = 1, \ldots, r$.
- The vectors $(c_{11}, \ldots, c_{1r}), \ldots, (c_{r1}, \ldots c_{rr})$ span $\mathbf{Z}^{\mathrm{rank}(Q_\mathbf{y})}$.

Next, because $|I_\mathbf{y}| \geq n - n^\epsilon$ for each $\mathbf{y} \in G''$, by an averaging argument, there exists a set I of size $n - 2n^\epsilon$ such that for each $i \in I$ we have

$$(8) \qquad \mathbf{P_y}(i \in I_\mathbf{y}, \mathbf{y} \in G'') \geq \frac{\mathbf{P_y}(\mathbf{y} \in G'')}{2}.$$

From now on we fix an arbitrary row $\mathbf{a}$ of index from I. We will focus on those $\mathbf{y} \in G''$ where the index of $\mathbf{a}$ belongs to $I_\mathbf{y}$.

Common coefficient tuple for each individual. Because $q(\mathbf{y}) \in Q_\mathbf{y}$ ($q(\mathbf{y})$ is the element of $Q_\mathbf{y}$ that is β-close to $\mathbf{a} \cdot \mathbf{y}$), we can write

$$q(\mathbf{y}) = c_1(\mathbf{y})g_1(\mathbf{y}) + \cdots + c_r(\mathbf{y})g_r(\mathbf{y})$$

where $c_i(\mathbf{y})$ are integers bounded by $n^{O_{B,\epsilon}(1)}$. For short, for each i we denote by $\mathbf{v}_i$ the vector $(c_{i1}, \ldots, c_{ir})$, we will also denote by $\mathbf{v}_{\mathbf{a},\mathbf{y}}$ the vector $(c_1(\mathbf{y}), \ldots, c_r(\mathbf{y}))$. Because $Q_{\mathbf{y}}$ is spanned by $q_1(\mathbf{y}), \ldots, q_r(\mathbf{y})$, we have $k = \det(\mathbf{v}_1, \ldots, \mathbf{v}_r) \neq 0$, and that

(9)
$$kq(\mathbf{y}) + \det(\mathbf{v}_{\mathbf{a},\mathbf{y}}, \mathbf{v}_2, \ldots, \mathbf{v}_r)q_1(\mathbf{y}) + \cdots + \det(\mathbf{v}_{\mathbf{a},\mathbf{y}}, \mathbf{v}_1, \ldots, \mathbf{v}_{r-1})q_r(\mathbf{y}) = 0.$$

It is crucial to note that k is independent of the choice of $\mathbf{a}$ and $\mathbf{y}$.

Next, because each coefficient of (9) is bounded by $n^{O_{B,\epsilon}(1)}$, there exists a subset $G_{\mathbf{a}}'''$ of G'' such that all $\mathbf{y} \in G_{\mathbf{a}}'''$ correspond to the same identity, and

(10) $$\mathbf{P}_{\mathbf{y}}\left(\mathbf{y} \in G_{\mathbf{a}}'''\right) \geq \frac{\mathbf{P}_{\mathbf{y}}(\mathbf{y} \in G'')/2}{(n^{O_{B,\epsilon}(1)})^r} = \frac{\rho}{n^{O_{B,\epsilon}(1)}} = n^{-O_{B,\epsilon}(1)}.$$

In other words, there exist integers $k_1, \ldots, k_r$ depending on $\mathbf{a}$, all bounded by $n^{O_{B,\epsilon}(1)}$, such that

(11) $$kq(\mathbf{y}) + k_1q_1(\mathbf{y}) + \cdots + k_rq_r(\mathbf{y}) = 0$$

for all $\mathbf{y} \in G_{\mathbf{a}}'''$.

Passing back to A. Because $q_i(\mathbf{y})$ are β-close to $\mathbf{a}_i \cdot \mathbf{y}$, it follows from (11) that

(12)
$$\|k\mathbf{a} \cdot \mathbf{y} + k_1\mathbf{a}_1 \cdot \mathbf{y} + \cdots + k_r\mathbf{a}_r \cdot \mathbf{y}\|_2 = \|(k\mathbf{a} + k_1\mathbf{a}_1 + \cdots + \mathbf{a}_r) \cdot \mathbf{y}\|_2 \leq n^{O_{B,\epsilon}(1)}\beta.$$

Furthermore, as $\mathbf{P}_{\mathbf{y}}(\mathbf{y} \in G_{\mathbf{a}}''') = n^{-O_{B,\epsilon}(1)}$, we have

(13) $$\mathbf{P}_{\mathbf{y}}\left(\|(k\mathbf{a} + k_1\mathbf{a}_1 + \cdots + k_r\mathbf{a}_r) \cdot \mathbf{y}\|_2 \leq n^{O_{B,\epsilon}(1)}\beta\right) = n^{-O_{B,\epsilon}(1)}.$$

Because (13) holds for any row $\mathbf{a}$ indexing from I, we have obtained the conclusion of Theorem 3.2.

4. PROOF OF THEOREM 1.6

By the definition of ξ, it is clear that the random variable $\xi - \xi'$ also satisfies Condition 1 (with different positive parameters). We next apply Theorem 3.2 to (4) to obtain the following lemma.

Lemma 4.1. *There exist a set $I_0(U)$ of size $O_{B,\epsilon}(1)$ and a set $I(U)$ of size at least $n - n^\epsilon$, and a nonzero integer $k(U)$ bounded by $n^{O_{B,\epsilon}(1)}$ such that for any $i \in I$, there are integers $k_{ii_0}(U), i_0 \in I_0(U)$, all bounded by $n^{O_{B,\epsilon}(1)}$, such that*

$$\mathbf{P}_{\mathbf{y}}\left(\left\|(k(U)\mathbf{a}_i(A_U) + \sum_{i_0 \in I_0} k_{ii_0}(U)\mathbf{a}_{i_0}(A_U)) \cdot \mathbf{y}\right\|_2 \leq \beta n^{O_{B,\epsilon}(1)}\right) = n^{-O_{B,\epsilon}(1)},$$

where $\mathbf{y} = (y_1, \ldots, y_n)$ and y_i are iid copies of $\xi - \xi'$.

Note that this lemma holds for all $U \subset [n]$. In what follows we will gather this information.

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

$$|\mathcal{U}| \geq \frac{2^n}{n^{O_{B,\epsilon}(1)}}.$$

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

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

From now on we fix an $i \in I$. Consider the tuples $(k_{ii_0}(U), i_0 \in I_0)$ over all U where $i \in I_U$. Because there are only $n^{O_{B,\epsilon}(1)}$ 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 $|\mathcal{U}|/2n^{O_{B,\epsilon}(1)} = 2^n/n^{O_{B,\epsilon}(1)}$ sets U. Because $|I_0| = O_{B,\epsilon}(1)$, there is a way to partition I_0 into $I_0' \cup I_0''$ such that there are $2^n/n^{O_{B,\epsilon}(1)}$ sets among the U above that satisfy $U \cap I_0 = I_0''$.

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. (The other case can be treated similarly).

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

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

Also, because $i \notin U$, we have $\mathbf{a}_i(A_U) \cdot \mathbf{y} = \sum_{j=1}^n a_{ij} u_j y_j$. Thus,

$$k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0 \in I_0} k_{ii_0} \mathbf{a}_{i_0}(A_U) \cdot \mathbf{y}$$

$$= k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0' \in I_0'} k_{ii_0'} \mathbf{a}_{i_0'}(A_U) \cdot \mathbf{y} + \sum_{i_0'' \in I_0''} k_{ii_0''} \mathbf{a}_{i_0''}(A_U) \cdot \mathbf{y}$$

$$= \sum_{j=1}^n k a_{ij} u_j y_j + \sum_{j=1}^n \sum_{i_0' \in I_0'} k_{ii_0'} a_{i_0' j} u_j y_j + \sum_{j=1}^n \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0'' j}(1 - u_j) y_j$$

$$= \sum_{j=1}^n \left(k a_{ij} + \sum_{i_0' \in I_0'} k_{ii_0'} a_{i_0' j} - \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0'' j} \right) u_j y_j + \sum_{j=1}^n \sum_{i_0'' \in I_0''} k_{ii_0''} a_{i_0'' j} y_j.$$

Next, by Lemma 4.1, the following holds for each $U \in \mathcal{U}_{I_0', I_0''}$,

$$\mathbf{P_y}\left(\left\|k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0 \in I_0} k_{ii_0}\mathbf{a}_{i_0}(A_U) \cdot \mathbf{y}\right\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right) = n^{-O_{B,\epsilon}(1)}.$$

Also, recall that

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

Hence,

$$\mathbf{E_y}\mathbf{E}_U\left(\left\|k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0 \in I_0} k_{ii_0}\mathbf{a}_{i_0}(A_U) \cdot \mathbf{y}\right\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right) \geq n^{-O_{B,\epsilon}(1)}.$$

By applying the Cauchy-Schwarz inequality, we obtain

$$n^{-O_{B,\epsilon}(1)} \leq \left[\mathbf{E_y}\mathbf{E}_U\left(\left\|k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0 \in I_0} k_{ii_0}\mathbf{a}_{i_0}(A_U) \cdot \mathbf{y}\right\|_2\right.\right.$$
$$\left.\left. = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right)\right]^2$$

$$\leq \mathbf{E_y}\left[\mathbf{E}_U\left(\left\|k\mathbf{a}_i(A_U) \cdot \mathbf{y} + \sum_{i_0 \in I_0} k_{ii_0}\mathbf{a}_{i_0}(A_U) \cdot \mathbf{y}\right\|_2\right.\right.$$
$$\left.\left. = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right)\right]^2$$

$$= \mathbf{E_y}\left[\mathbf{E_u}\left(\left\|\sum_{j=1}^{n}\left(ka_{ij} + \sum_{i_0' \in I_0'} k_{ii_0'}a_{i_0'j} - \sum_{i_0'' \in I_0''} k_{ii_0''}a_{i_0''j}\right)u_jy_j\right.\right.\right.$$
$$\left.\left.\left. + \sum_{j=1}^{n}\sum_{i_0'' \in I_0''} k_{ii_0''}a_{i_0''j}y_j\right\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right)\right]^2$$

$$\leq \mathbf{E_y}\mathbf{E}_{\mathbf{u},\mathbf{u}'}\left(\left\|\sum_{j=1}^{n}\left(k_{ij}a_{ij} + \sum_{i_0' \in I_0'} k_{ii_0'}a_{i_0'j} - \sum_{i_0'' \in I_0''} k_{ii_0''}a_{i_0''j}\right)\right.\right.$$
$$\left.\left. \times (u_j - u_j')y_j\right\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right)$$

$$(14) \qquad = \mathbf{E_z}\left(\left\|\sum_{j=1}^{n}\left(ka_{ij} + \sum_{i_0' \in I_0'} k_{ii_0'}a_{i_0'j} - \sum_{i_0'' \in I_0''} k_{ii_0''}a_{i_0''j}\right)z_j\right\|_2\right.$$
$$\left. = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right),$$

where $z_j := (u_j - u'_j)y_j$, and in the last inequality we used the fact that

$$\mathbf{E}_{\mathbf{u},\mathbf{u}'}\left(\|f(\mathbf{u})\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right), \|f(\mathbf{u}')\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right)$$
$$\leq \mathbf{E}_{\mathbf{u},\mathbf{u}'}\left(\|f(\mathbf{u}) - f(\mathbf{u}')\|_2 = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right).$$

Note that $u_j - u'_j$ are iid copies of the Bernoulli random variable $2\eta^{(1/2)}$. Hence z_j are iid copies of $2\eta^{(1/2)}(\xi - \xi')$, where $\eta^{(1/2)}$ is independent of ξ and ξ'.

In conclusion, the following holds for any $i \in I$,

$$\mathbf{P_z}\left(\left\|\sum_{j=1}^{n}\left(ka_{ij} + \sum_{i'_0 \in I'_0} k_{ii'_0} a_{i'_0 j} - \sum_{i''_0 \in I''_0} k_{ii''_0} a_{i''_0 j}\right)z_j\right\|_2\right.$$
$$\left. = O\left(\beta n^{O_{B,\epsilon}(1)}\right)\right) \geq n^{-O_{B,\epsilon}(1)}.$$

Note that k and I_0 are independent of the choice of i. By changing the sign of $k_{ii''_0}$, we are done with the proof of Theorem 1.6.

A. Proof of Lemma 3.1

The goal of this section is to establish the inequality

$$\mathbf{P_{v,w}}\left(\left\|\sum_{i,j} A_U(ij)v_i w_j\right\|_2 = O_B\left(\beta\sqrt{\log n}\right)\right) \geq \frac{\rho^8}{2(2\pi)^{7d/2}\exp(8\pi)},$$

under the assumption

$$\sup_{a,b_1,\ldots,b_n}\mathbf{P_x}\left(\left\|\sum_{i,j} a_{ij}x_i x_j + \sum_i b_i x_i - a\right\| \leq \beta\right) = \rho \geq n^{-B}.$$

Setting $a'_{ij} := a_{ij}/\beta$, we have

$$\sup_{a',b'_i}\mathbf{P_x}\left(\left\|\sum_{i,j} a'_{ij}x_i x_j + \sum_i b'_i x_i - a'\right\|_2 \leq 1\right) \geq n^{-B}.$$

Next, by Markov's inequality,

$$\mathbf{P_x}\left(\left\|\sum_{i,j} a'_{ij}x_i x_j + \sum_i b'_i x_i - a'\right\|_2 \leq 1\right)$$
$$= \mathbf{P}\left(\exp\left(-\frac{\pi}{2}\right)\left\|\sum_{i,j} a'_{ij}x_i x_j + \sum_i a'_i x_i - a'\right\|_2^2 \geq \exp\left(-\frac{\pi}{2}\right)\right)$$
$$\leq \exp\left(\frac{\pi}{2}\right)\mathbf{E_x}\exp\left(-\frac{\pi}{2}\left\|\sum_{i,j} a'_{ij}x_i x_j + \sum_i b'_i x_i - a'\right\|_2^2\right).$$

Note that

$$\exp\left(-\frac{\pi}{2}\|x\|_2^2\right) = \int_{\mathbf{R}^d} e(x \cdot t) \exp\left(-\frac{\pi}{2}\|t\|_2^2\right) dt.$$

Thus

$$\mathbf{P_x}\left(\left\|\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i - a'\right\|_2 \leq 1\right)$$

$$\leq \exp\left(\frac{\pi}{2}\right) \int_{\mathbf{R}^d} \left|\mathbf{E_x} e\left[\left(\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i\right) \cdot t\right]\right| \exp\left(-\frac{\pi}{2}\|t\|^2\right) dt$$

$$\leq \exp\left(\frac{\pi}{2}\right) \left(\sqrt{2\pi}\right)^d \int_{\mathbf{R}^d} \left|\mathbf{E_x} e\left[\left(\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i\right) \cdot t\right]\right|$$

$$\times \frac{\exp\left(-\pi\|t\|_2^2/2\right)}{\left(\sqrt{2\pi}\right)^d} dt.$$

Consider $\mathbf{x}$ as $(\mathbf{x}_U, \mathbf{x}_{\bar{U}})$, where $\mathbf{x}_U, \mathbf{x}_{\bar{U}}$ are the vectors corresponding to $i \in U$ and $i \notin U$ respectively. By the Cauchy-Schwarz inequality we have

$$\left[\int_{\mathbf{R}^d} \left|\mathbf{E_x} e\left(\left(\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i\right) \cdot t\right)\right| \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} dt\right]^4$$

$$\leq \left[\int_{\mathbf{R}^d} \left|\mathbf{E_x} e\left(\left(\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i\right) \cdot t\right)\right|^2 \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} dt\right]^2$$

$$\leq \left[\int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{x}_U} \left|\mathbf{E}_{\mathbf{x}_{\bar{U}}} e\left(\left(\sum_{i,j} a'_{ij} x_i x_j + \sum_i b'_i x_i\right) \cdot t\right)\right|^2\right.$$

$$\left. \times \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} dt\right]^2$$

$$= \left[\int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{x}_U} \mathbf{E}_{\mathbf{x}_{\bar{U}}, \mathbf{x}'_{\bar{U}}} e\left(\left(\sum_{i \in U, j \in \bar{U}} a'_{ij} x_i (x_j - x'_j) + \sum_{j \in \bar{U}} b'_j (x_j - x'_j)\right.\right.\right.$$

$$\left.\left.\left. + \sum_{i \in \bar{U}, j \in \bar{U}} a'_{ij} (x_i x_j - x'_i x'_j)\right) \cdot t\right) \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} dt\right]^2$$

$$\leq \int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{x}_{\bar{U}}, \mathbf{x}'_{\bar{U}}} \left|\mathbf{E}_{\mathbf{x}_U} e\left(\left(\sum_{i \in U, j \in \bar{U}} a'_{ij} x_i (x_j - x'_j) + \sum_{j \in \bar{U}} b'_j (x_j - x'_j)\right.\right.\right.$$

$$\left.\left.\left. + \sum_{i \in \bar{U}, j \in \bar{U}} a'_{ij} (x_i x_j - x'_i x'_j)\right) \cdot t\right)\right|^2 \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} dt$$

$$= \int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{x}_U,\mathbf{x}'_U,\mathbf{x}_{\bar{U}},\mathbf{x}'_{\bar{U}}} e\left(\left(\sum_{i\in U, j\in \bar{U}} a'_{ij}(x_i - x'_i)(x_j - x'_j)\right)\cdot t\right)$$

$$\times \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt$$

$$= \int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{y}_U,\mathbf{z}_{\bar{U}}} e\left(\left(\sum_{i\in \bar{U}, j\in U} a'_{ij} y_i z_j\right)t\right) \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt,$$

where $\mathbf{y}_U = \mathbf{x}_U - \mathbf{x}'_U$ and $\mathbf{z}_{\bar{U}} = \mathbf{x}_{\bar{U}} - \mathbf{x}'_{\bar{U}}$, whose entries are iid copies of $\xi - \xi'$. Thus we have

$$\left[\int_{\mathbf{R}^d} \left|\mathbf{E}_{\mathbf{x}} e\left(\left(\sum_{i,j} a'_{ij} x_i x_j\right)\cdot t\right)\right| \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt\right]^8$$

$$\leq \left[\int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{y}_U,\mathbf{z}_{\bar{U}}} e\left(\left(\sum_{i\in U, j\in \bar{U}} a'_{ij} y_i z_j\right)\cdot t\right) \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt\right]^2$$

$$\leq \int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{y}_U,\mathbf{z}_{\bar{U}},\mathbf{y}'_U,\mathbf{z}'_{\bar{U}}} e\left(\left(\sum_{i\in U, j\in \bar{U}} a'_{ij} y_i z_j - \sum_{i\in U, j\in \bar{U}} a'_{ij} y'_i z'_j\right)\cdot t\right)$$

$$\times \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt.$$

Because $a'_{ij} = a'_{ji}$, we can write the last term as

$$\int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{y}_U,\mathbf{z}'_{\bar{U}},\mathbf{y}'_U,\mathbf{z}_{\bar{U}}} e\left(\left(\sum_{i\in U, j\in \bar{U}} a'_{ij} y_i z_j + \sum_{j\in \bar{U}, i\in U} a_{ji}(-z'_j) y'_i\right)\cdot t\right)$$

$$\times \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt$$

$$= \int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{v},\mathbf{w}} e\left(\left(\sum_{i\in U, j\in \bar{U}} a'_{ij} v_i w_j + \sum_{i\in \bar{U}, j\in U} a'_{ij} v_i w_j\right)\cdot t\right)$$

$$\times \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt,$$

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

Next, recall that $A_U(ij) = a_{ij}$ if either $i \in U, j \notin U$ or $i \notin U, j \in U$. We have

$$\int_{\mathbf{R}^d} \mathbf{E}_{\mathbf{v},\mathbf{w}} e\left(\left(\sum_{i\in U j\in \bar{U}} a'_{ij} v_i w_j + \sum_{i\in \bar{U}, j\in U} a'_{ij} v_i w_j\right)\cdot t\right) \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d}\, dt$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^d \mathbf{E}_{\mathbf{v},\mathbf{w}} \exp\left(-\frac{\pi}{2}\left\|\sum_{i,j} A_U(ij)' v_i w_j\right\|_2^2\right),$$

where $A_U(ij)' := A_U(ij)/\beta$. Thus

$$\rho^8 = \left(\mathbf{P_x} \left(\Big| \sum_{i,j} a'_{ij} x_i, x_j + \sum_i b'_i x_i - a' \Big| \le 1 \right) \right)^8$$

$$\le \exp(4\pi)(2\pi)^{4d} \left(\int_{\mathbf{R}^d} \Big| \mathbf{E_x} e \Big(\Big(\sum_{i,j} a'_{ij} x_i x_j \Big) \cdot t \Big) \Big| \frac{\exp\left(-\frac{\pi}{2}\|t\|_2^2\right)}{\left(\sqrt{2\pi}\right)^d} \, dt \right)^8$$

$$\le \exp(4\pi)(2\pi)^{7d/2} \, \mathbf{E_{v,w}} \exp\left(-\frac{\pi}{2} \Big\| \sum_{i,j} A_U(ij)' v_i w_j \Big\|_2^2 \right).$$

Because $\rho \ge n^{-B}$, the inequality above implies that

$$\mathbf{P_{v,w}} \left(\Big\| \sum_{i,j} A_U(ij)' v_i w_j \Big\|_2 = O_B\left(\sqrt{\log n}\right) \right) \ge \frac{\rho^8}{2(2\pi)^{7d/2} \exp(4\pi)}.$$

Scaling back to A_{ij}, we obtain

$$\mathbf{P_{v,w}} \left(\Big\| \sum_{i,j} A_U(ij) v_i w_j \Big\|_2 = O_B\left(\beta\sqrt{\log n}\right) \right) \ge \frac{\rho^8}{2(2\pi)^{7d/2} \exp(4\pi)},$$

completing the proof.

References

1. K. Costello, *Bilinear and quadratic variants on the Littlewood-Offord problem*, Submitted, `arxiv.org/abs/0902.1538`.
2. K. Costello, T. Tao, and V. Vu, *Random symmetric matrices are almost surely non-singular*, Duke Math. J. **135** (2006).
3. P. Erdős, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51** (1945), 898–902.
4. C. G. Esséen, *On the Kolmogorov-Rogozin inequality for the concentration function*, Z. Wahrsch. Verw. Gebiete **5** (1966), 210–216.
5. G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Period. Math. Hungar. **8** (1977), no. 3/4, 197–211.
6. D. Kleitman, *On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors*, Advances in Math. **5** (1970), 155–157.
7. 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**.
8. H. Nguyen, *Inverse Littlewood-Offord problems and the singularity of random symmetric matrices*, to appear in Duke Math. J., `arxiv.org/abs/1101.3074`.
9. _____, *On the condition number of randomly perturbed Hermitian matrices*, in preparation.
10. _____, *On the least singular value of random symmetric matrices*, to appear in Electric J. of Prob., `arxiv.org/abs/1102.1476`.
11. H. Nguyen and V. Vu, *Optimal Littlewood-Offord theorems*, Advances in Math. **226** (2011), no. 6, 5298–5319.
12. J. Rosiński and G. Samorodnitsky, *Symmetrization and concentration inequality for multilinear forms with applications to zero-one laws for Lévy chaos*, Annals of Probability.
13. T. Tao and V. Vu, *Additive combinatorics*, Cambridge Univ. Press, 2006.

14. _____ , *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. **20** (2007), 603–628.
15. _____ , *Random matrices: The circular law*, Communication in Contemporary Mathematics **10** (2008), 261–307.
16. _____ , *From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices*, Bull. Amer. Math. Soc. **46** (2009), no. 3, 377–396.
17. _____ , *Inverse Littlewood-Offord theorems and the condition number of random matrices*, Annals of Mathematics (2) **169** (2009), no. 2, 595–632.
18. _____ , *Smooth analysis of the condition number and the least singular value*, Mathematics of Computation **79** (2010).

Department of Mathematics, University of Pennsylvania
209 South 33rd Street, Philadelphia, PA 19104, USA
E-mail address: hoing@math.upenn.edu