STUDIA MATHEMATICA

Online First version

The small ball probability for singular values of the inhomogeneous subgaussian random matrices

by

Guozheng Dai, Zeyan Song, and Hanchao Wang

Abstract. Let A be an $n \times n$ random matrix with independent centered nonconstant subgaussian entries. We give the small ball probability for the k-th smallest singular value of A when $k \in (c \log n, c\sqrt{n})$. Furthermore, we prove that for any $k < c\sqrt{n}$, the probability of the matrix having a rank lower than $n - k$ is sub-exponential. This extends the earlier results of Rudelson's 2024 paper by removing the assumption of the identical distribution of the entries across the matrix. Our model covers inhomogeneous matrices, allowing different subgaussian moments for the entries as long as their subgaussian moments have a standard upper bound. In past work, the assumption of i.i.d. entries was required due to the lack of least common denominators of the non-i.i.d. random matrix. We overcome this problem using the randomized least common denominator (RLCD) of Livshyts (2021).

1. Introduction. Let A be an $n \times n$ random matrix. A classical problem is to estimate the probability that the random matrix A is singular, i.e., $\mathsf{P}(\det(A) = 0)$. Consider a random matrix A with i.i.d. Rademacher entries (taking values ± 1 with probabilities $1/2$). A remarkable work [KKS95] of Kahn, Komlós, and Szemerédi states that

$$\mathsf{P}(\det(A) = 0) \le (0.998 + o(1))^n.$$

Subsequently, a lot of work was done to explore the asymptotically optimal upper bound. For example, the singularity probability was bounded by $(3/4 + O(1))^n$ by Tao and Vu [TV06, TV07], and further developed by Bourgain, Vu, and Wood [BVW10], who established a bound of $(2^{-1/2} + o(1))^n$. Recently, Tikhomirov [Tik20] proved the following optimal bound:

$$\mathsf{P}(\det(A) = 0) = \left(\frac{1}{2} + o(1)\right)^n.$$

2020 *Mathematics Subject Classification*: Primary 15B52.
Key words and phrases: inhomogeneous random matrix, rank, singular value.
Received 19 January 2025; revised 21 May 2026.
Published online 4 November 2026.

DOI: 10.4064/sm260119-15-6 © Instytut Matematyczny PAN, 2026

Furthermore, an interesting extension of the problem of estimating the probability of a random matrix's singularity is investigating the distribution of the smallest singular value; in particular, we want to give a small ball probability. Recall that the singular values of A are defined as square roots of the eigenvalues of the positive semidefinite $n \times n$ matrix $A^\top A$:

$$s_i(A) = \sqrt{\lambda_i(A^\top A)}, \quad i = 1, \dots, n,$$

where we assume the non-increasing ordering $\lambda_1(A^\top A) \geq \cdots \geq \lambda_n(A^\top A)$. Spielman and Teng [ST02] conjectured that, when A is an i.i.d. Rademacher random matrix, then for $\varepsilon \geq 0$,

$$\mathsf{P}(s_n(A) \leq \varepsilon n^{-1/2}) \leq \varepsilon + \exp(-cn).$$

In the past 20 years, much work has been done around this conjecture. In particular, Rudelson and Vershynin [RV08] proved the Spielman–Teng conjecture to a constant. In particular, they obtained, for $\varepsilon \geq 0$,

$$\mathsf{P}(s_n(A) \leq \varepsilon n^{-1/2}) \leq C\varepsilon + \exp(-cn), \tag{1.1}$$

which is the first exponentially small ball probability. In fact, they gave a more general result, which states that the above deviation inequality is valid when the entries of A are i.i.d. subgaussian random variables with mean 0 and variance 1. Recall that a random variable X is called *subgaussian* if

$$\mathsf{E} \exp((X/K)^2) < \infty \tag{1.2}$$

for some $K > 0$ and we define the *subgaussian norm* by

$$\|X\|_{\psi_2} = \inf (t > 0 : \mathsf{E}[\exp((X/t)^2)] \leq 2).$$

There is also considerable interest in computing the distribution of $s_n(A)$ in a general ensemble. Rebrova and Tikhomirov [RT18] recovered the full strength of (1.1), assuming only i.i.d. mean zero, variance one entries. A recent paper [LTV21] by Livshyts, Tikhomirov, and Vershynin established the deviation inequality of $s_n(A)$ when A has independent but not identically distributed heavy-tailed entries, which may be the most mild assumption in this context. There are many other interesting works (e.g., see [Liv21a, RV09]) computing the distribution of $s_n(A)$, which we will not cover here.

Considering the behavior of the intermediate singular values is also an interesting and natural problem. Nguyen [Ngu18] considered the distribution $s_{n-k+1}(A)$ where A has i.i.d. subgaussian entries with mean 0 and variance 1 and proved that for any $\varepsilon > 0$, $\gamma \in (0, 1)$, and $k \in (c_0, n/c_0)$,

$$\mathsf{P}(s_{n-k+1}(A) \leq \varepsilon n^{-1/2}) \leq (C\varepsilon/k)^{(1-\gamma)k^2} + \exp(-cn).$$

Recently, Dai et al. [DSW24] improved the above decay rate. In particular, they proved for $\varepsilon > 0$, $\gamma \in (0, 1/2)$, and $\log n \leq k \leq c_0\sqrt{n}$,

$$\mathsf{P}(s_{n-k+1}(A) \leq \varepsilon n^{-1/2}) \leq (C\varepsilon/k)^{(1-\gamma)k^2} + \exp(-ckn).$$

As we have seen, all the results mentioned above revolve around i.i.d. random matrices. What about inhomogeneous random matrices? The first main result of this paper shows that the small ball probability for the k-th singular value of a random matrix whose entries follow different distributions behaves similarly to the i.i.d. case. Assume that a random variable X satisfies

$$\mathsf{E}X = 0, \quad \mathsf{E}X^2 = 1, \quad \|X\|_{\psi_2} \le K. \tag{1.3}$$

THEOREM 1.1. *Let A be an $n \times n$ random matrix with independent entries satisfying* (1.3). *For any fixed $\gamma \in (0,1)$ we have, for $\varepsilon > 0$ and $\log n \le k \le d_{1.1}\sqrt{n}$,*

$$\mathsf{P}\left(s_{n-k+1}(A) \le \frac{\varepsilon}{\sqrt{n}} \right) \le \left(\frac{C_{1.1}\varepsilon}{k} \right)^{\gamma k^2} + e^{-c_{1.1}kn},$$

where $d_{1.1}, C_{1.1}, c_{1.1}$ are some constants depending only on γ and K.

Similarly to the smallest singular value, we note that if the k-th smallest singular value is equal to zero, then the rank of the matrix is at most $n-k$. Consequently, the small ball probability of the k-th smallest singular value can be used to estimate the rank of the matrix.

Kahn, Komlós, and Szemerédi showed that the probability that the Rademacher random matrix A has a lower rank than $n-k$ is $O(f(k))^n$, where $\lim_{k\to\infty} f(k) = 0$. When the entries of A are i.i.d. subgaussian variables (including the Rademacher variables), Rudelson [Rud24] very recently proved for $k < c\sqrt{n}$,

$$\mathsf{P}(\operatorname{rank}(A) \le n-k) \le e^{-ckn}.$$

Our second main result is the following.

THEOREM 1.2. *Let $k, n \in \mathbb{N}$ be such that $k < d_{1.1}\sqrt{n}$ and let A be an $n \times n$ matrix with independent entries satisfying* (1.3). *Then*

$$\mathsf{P}(\operatorname{rank}(A) \le n-k) \le \exp(-c_{1.2}kn),$$

where $d_{1.1}$, $c_{1.2} > 0$ are some constants depending only on K.

REMARK 1.3. In fact, a direct application of Theorem 1.1 only yields the part $k > c\log n$. In the subsequent proof we will instead establish Theorem 1.2 first, because the case of sufficiently small n in Theorem 1.1 will be derived from Theorem 1.2; see the proof of Theorem 1.1 in Section 4 for details.

To conclude the introduction, we briefly describe our main techniques. Firstly, we will use the geometric approach for the k-th singular value of a random matrix, which allows us to estimate the small ball probability via the distance from a random vector to a random subspace. Next, we focus on the properties of the random subspace. The main technique is the "almost orthogonal system" from [Rud24]. Ultimately, we reduce the problem to the

small ball probabilities of vectors in "k-dimensional space", and we must simultaneously characterize both the size of vectors and the magnitude of these small ball probabilities. On the one hand, we introduce a *randomized log-least common denominator* (RLCD) to study the small ball probability of random matrices. The RLCD is used to estimate the small ball probability in linear spaces, which overcomes the issues caused by random variables with different distributions. On the other hand, we use the random rounding method to discretize certain closed sets within linear spaces. During the discretization process, we need to estimate the size of the discretized network. Due to the influence of RLCD, we need to calculate the number of points in the discretized network that are not too far away in "distance" from the "integer lattice". RLCD is the key to proving our result and will be detailed in Section 3.

The rest of this paper is organized as follows. In Section 2, we will introduce some notation and necessary lemmas. In Section 3, we shall prove some properties of subspaces in $\mathbb{R}^n$, which are important for the proof of our main results. Finally, we complete the proofs of Theorems 1.2 and 1.1 in Section 4.

2. Preliminaries

2.1. Notation. We denote by $[n]$ the set of natural numbers from 1 to n. For a vector $x \in \mathbb{R}^n$, its *Euclidean norm* is defined as $\|x\|_2 = (\sum_{j\in[n]} x_j^2)^{1/2}$, and its *supnorm* is defined as $\|x\|_\infty = \max_i |x_i|$. The unit sphere in $\mathbb{R}^n$ is denoted by S^{n-1}, and the cardinality of a finite point set $I \subset \mathbb{R}^n$ is denoted by $|I|$.

For on $m \times l$ matrix V, we write $\mathrm{Row}_i(V)$ for its i-th row and $\mathrm{Col}_j(V)$ for its j-th column. Its singular values are denoted by

$$s_1(V) \geq \cdots \geq s_m(V) \geq 0.$$

The *Euclidean operator norm* of V is

$$\|V\| = \max_{x\in S^{l-1}} \|Vx\|_2,$$

and the *Hilbert–Schmidt* norm is

$$\|V\|_{\mathrm{HS}} = \Big(\sum_{i=1}^{m}\sum_{j=1}^{l} v_{i,j}^2\Big)^{1/2}.$$

Note that $\|V\| = s_1(V)$ and $\|V\|_{\mathrm{HS}} = (\sum_{j=1}^m s_j(V)^2)^{1/2}$.

For a random variable X we denote by $\overline{X}$ the symmetrization of X defined as $\overline{X} = X - X'$, where X' is an independent copy of X. Note that

$$\mathsf{E}|\overline{X}|^2 = 2\,\mathrm{Var}(X). \tag{2.1}$$

We denote by $\mathcal{L}(X,t)$ the *Lévy concentration function* of a random vector $X \in \mathbb{R}^m$:

$$\mathcal{L}(X,t) = \sup_{y\in\mathbb{R}^m} \mathsf{P}(\|X-y\|_2 \le t).$$

For $x, y \in \mathbb{R}^n$, we define the *Schur product* of x and y as

$$x \star y = (x_1y_1, \dots, x_ny_n)^\top.$$

Denote by c, $c', \dots$ the universal constants and by $c(u)$, $C(u)$ the constants depending only on u. Their values may change from line to line.

2.2. Decomposition of the sphere. We will first introduce the classic space partition in this field, as specifically defined by Rudelson and Vershynin [RV08].

Definition 2.1. Let $\delta, \rho \in (0,1)$. We define the sets of *sparse*, *compressible* and *incompressible vectors* as follows:

- $\mathrm{Sparse}(\delta) := \{x \in \mathbb{R}^n : |\mathrm{supp}(x)| \le \delta n\}$;
- $\mathrm{Comp}(\delta,\rho) := \{x \in S^{n-1} : \mathrm{dist}(x, \mathrm{Sparse}(\delta)) \le \rho\}$;
- $\mathrm{Incomp}(\delta,\rho) := S^{n-1} \setminus \mathrm{Comp}(\delta,\rho)$.

2.3. The norm of random matrices and tensorization. In this subsection, we assume that the entries in the matrix A are independent and satisfy (1.3). Without loss of generality, we may assume that $K \ge 1$.

We first introduce the following tensorization property.

Lemma 2.2 ([RV08, Lemma 2.2]). *Let $X_1, \dots, X_n$ be independent nonnegative random variables, and let M, $m > 0$ with $\mathsf{P}(X_j \le s) \le (Ms)^m$ for all $s \ge s_0$. Then*

$$\mathsf{P}\Big(\sum_{j=1}^n X_j \le nt\Big) \le (C_{2.2}Mt)^{mn} \quad \text{for all } t \ge s_0,$$

where $C_{2.2}$ is a constant.

Next, we will give some inequalities for the norm of random matrices.

Lemma 2.3 ([RV08, Proposition 2.5]). *Let $m \le n$, and Q be an $m \times n$ random matrix with centered independent entries $q_{i,j}$ such that $|q_{i,j}| \le 1$. Then*

$$\mathsf{P}(\|Q\| \ge C_{2.3}\sqrt{n}) \le \exp(-c_{2.3}n),$$

where $C_{2.3}$ and $c_{2.3}$ are constants.

Lemma 2.4 ([Rud24, Lemma 3.6]). *Let $m \le n$ and let A be an $m \times n$ matrix whose entries are independent and satisfy* (1.3)*. We have*

$$\mathsf{P}(\|A\|_{\mathrm{HS}} \ge 2Kn) \le \exp(-c_{2.4}n^2),$$

where $c_{2.4}$ is a constant depending only on K.

2.4. Randomized least common denominators. The least common denominator of a vector of $\mathbb{R}^n$ was first introduced in Rudelson and Vershynin [RV08] as a valuable tool for estimating the small ball probability (the Lévy function of the inner product of a real-valued vector and a random vector with independent entries). To prove the main result, we need to estimate the small ball probability of orthogonal projection, so we need the small ball probability for a linear subspace similar to [RV16, Section 7]. However, we know that the property of identically distributed entries in a random vector is the key to estimating small ball probability. Thus, we need a new anti-concentration inequality to characterize the small ball probabilities with different distributions. In the following, we will give the randomized least common denominators introduced in [LTV21] to overcome this problem.

DEFINITION 2.5. Let V be an $m \times n$ (deterministic) matrix, $\xi = (\xi_1, \dots, \xi_n)$ let be a real-valued random vector with independent entries satisfying (1.3) and let $L > 0$, $\alpha \in (0, 1)$. Define the *randomized log-least common denominator* (RLCD) of V and ξ by

$$RD^{\xi}_{L,\alpha}(V) = \inf\left\{\|\theta\|_2 : \theta \in \mathbb{R}^m, \mathsf{E}\,\mathrm{dist}^2(V^T\theta \star \overline{\xi}, \mathbb{Z}^n) < L^2 \log_+ \frac{\alpha\|V^T\theta\|_2}{L}\right\},$$

where $\log_+ x = \max\{0, \log x\}$.

If $E \subset \mathbb{R}^n$ is a linear subspace, we can adapt this definition to the orthogonal projection P_E to E setting

$$RD^{\xi}_{L,\alpha}(E) = \inf\left\{\|y\|_2 : y \in E,\, \mathsf{E}\,\mathrm{dist}^2(y \star \overline{\xi}, \mathbb{Z}^n) < L^2 \log_+ \frac{\alpha\|y\|_2}{L}\right\}.$$

Moreover, let A be an $n \times t$ matrix with columns $A_1, \dots, A_t$, and A_j having independent entries satisfying (1.3). Then, we define the RLCD of V and A by

$$RD^{A}_{L,\alpha}(V) = \min_{j\in[t]} RD^{A_j}_{L,\alpha}(V).$$

We give the following result to estimate the small ball probability using RLCD.

PROPOSITION 2.6 (Small ball probability via RLCD). *Consider a real-valued random vector $\xi = (\xi_1, \dots, \xi_n)$ with independent entries satisfying* (1.3) *and $V \in \mathbb{R}^{m\times n}$. Then there exists a universal constant $c_{2.6} > 0$ such that for any $L \geq c_{2.6}\sqrt{m}$, we have*

$$\mathcal{L}(V\xi, t\sqrt{m}) \leq \frac{(C_{2.6}L/(\alpha\sqrt{m}))^m}{\det(VV^T)^{1/2}}\left(t + \frac{\sqrt{m}}{RD^{\xi}_{L,\alpha}(V)}\right)^m, \qquad t \geq 0, \tag{2.2}$$

where $C_{2.6}$ is an absolute constant.

Proof. By Esseen's inequality for the Lévy concentration function of a general random vector Y,

$$\mathcal{L}(Y, \sqrt{m}) \le C^m \int\limits_{B(0,\sqrt{m})} |\phi_Y(\theta)| \, \mathrm{d}\theta, \tag{2.3}$$

where $\phi_Y(\theta) = \mathsf{E} \exp(2\pi \mathrm{i} \langle \theta, Y \rangle)$ is the characteristic function of Y and $B(0, \sqrt{m})$ is the ball of radius $\sqrt{m}$ centered at 0. Set $Y = t^{-1} V\xi$ and assume that $\mathrm{Col}_k(V) = V_k$. Then

$$\langle \theta, Y \rangle = \sum_{k=1}^n t^{-1} \langle \theta, V_k \rangle \xi_k \quad \text{and} \quad \phi_Y(\theta) = \prod_{k=1}^n \phi_k(t^{-1} \langle \theta, V_k \rangle),$$

where $\phi_k(s) := \mathsf{E} \exp(2\pi \mathrm{i} s \overline{\xi}_k)$.

Applying Esseen's inequality and (2.3) yields

$$\mathcal{L}(V\xi, t\sqrt{m}) \le C^m \int\limits_{B(0,\sqrt{m})} \prod_{k=1}^n |\phi_k(t^{-1} \langle \theta, V_k \rangle)| \, \mathrm{d}\theta. \tag{2.4}$$

Note that for any $s \in \mathbb{R}$,

$$|\phi_k(s)|^2 = \mathsf{E} \exp(2\pi \mathrm{i} s \overline{\xi}_k) = \mathsf{E} \cos(2\pi s \overline{\xi}_k).$$

Then for each $k \le n$, we have

$$|\phi_k(s)| \le \exp\biggl(-\frac{1}{2} \mathsf{E}[1 - \cos(2\pi s \overline{\xi}_k)] \biggr),$$

using the inequality $|x| \le \exp\bigl(-\frac{1}{2}(1 - x^2)\bigr)$, valid for any $x \in \mathbb{R}$.

Moreover, for any $s \in \mathbb{R}$ and ξ_k satisfying (1.3), we have

$$\mathsf{E}[1 - \cos(2\pi s \overline{\xi}_k)] \ge \tilde{c} \, \mathsf{E} \, \mathrm{dist}^2(s \overline{\xi}_k, \mathbb{Z}).$$

Now, assume that

$$t \ge t_0 = \frac{\sqrt{m}}{R D^{\xi}_{L,\alpha}(V)}.$$

For $\theta \in B(0, \sqrt{m})$,

$$\|\theta / t\|_2 \le R D^{\xi}_{L,\alpha}(V).$$

Then (2.4) yields

$$\begin{aligned}
\mathcal{L}(V\xi, t\sqrt{m}) &\le C^m \int\limits_{B(0,\sqrt{m})} \prod_{k=1}^n |\phi_k(t^{-1} \langle \theta, V_k \rangle)| \, \mathrm{d}\theta \\
&\le C^m \int\limits_{B(0,\sqrt{m})} \exp\biggl(-\frac{\tilde{c}}{2} \mathsf{E} \, \mathrm{dist}^2(V^T \theta \star \overline{\xi}/t, \mathbb{Z}^n) \biggr) \, \mathrm{d}\theta \\
&\le C^m \int\limits_{B(0,\sqrt{m})} \exp\biggl(-cL^2 \log_+ \frac{\alpha \|V^T \theta\|_2}{Lt} \biggr) \, \mathrm{d}\theta.
\end{aligned}$$

Set $z = V^T\theta$. Then

$$\mathcal{L}(V\xi, t\sqrt{m}) \le \frac{(CLt/\alpha)^m}{\det(VV^T)^{1/2}} \int\limits_{\mathbb{R}^m} \exp(-cL^2 \log_+ \|z\|_2)\,\mathrm{d}\theta.$$

Since

$$\int\limits_{\mathbb{R}^m} \exp(-cL^2 \log_+ \|z\|_2)\,\mathrm{d}z \le \left(\frac{C}{\sqrt{m}}\right)^m,$$

we have

$$\mathcal{L}(V\xi, t\sqrt{m}) \le \frac{(CL/\alpha\sqrt{m})^m}{\det(VV^T)^{1/2}} t^m,$$

which completes the proof. ∎

Moreover, we have the following corollary for orthogonal projection onto a linear subspace.

COROLLARY 2.7. *Consider a real-valued random vector ξ as in Proposition* 2.6*, let E be a subspace of $\mathbb{R}^n$ with $\dim E = m$, and let P_E denote the orthogonal projection to E. Then there exists a universal constant $c_{2.7} > 0$ and for every $L \ge c_{2.7}\sqrt{m}$ we have*

$$\mathcal{L}(P_E\xi, t\sqrt{m}) \le \left(\frac{C_{2.7}L}{(\alpha\sqrt{m})}\right)^m \left(t + \frac{\sqrt{m}}{RD^{\xi}_{L,\alpha}(E)}\right)^m, \quad t \ge 0,$$

where $C_{2.7}$ is an absolute constant.

We will first present some properties of the random variables that satisfy (1.3).

LEMMA 2.8. *Let ξ be a random variable satisfying* (1.3)*. Then*

$$\mathsf{P}(|\overline{\xi}| \ge 1) \ge p_{2.8}, \tag{2.5}$$

where $p_{2.8} := p(K)$ is a positive constant depending only on K.

Proof. Let $p_{2.8} := (6 + 4K^4)^{-1}$ and define $X := \overline{\xi}^2$. Note that $\mathsf{E}X = 2$. Applying the Paley–Zygmund inequality yields

$$\mathsf{P}(X \ge 1) \ge \left(1 - \frac{1}{2}\right)^2 \frac{(\mathsf{E}X)^2}{\mathsf{E}X^2} \ge (\mathsf{E}X^2)^{-1}.$$

Note that $\mathsf{E}X^2 = 6 + 2\mathsf{E}\xi^4$ and

$$1 + \frac{\mathsf{E}\xi^4}{2K^4} \le \mathsf{E}e^{\xi^2/K^2} \le 2.$$

Combining the inequalities above, we can derive

$$\mathsf{P}(|\overline{\xi}| \ge 1) \ge (6 + 4K^4)^{-1} = p_{2.8}.$$ ∎

Later, we will fix $p = p_{2.8}$, and with the help of the above lemma, we can prove that the RLCD of the incompressible part is relatively large.

LEMMA 2.9. *Let δ, $\rho \in (0,1)$, let A be an $n \times n$ random matrix as in Theorem* 1.1 *and let U be an $n \times l$ deterministic matrix such that $(U\mathbb{R}^l) \cap S^{n-1} \subset \mathrm{Incomp}(\delta, \rho)$. Then there exists $h_{2.9} = h_{2.9}(\delta, \rho, K) > 0$ such that for any $\theta \in \mathbb{R}^l$ with $\|U\theta\|_2 \le h_{2.9}\sqrt{n}$ and $j \in [n]$, satisfies*

$$\mathsf{E}\,\mathrm{dist}^2(U\theta \star \overline{\mathrm{Col}_j(A)}, \mathbb{Z}^n) \ge L^2 \log_+ \frac{\alpha\|U\theta\|_2}{L}, \tag{2.6}$$

where $\alpha \le \alpha_{2.9} = \alpha_{2.9}(\delta, \rho, K)$.

Proof. Assume that there exists $j \in [n]$ such that

$$\mathsf{E}_A\,\mathrm{dist}^2(U\theta \star \overline{\mathrm{Col}_j(A)}, \mathbb{Z}^n) < L^2 \log_+ \frac{\alpha\|U\theta\|_2}{L},$$

for some $\theta \in \mathbb{R}^l$; we want to prove that $\|U\theta\|_2 > h\sqrt{n}$.

Set $\|U\theta\|_2 = t$, $U\theta/t = u$, $\mathrm{Col}_j(A) = X = (X_1, \dots, X_n)$. Applying $\log_+ s < s^2$, we have

$$\mathsf{E}_X\,\mathrm{dist}^2(tu \star \overline{X}, \mathbb{Z}^n) < \alpha^2 t^2.$$

Let $q \in \mathbb{Z}^n$ denote the closest integer vector to $tu \star \overline{X}$ and rewrite the above inequality as

$$\mathsf{E}|u \star \overline{X} - q/t|^2 < \alpha^2.$$

By Markov's inequality,

$$|u \star \overline{X} - q/t|^2 < c^2\alpha^2$$

with probability at least $1 - 1/c^2$. Using Markov's inequality again shows that

$$|u_i\overline{X}_i - q_i/t| < \frac{c\alpha}{c_1\sqrt{n}} \quad \text{for any } i \in J_1,$$

where J_1 is some subset of $[n]$ and $|J_1| \ge n - c_1 n$.

Furthermore, since $\mathsf{E}|\overline{X}_j|^2 = 2\,\mathrm{Var}\,|X_j| \le 2K$, applying Markov's inequality shows that, with probability at least $1 - 1/c^2$,

$$|\overline{X}_i| \le \frac{c}{c_2}\sqrt{2K} \quad \text{for any } i \in J_2,$$

where J_2 is some subset of $[n]$ and $|J_2| \ge n - c_2 n$.

Moreover, applying Lemma 3.4 of Rudelson and Vershynin [RV08], there exists a set

$$J_3 := \left\{ i : \frac{\rho}{\sqrt{2n}} \le |u_i| \le \frac{1}{\sqrt{\delta n}} \right\},$$

satisfying $|J_3| \ge \frac{1}{2}\rho^2\delta n$.

Finally, Lemma 2.8 shows that $\mathsf{P}(|\overline{X}_i| \ge 1) \ge p$. Let $J_4 \subseteq J_3$ be such that

$$J_4 := \{i \in J_3 : |\overline{X}_i| \ge 1\}.$$

Set $|J_3| = m$ and apply the Paley–Zygmund inequality. Then we have

$$\mathsf{P}(|J_4| \geq pm/2) \geq \mathsf{P}(|J_4| \geq \mathsf{E}|J_4|/2) \geq \frac{(\mathsf{E}|J_4|)^2}{4\mathsf{E}|J_4|^2} \geq p^2/4,$$

where $|J_4| \leq m$ and $\mathsf{E}|J_4| \geq pm$.

This means that there exists $J_4 \subset J_3$, which satisfies $|J_4| \geq \frac{p}{2}|J_3|$ with at least $p^2/4$ probability

$$|\overline{X}_i| \geq 1 \quad \text{for any } i \in J_4.$$

We can choose the constants c, c_1, c_2 that depend only on δ, ρ, p such that

$$2(1 - 1/c^2) + p^2/4 > 2, \qquad n - c_1 n + n - c_2 n + \frac{p}{4}\rho^2\delta n > 2n.$$

Then, there exists a coordinate i for which we have simultaneously the following three bounds:

$$|u_i\overline{X}_i - q_i/t| \leq \frac{c\alpha}{c_1\sqrt{n}}, \quad 1 \leq |\overline{X}_i| \leq \frac{c}{c_2}\sqrt{2K}, \quad \frac{\rho}{\sqrt{2n}} \leq |u_i| \leq \frac{1}{\sqrt{\delta n}}.$$

Furthermore, using the triangle inequality, we get

$$\left|\frac{q_i}{t}\right| \geq |u_i\overline{X}_i| - \frac{c\alpha}{c_1\sqrt{n}} \geq \frac{\rho}{\sqrt{2n}} - \frac{c\alpha}{c_1\sqrt{n}} > 0,$$

where $0 < \alpha < \alpha_{2.9} = \frac{c_1\rho}{c\sqrt{2}}$.

Thus $q_i \neq 0$, $q_i \in \mathbb{Z}$, we necessarily have $|q_i| \geq 1$. Furthermore

$$\left|\frac{q_i}{t}\right| \leq |u_i\overline{X}_i| + \frac{c\alpha}{c_1\sqrt{n}} < \frac{c\sqrt{2K}}{c_2\sqrt{\delta n}} + \frac{c\alpha}{c_1\sqrt{n}} < \frac{c}{c_2}\frac{\sqrt{2K}}{\sqrt{\delta n}} + \frac{\rho}{\sqrt{2n}},$$

where $h_{2.9}^{-1} = \frac{c}{c_2}\sqrt{2K/\delta} + \frac{\rho}{\sqrt{2}} > 0$. Then $|t| \geq |q_i| h_{2.9}\sqrt{n} \geq h_{2.9}\sqrt{n}$. ∎

2.5. The number of integer points inside a ball. We will need to estimate the number of integer points in a ball in $\mathbb{R}^n$. The set $B(0, R)$ is the ball of radius R centered at 0.

LEMMA 2.10 ([Rud24, Lemma 3.12]). *For any $R > 0$,*

$$|\mathbb{Z}^n \cap B(0, R)| \leq (2 + C_{2.10}R/\sqrt{n})^n,$$

where $C_{2.10} > 0$ is an absolute constant.

2.6. Almost orthogonal systems of vectors. In our paper, we need to control the arithmetic structure of the kernel of some random matrix B. This structure is essentially derived from the randomized least common denominator (RLCD). This forces us to look for a way to divide the subspace. At the same time, we want to find a suitable representative vector in the subspace to estimate the probability that such vectors are in the kernel. We will employ the result of Section 3.1 in Rudelson [Rud24], which solves the problem above.

DEFINITION 2.11. Let $\nu \in (0,1)$. An l-tuple of vectors $(v_1, v_2, \dots, v_l) \subset \mathbb{R}^n \setminus \{0\}$ is called *ν-almost orthogonal* if the $n \times l$ matirx V_0 with $\mathrm{Col}_j(V_0) = v_j/\|v_j\|_2$ satisfies

$$1 - \nu \le s_l(V_0) \le s_1(V_0) \le 1 + \nu.$$

The following lemma shows how to divide the linear subspace E into $E \cap W$, where $W \subset \mathbb{R}^n \setminus \{0\}$ is a closed set, and the linear subspace $F \subset E$ of high dimension. This lemma is critical for estimating the arithmetic structure.

LEMMA 2.12 ([Rud24, Lemma 3.3]). *Let $W \subset \mathbb{R}^n \setminus \{0\}$ be a closed set. Let $l < k \le n$, and let $E \subset \mathbb{R}^n$ be a linear subspace of dimension k. Then at least one of the following holds:*

(1) *There exist vectors $v_1, \dots, v_l \in E \cap W$ such that*
   - *the l-tuple $(v_1, \dots, v_l)$ is $\frac{1}{8}$-almost orthogonal;*
   - *for any $\theta \in \mathbb{R}^l$ such that $\|\theta\|_2 \le \frac{1}{20\sqrt{l}}$,*

$$\sum_{i=1}^{l} \theta_i v_i \notin W$$

(2) *There exists a subspace $F \subset E$ of dimension $k - l$ such that $F \cap W = \varnothing$.*

REMARK 2.13. The assumption $k < c\sqrt{n}$ in Theorem 1.2 is based on the condition $\|\theta\|_2 \le \frac{1}{20\sqrt{l}}$ in Lemma 2.12 and Proposition 3.1 (see more details in Section 3).

2.7. Restricted invertibility phenomenon. In this subsection, we will introduce some properties that will be used to estimate singular values of deterministic matrices.

LEMMA 2.14 ([NY17, Theorem 6]). *Assume that M is a full-rank matrix of size $k \times d$ with $k \le d$. Then for $1 \le l \le k-1$, there exist l different indices $i_1, \dots, i_l$ such that the matrix $M_{i_1,\dots,i_l}$ with columns $\mathrm{Col}_{i_1}(M), \dots, \mathrm{Col}_{i_l}(M)$ has the smallest non-zero singular value $s_l(M_{i_1,\dots,i_l})$ satisfying*

$$s_l(M_{i_1,\dots,i_l})^{-1} \le C_{2.14} \min_{r \in \{l+1,\dots,k\}} \sqrt{\frac{dr}{(r-l)\sum_{i=r}^{k} s_i(M)^2}},$$

where $C_{2.14}$ is an absolute constant.

2.8. Estimate of the compressible case. Here we aim to prove that the kernel of B is unlikely to contain an extensive, almost orthogonal system of compressible vectors.

LEMMA 2.15 ([Rud24, Proposition 4.2]). *Let $k, n \in \mathbb{N}$ be such that $k < n/2$ and let B be an $(n-k) \times n$ matrix whose entries are independent random variables satisfying* (1.3). *There exists $\tau = \tau_{2.15} > 0$ such that the probability*

that there exists a $\frac{1}{4}$*-almost orthogonal* l*-tuple* $x_1, \dots, x_l \in \mathrm{Comp}(\tau^2, \tau^4)$ *with* $l \le \tau^3 n$ *and*

$$\|Bx_j\|_2 \le \tau\sqrt{n} \quad \text{for all } j \in [l]$$

is less than $\exp(-c_{2.15} l n)$*, where* $c_{2.15}$ *is a constant.*

REMARK 2.16. The proof of Proposition 4.2 in [Rud24] does not rely on the variables being identically distributed, so the lemma can be invoked as stated. What it does require is a uniform lower bound on the second moments; this is why we take the variances of the matrix entries to be 1. Any other positive lower bound would work equally well.

3. Incompressible vectors. Before stating the main result of this section, we single out the linear subspace spanned by incompressible vectors in $\ker B$. Inside this subspace, "large RLCD" should be typical behavior; more precisely, we want the probability that a vector has small RLCD to be super-exponentially small. As the discussion in Section 4 shows, the key obstruction is the existence of an approximately orthonormal l-tuple of vectors, all with small RLCD. The main result below asserts that this event is indeed super-exponentially unlikely.

We now introduce several important constants before presenting our main results.

$$L = c_{2.6}\sqrt{k}, \qquad \alpha = \min\{\alpha_{2.9}(\tau^2, \tau^4, K), \alpha_{2.9}(\tau^2, \tau^4/2, K)\},$$

and

$$r = \min\{h_{2.9}(\tau^2, \tau^4, K)/2, h_{2.9}(\tau^2, \tau^4/2, K)/8\},$$

where $c_{2.6}$ is an absolute constant from Proposition 2.6, k appears in Theorem 1.2, K is a parameter from (1.3), and τ is chosen by Lemma 2.15.

For ease of writing, we set

$$\mathrm{d}_A(x, \mathbb{Z}^n) = \min_{1 \le i \le n} \sqrt{\mathsf{E}\,\mathrm{dist}^2(x \star \overline{\mathrm{Col}_i(A)}, \mathbb{Z}^n)}.$$

PROPOSITION 3.1. *Let* $\rho \in (0, \rho_0)$*, where* $\rho_0 = \rho_0(\tau, K)$ *is some positive number. Assume that* $l \le k \le \frac{\rho}{2}\sqrt{n}$*. Let* B *be an* $(n-k) \times n$ *matrix with independent entries satisfying* (1.3)*. Consider the event* $\mathcal{E}_{3.1}$ *that there exist vectors* $v_1, \dots, v_l \in \ker B$ *having the following properties:*

(1) $2r\sqrt{n} \le \|v_j\|_2 \le R := \exp\left(\frac{\rho^2 n}{4L^2}\right)$ *for all* $j \in [l]$*;*
(2) $\mathrm{span}(v_1, \dots, v_l) \cap S^{n-1} \subset \mathrm{Incomp}(\tau^2, \tau^4)$*;*
(3) *the vectors* $v_1, \dots, v_l$ *are* $\frac{1}{8}$*-almost orthogonal system;*
(4) $\mathrm{d}_A(v_j, \mathbb{Z}^n) \le \rho\sqrt{n}$ *for all* $j \in [l]$*;*
(5) *the* $n \times l$ *matrix* V *with columns* $v_1, \dots, v_l$ *satisfies*

$$\mathrm{d}_A(V\theta, \mathbb{Z}^n) > \rho\sqrt{n}$$

for all $\theta \in \mathbb{R}^l$ *such that* $\|\theta\|_2 \le \frac{1}{20\sqrt{l}}$ *and* $\|V\theta\|_2 \ge 2r\sqrt{n}$*.*

Then

$$\mathsf{P}(\mathcal{E}_{3.1}) \le \exp(-ln).$$

Let

$$\mathcal{W} = \{(w_1, \dots, w_l) : w_i \text{ satisfy (1)–(5)}\}.$$

Therefore, we can obtain Proposition 3.1 by proving that:

$$\mathsf{P}\big(\exists w_1, \dots, w_l \in \ker B \text{ with } (w_1, \dots, w_l) \in \mathcal{W}\big) \le \exp(-ln).$$

To prove the above inequality, first consider a vector $d = (d_1, \dots, d_l) \in [r\sqrt{n}, R]^l$ with $d_i = 2^{s_i}$ for some integers s_i. Then, define $\mathcal{W}_d$ to be the subset of $\mathcal{W}$ consisting of all $(v_1, \dots, v_l)$ satisfying

$$\|v_j\|_2 \in [d_j, 2d_j] \quad \text{for all } j \in [l].$$

Now, we only need to estimate the intersection of $\mathcal{W}_d$ and $\ker B$, then we divide the proof of Proposition 3.1 into three parts. The first part estimates the size of the net that discretizes $\mathcal{W}_d$ and the second part begins with a detailed introduction on how to discretize $\mathcal{W}_d$, and the final part completes the proof of the proposition.

3.1. Size of the net. In this subsection, our main result is that the size of the net is super-exponentially large. First, we set a temporarily undetermined constant δ, to be determined in Section 3.2, which satisfies

$$0 < \delta \le \rho. \tag{3.1}$$

Proposition 3.2. *Let $d = (d_1, \dots, d_l)$ be a vector such that $d_j \in [r\sqrt{n}, R]$ for all $j \in [l]$. Let δ be as in* (3.1) *and $\mathcal{N}_d \subset (\delta\mathbb{Z}^n)^l$ be the set of all l-tuples of vectors $u_1, \dots, u_l$ such that for all $j \in [l]$,*

$$\|u_j\|_2 \in \left[\tfrac{1}{2}d_j, 4d_j\right] \quad \text{with } \mathrm{d}_A(u_j, \mathbb{Z}^n) < 2\rho\sqrt{n}$$

and

$$\operatorname{span}(u_1, \dots, u_l) \cap S^{n-1} \subset \operatorname{Incomp}(\tau^2, \tau^4/2).$$

Then

$$|\mathcal{N}_d| \le \left(\frac{C_{3.2}\rho^{c_{3.2}}}{r\delta}\right)^{ln} \left(\prod_{j=1}^{l} \frac{d_j}{\sqrt{n}}\right)^n,$$

where $C_{3.2}, c_{3.2} > 0$ depend only on τ, K.

To prove this proposition, we need to estimate the number of vectors in the set that satisfy the property $\mathrm{d}_A(x, \mathbb{Z}^n) < 2\rho\sqrt{n}$. For technical convenience in the proof, we slightly modify the above set. Let

$$\Lambda_{d_j}^J = \{\|u_{J^c}\|_2 \le 4d_j\} \cap \left\{u : \frac{4d_j}{\tau\sqrt{n}} \ge |u_i| \ge \frac{\tau^4 d_j}{8\sqrt{2n}}, \text{ for all } i \in J\right\} \cap \delta\mathbb{Z}^n,$$

where u_{J^c} is the vector obtained by zeroing out the coordinates of u in J and keeping the others unchanged. In fact, $\Lambda_{d_j}^J$ can be regarded as possessing

an approximate product structure. More precisely, we decompose a vector into two parts according to the index sets J and J^c. The component indexed by J^c can be viewed as a discrete lattice on the sphere, while the component indexed by J resembles a high-dimensional polytope. In particular, the coordinates in J are essentially decoupled, which leads to a product-type structure.

Note that, by [RV08, Lemma 3.4], there exists a constant $c_\tau > 0$ such that for any vector $v \in \mathrm{Incomp}(\tau^2, \tau^4/2)$, there exists a subset $I \subset [n]$ with $|I| \geq c_\tau n$ such that

$$\frac{\tau^4}{2\sqrt{2n}} \leq |v_i| \leq \frac{1}{\tau\sqrt{n}} \quad \text{for all } i \in I.$$

Therefore, if we ignore the condition $d_A(u, \mathbb{Z}^n) < 2\rho\sqrt{n}$ in $\mathcal{N}_d$, then each vector is contained in some $\Lambda_{d_j}^J$, where $J \subset [n]$ is a subset with cardinality at least $c_\tau n$.

Therefore, it suffices to estimate the number of vectors u in $\Lambda_{d_j}^J$ that satisfy $d_A(u, \mathbb{Z}^n) < 2\rho\sqrt{n}$, which is precisely the content of the following lemma.

LEMMA 3.3. *Let W be a vector uniformly distributed in the set $\Lambda_{d_j}^J$ for some $J \subset [n]$ with $|J| \geq c_\tau n$ and $d_j \in [r\sqrt{n}, R]$. Then for each $\mathrm{Col}_i(A)$, $i \in [n]$,*

$$\mathsf{P}_W\{\mathsf{E}_A \operatorname{dist}^2(W \star \overline{\mathrm{Col}_i(A)}, \mathbb{Z}^n) \leq (2\rho\sqrt{n})^2\} < (C_{3.3}\rho)^{c_{3.3}n}.$$

where $C_{3.3}, c_{3.3} > 0$ depend only on τ, K.

LEMMA 3.4. *For any $t \in (0,1)$ and $u > 1$, there exists $n_{3.4} = n(t, u, \tau, K)$ such that for $n \geq n_{3.4}$ and $|J| \geq c_\tau n$. If $x \in \mathbb{R}^n$ satisfies*

$$\|x\|_2^2 \leq \frac{c_\tau t}{2} u^2 n, \quad |\{i \in J : |x_i| \geq 1\}| \geq t|J|,$$

and if W is a vector uniformly distributed on the set $\Lambda_{d_j}^J$, then, for $\rho \geq \delta$, we have

$$\mathsf{P}_W\{\operatorname{dist}(W \star x, \mathbb{Z}^n)^2 \leq \rho^2 n\} \leq (C_{3.4}\rho)^{c_{3.4}n},$$

where $C_{3.4}, c_{3.4} > 0$ depend only on t, u, τ, K.

Proof. Assume that $I_0 = \{i \in J : |x_i| \geq 1\}$ and $I = \{i \in I_0 : |x_i| \leq u\}$. Then

$$u^2(|I_0| - |I|) \leq \sum_{i \in I_0 \setminus I} |x_i|^2 \leq \sum_{i \in I_0} |x_i|^2 \leq \frac{c_\tau t}{2} u^2 n.$$

We obtain $|I| \geq \frac{c_\tau t}{2} n$. Furthermore, set $Y = \delta^{-1} W$. Then for $i \in I$,

$$T_1 := \delta^{-1} \frac{\tau^4 d_j}{8\sqrt{2n}} \leq |Y_i| \leq T_2 := \delta^{-1} \frac{4 d_j}{\tau\sqrt{n}}.$$

Recall that W is uniformly distributed on the set $\Lambda_{d_j}^J$ and can be written as $W = \delta Y$. Then, for all $i \in I$, the random variables Y_i are independent and each is uniformly distributed on the integer set $[-T_2, -T_1] \cup [T_1, T_2]$.

Note that $|W_i x_i| \le T := \delta u T_2$ and we have

$$\begin{aligned}\mathsf{P}(\operatorname{dist}(W_i x_i, \mathbb{Z}) < \epsilon) &\le \sum_{q_i=-T}^{T} \mathsf{P}(|\delta Y_i x_i - q_i| \le \varepsilon) \\ &\le \sum_{q_i=-T}^{T} \mathsf{P}(|Y_i - \delta^{-1} q_i x_i^{-1}| \le \varepsilon/\delta x_i) \\ &\le \frac{(2T+1)(\varepsilon/\delta + 1)}{2(T_2 - T_1)} \le C\varepsilon,\end{aligned}$$

where $\varepsilon \ge \delta$ and $C = C(u, \tau) > 0$ is a constant depending only on τ and u.

Thus, for $s \ge \delta^2$,

$$\mathsf{P}(\operatorname{dist}(w_j x_j, \mathbb{Z})^2 \le s) \le C\sqrt{s}.$$

Applying Lemma 2.2 with $X_j = \operatorname{dist}(w_j x_j, \mathbb{Z})^2$ for all $j \in I$ and $m = 1/2$, we obtain, for any $\epsilon \ge \delta$,

$$\mathsf{P}(\operatorname{dist}(W \star x, \mathbb{Z}^n)^2 < \epsilon^2 |I|) \le \mathsf{P}\Big(\sum_{i \in I} \operatorname{dist}(w_i x_i, \mathbb{Z})^2 \le \epsilon^2 |I|\Big) \le (\tilde{c}\epsilon)^{|I|},$$

where $\tilde{c}$ is an absolute constant.

Lettig $\epsilon = \max\big\{\rho, \sqrt{\frac{2}{c_\tau t}}\rho\big\}$, we obtain

$$\mathsf{P}(\operatorname{dist}(W \star x, \mathbb{Z}^n)^2 \le \rho^2 n) \le (C\rho)^{cn}. \ \blacksquare$$

Proof of Lemma 3.3. It suffices to prove that, if W is a vector uniformly distributed on the set of $\Lambda_{d_j}^J$, then

$$\mathsf{P}_W(\mathsf{E}_{\overline{X}} \operatorname{dist}^2(W \star \overline{X}, \mathbb{Z}^n) \le 4\rho^2 n) \le (C\rho)^{cn},$$

where $X = (X_1, \dots, X_n)$ with X_i satisfying (1.3) for all $i \in [n]$.

Applying Lemma 2.8, $\mathsf{P}\{|\overline{X}_i| \ge 1\} \ge p$. On the one hand, with at least $1 - \frac{1-p}{1-t}$ probability,

$$|\{i \in J : |\overline{x}_i| \ge 1\}| \ge t|J|. \tag{3.2}$$

On the other hand, since $\|X_i\|_{\psi_2} \le K$,

$$\mathsf{P}(\|\overline{X}\|_2^2 \ge Cn) \le \exp((2\log 2 - C/K)n).$$

Thus, setting $C = \frac{c_\tau t}{2} u^2 > 2K \log 2$, we have

$$\|X\|_2^2 \le \frac{c_\tau t}{2} u^2 n \tag{3.3}$$

with probability at least $1 - \exp\big(\big(2\log 2 - \frac{c_\tau t u^2}{2K}\big)n\big)$.

Denote by $\mathcal{E}_{3.3}$ the event that both (3.2) and (3.3) occur. Then we can choose $t = t(\tau, K)$ and $u = u(\tau, K, t)$ such that

$$\mathsf{P}(\mathcal{E}_{3.3}) > p/2.$$

Thus, for every fixed X such that $X \in \mathcal{E}_{3.3}$, applying Lemma 3.4 with $\varepsilon = 4p^{-1/2}\rho$, we obtain

$$\mathsf{P}_W\big(\operatorname{dist}(W \star \overline{X}, \mathbb{Z}^n)^2 \le 16p^{-1}\rho^2 n\big) \le (C\rho)^{cn}. \tag{3.4}$$

Let

$$D(W, X) := \operatorname{dist}(W \star \overline{X}, \mathbb{Z}^n)^2 \quad \text{and} \quad \mathcal{E} := \mathcal{E}_{3.3}.$$

Since (3.4) holds uniformly for all $X \in \mathcal{E}$, and since W and X are independent while $\mathcal{E}$ is determined only by X, Fubini's theorem gives

$$\begin{aligned}\mathsf{E}_W\mathsf{P}_X(D(W,X) \le 16p^{-1}\rho^2 n \mid \mathcal{E}) &= \mathsf{P}_{W,X}(D(W,X) \le 16p^{-1}\rho^2 n \mid \mathcal{E}) \\ &\le (C\rho)^{cn}.\end{aligned}$$

Hence, by Markov's inequality,

$$\mathsf{P}_W\left(\mathsf{P}_X\big(D(W,X) \le 16p^{-1}\rho^2 n \mid \mathcal{E}\big) \ge \frac{1}{2}\right) \le 2(C\rho)^{cn}.$$

Absorbing the factor 2 into the absolute constant C, we find that, with probability at least $1 - (C\rho)^{cn}$,

$$\mathsf{P}_X(D(W,X) > 16p^{-1}\rho^2 n \mid \mathcal{E}) \ge \frac{1}{2}.$$

On this event,

$$\mathsf{E}_X[D(W,X) \mid \mathcal{E}] \ge 8p^{-1}\rho^2 n.$$

Moreover, since $\mathsf{P}_X(\mathcal{E}_{3.3}) \ge p/2$, we have

$$\begin{aligned}\mathsf{E}_X D(W,X) &\ge \mathsf{P}_X(\mathcal{E}_{3.3})\mathsf{E}_X[D(W,X) \mid \mathcal{E}_{3.3}] \\ &\ge \frac{p}{2} 8p^{-1}\rho^2 n = 4\rho^2 n.\end{aligned}$$

Therefore,

$$\mathsf{P}_W(\mathsf{E}_X \operatorname{dist}(W \star \overline{X}, \mathbb{Z}^n)^2 \ge 4\rho^2 n) \ge 1 - (C\rho)^{cn}.$$

This concludes the proof. ∎

Proof of Proposition 3.2. Applying Lemma 2.10 yields

$$|\Lambda_{d_j}^J| \le \left|\delta\mathbb{Z}^n \cap B\left(0, \left(4 + \frac{4}{\tau}\right)d_j\right)\right| \le \left(\frac{Cd_j}{r\tau\delta\sqrt{n}}\right)^n.$$

Furthermore, by the discussion preceding Lemma 3.3, we have

$$|\mathcal{N}_d| \le \sum_{|J_j|\ge c_\tau n} \Big[\sum_{j=1}^{n} \mathsf{P}_W(\mathsf{E}\,\mathrm{dist}^2\left(W \star \overline{\mathrm{Col}_j(A)}, \mathbb{Z}^n\right) < 4\rho^2 n)\Big]^l \prod_{j=1}^{l} |\Lambda_{d_j}^{J_j}|$$
$$\le 2^{nl} n^l \Big(\frac{C\rho^c}{r\tau\delta}\Big)^{ln} \Big(\prod_{j=1}^{l} \frac{d_j}{\sqrt{n}}\Big)^n \le \Big(\frac{C\rho^c}{r\delta}\Big)^{ln} \Big(\prod_{J=1}^{l} \frac{d_j}{\sqrt{n}}\Big)^n.$$

This completes the proof of Proposition 3.2. ∎

3.2. Approximating. The following lemma is about random rounding, whose use in random matrix theory was pioneered by Livshyts [Liv21a].

PROPOSITION 3.5. *Let $k \le c_{3.5}n$ and $d = (d_1, \dots, d_l) \in [r\sqrt{n}, R]^l$. Let $\delta > 0$ be a small constant satisfying* (3.1)*. Let B be an $(n-k)\times n$ matrix such that $\|B\|_{\mathrm{HS}} \le 2kn$. For any $(v_1, \dots, v_l) \in \mathcal{W}_d$ with $v_j \in \ker B$, for all $j \in [l]$, there exist $u_1, \dots, u_l \in \mathcal{N}_d$ with the following properties:*

(1) $\|u_j - v_j\|_\infty \le \delta$ *for all* $j \in [l]$*;*
(2) *if U and V are $n \times l$ matrices with columns $u_1, \dots, u_l$ and $v_1, \dots, v_l$ respectively, then*
$$\|U - V\| \le C_{2.3}\delta\sqrt{n};$$
(3) $(u_1, \dots, u_l)$ *is $\frac{1}{4}$-almost orthogonal;*
(4) $\mathrm{span}(u_1, \dots, u_l) \cap S^{n-1} \subset \mathrm{Incomp}(\tau^2, \tau^4/2)$*;*
(5) $\mathrm{d}_A(u_j, \mathbb{Z}^n) < 2\rho\sqrt{n}$ *for all* $j \in [l]$*;*
(6) *if U is a matrix as in* (2)*, then*
$$\mathrm{d}_A(U\theta, \mathbb{Z}^n) > \frac{\rho}{2}\sqrt{n}$$
for any $\theta \in \mathbb{R}^l$ satisfying
$$\|\theta\|_2 \le \frac{1}{20\sqrt{l}} \quad \text{and} \quad \|U\theta\|_2 \ge 8r\sqrt{n};$$
(7) $\|Bu_j\|_2 \le 2K\delta n$ *for all* $j \in [l]$.

Proof. Let $(v_1, \dots, v_l) \in W_d$ and $\overline{v}_1, \dots, \overline{v}_l \in \delta\mathbb{Z}^n$ be such that for all $j \in [l]$,
$$v_j \in \overline{v}_j + \delta[0,1]^n.$$
Define independent random variables ε_{ij}, $i \in [n]$, $j \in [l]$, as follows:
$$\mathsf{P}(\varepsilon_{ij} = \overline{v}_j(i) - v_j(i)) = 1 - \frac{v_j(i) - \overline{v}_j(i)}{\delta},$$
$$\mathsf{P}(\varepsilon_{ij} = \overline{v}_j(i) - v_j(i) + \delta) = \frac{v_j(i) - \overline{v}_j(i)}{\delta}.$$

Then we have $|\varepsilon_{ij}| \le \delta$ and $\mathsf{E}\varepsilon_{ij} = 0$. Let

$$u_j = v_j + \sum_{i=1}^{n} \varepsilon_{ij} e_i \in \delta\mathbb{Z}^n.$$

Lemma 5.3 in [Rud24] already gives (1) and (2) with probability at least $1-\exp(-c_{2.3}n)$; (3) and (4) are immediate consequences. It remains to verify that (5) and (6) also follow from (1) and (2).

Note that for any $i \in [n]$,

$$\begin{aligned}\operatorname{dist}(u_j \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n) &\le \operatorname{dist}(v_j \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n) + \|u_j - v_j\|_\infty \|\overline{\operatorname{Col}_i(A)}\|_2 \\ &\le \rho\sqrt{n} + \delta\|\overline{\operatorname{Col}_i(A)}\|_2.\end{aligned}$$

Then

$$\begin{aligned}\mathrm{d}_A(u_j, \mathbb{Z}^n) &= \sqrt{\min_i \mathsf{E}\operatorname{dist}^2(u_j \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n)} \\ &\le \sqrt{\min_i \mathsf{E}(\rho\sqrt{n} + \delta\|\overline{\operatorname{Col}_i(A)}\|_2)^2} \\ &< 2\rho\sqrt{n}.\end{aligned}$$

Furthermore, since $(u_1, \dots, u_l)$ and $(v_1, \dots, v_l)$ are $\frac{1}{2}$-almost orthogonal, and $\|v_j\|_2 \ge \frac{1}{2}\|u_j\|_2$, it follows that

$$\|V\theta\|_2^2 \ge \frac{1}{4}\sum_{j=1}^{l} \theta_j^2 \|v_j\|_2^2 \ge \frac{1}{16}\sum_{j=1}^{l} \theta_j^2 \|u_j\|_2^2 \ge \frac{1}{64}\|U\theta\|_2^2 \ge r^2 n.$$

As $(v_1, \dots, v_l) \in W_d$, we have $\mathrm{d}_A(V\theta, \mathbb{Z}^n) \ge \rho\sqrt{n}$. Since

$$\|(U-V)\theta\|_\infty = \max_{i\in[n]} \Big|\sum_{j=1}^{l} \varepsilon_{ij}\theta_j\Big| \le \max_{i\in[n]} \sqrt{\sum_{j=1}^{l} \varepsilon_{ij}^2}\sqrt{\sum_{j=1}^{l} \theta_j^2} \le \frac{\delta}{20},$$

for any $i \in [n]$ we have

$$\begin{aligned}\operatorname{dist}(U\theta \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n) &\ge \operatorname{dist}(V\theta \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n) - \|(U-V)\theta\|_\infty \|\overline{\operatorname{Col}_i(A)}\|_2 \\ &\ge \rho\sqrt{n} - \frac{\delta}{20}\|\overline{\operatorname{Col}_j(A)}\|_2.\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{d}_A(U\theta, \mathbb{Z}^n) &= \sqrt{\min_i \mathsf{E}\operatorname{dist}^2(U\theta \star \overline{\operatorname{Col}_i(A)}, \mathbb{Z}^n)} \\ &\ge \sqrt{\min_i \mathsf{E}(\rho\sqrt{n} - \delta\|\overline{\operatorname{Col}_i(A)}\|_2/20)^2} \ge \frac{\rho}{2}\sqrt{n}.\end{aligned}$$

Indeed, once (1) and (2) are secured, we can make (5) and (6) hold by choosing δ sufficiently small.

So far we have shown that (1)–(6) hold simultaneously with probability at least $1-\exp(-c_{2.3}n)$. Because

$$\mathsf{P}\big(\|Bu_j\|_2 \le 2K\delta n \ \text{ for all } j\in[l]\big) \ge 2^{-l}$$

and $1-\exp(-c_{2.3}n)+2^{-l}>1$, there exists a fixed tuple $(u_1,\dots,u_l)\in\mathcal{N}_d$ satisfying (1)–(7), and the lemma follows. ∎

3.3. Proof of Proposition 3.1. Before proving Proposition 3.1, we record one more lemma. Throughout its statement and proof, δ is the constant already fixed in Proposition 3.5.

LEMMA 3.6. *Let $d=(d_1,\dots,d_l)\in[r\sqrt{n},R]^l$ and $c_{3.6}k\le l\le k\le\frac{\delta}{20}\sqrt{n}$. Then*

$$\mathsf{P}(\mathcal{W}_d\cap\ker B\ne\emptyset)\le\exp(-2ln).$$

Proof. Let $\tilde{\mathcal{N}}_d\subset\mathcal{N}_d$ be the set of $(u_1,\dots,u_l)\in\mathcal{N}_d$ that satisfy (3)–(7) of Proposition 3.5. Let U be the $n\times l$ matrix with columns $u_1,\dots,u_l$. Firstly, we shall show that

$$\mathrm{RD}^A_{L,\alpha}(U^T)\ge\frac{1}{20\sqrt{l}}. \tag{3.5}$$

Take $\theta\in\mathbb{R}^l$ such that $\|\theta\|_2\le\frac{1}{20\sqrt{l}}$. Assume that

$$\|U\theta\|_2\le 8r\sqrt{n}\le h_{2.9}(\tau^2,\tau^4/2,K)\sqrt{n}.$$

By applying Lemma 2.9 with $\alpha<\alpha_{2.9}(\tau^2,\tau^4/2,K)$, we have

$$\mathrm{d}_A(U\theta,\mathbb{Z}^n)\ge L\sqrt{\log_+\frac{\alpha\|U\theta\|_2}{L}}.$$

Assume that $\|U\theta\|_2\le\sqrt{l}\max_{j\in[l]}\|u_j\|_2\le\sqrt{l}R$; then we obtain

$$L\sqrt{\log_+\frac{\alpha\|U\theta\|_2}{L}}\le\frac{\rho}{2}\sqrt{n}.$$

By virtue of Proposition 3.5(6), we get

$$\mathrm{d}_A(U\theta,\mathbb{Z}^n)>\frac{\rho}{2}\sqrt{n}, \tag{3.6}$$

whenever $\theta\in\mathbb{R}^l$ satisfies

$$\|\theta\|_2\le\frac{1}{20\sqrt{l}}\quad\text{ and }\quad\|U\theta\|_2\ge 8r\sqrt{n}.$$

Combining the preceding cases, we conclude that for every $\theta\in\mathbb{R}^l$ with $\|\theta\|_2\le\frac{1}{20\sqrt{l}}$,

$$\mathrm{d}_A(U\theta,\mathbb{Z}^n)\ge L\sqrt{\log_+\frac{\alpha\|U\theta\|_2}{L}}.$$

Hence, by definition,

$$\mathrm{RD}^A_{L,\alpha}\big(U^T\big)\ge\frac{1}{20\sqrt{l}}.$$

Item (3) of Proposition 3.5 yields

$$\det{}^{1/2}(U^TU) \ge 4^{-l}\prod_{j=1}^{l}\|u_j\|_2 \ge 8^{-l}\prod_{j=1}^{l} d_j.$$

Write $B_i = \mathrm{Row}_i(B)^T$. For every $t \ge \delta\sqrt{n}$ we have $t \ge 20l \ge \sqrt{l}/\mathrm{RD}^A_{L,\alpha}(U^T)$, so Proposition 2.6 gives

$$\mathsf{P}\big(\|U^TB_i\|_2 \le t\sqrt{l}\big) \le \frac{\big(C_{2.6}L/\alpha\sqrt{l}\big)^l}{\det^{1/2}(U^TU)}\left(t + \frac{\sqrt{l}}{\mathrm{RD}^A_{L,\alpha}(U^T)}\right)^l \le \frac{C^l}{\prod_{j=1}^l d_j}\,t^l,$$

where $C = C(\tau, K)$ depends only on τ and K.

Put $Y_i = \frac{1}{l}\|U^TB_i\|_2^2$ and $M = C^2\big(\prod_{j=1}^l d_j\big)^{-2/l}$. Then

$$\mathsf{P}(Y_i \le s) \le (Ms)^{l/2} \quad \text{for all } s \ge s_0 := \delta^2 n.$$

Apply Lemma 2.2 with $m = l/2$ and $t = 4K^2s_0$:

$$\begin{aligned}
\mathsf{P}\big(\|Bu_j\|_2 \le 2K\delta n\ \forall j \in [l]\big) &\le \mathsf{P}\Big(\sum_{j=1}^{l}\|Bu_j\|_2^2 \le 4K^2\delta^2 l n^2\Big)\\
&= \mathsf{P}\Big(\sum_{i=1}^{n-k} Y_i \le n\cdot 4K^2 s_0\Big)\\
&\le (C_{2.2}Mt)^{m(n-k)}\\
&= \big(\tilde{C}\delta\big)^{l(n-k)}\left(\prod_{j=1}^{l}\frac{\sqrt{n}}{d_j}\right)^{n-k}.
\end{aligned}$$

Hence

$$\begin{aligned}
P_d &:= \mathsf{P}\big(\exists\,(u_1,\dots,u_l) \in \tilde{\mathcal{N}}_d : \|Bu_j\|_2 \le 2K\delta n\ \forall j\big)\\
&\le |\tilde{\mathcal{N}}_d|\big(\tilde{C}\delta\big)^{l(n-k)}\left(\prod_{j=1}^{l}\frac{\sqrt{n}}{d_j}\right)^{n-k}\\
&\le \left(\frac{C_{3.2}\rho^{c_{3.2}}}{r\delta}\right)^{ln}\left(\prod_{j=1}^{l}\frac{d_j}{\sqrt{n}}\right)^{n}(\tilde{C}\delta)^{l(n-k)}\left(\prod_{j=1}^{l}\frac{\sqrt{n}}{d_j}\right)^{n-k}\\
&= \left(\frac{C_1\rho^{c_{3.2}}}{r}\right)^{ln}\delta^{-lk}\left(\prod_{j=1}^{l}\frac{d_j}{\sqrt{n}}\right)^{k}\\
&\le \left(\frac{C_1\rho^{c_{3.2}}}{r}\right)^{ln}\left(\frac{R}{\delta\sqrt{n}}\right)^{lk}\\
&< \left(\frac{\tilde{c}\rho^{c_{3.2}}}{r}e^{c\rho^2}\right)^{ln} < e^{-2ln},
\end{aligned}$$

where the last line holds since $\rho^{c_{3.2}} < ar$ for a small enough absolute constant $a > 0$.

Finally, Lemma 2.4 gives

$$\begin{aligned}\mathsf{P}(\mathcal{W}_d \cap \ker B \neq \varnothing) &\leq \mathsf{P}(\mathcal{W}_d \cap \ker B \mid \|B\|_{\mathrm{HS}} \leq 2Kn) + \mathsf{P}(\|B\|_{\mathrm{HS}} > 2Kn) \\ &\leq P_d + e^{-c_{2.4}n^2} \leq e^{-2ln}. \ \blacksquare\end{aligned}$$

Summing up, we can prove Proposition 3.1 by applying Lemma 3.6 for every $d \in [r\sqrt{n}, R]^l$.

Proof of Proposition 3.1. Let $\mathcal{E}_d$ be the event that $W_d \cap \ker(B) \neq \emptyset$. Then $\mathcal{E}_{3.1} = \bigcup \mathcal{E}_d$, where $d = (d_1, \dots, d_l) \in [r\sqrt{n}, R]^l \cap 2^{\mathbb{Z}}$. Note that $d_j = 2^{s_j}$ and $s_j \in \mathbb{N}$; by estimating the cardinality of d, we have

$$\mathsf{P}(\mathcal{E}_{3.1}) \leq \left[\log\left(\frac{R}{r\sqrt{n}}\right)\right]^l \exp(-2ln) \leq \exp(-ln). \ \blacksquare$$

4. Proof of Theorem 1.2. In this section, we prove Theorem 1.2 by combining Lemma 2.15 with Proposition 3.1, with Lemma 2.12 being the key additional ingredient. Furthermore, Theorem 1.1 is proved via a linear algebra argument combined with Theorem 1.2. We first show that whenever a subspace of $\ker B$ has a sufficiently large RLCD, it automatically carries a strong exponential lower bound.

LEMMA 4.1. *Let A be an $n \times n$ random matrix with independent entries that satisfy* (1.3). *For $k < c\sqrt{n}$, let $J \subset [n]$ with $|J| = n - k$ and define $\mathcal{E}_J^k$ as the event that there exists a linear subspace $E \subset \big(\mathrm{span}(\mathrm{Col}_i(A),\, i \in J)\big)^{\perp}$ such that $\dim E \geq k/2$ and*

$$\mathrm{RD}_{L,\alpha}^{A}(E) \geq \exp\left(C_{4.1}\frac{n}{k}\right).$$

Then

$$\mathsf{P}\big(\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n] \text{ and } \mathcal{E}_J^k\big) \leq \exp\big(-c_{4.1}nk\big),$$

where $c_{4.1}$ and $C_{4.1}$ are constants depending only on K.

The proof follows by adapting the argument for [Rud24, Lemma 6.1], and is therefore omitted.

Proof of Theorem 1.2. Assume that $\mathrm{rank}(A) \leq n-k$. There exists $J \subset [n]$ with $|J| = n - k$ such that $\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A),\, i \in J)$ for all $j \in [n]$. Then

$$\begin{aligned}&\mathsf{P}(\mathrm{rank}(A) \leq n-k) \\ &\quad= \mathsf{P}\big(\exists J \subset [n], |J| = n-k : \mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n]\big) \\ &\quad\leq \sum_{J \subset [n],\, |J| = n-k} \mathsf{P}\big(\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n] \setminus J\big) \\ &\quad\leq \binom{n}{k} \sup_{J \subset [n],\, |J| = n-k} \mathsf{P}\big(\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n] \setminus J\big).\end{aligned}$$

Note that

$$\binom{n}{k} \le \exp\left(k \log\left(\frac{en}{k}\right)\right) < \exp(ckn),$$

we only need to prove

$$\mathsf{P}\big(\operatorname{Col}_j(A) \in \operatorname{span}(\operatorname{Col}_i(A), i \in J) \text{ for all } j \in [n]\big) \le \exp(-2ckn).$$

To this end, we shall show that $\mathcal{E}_J^k$, defined in Lemma 4.1, occurs with probability at least $1 - \exp(-ckn)$.

Let B_J be an $(n-k)\times n$ random matrix with rows $\operatorname{Row}_j(B) = \operatorname{Col}_{n_j}(A)^T$, where n_j is the j-th entry of J. Let $E_0 = \ker B_J$ and denote by P_{E_0} the orthogonal projection onto E_0. Let τ be chosen in Lemma 2.15, and denote $W_0 := \operatorname{Comp}(\tau^2, \tau^4)$.

Applying Lemma 2.12 with E_0, W_0 and $l = k/4$, we observe that at least one of the events described by (1) and (2) of the lemma occurs. Denote these events by $\mathcal{E}_{2.12}^{(1)}$ and $\mathcal{E}_{2.12}^{(2)}$ respectively. Using Lemma 2.15,

$$\mathsf{P}(\mathcal{E}_{2.12}^{(1)}) \le \exp\left(-\frac{c_{2.15}k}{4}n\right). \tag{4.1}$$

Assume $\mathcal{E}_{2.12}^{(2)}$ occurs, so there is a subspace $F \subset E_0$ with $\dim F = \frac{3}{4}k$ and $F \cap W_0 = \varnothing$. Fix the constant ρ chosen in Proposition 3.1 and set

$$W_1 = \left\{v \in F : 2r\sqrt{n} \le \|v\|_2 \le \exp\left(\frac{\rho^2 n}{4L^2}\right),\ \mathrm{d}_A(v, \mathbb{Z}^n) \le \rho\sqrt{n}\right\}.$$

Apply Lemma 2.12 to F, W_1 and l; one of its two conclusions must hold. Denote the corresponding events by $\mathcal{V}_{2.12}^{(1)}$ and $\mathcal{V}_{2.12}^{(2)}$. Proposition 3.1 yields

$$\mathsf{P}(\mathcal{V}_{2.12}^{(1)}) \le \exp\left(-\frac{k}{4}n\right). \tag{4.2}$$

Assume $\mathcal{V}_{2.12}^{(2)}$ occurs, so there exists a subspace $\overline{F} \subset F$ with $\dim \overline{F} = k/2$ and $\overline{F} \cap W_1 = \varnothing$. We shall show that this forces

$$\operatorname{RD}_{L,\alpha}^A(\overline{F}) \ge R := \exp\left(\frac{\rho^2 n}{4L^2}\right).$$

Choose an isometry $U\colon \mathbb{R}^{k/2} \to \mathbb{R}^n$ with $U\mathbb{R}^{k/2} = \overline{F}$; then $\operatorname{RD}_{L,\alpha}^A(U^T) = \operatorname{RD}_{L,\alpha}^A(\overline{F})$. Fix $\theta \in \mathbb{R}^{k/2}$ and suppose

$$\mathrm{d}_A(U\theta, \mathbb{Z}^n) < L\sqrt{\log_+\left(\frac{\alpha\|\theta\|_2}{L}\right)}.$$

Because $U(\mathbb{R}^{k/2} \cap S^{n-1}) \subseteq \mathrm{Incomp}(\tau^2, \tau^4)$, Lemma 2.9 applies once $\alpha < \alpha_{2.9}(\tau^2, \tau^4, K)$ and $\|\theta\|_2 \le 2r\sqrt{n} \le h_{2.9}(\tau^2, \tau^4, K)\sqrt{n}$, giving

$$\mathrm{d}_A(U\theta, \mathbb{Z}^n) \ge L\sqrt{\log_+\biggl(\frac{\alpha\|\theta\|_2}{L}\biggr)}.$$

Hence the strict inequality above can never occur in that range, so any θ violating it must satisfy $\|\theta\|_2 \ge 2r\sqrt{n}$. Since $U\theta \notin W_1$, the definition of W_1 forces $\|\theta\|_2 \ge R$. Consequently, $\mathrm{RD}^A_{L,\alpha}(\overline{F}) \ge R$ with probability at least $1 - e^{-ckn}$. Furthermore, we have

$$\begin{aligned}
&\mathsf{P}\bigl(\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n]\bigr)\\
&\le \exp(-ckn) + \mathsf{P}\bigl(\mathrm{Col}_j(A) \in \mathrm{span}(\mathrm{Col}_i(A), i \in J) \text{ for all } j \in [n] \text{ and } \mathcal{V}^{(2)}_{2.12}\bigr)\\
&\le \exp(-\tilde{c}kn). \ \blacksquare
\end{aligned}$$

Proof of Theorem 1.1. We first prove Theorem 1.1 for all sufficiently large n.

By the Courant–Fischer–Weyl min-max principle,

$$s_{n-k+1}(A) = \min_{\dim H = k} \max_{x \in H \cap S^{n-1}} \|Ax\|_2.$$

Hence, if $s_{n-k+1}(A) \le \varepsilon/\sqrt{n}$, there exists an orthonormal set $(z_1, \dots, z_k)$ such that

$$\|Az_i\|_2 \le \frac{\varepsilon}{\sqrt{n}}, \qquad i = 1, \dots, k.$$

Let $Z^\top = (z_1, \dots, z_k)$ be an $n \times k$ full-rank matrix. Applying Lemma 2.14 shows that there exist $i_1, \dots, i_l \in [n]$ such that

$$\begin{aligned}
s_l(Z_{i_1,\dots,i_l})^{-1} &\le C_{2.14} \min_{r \in \{l+1,\dots,k\}} \sqrt{\frac{rn}{(r-l)\sum_{i=r}^k s_i(Z)^2}}\\
&\le C_{2.14} \min_{r \in \{l+1,\dots,k\}} \sqrt{\frac{rn}{(r-l)(k-r+1)}}\\
&\le C_1\sqrt{\frac{kn}{(k-l)^2}}.
\end{aligned}$$

Set $F = (Z_{i_1,\dots,i_l})^\top$, $\tilde{F} = (Z_{i_{l+1},\dots,i_n})^\top$, and $A_i = \mathrm{Col}_i(A)$. Then

$$M := AZ^\top = (A_{i_1}, \dots, A_{i_l})F + (A_{i_{l+1}}, \dots, A_{i_n})\tilde{F}.$$

Let $F_1 = F^\top(FF^\top)^{-1}$ be the right inverse of F. Thus, we have

$$MF_1 := (A_{i_1}, \dots, A_{i_l}) + (A_{i_{l+1}}, \dots, A_{i_n})\tilde{F}F_1.$$

Let H be the linear space spanned by $A_{i_{l+1}}, \dots, A_{i_n}$. Let P be the orthogonal projection in $\mathbb{R}^n$ onto $H^\perp$. Then

$$PMF_1 = P(A_{i_1}, \dots, A_{i_l}),$$

which yields

$$\begin{aligned}&\operatorname{dist}^2(A_{i_1},H)+\cdots+\operatorname{dist}^2(A_{i_l},H)\\&\qquad=\|PMF_1\|_{\mathrm{HS}}^2\le\|MF_1\|_{\mathrm{HS}}^2\le\|F_1\|_2^2\|M\|_{\mathrm{HS}}^2.\end{aligned}$$

On the one hand, for any $x\in\mathbb{R}^l$,

$$1=\|FF_1\|\ge\frac{\|FF_1x\|_2}{\|x\|_2}\ge s_l(F)\frac{\|F_1x\|_2}{\|x\|_2}.$$

This implies that

$$s_1(F_1)\le s_l(F)^{-1}\le C_1\sqrt{\frac{kn}{(k-l)^2}}. \tag{4.3}$$

On the other hand,

$$\|M\|_{\mathrm{HS}}^2\le\sum_{i=1}^k\|\operatorname{Col}_i(M)\|_2^2=\sum_{i=1}^k\|Az_i\|_2^2\le\frac{k\varepsilon^2}{n}. \tag{4.4}$$

By virtue of (4.3) and (4.4), we have

$$\|MF_1\|_{\mathrm{HS}}\le\|F_1\|_2\|M\|_{\mathrm{HS}}\le C_1\sqrt{\frac{kn}{(k-l)^2}}\sqrt{\frac{k\varepsilon^2}{n}}\le C_1\frac{k\varepsilon}{k-l}.$$

Furthermore, there exist $i_1,\dots,i_l\in[n]$ such that

$$\operatorname{dist}^2(A_{i_1},H)+\cdots+\operatorname{dist}^2(A_{i_l},H)\le\left(\frac{C_1k\varepsilon}{k-l}\right)^2.$$

Now, we have

$$\begin{aligned}&\mathsf{P}\left(s_{n-k+1}(A)\le\frac{\varepsilon}{\sqrt{n}}\right)\\&\qquad\le\mathsf{P}\left(\exists\{i_1,\cdots,i_l\}\subset[n]:\sum_{m=1}^l\|P_{H^\perp}A_{i_m}\|_2^2\le\left(\frac{C_1k\varepsilon}{k-l}\right)^2\right)\\&\qquad\le\sum_{i_1,\dots,i_l\in[n]}\mathsf{P}\left(\sum_{m=1}^l\|P_{H^\perp}A_{i_m}\|_2^2\le\left(\frac{C_1k\varepsilon}{k-l}\right)^2\right).\end{aligned}$$

Similar to the proof of Theorem 1.2, for all $\omega\in(0,1)$, there exists ρ depending on K and ω such that

$$\mathsf{P}\left(\mathcal{H}:=\exists E\subset H^\perp,\dim E=\omega l,\mathrm{RD}_{\tilde L,\alpha}^A(E)\ge R_1:=\exp\left(\frac{\rho^2n}{4\tilde L^2}\right)\right)\ge1-e^{-cln},$$

where $\tilde L=c_{2.6}\sqrt{l}$ and $c>0$ is a constant depending on K and ω.

Then

$$\begin{aligned}\mathsf{P}\Big(\sum_{m=1}^{l}\|P_{H^\perp}A_{i_m}\|_2^2 \le \Big(\frac{C_1k\varepsilon}{k-l}\Big)^2\Big)&\\ \le \mathsf{P}\Big(\sum_{m=1}^{l}\|P_{H^\perp}A_{i_m}\|_2^2 \le \Big(\frac{C_1k\varepsilon}{k-l}\Big)^2, \mathcal{H}\Big) + 2e^{-cln}&\\ \le \mathsf{P}\Big(\sum_{m=1}^{l}\|P_{E}A_{i_m}\|_2^2 \le \Big(\frac{C_1k\varepsilon}{k-l}\Big)^2, \mathcal{H}\Big) + 2e^{-cln}.&\end{aligned}$$

By Corollary 2.7,

$$\mathsf{P}\big(\|P_EA_{i_m}\|_2 \le t\sqrt{\omega l},\, \mathcal{H}\big) \le (C't)^{\omega l} + e^{-c'n}.$$

A tensorization argument identical to the one used in Lemma 2.2 gives

$$\mathsf{P}\Big(\sum_{m=1}^{l}\|P_EA_{i_m}\|_2^2 \le l^2t^2,\, \mathcal{H}\Big) \le (Ct)^{\omega l^2} + e^{-cln}.$$

Substituting $t = \frac{Ck\varepsilon}{l(k-l)}$ we obtain

$$\mathsf{P}\Big(\sum_{m=1}^{l}\|P_EA_{i_m}\|_2^2 \le \Big(\frac{Ck\varepsilon}{k-l}\Big)^2,\, \mathcal{H}\Big) \le \Big(\frac{Ck\varepsilon}{l(k-l)}\Big)^{\omega l^2} + e^{-cln}.$$

Furthermore, we have

$$\begin{aligned}\mathsf{P}\Big(s_{n-k+1}(A) \le \frac{\varepsilon}{\sqrt{n}}\Big) &\le \binom{n}{l}\Big(\Big(\frac{Ck\varepsilon}{l(k-l)}\Big)^{\omega l^2} + \exp(-cln)\Big)\\ &\le n^l\Big(\Big(\frac{Ck\varepsilon}{l(k-l)}\Big)^{\omega l^2} + \exp(-cln)\Big).\end{aligned}$$

Let $l = \lceil\sqrt{\gamma/\omega}k\rceil$, $\gamma \in (0,\omega)$. Then

$$\mathsf{P}\Big(s_{n-k+1}(A) \le \frac{\varepsilon}{\sqrt{n}}\Big) \le \Big(\frac{C\varepsilon}{k}\Big)^{\gamma k^2} + \exp(-ckn).$$

Then Theorem 1.1 is valid for any $n > n_{3.4}$. Note that

$$\lim_{\varepsilon\to 0}\mathsf{P}\Big(s_{n-k+1}(A) \le \frac{\varepsilon}{\sqrt{n}}\Big) = \mathsf{P}(s_{n-k+1}(A) = 0) \le e^{-ckn},$$

where the last inequality is due to Theorem 1.2. Hence, for any fixed n, there exists an $\varepsilon_0(n, n_{3.4})$ such that for any $\varepsilon < \varepsilon_0(n, n_{3.4})$

$$\mathsf{P}(s_{n-k+1}(A) \le \varepsilon/\sqrt{n}) - \mathsf{P}(s_{n-k+1}(A) = 0) \le e^{-cn_{3.4}^2}. \tag{4.5}$$

Set $\varepsilon_0 = \min\{\varepsilon(1, n_{3.4}), \dots, \varepsilon(n_{3.4}-1, n_{3.4})\}$. Applying (4.5), we deduce that

for any $n < n_{3.4}$ and $\varepsilon < \varepsilon_0$,

$$\mathsf{P}(s_{n-k+1}(A) \le \varepsilon/\sqrt{n}) \le 2e^{-ckn}.$$

For any $n < n_{3.4}$, $\varepsilon > \varepsilon_0$, we have

$$\mathsf{P}(s_{n-k+1}(A) \le \varepsilon/\sqrt{n}) \le 1 \le (C\varepsilon/k)^{\gamma k^2}.$$

where $C = c\sqrt{n_{3.4}}/\varepsilon_0$. Hence Theorem 1.1 is valid for any n and $\varepsilon \ge 0$. ∎

Acknowledgements. The authors sincerely thank the editorial board members and the reviewer for their careful reading of the manuscript and their insightful comments.

Funding. This work was supported by the National Key R&D Program of China (No.2024YFA1013501), Shandong Provincial Natural Science Foundation (No. ZR2024MA082), and the Youth Student Fundamental Research Funds of Shandong University (No. SDU-QM-B202407).

References

[BVW10] J. Bourgain, V. Vu, and P. Wood, *On the singularity probability of discrete random matrices*, J. Funct. Anal. 258 (2010), 559–603.

[DSW24] G. Dai, Z. Su, and H. Wang, *Quantitative estimates of the singular values of random i.i.d. matrices*, arXiv:2412.18912v1 (2024).

[KKS95] 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.

[Liv21a] G. V. Livshyts, *The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding*, J. Anal. Math. 145 (2021), 257–306.

[LTV21] G. V. Livshyts, K. Tikhomirov, and R. Vershinin, *The smallest singular value of inhomogeneous square random matrices*, Ann. Probab. 49 (2021), 1286–1309.

[NY17] A. Naor and P. Youssef, *Restricted invertibility revisited*, in: A Journey Through Discrete Mathematics, Springer, Cham, 2017, 657–691.

[Ngu18] H. Nguyen, *Random matrices: overcrowding estimates for the spectrum*, J. Funct. Anal. 275 (2018), 2197–2224.

[RT18] E. Rebrova and K. Tikhomirov, *Coverings of random ellipsoids, and invertibility of matrices with i.i.d. heavy-tailed entries*, Israel J. Math. 227 (2018), 507–544.

[Rud24] M. Rudelson, *A large deviation inequality for the rank of a random matrix*, Ann. Probab. 52 (2024), 1992–2018.

[RV08] M. Rudelson and R. Vershynin, *The Littlewood–Offord problem and invertibility of random matrices*, Adv. Math. 218 (2008), 600–633.

[RV09] M. Rudelson and R. Vershynin, *Smallest singular value of a random rectangular matrix*, Comm. Pure Appl. Math. 62 (2009), 1707–1739.

[RV16] M. Rudelson and R. Vershynin, *No-gaps delocalization for general random matrices*, Geom. Funct. Anal. 26 (2016), 1716–1776.

[ST02] D. Spielman and S. Teng, *Smoothed analysis of algorithms*, in: Proceedings of the International Congress of Mathematicians (Beijing, 2002), Vol. I, 597–606.

[TV06] T. Tao and V. Vu, *On random ±1 matrices: singularity and determinant*, Random Structures Algorithms 28 (2006), 1–23.

[TV07] T. Tao and V. Vu, *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. 20 (2007), 603–628.

[Tik20] K. Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) 191 (2020), 593–634.

Guozheng Dai
Hong Kong University of Science and Technology
Kowloon, 999077, Hong Kong, P. R. China
E-mail: guozhengdai1997@gmail.com

Zeyan Song, Hanchao Wang
Shandong University
Jinan, 250100, P. R. China
E-mail: zeyansong8@gmail.com
wanghanchao@sdu.edu.cn