# ON RANDOM MATRICES WITH LARGE CORANK

ZACH HUNTER, MATTHEW KWAN, LISA SAUERMANN, AND MEHTAAB SAWHNEY

Abstract. Let $1 \le k \le n$ and $M$ be a random $n \times n$ matrix with independent uniformly random $\{\pm 1\}$-entries. We show that there exists an absolute constant $c > 0$ such that

$$\mathbf{P}[\operatorname{rank}(M) \le n-k] \le \exp(-cnk).$$

## 1. Introduction

The main result of this paper is the following bound for the probability that a random $\{\pm 1\}$-matrix has large corank.

**Theorem 1.1.** *There exists an absolute constant $c > 0$ such that the following holds. Take $1 \le k \le n$, and let $M$ be a random $n \times n$ matrix with independent entries uniformly random in $\{\pm 1\}$. We have*

$$\mathbf{P}[\operatorname{rank}(M) \le n-k] \le \exp(-cnk).$$

For the sake of simplicity we have restricted to the case of $\{\pm 1\}$ entries, but using the techniques of Bourgain, Vu and Wood [1], it is plausible that our proof can be adapted to allow the entries of $M$ to come from more flexible entry distributions.

### 1.1. **History.**

The singularity probability of random $\{\pm 1\}$-matrices has been intensively studied. After a series of works of Komlós [6], Kahn, Komlós and Szemerédi [4], Tao and Vu [12, 13], and Bourgain, Vu and Wood [1], celebrated work of Tikhomirov [14] established that

$$\mathbf{P}[\operatorname{rank}(M) < n] \le \left(\frac{1}{2} + o(1)\right)^n.$$

The base of $1/2 + o(1)$ is seen to be sharp by considering the probability that two rows match.

By considering the probability that $k+1$ rows match, it is natural to conjecture that in fact for all $k$ we have

$$\mathbf{P}[\operatorname{rank}(M) \le n-k] = \left(\frac{1}{2} + o(1)\right)^{nk}.$$

When $k$ is constant, this was verified by Jain, Sah and the final author [3] (building heavily on aforementioned work of Tikhomirov [14]). However, the methods in that work are not suitable when $k$ grows faster than say $\log n$. More recently, Rudelson [9] managed to prove Theorem 1.1 for $k \le \sqrt{n}$, via an ingenious geometric argument.

### 1.2. **Proof ideas.**

As the proof is rather short, we give only brief comments. The previously-mentioned works [14], [3] and [9] (along with a whole host of other works) fall broadly within the so-called "geometric" approach of Rudelson and Vershynin [10]. Approaching Theorem 1.1 in this manner appears to provide substantial resistance. Therefore, our work essentially reverts to the previous strategy of Kahn, Komlós and Szemerédi [4]. In fact, in some ways our proof is even simpler than in [4]: as we may assume $k$ is larger than an absolute constant, we can avoid certain technicalities about "structured subspaces" which makes our argument more streamlined.

The key new ingredient is a "high-dimensional relative anticoncentration inequality" (see Proposition 2.1), which provides a comparison between the probability that a random $\{\pm 1\}^n$-vector lies in a $(n-k)$-dimensional subspace, and the probability that a suitably "lazy" random vector lies in the same subspace (here "lazy" means that each component has some probability of being zero instead of $\pm 1$). The key technical innovation is to win a factor exponential in $k$ (i.e., a factor of $\gamma^k$ for an absolute constant $\gamma < 1$) when comparing the probabilities of these two events. We remark that earlier work of Kahn, Komlós and Szemerédi [4] proceeds via a similar inequality for the case $k = 1$ (which in turns builds on techniques of Halàsz [2]). Actually, this earlier inequality in some ways is quantitatively stronger than ours (it wins an arbitrarily large constant factor as the lazy random vector becomes more lazy), but it can only be effectively applied to subspaces which are in a certain sense "unstructured". A key simplifying feature of our proof is that we do not need to worry about such "unstructuredness" properties, and can work with completely arbitrary subspaces of codimension $k$.

The proof of Proposition 2.1 is closely modeled on the work of Tao and Vu [12], however at a crucial juncture we need to replace an inequality[1] due to Raikov [8] (which may be viewed as a continuous version of the Cauchy–Davenport inequality) with a suitable "high-dimensional" variant. This is achieved via a combination of pigeonholing and compression techniques to reduce to a case where the Brunn–Minkowski inequality in $\mathbf{R}^d$ may be invoked.

1.3. **Acknowledgements.** This work began when the authors were visiting Mathematisches Forschungsinstitut Oberwolfach, which provided ideal working conditions. MS thanks Vishesh Jain for initial discussions regarding the problem.

ZH was supported by SNSF grant 200021-228014. MK was supported by ERC Starting Grant "RANDSTRUCT" No. 101076777. LS was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – CRC 1720 – 539309657. This research was conducted during the period MS served as a Clay Research Fellow.

## 2. Reduction to high-dimensional relative concentration result

Our key technical ingredient is the following relative concentration result.

**Proposition 2.1.** *There exist absolute constants $\gamma = \gamma_{2.1} < 1$ and $k_{2.1} \geq 1$ such that the following holds. Let $p \in (0, 1/32]$, and let $\mu_p$ be the distribution which takes on values $1$ and $-1$ each with probability $p$, and the value $0$ with probability*[2] *$1 - 2p$. For some $k \geq k_{2.1}$, let $V \subseteq \mathbf{Q}^n$ be a linear subspace of dimension $\dim(V) = n - k$. Then for random vectors $X \sim \mu_{1/2}^{\otimes n}$ and $Y \sim \mu_p^{\otimes n}$, we have*

$$\sup_{t \in \mathbf{Q}^n} \mathbf{P}[X + t \in V] \leq \gamma^k \cdot \mathbf{P}[Y \in V].$$

*Remark.* Note that $\mu_{1/2}$ is the uniform distribution on $\{\pm 1\}$. We only need the case where $t = \vec{0}$ in our proof of Theorem 1.1. Also, we note that the condition $k \geq k_{2.1}$ can likely be removed at the cost of slightly complicating the proof.

It would be of interest to understand the quantitative behaviour of $\gamma$ in Proposition 2.1 when taking $p \to 0$. By considering $V = \{\vec{0}\} \subseteq \mathbf{Q}^n$ we see that $\gamma \geq 1/2$ for every $p$; it appears plausible that $\gamma \to 1/2$ as $p \to 0$.

We postpone the proof of Proposition 2.1 to the next section and first proceed with the proof of Theorem 1.1. The proof is closely modeled after the proof of Kahn, Komlós and Szemerédi [4] (following the presentation of Tao and Vu [12]).

*Proof of Theorem 1.1.* We may assume that $k$ is sufficiently large (larger than any absolute constant), recalling that $\mathbf{P}[\operatorname{rank}(M) < n] \leq \exp(-c'n)$ for an absolute constant $c' > 0$ by the work of [4].

We fix $p = 1/32$, and let $\gamma < 1$ be an absolute constant as in Proposition 2.1. Now, fix $c > 0$ such that $e^{-4c} > \max(1 - 2p, \gamma)$.

Let $\mathcal{V}$ denote the set of $(n-k)$-dimensional linear subspaces $V \subseteq \mathbf{Q}^n$. Consider independent random vectors $X_1, \dots, X_n \sim \mu_{1/2}^{\otimes n}$, and note that (since the columns of $M$ also have the same distribution $\mu_{1/2}^{\otimes n}$)

$$\mathbf{P}[\operatorname{rank}(M) = n - k] = \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}].$$

For any $V \in \mathcal{V}$, we define

$$\rho_V = \mathbf{P}_{X \sim \mu_{1/2}^{\otimes n}}[X \in V].$$

[1] This inequality asserts that for any subsets $A, B \subseteq \mathbf{T}$ of the circle $\mathbf{T} = \mathbf{R}/\mathbf{Z}$, writing $\mu(\cdot)$ for the Lebesgue measure, we have $\mu(A + B) \geq \min(1, \mu(A) + (B))$. This inequality, and generalizations thereof, are also known under various other names in the literature (in particular, two well-known generalizations were made by Macbeath [7] and Kneser [5]).

[2] Tao and Vu [12] and Bourgain, Vu and Wood [1] use the slightly differing convention that $\mu_p$ is $0$ with probability $1 - p$ and $\pm 1$ each with probability $p/2$.

Furthermore, define a subspace $V \in \mathcal{V}$ to be *thin* if $\rho_V \le (1-p)^{n/2}$ and to be *thick* if $\rho_V > (1-p)^{n/2}$. Denoting by $\mathcal{V}_{\text{thin}}$ and $\mathcal{V}_{\text{thick}}$ the sets of thin and thick subspaces, respectively, we obtain a partition $\mathcal{V} = \mathcal{V}_{\text{thin}} \sqcup \mathcal{V}_{\text{thick}}$. We bound the probability that $\operatorname{span}(X_1, \dots, X_n)$ belongs to $\mathcal{V}_{\text{thin}}$ and to $\mathcal{V}_{\text{thick}}$, respectively:

**Claim 2.2.** *We have that*

$$\mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}_{\text{thin}}] \le 2^n \cdot (1-p)^{nk/2}.$$

**Claim 2.3.** *We have that*

$$\mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}_{\text{thick}}] \le n2^{2n} \cdot \gamma^{nk/2}.$$

Combining these two claims, we obtain

$$\begin{aligned}\mathbf{P}[\operatorname{rank}(M) = n-k] &= \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}_{\text{thin}}] + \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}_{\text{thick}}] \\ &\le 2^n \cdot (1-2p)^{nk/2} + n2^{2n} \cdot \gamma^{nk/2} \le n2^{2n+1}. \exp(-2cnk)\end{aligned}$$

Summing this for all $k' = k, k+1, \dots, n$, we can conclude that (using that $k$ is sufficiently large with respect to $c$)

$$\mathbf{P}[\operatorname{rank}(M) \le n-k] = \sum_{k'=k}^{n} \mathbf{P}[\operatorname{rank}(M) = n-k'] \le n^2 2^{2n+1} \cdot \exp(-2cnk) \le \exp(-cnk) \qquad \square$$

It remains to prove the two claims. We start with Claim 2.2 handling thin subspaces.

*Proof of Claim 2.2.* Note that whenever $\operatorname{span}(X_1, \dots, X_n) = V$ for some $V \in \mathcal{V}_{\text{thin}}$, we can find a subset $S \subseteq [n]$ of size $|S| = n-k$ such that the vectors $X_s$ for $s \in S$ form a basis of $V$. In other words, we have $\operatorname{span}((X_s)_{s\in S}) = V$ and $X_i \in V$ for all $i \in [n] \setminus S$. By symmetry, the probability for this happening is the same for all subsets $S \subseteq [n]$ of size $|S| = n-k$, and so we can conclude

$$\begin{aligned}\mathbf{P}[\operatorname{span}(X_1, \dots, X_n) \in \mathcal{V}_{\text{thin}}] &= \sum_{V \in \mathcal{V}_{\text{thin}}} \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) = V] \\ &\le \binom{n}{n-k} \sum_{V \in \mathcal{V}_{\text{thin}}} \mathbf{P}[\operatorname{span}(X_1, \dots, X_{n-k}) = V \text{ and } X_{n-k+1}, \dots, X_n \in V] \\ &\le 2^n \sum_{V \in \mathcal{V}_{\text{thin}}} \mathbf{P}[\operatorname{span}(X_1, \dots, X_{n-k}) = V] \cdot \rho_V^k \\ &\le 2^n \cdot (1-p)^{nk/2} \sum_{V \in \mathcal{V}_{\text{thin}}} \mathbf{P}[\operatorname{span}(X_1, \dots, X_{n-k}) = V] \le 2^n \cdot (1-p)^{nk/2}. \qquad \square\end{aligned}$$

We now prove Claim 2.3 about thick subspaces using Proposition 2.1.

*Proof of Claim 2.3.* Let $m = \lceil n/2 \rceil$ and consider independent random vectors $Y_1, \dots, Y_m \sim \mu_p^{\otimes n}$ (also independent from $X_1, \dots, X_n$). For any $V \in \mathcal{V}_{\text{thick}}$ and any $i \in [m]$, by Proposition 2.1 we have that

$$\mathbf{P}[Y_i \in V] \ge \gamma^{-k} \rho_V.$$

Therefore we obtain

$$\begin{aligned}\mathbf{P}[\operatorname{span}(X_1, \dots, X_n, Y_1, \dots, Y_m) = V] &\ge \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) = V \text{ and } Y_1, \dots, Y_m \in V] \\ &\ge \gamma^{-km} \rho_V^m \cdot \mathbf{P}[\operatorname{span}(X_1, \dots, X_n) = V] \qquad (2.1)\end{aligned}$$

for any $V \in \mathcal{V}_{\text{thick}}$. Note that whenever $\operatorname{span}(X_1, \dots, X_n, Y_1, \dots, Y_m) = V$, there exist subsets $J \subseteq [m]$ and $I \subseteq [n]$ with $|I| + |J| = n-k$ such that $\operatorname{span}((X_i)_{i\in I}, (Y_j)_{j\in J}) = V$ and $Y_s \in \operatorname{span}((Y_j)_{j\in J})$ for all $s \in [m] \setminus J$ (and $X_s \in V$ for all $s \in [n] \setminus I$). Indeed, one can take $(Y_j)_{j\in J}$ to be a basis of $\operatorname{span}(Y_1, \dots, Y_m)$ and then extend to a basis of $\operatorname{span}(X_1, \dots, X_n, Y_1, \dots, Y_m) = V$. Therefore, using symmetry, we conclude that

$$\begin{aligned}&\mathbf{P}[\operatorname{span}(X_1, \dots, X_n, Y_1, \dots, Y_m) = V] \\ &\le \sum_{r=0}^{m} \binom{m}{r} \binom{n}{n-k-r} \cdot \mathbf{P}[\operatorname{span}(Y_1, \dots, Y_r, X_1, \dots, X_{n-k-r}) = V \text{ and } Y_{r+1}, \dots, Y_m \in \operatorname{span}(Y_1, \dots, Y_r) \\ &\qquad\qquad \text{and } X_{n-k-r+1}, \dots, X_n \in V]\end{aligned}$$

$$\leq \sum_{r=0}^{m} 2^m 2^n \cdot \mathbf{P}[\operatorname{span}(Y_1,\ldots,Y_r,X_1,\ldots,X_{n-k-r}) = V] \cdot (1-p)^{(n-r)(m-r)} \cdot \rho_V^{k+r}.$$

Here, we used that for any given outcomes of $Y_1,\ldots,Y_r$ we have $\mathbf{P}[Y_i \in \operatorname{span}(Y_1,\ldots,Y_r) \mid Y_1,\ldots,Y_r] \leq (1-p)^{n-r}$ for each $i = r+1,\ldots,m$, by the weighted Odlyzko lemma (see [13, Lemma 4.3]). Indeed, this lemma shows that $\mathbf{P}_{Y\sim\mu_p^{\otimes n}}[Y \in W] \leq (1-2p)^{n-\dim W} \leq (1-p)^{n-\dim W}$ for any linear subspace $W \subseteq \mathbf{Q}^n$.

Furthermore, observing that $(1-p)^{(n-r)(m-r)} \leq \rho_V^{m-r}$ for any $r = 0,\ldots,m$ and any $V \in \mathcal{V}_{\text{thick}}$ (for $r = m$ this holds trivially, and for $r \leq m-1 \leq n/2$ we can observe that $(1-p)^{n-r} \leq (1-p)^{n/2} \leq \rho_V$), we obtain

$$\begin{aligned}\mathbf{P}[\operatorname{span}(X_1,\ldots,X_n,Y_1,\ldots,Y_m) = V] &\leq \sum_{r=0}^{m} 2^{2n}\mathbf{P}[\operatorname{span}(Y_1,\ldots,Y_r,X_1,\ldots,X_{n-k-r}) = V] \cdot \rho_V^{m-r} \cdot \rho_V^{k+r} \\ &\leq 2^{2n} \cdot \rho_V^{m} \cdot \sum_{r=0}^{m} \mathbf{P}[\operatorname{span}(Y_1,\ldots,Y_r,X_1,\ldots,X_{n-k-r}) = V].\end{aligned}$$

Combining this with (2.1), we can conclude

$$\mathbf{P}[\operatorname{span}(X_1,\ldots,X_n) = V] \leq 2^{2n} \cdot \gamma^{km} \cdot \sum_{r=0}^{m} \mathbf{P}[\operatorname{span}(Y_1,\ldots,Y_r,X_1,\ldots,X_{n-k-r}) = V]$$

for every $V \in \mathcal{V}_{\text{thick}}$. Summing this for all $V \in \mathcal{V}_{\text{thick}}$ gives

$$\begin{aligned}\mathbf{P}[\operatorname{span}(X_1,\ldots,X_n) \in \mathcal{V}_{\text{thick}}] &= \sum_{V\in\mathcal{V}_{\text{thick}}} \mathbf{P}[\operatorname{span}(X_1,\ldots,X_n) = V] \\ &\leq 2^{2n} \cdot \gamma^{km} \cdot \sum_{r=0}^{m} \sum_{V\in\mathcal{V}_{\text{thick}}} \mathbf{P}[\operatorname{span}(Y_1,\ldots,Y_r,X_1,\ldots,X_{n-k-r}) = V] \\ &\leq 2^{2n} \cdot \gamma^{km} \cdot \sum_{r=0}^{m} 1 \leq n2^{2n} \cdot \gamma^{km} \leq n2^{2n} \cdot \gamma^{nk/2}.\end{aligned}$$

□

## 3. Proof of Proposition 2.1

It remains to prove Proposition 2.1. The proof is modeled after the Fourier comparison argument in [12]; the crucial trick is replacing a certain "doubling" inequality of Raikov [8] (which morally is the continuous analogue of the Cauchy–Davenport theorem, a *one-dimensional* result about additive doubling) with a suitable "high-dimensional" variant. More precisely, we will use the following "doubling" property for subsets of the torus $\mathbf{T}^k := (\mathbf{R}/\mathbf{Z})^k$ satisfying an appropriate coordinate restriction. Throughout the remainder of the paper, we will use $\mu(\cdot)$ to denote Lebesgue measure on $\mathbf{T}$, $\mathbf{T}^d$ and $\mathbf{R}^d$.

**Lemma 3.1.** *Let $A_1,\ldots,A_k \subseteq \mathbf{T}$ be closed subsets with $\mu(A_i) \leq 1/2$ for $i = 1,\ldots,k$. Then for any closed set $S \subseteq A_1 \times \cdots \times A_k \subseteq \mathbf{T}^k$, we have*

$$\mu(S+S) \geq 2^k \cdot \mu(S).$$

In the case where $A_1,\ldots,A_k = [0,1/2]$, the sumset $S+S$ has no "wraparound" and therefore Lemma 3.1 follows from the Brunn–Minkowski inequality (see e.g. [11, Theorem 3.16]). The general result (proven in the next section) reduces to this case via compressions.

We are now ready to prove Proposition 2.1.

*Proof of Proposition 2.1.* As in the statement of the proposition, let $p \in (0, 1/32]$ and consider an $(n-k)$-dimensional linear subspace $V \subseteq \mathbf{Q}^n$. Let $L$ be a $k \times n$ matrix whose rows form a basis of the orthogonal complement of $V$. By permuting columns (which does not change the probability of the underlying event), we may assume that the first $k \times k$ block of $L$ is nonsingular. By performing row operations we may assume that this $k$ by $k$ block is diagonal. By rescaling the rows of $L$ we may assume that all the entries of $L$ are integral and furthermore that the diagonal entries are equal to some integer $Z$ (but note that we have no control over the size of $Z$). Finally we let $w_1,\ldots,w_n$ denote the columns of $L$.

Recall that $\mathbf{T}^k$ is the $k$-dimensional torus $(\mathbf{R}/\mathbf{Z})^k$. For a random vector $X \sim \mu_{1/2}^{\otimes n}$ we have (by the Fourier inversion formula)

$$\begin{aligned}
\sup_{t \in \mathbf{Q}^n} \mathbf{P}[X + t \in V] &= \sup_{t' \in \mathbf{Z}^k} \mathbf{P}[LX = t'] \\
&= \sup_{t' \in \mathbf{Z}^k} \int_{\mathbf{T}^k} \mathbf{E}[\exp(2\pi i \theta^T (LX - t'))]\, d\theta \\
&\le \int_{\mathbf{T}^k} \Big| \mathbf{E}[\exp(2\pi i \theta^T LX)] \Big|\, d\theta \\
&= \int_{\mathbf{T}^k} \prod_{j=1}^n \Big| \frac{1}{2}\exp(2\pi i \theta^T w_j) + \frac{1}{2}\exp(-2\pi i \theta^T w_j) \Big|\, d\theta \\
&= \int_{\mathbf{T}^k} \prod_{j=1}^n |\cos(2\pi \theta^T w_j)|\, d\theta \\
&= \int_{\mathbf{T}^k} \prod_{j=1}^n |\cos(\pi \theta^T w_j)|\, d\theta.
\end{aligned}$$

In the final line, we have applied the change of variable $\theta \to \theta/2$ and noted that $|\cos(\theta + \pi)| = |\cos(\theta)|$ to rewrite the integral.

Via a similar computation, for a random vector $Y \sim \mu_p^{\otimes n}$ we have

$$\mathbf{P}[Y \in V] = \mathbf{P}[LY = 0] = \int_{\mathbf{T}^k} \mathbf{E}[\exp(2\pi i \theta^T LY)]\, d\theta = \int_{\mathbf{T}^k} \prod_{j=1}^n (1 - 2p + 2p\cos(2\pi\theta^T w_j))\, d\theta.$$

Note that $1 - 2p + 2p\cos(\varphi) \ge 1 - 4p \ge 0$ for all $\varphi \in \mathbf{R}$, so all factors in this integral are nonnegative everywhere. Thus, it suffices to show that

$$\int_{\mathbf{T}^k} \prod_{j=1}^n |\cos(\pi\theta^T w_j)|\, d\theta \le \gamma^k \cdot \int_{\mathbf{T}^k} \prod_{j=1}^n (1 - 2p + 2p\cos(2\pi\theta^T w_j))\, d\theta \tag{3.1}$$

for some absolute constant $\gamma < 1$. To this end, we use the elementary trigonometric inequality (which is essentially [12, Lemma 7.1 (19)]), whose proof can be found at the end of this section.

**Lemma 3.2.** *For $p \in (0, 1/32]$, and any $\varphi, \varphi' \in \mathbf{R}$, we have*

$$|\cos(\varphi)| \cdot |\cos(\varphi')| \le (1 - 2p + 2p\cos(2\varphi + 2\varphi'))^2.$$

Let us now fix a small absolute constant $\beta > 0$ (small enough to satisfy certain inequalities later in the proof), and define $\tau = e^{-\beta k}$. Note that we can rewrite the left hand side of (3.1) as

$$\int_{\mathbf{T}^k} \prod_{j=1}^n |\cos(\pi\theta^T w_j)|\, d\theta = \int_{\mathbf{T}^k} \min\Big(\prod_{j=1}^n |\cos(\pi\theta^T w_j)|, \tau\Big)\, d\theta + \int_{\mathbf{T}^k} \max\Big(\prod_{j=1}^n |\cos(\pi\theta^T w_j)| - \tau, 0\Big)\, d\theta.$$

To bound the first summand, note that Lemma 3.2 (applied, for $j = 1, \dots, n$, to $\varphi = \pi\theta^T w_j$ and $\varphi' = 0$) yields

$$\prod_{j=1}^n |\cos(\pi\theta^T w_j)| \le \prod_{j=1}^n (1 - 2p + 2p\cos(2\pi\theta^T w_j))^2$$

for all $\theta \in \mathbf{T}^k$. Therefore, we obtain the bound

$$\int_{\mathbf{T}^k} \min\Big(\prod_{j=1}^n |\cos(\pi\theta^T w_j)|, \tau\Big)\, d\theta \le \int_{\mathbf{T}^k} \tau^{1/2} \prod_{j=1}^n |\cos(\pi\theta^T w_j)|^{1/2}\, d\theta \le \tau^{1/2} \int_{\mathbf{T}^k} \prod_{j=1}^n (1 - 2p + 2p\cos(2\pi\theta^T w_j))\, d\theta$$

for the first summand. To handle the second summand, we define

$$S_\eta := \Big\{\theta \in \mathbf{T}^k : \prod_{j=1}^n |\cos(\pi\theta^T w_j)| \ge \eta\Big\}$$

for any $\eta \in [\tau, 1]$. Then we have

$$\int_{\mathbf{T}^k} \max\Big(\prod_{j=1}^{n} |\cos(\pi\theta^T w_j)| - \tau, 0\Big)\, d\theta = \int_{\tau}^{1} \mu(S_\eta)\, d\eta.$$

Recall that the first $k$ columns of $L$ form the $k \times k$ matrix $ZI_k$ (i.e., a diagonal matrix with $Z$ everywhere on the diagonal). Therefore each $\theta \in S_\eta \subseteq \mathbf{T}^k$ satisfies

$$\prod_{i=1}^{k} |\cos(\pi Z\theta_i)| = \prod_{j=1}^{k} |\cos(\pi\theta^T w_j)| \geq \prod_{j=1}^{n} |\cos(\pi\theta^T w_j)| \geq \eta \geq \tau = e^{-\beta k}.$$

Note that whenever $Z\theta_i \notin \mathbf{Z} + [-1/4, 1/4]$, we have $|\cos(\pi Z\theta_i)| \leq 2^{-1/2}$. Therefore, for any $\theta \in S_\eta$, there are at most $M := \lfloor 2\beta k/\log(2)\rfloor$ coordinates $i \in [k]$ with $Z\theta_i \notin \mathbf{Z} + [-1/4, 1/4]$. For each subset $I \subseteq [k]$ of size $|I| \leq M$, let $B_I \subseteq \mathbf{T}^k$ denote the "box" of all points $\theta \in \mathbf{T}^k$ with $Z\theta_i \notin \mathbf{Z} + [-1/4, 1/4]$ for all indices $i \in I$ and $Z\theta_i \in \mathbf{Z} + [-1/4, 1/4]$ for all indices $i \in [k] \setminus I$. Then for each $\eta \in [\tau, 1]$, we have $S_\eta \subseteq \bigcup_I B_I$, where the union is taken over all subsets $I \subseteq [k]$ of size $|I| \leq M$. The number of such subsets $I$ is

$$\sum_{j=0}^{M} \binom{k}{j} \leq k \cdot (ek/M)^M \leq (3/2)^k,$$

provided that the absolute constant $\beta > 0$ was chosen to be sufficiently small in the beginning of the proof (and provided that $k$ is larger than a suitable absolute constant). Thus, by the pigeonhole principle, there exists a subset $I \subseteq [k]$ such that $\mu(S_\eta \cap B_I) \geq (2/3)^k \mu(S_\eta)$. Now, the set $S_\eta \cap B_I$ satisfies the assumption of Lemma 3.1, and applying the lemma we obtain

$$\mu(S_\eta + S_\eta) \geq \mu((S_\eta \cap B_I) + (S_\eta \cap B_I)) \geq 2^k \mu(S_\eta \cap B_I) \geq (4/3)^k \cdot \mu(S_\eta).$$

Now, for any $\theta, \theta' \in S_\eta$, by Lemma 3.2, we have

$$\prod_{j=1}^{n}(1 - 2p + 2p\cos(2\pi(\theta+\theta')^T w_j)) \geq \prod_{j=1}^{n} |\cos(\pi\theta^T w_j)|^{1/2} \cdot |\cos(\pi\theta'^T w_j)|^{1/2} \geq \eta^{1/2} \cdot \eta^{1/2} = \eta.$$

Thus, we can conclude that

$$\mu(S_\eta) \leq (3/4)^k \cdot \mu(S_\eta + S_\eta) \leq (3/4)^k \cdot \mu\Big(\Big\{\theta \in \mathbf{T}^k : \prod_{j=1}^{n}(1 - 2p + 2p\cos(2\pi\theta^T w_j)) \geq \eta\Big\}\Big)$$

for all $\eta \in [\tau, 1]$. Integrating this over the interval $[\tau, 1]$ yields

$$\int_{\mathbf{T}^k} \max\Big(\prod_{j=1}^{n} |\cos(\pi\theta^T w_j)| - \tau, 0\Big)\, d\theta = \int_{\tau}^{1} \mu(S_\eta)\, d\eta \leq (3/4)^k \cdot \int_{\mathbf{T}^k} \prod_{j=1}^{n}(1 - 2p + 2p\cos(2\pi\theta^T w_j))\, d\theta.$$

All in all, we can conclude

$$\int_{\mathbf{T}^k} \prod_{j=1}^{n} |\cos(\pi\theta^T w_j)|\, d\theta \leq (\tau^{1/2} + (3/4)^k) \cdot \int_{\mathbf{T}^k} \prod_{j=1}^{n}(1 - 2p + 2p\cos(2\pi\theta^T w_j))\, d\theta.$$

This shows the desired inequality (3.1), setting $\gamma = e^{-\beta/4}$, and observing that then we have $\tau^{1/2} + (3/4)^k \leq e^{-\beta k/2} + (3/4)^k \leq \gamma^k$ (assuming that $\beta$ was chosen to be sufficiently small and $k$ is sufficiently large). □

We end this section with the proof of Lemma 3.2, and postpone the proof of Lemma 3.1 to the next section.

*Proof of Lemma 3.2.* First, note that both sides of the inequality are $\pi$-periodic, so we may assume without loss of generality that $\varphi, \varphi' \in [-\pi/2, \pi/2]$. Since $\partial^2/\partial\varphi^2 \log(\cos(\varphi)) = -\cos(\varphi)^{-2} \leq 0$ for $\varphi \in (-\pi/2, \pi/2)$, the function $\log\cos(\varphi)$ is concave on $(-\pi/2, \pi/2)$ and therefore by Jensen's inequality we have

$$|\cos(\varphi)| \cdot |\cos(\varphi')| = \cos(\varphi) \cdot \cos(\varphi') \leq (\cos(\varphi/2 + \varphi'/2))^2.$$

Noting that $\cos(\varphi/2+\varphi'/2) \geq 0$ (since $\varphi/2+\varphi'/2 \in [-\pi/2,\pi/2]$), and furthermore $1-2p+2p\cos(2\varphi+2\varphi') \geq 15/16+(1/16)\cdot\cos(2\varphi+2\varphi')$ (since $p \leq 1/32$), it now suffices to prove that

$$\cos(\varphi/2+\varphi'/2) \leq \frac{15}{16}+\frac{1}{16}\cdot\cos(2\varphi+2\varphi').$$

We define $x=\cos(\varphi/2+\varphi'/2)$. Then, recalling that $\cos(2\alpha)=(\cos\alpha)^2-(\sin\alpha)^2=2(\cos\alpha)^2-1$ for all $\alpha \in \mathbf{R}$, we have $\cos(2\varphi+2\varphi')=2(\cos(\varphi+\varphi'))^2-1=2(2x^2-1)^2-1=8x^4-8x^2+1$. Now, it suffices to check that

$$x \leq x+\frac{(x-1)^2\cdot((x+1)^2+1)}{2}=x+\frac{(x^2-2x+1)\cdot(x^2+2x+2)}{2}=\frac{x^4}{2}-\frac{x^2}{2}+1=\frac{15}{16}+\frac{1}{16}\cdot(8x^4-8x^2+1)$$

to finish the proof of the lemma. □

## 4. Proof of Lemma 3.1

We first recall an inequality due to Raikov [8]. This may also be derived from the Cauchy–Davenport inequality (see e.g. [11, Theorem 5.4]) via a limiting argument.

**Theorem 4.1.** *Consider closed sets $A,B \subseteq \mathbf{T}$. Then*

$$\mu(A+B) \geq \min(\mu(A)+\mu(B),1).$$

We next define the *compression* of a closed set $S \subseteq \mathbf{T}^k$ in the $i$-th coordinate direction, for $i \in [k]$: For $(\theta_1,\ldots,\theta_k) \in \mathbf{T}^k$, where $\theta_1,\ldots,\theta_k \in [0,1)$, let us say that $(\theta_1,\ldots,\theta_k) \in \pi_i(S)$ if and only if

$$\theta_i \leq \int_{\mathbf{T}} \mathbf{1}_{(\theta_1,\ldots,\theta_{i-1},z,\theta_{i+1},\ldots,\theta_k)\in S}\, dz.$$

We observe some properties which are immediate by construction.

**Fact 4.2.** *Consider any $i \in [k]$. Then for any closed set $S \subseteq \mathbf{T}^k$, the set $\pi_i(S)$ is closed, and we have $\mu(\pi_i(S))=\mu(S)$. Furthermore, for any closed sets $S \subseteq S' \subseteq \mathbf{T}^k$, we have $\pi_i(S) \subseteq \pi_i(S')$.*

The crucial property is that $\pi_i(S)$ has smaller sumset than $S$. This is a continuous analogue of certain standard facts about compressions of discrete sets.

**Lemma 4.3.** *Consider any closed set $S \subseteq \mathbf{T}^k$. We have that*

$$\mu(\pi_i(S)+\pi_i(S)) \leq \mu(S+S).$$

*Proof.* By symmetry, it suffices to prove the case when $i=k$. By Fact 4.2, observe it suffices to show that $\pi_k(S)+\pi_k(S) \subseteq \pi_k(S+S)$. So, fix any $\theta,\psi \in \pi_k(S)$; we will prove that $\theta+\psi \in \pi_k(S+S)$.

First, it is convenient to introduce some notation: given a set $U \subseteq \mathbf{T}^k$ and $\chi \in \mathbf{T}^k$, let us define $U_\chi=\{z \in \mathbf{T}:(\chi_1,\ldots,\chi_{k-1},z) \in U\}$. Note that $U_\chi$ is independent of the $k$-th coordinate of $\chi$.

Now, note that $\mu(S_\theta) \geq \theta_k$ and $\mu(S_\psi) \geq \psi_k$. By Theorem 4.1 and the inclusion $(S+S)_{\theta+\psi} \subseteq S_\theta+S_\psi$, we have that $\mu((S+S)_{\theta+\psi}) \geq \mu(S_\theta+S_\psi) \geq \min(1,\theta_k+\psi_k)$. This implies that $\theta_k+\psi_k \bmod 1 \in (\pi_k(S+S))_{\theta+\psi}$. Therefore $\theta+\psi \in \pi_k(S+S)$, which completes the proof. □

**Lemma 4.4.** *Let $A_1,\ldots,A_k \subseteq \mathbf{T}$ be closed sets. Then $\pi_1\circ\cdots\circ\pi_k(A_1\times\cdots\times A_k)=[0,\mu(A_1)]\times\cdots\times[0,\mu(A_k)]$.*

*Proof.* Note that for arbitrary closed sets $B_1,\ldots,B_k \subset \mathbf{T}$, we have

$$\pi_\ell(B_1\times\cdots\times B_k)=B_1\times\cdots\times B_{\ell-1}\times[0,\mu(B_\ell)]\times B_{\ell+1}\times\cdots\times B_k.$$

Whence by a simple inductive argument, we have that

$$\pi_\ell\circ\cdots\circ\pi_k(A_1\times\cdots\times A_k)=A_1\times\cdots\times A_{\ell-1}\times[0,\mu(A_\ell)]\times\cdots\times[0,\mu(A_k)].$$

The desired result follows by the case $\ell=1$. □

We now give the proof of Lemma 3.1.

*Proof of Lemma 3.1.* By Lemma 4.3 and Fact 4.2 (both applied $k$ times), it suffices to prove that

$$\mu(\pi_1 \circ \cdots \circ \pi_k(S) + \pi_1 \circ \cdots \circ \pi_k(S)) \geq 2^k \mu(\pi_1 \circ \cdots \circ \pi_k(S)).$$

Note by Lemma 4.4 and the second part of Fact 4.2, we have that

$$\pi_1 \circ \cdots \circ \pi_k(S) \subseteq [0, 1/2]^k.$$

Therefore we may identify $\pi_1 \circ \cdots \circ \pi_k(S)$ as a subset of $\mathbf{R}^k$ (i.e., there can be no "wrap-around" when forming $\pi_1 \circ \cdots \circ \pi_k(S) + \pi_1 \circ \cdots \circ \pi_k(S))$ and conclude with the Brunn–Minkowski inequality[3] (see e.g. [11, Theorem 3.16]). □

## References

[1] Jean Bourgain, Van H. Vu, and Philip Matchett Wood, *On the singularity probability of discrete random matrices*, J. Funct. Anal. **258** (2010), 559–603. 1, 2

[2] Gábor Halász, *On the distribution of additive arithmetic functions*, Acta Arith. **27** (1975), 143–152. 1

[3] Vishesh Jain, Ashwin Sah, and Mehtaab Sawhney, *Rank deficiency of random matrices*, Electronic Communications in Probability **27** (2022), 1–9. 1

[4] Jeff Kahn, János Komlós, and Endre Szemerédi, *On the probability that a random* ±1*-matrix is singular*, Journal of the American Mathematical Society **8** (1995), 223–240. 1, 2

[5] Martin Kneser, *Summenmengen in lokalkompakten abelschen Gruppen*, Math. Z. **66** (1956), 88–110. 2

[6] János Komlós, *On the determinant of* (0, 1) *matrices*, Studia Sci. Math. Hungar. **2** (1967), 7–21. 1

[7] A. M. Macbeath, *On measure of sum sets. II. The sum-theorem for the torus*, Proc. Cambridge Philos. Soc. **49** (1953), 40–43. 2

[8] D. Raikov, *On the addition of point-sets in the sense of Schnirelmann*, Rec. Math. [Mat. Sbornik] N.S. **5/47** (1939), 425–440. 2, 4, 7

[9] Mark Rudelson, *A large deviation inequality for the rank of a random matrix*, Ann. Probab. **52** (2024), 1992–2018. 1

[10] Mark Rudelson and Roman Vershynin, *The Littlewood-Offord problem and invertibility of random matrices*, Adv. Math. **218** (2008), 600–633. 1

[11] Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006. 4, 7, 8

[12] Terence Tao and Van Vu, *On random* ±1 *matrices: singularity and determinant*, Random Structures Algorithms **28** (2006), 1–23. 1, 2, 4, 5

[13] Terence Tao and Van Vu, *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. **20** (2007), 603–628. 1, 4

[14] Konstantin Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), 593–634. 1

Department of Mathematics, ETH Zürich, Zürich, Switzerland.
*Email address*: zach.hunter@math.ethz.ch

Institute of Science and Technology Austria (ISTA). Am Campus 1, 3400 Klosterneuburg, Austria
*Email address*: matthew.kwan@ist.ac.at

Institute for Applied Mathematics, University of Bonn, Germany
*Email address*: sauermann@iam.uni-bonn.de

Department of Mathematics, Columbia University, New York, NY 10027
*Email address*: m.sawhney@columbia.edu

[3] Formally, there is a possibility of wraparound at the boundary of $[0, 1/2]^k$, but the boundary has measure zero so this causes no problems. Specifically, we may consider $(\pi_1 \circ \cdots \circ \pi_k(S)) \cap [0, 1/2)^k$ and only then apply the Brunn–Minkowski inequality in $\mathbf{R}^k$, noting that $\mu((\pi_1 \circ \cdots \circ \pi_k(S)) \cap [0, 1/2)^k) = \mu(\pi_1 \circ \cdots \circ \pi_k(S))$.