Combinatorics, Probability and Computing (2021), **30**, pp. 163–174
doi:10.1017/S0963548320000413

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Resilience of the rank of random matrices

Asaf Ferber[1], Kyle Luh[2] and Gweneth McKinley[3]*

[1]Department of Mathematics, University of California, Irvine, CA 92697, USA, [2]Department of Mathematics, University of Colorado Boulder, Campus Box 395, 2300 Colorado Avenue, Boulder, CO 80309-0395, USA and [3]Department of Mathematics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0112, USA
*Corresponding author. Email: gmckinley@ucsd.edu

(Received 8 October 2019; revised 18 June 2020; first published online 28 August 2020)

Abstract

Let M be an $n \times m$ matrix of independent Rademacher (± 1) random variables. It is well known that if $n \leqslant m$, then M is of full rank with high probability. We show that this property is resilient to adversarial changes to M. More precisely, if $m \geqslant n + n^{1-\varepsilon/6}$, then even after changing the sign of $(1-\varepsilon)m/2$ entries, M is still of full rank with high probability. Note that this is asymptotically best possible as one can easily make any two rows proportional with at most $m/2$ changes. Moreover, this theorem gives an asymptotic solution to a slightly weakened version of a conjecture made by Van Vu in [17].

2020 MSC Codes: 60B20 - Random Matrices (primary classification); 05 - Combinatorics; 60 - Probability

1. Introduction

Random discrete matrices, in particular 0/1 and ± 1 random matrices, have a distinguished history in random matrix theory. They have applications in computer science, physics and random graph theory, among others, and numerous investigations have been tailored to this class of random matrices [1, 8, 9, 11, 13, 14, 15]. Discrete random matrices are often of interest in their own right as they pose combinatorial questions that are vacuous or trivial for other models such as the Gaussian ensembles (*e.g.* singularity and simpleness of spectrum). For example, it is already non-trivial to show that a Bernoulli (0/1) random matrix is non-singular with probability $1 - o(1)$ (this was first proved by Komlós in [9]). For an $n \times n$ Bernoulli random matrix M_n, it was a long-standing conjecture that

$$p_n := \mathbb{P}(M_n \text{ is singular}) = \left(\frac{1}{2} + o(1)\right)^n,$$

which corresponds to the probability that any two rows or columns are identical. This problem has stimulated much activity [1, 8, 13], culminating in Tikhomirov's recent resolution of the above conjecture [15].

In this work, we examine another aspect of the singularity problem for discrete random matrices. We will be concerned with robustness of the non-singularity, meaning how many changes to the entries of the matrix need to be performed to make a typical random matrix singular. This has been called the 'resilience' of a random matrix with respect to singularity [17]. Note that an $n \times n$

matrix is singular if and only if its rank is less than n. Therefore we can extend the above notion for general matrices (not necessarily square) as follows.

Definition 1.1. Given an $n \times m$ matrix M with entries in $\{\pm 1\}$, we let $\mathrm{Res}(M)$ denote the minimum number of sign flips necessary in order to make M of rank less than n.

Note that for every two ± 1 vectors $a, b \in \{\pm 1\}^m$, one can easily achieve either $a = b$ or $a = -b$ by changing at most $m/2$ entries. So, in particular, for an $n \times m$ matrix M we have the deterministic upper bound

$$\mathrm{Res}(M) \leqslant m/2.$$

Indeed, for the case $n = m$ it is conjectured by Vu that

$$\mathrm{Res}(M_n) = \left(\frac{1}{2} + o(1)\right) n$$

with probability $1 - o(1)$ [17, Conjecture 7.4]. Note that by a simple union bound, using any exponential upper bound on p_n, one can easily show that a.a.s. we have

$$\mathrm{Res}(M_n) \geqslant cn/\log n$$

for some appropriate choice of $c > 0$. Surprisingly, no better lower bound is known.

In this paper we prove that for $m \geqslant (1 + o(1))n$ the trivial upper bound $m/2$ is asymptotically tight. Before stating our main result we define the following notation: given $n, m \in \mathbb{N}$, we let $M_{n,m}$ be an $n \times m$ matrix with independent entries chosen uniformly from $\{\pm 1\}$.

Theorem 1.1. *For every $\varepsilon > 0$ and $m \geqslant n + n^{1-\varepsilon/6}$, a.a.s. we have*

$$\mathrm{Res}(M_{n,m}) \geqslant (1 - \varepsilon)m/2.$$

Our proof strategy goes roughly as follows. Consider an outcome M of $M_{n,m}$. Note that if the rank of M is less than n, then in particular, writing $m' = m - n^{1-\varepsilon/6}$, there exists an $n \times m'$ submatrix M' of M with rank less than n. Moreover, as M' is not of full rank, there exists $a \in \mathbb{R}^n \setminus \{0\}$ which lies in the *left kernel of* M' (*i.e.* with $a^T M' = 0$). Our main goal is to show that for each such a (if it exists) and for a randomly chosen $x \in \{\pm 1\}^n$, the probability

$$\rho(a) \coloneqq \mathbb{P}[a^T x = 0]$$

is typically very small.

Next, observe that a vector a will be in the left kernel of M if and only if it is in the left kernel of M' and is also orthogonal to the remaining $n^{1-\varepsilon/6}$ columns of M. Therefore, using the bound on $\rho(a)$ and the extra $n^{1-\varepsilon/6}$ columns of M, we want to 'boost' the probability and show that

$$\mathbb{P}[\exists a \text{ such that } a^T M = 0] = n^{-(1/2 - o(1))m}. \tag{1.1}$$

Note that since there are at most

$$\binom{nm}{(1/2 - o(1))m} \approx n^{(1/2 + o(1))m}$$

matrices that can be obtained from M by changing $s \leqslant (1/2 - o(1))m$ entries, and there are at most 2^m such choices for M', using the above bound we complete the proof by a simple union bound (and, of course, showing that the $o(1)$ terms in (1.1) work in our favour).

The main challenge is to prove (1.1), as it involves a union bound over all possible $a \in \mathbb{R}^n$. In order to overcome this difficulty, we use some recently developed machinery introduced in [4]. Roughly speaking, we embed the problem into a sufficiently large finite field $\mathbb{F}_p$. Then, as there are finitely many options for $a \in \mathbb{F}_p$ in the left kernel of M, we can use a counting argument

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

from [4] to bound the probability of encountering each possible kernel vector a according to the corresponding value of $\rho(a)$.

We mention that the approach of bounding $\rho(a)$ for possible null-vectors in the context of singularity is not new (see *e.g.* [8], [11], [12], [14] and [16]). The novelty of our argument is that we utilize the methods in [4] to obtain the bound (1.1). Most of the previously used arguments yield exponential or polynomial probabilities which would only tolerate a sublinear number of modifications to the matrix. Although it is possible to modify the previous arguments to generate super-exponential bounds, the exact constant of $1/2$ in (1.1) seems to be difficult to achieve via other arguments.

Lastly, we mention that the method in [4] has already been successfully applied to a variety of combinatorial problems in random matrix theory [2, 3, 6, 7, 10].

The remainder of this paper is organized as follows. In Section 2 we provide the necessary background to state the counting lemma from [4]. In Section 2.3 we provide a convenient interface to apply the counting lemma. This is drawn from [4] as well. Finally, in Section 3 we provide the short proof of Theorem 1.1.

2. Auxiliary results

Here we review some auxiliary results and introduce convenient notation to be used in the proof of our main result.

2.1 Halász inequality in $\mathbb{F}_p$

Let $a \coloneqq (a_1, \dots, a_n) \in (\mathbb{Z} \setminus \{0\})^n$ and let $\varepsilon_1, \dots, \varepsilon_n$ be independent and identically distributed (i.i.d.) Rademacher random variables; that is, each ε_i independently takes values ± 1 with probability $1/2$ each. We define the largest atom probability $\rho(a)$ by

$$\rho(a) \coloneqq \sup_{x \in \mathbb{Z}} \mathbb{P}(\varepsilon_1 a_1 + \cdots + \varepsilon_n a_n = x).$$

Similarly, if we are working over some finite field $\mathbb{F}_p$, let

$$\rho_{\mathbb{F}_p}(a) \coloneqq \sup_{x \in \mathbb{F}_p} \mathbb{P}(\varepsilon_1 a_1 + \cdots + \varepsilon_n a_n = x),$$

where, of course, the arithmetic is done over $\mathbb{F}_p$.

Now, let $R_k(a)$ denote the number of solutions to $\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} \equiv 0$, where repetitions are allowed in the choice of $i_1, \dots, i_{2k} \in [n]$. A classical theorem of Halász [5] gives an estimate on the atom probability based on $R_k(a)$. Here we need the following, slightly different version of this theorem, which can be applied to the finite field setting.

Theorem 2.1. Halász's inequality over $\mathbb{F}_p$; Theorem 1.4 in [4]. *There exists an absolute constant C such that the following holds for every odd prime p, integer n, and vector $a \coloneqq (a_1, \dots, a_n) \in \mathbb{F}_p^n \setminus \{0\}$. Suppose that an integer $k \geqslant 1$ and a positive real number M satisfy $30M \leqslant |\mathrm{supp}(a)|$ and $80kM \leqslant n$. Then*

$$\rho_{\mathbb{F}_p}(a) \leqslant \frac{1}{p} + \frac{CR_k(a)}{2^{2k} n^{2k} \cdot M^{1/2}} + e^{-M}.$$

2.2 Counting lemma

In this section we state a counting lemma from [4] which plays a key role in our proof. First we need the following definition.

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

Definition 2.1. Suppose that $a \in \mathbb{F}_p^n$ for an integer n and a prime p and let $k \in \mathbb{N}$. For every $\alpha \in [0,1]$, we define $R_k^\alpha(a)$ to be the number of solutions to

$$\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} = 0 \mod p$$

that satisfy $|\{i_1, \ldots, i_{2k}\}| \geqslant (1+\alpha)k$.

It is important that we introduce $R_k^\alpha(a)$ as this set is more combinatorially tractable. However, it is easily seen that $R_k(a)$ cannot be much larger than $R_k^\alpha(a)$. This is formalized in the following simple lemma, which is proved in [4].

Lemma 2.1. *For all $k, n \in \mathbb{N}$ with $k \leqslant n/2$, and any prime p, vector $a \in \mathbb{F}_p^n$ and $\alpha \in [0,1]$,*

$$R_k(a) \leqslant R_k^\alpha(a) + (40k^{1-\alpha}n^{1+\alpha})^k.$$

Proof. By definition, $R_k(a)$ is equal to $R_k^\alpha(a)$ plus the number of solutions to $\pm a_{i_1} \pm a_{i_2} \cdots \pm a_{i_{2k}} = 0$ that satisfy $|\{i_1, \ldots, i_{2k}\}| < (1+\alpha)k$. The latter quantity is bounded from above by the number of sequences $(i_1, \ldots, i_{2k}) \in [n]^{2k}$ with at most $(1+\alpha)k$ distinct entries times 2^{2k}, the number of choices for the $\pm$ signs. Thus

$$R_k(a) \leqslant R_k^\alpha(a) + \binom{n}{(1+\alpha)k}((1+\alpha)k)^{2k}2^{2k} \leqslant R_k^\alpha(a) + (4e^{1+\alpha}k^{1-\alpha}n^{1+\alpha})^k,$$

where the final inequality follows from the well-known bound $\binom{a}{b} \leqslant (ea/b)^b$. Finally, noting that $4e^{1+\alpha} \leqslant 4e^2 \leqslant 40$ completes the proof. □

Given a vector $a \in \mathbb{F}_p^n$ and a subset of coordinates $I \subseteq [n]$, we define a_I to be its restriction to the coordinates in I; that is, $a_I = (a_i)_{i \in I} \in \mathbb{F}_p^I$. We write $b \subseteq a$ if there exists an $I \subseteq [n]$ for which $b = a_I$. For $b \subseteq a$ we let $|b|$ be the size of the subset I determining b.

Now we are ready to state the counting lemma from [4].

Theorem 2.2. Theorem 1.7 in [4]. *Let p be a prime, let $k, n \in \mathbb{N}$, $s \in [n]$, $t \in [p]$, and let $\alpha \in (0,1)$. Denoting*

$$B_{k,s,\geqslant t}^\alpha(n) := \left\{ a \in \mathbb{F}_p^n \colon R_k^\alpha(b) \geqslant t \cdot \frac{2^{2k} \cdot |b|^{2k}}{p} \text{ for every } b \subseteq a \text{ with } |b| \geqslant s \right\},$$

we have

$$|B_{k,s,\geqslant t}^\alpha(n)| \leqslant \left(\frac{s}{n}\right)^{2k-1} (\alpha t)^{s-n} p^n.$$

2.3 'Good' and 'bad' vectors

The purpose of this section is to formulate easy-to-use versions of Halász's inequality (Theorem 2.1) and our counting theorem (Theorem 2.2). This follows [4] closely, but requires a more delicate choice of parameters as we need to achieve the bound in (1.1) (and crucially, the constant 1/2 in the exponent). We shall partition $\mathbb{F}_p^n$ into 'good' and 'bad' vectors. We shall then show that, on the one hand, every 'good' vector a has a small $\rho_{\mathbb{F}_p}(a)$ and that, on the other hand, there are relatively few 'bad' vectors.[a] The formal statements now follow. In order to simplify the notation, we suppress the implicit dependence of the defined notions on n, k, p and α.

[a] In fact, we shall only show that there are relatively few 'bad' vectors that have some number of non-zero coordinates. The number of remaining vectors (ones with very small support) is so small that even a very crude estimate will suffice for our needs.

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

Definition 2.2. Let p be a prime, let $n, k \in \mathbb{N}$ and let $\alpha \in (0,1)$. For any $t > 0$, define the set H_t of *t-good* vectors by

$$H_t := \left\{ a \in \mathbb{F}_p^n \colon \exists b \subseteq a \text{ with } |\mathrm{supp}(b)| \geqslant n^{1-\varepsilon/2} \text{ and } R_k^\alpha(b) \leqslant t \cdot \frac{2^{2k} \cdot |b|^{2k}}{p} \right\}.$$

The *goodness* of a vector $a \in \mathbb{F}_p^n$, denoted by $h(a)$, will be the smallest t such that $a \in H_t$. In other words

$$h(a) = \min \left\{ \frac{p \cdot R_k^\alpha(b)}{2^{2k} \cdot |b|^{2k}} \colon b \subseteq a \text{ and } |\mathrm{supp}(b)| \geqslant n^{1-\varepsilon/2} \right\}.$$

Intuitively, $h(a)$ captures the segment of a with the smallest amount of arithmetic structure. Therefore the smaller $h(a)$, the stronger the anti-concentration bound as quantified below in Lemma 2.2.

Notice also that if a vector $a \in \mathbb{F}_p^n$ has fewer than $n^{1-\varepsilon/2}$ non-zero coordinates, then it cannot be t-good for any t and thus $h(a) = \infty$. On the other hand, trivially $R_k^\alpha(b) \leqslant 2^{2k} \cdot |b|^{2k}$ for every vector b, as there are $2^{2k}|b|^{2k}$ total possible choices of a sequence $\pm b_{i_1} \pm b_{i_2} \pm \cdots \pm b_{i_{2k}}$. Thus every $a \in \mathbb{F}_p^n$ with at least $n^{1-\varepsilon/2}$ non-zero coordinates must be p-good, that is, $h(a) \leqslant p$ for each such a.

Having formalized the notion of a 'good' vector, we are now ready to state and prove two corollaries of Theorems 2.1 and 2.2 that lie at the heart of our approach. (Note: the particular choice of parameters in Lemma 2.2 is made for convenience in a later application.)

Lemma 2.2. *Let $a \in H_t$, let $\alpha \in (0,1)$ and let $\varepsilon < 1/100$. Suppose that $p = \Theta(2^{n^{\varepsilon/3}})$ is a prime, $t \geqslant n$, and $k = \Theta(n^{\varepsilon/3})$. Then, for sufficiently large n, we have*

$$\rho_{\mathbb{F}_p}(a) \leqslant \frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}},$$

where $C = C(\alpha, \varepsilon)$ is a constant depending only on α and ε.

Proof. As $a \in H_t$, we can find a subvector b of a such that $|\mathrm{supp}(b)| \geqslant n^{1-\varepsilon/2}$ and $R_k^\alpha(b) \leqslant t \cdot 2^{2k} \cdot |b|^{2k}/p$. We will focus only on the entries of b, and bound $\rho_{\mathbb{F}_p}(b)$; since $b \subseteq a$, it is straightforward to see via conditioning that $\rho_{\mathbb{F}_p}(a) \leqslant \rho_{\mathbb{F}_p}(b)$.

Set $M = \lfloor n^{1-\varepsilon/2}/(80k) \rfloor = \Theta(n^{1-5\varepsilon/6})$ so that

$$\max\{30M, 80Mk\} = 80Mk \leqslant n^{1-\varepsilon/2} \leqslant |\mathrm{supp}(b)| \leqslant |b|.$$

Thus we may apply Theorem 2.1 to obtain, for some absolute constant C_0,

$$\rho_{\mathbb{F}_p}(b) \leqslant \frac{1}{p} + \frac{C_0 R_k(b)}{2^{2k} \cdot |b|^{2k} \cdot M^{1/2}} + e^{-M}.$$

Now, using Lemma 2.1 we can upper-bound the right-hand side by

$$\begin{aligned}
\rho_{\mathbb{F}_p}(b) &\leqslant \frac{1}{p} + \frac{C_0 R_k^\alpha(b) + C_0(40k^{1-\alpha}|b|^{1+\alpha})^k}{2^{2k} \cdot |b|^{2k} \cdot M^{1/2}} + e^{-M} \\
&\leqslant \frac{1}{p} + \frac{C_0 t \cdot 2^{2k} \cdot |b|^{2k}/p + C_0(40k^{1-\alpha}|b|^{1+\alpha})^k}{2^{2k} \cdot |b|^{2k} \cdot M^{1/2}} + e^{-M} \\
&= \frac{1}{p}\left(1 + \frac{C_0 t}{M^{1/2}} + C_0(10(k/|b|)^{1-\alpha})^k \cdot \frac{p}{M^{1/2}}\right) + e^{-M}.
\end{aligned}$$

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

Now we wish to show that with the parameter assignments above, the dominant term in this sum is $C_0 t/(pM^{1/2})$. To this end, we bound each of the other terms as follows. First,

$$e^{-M} = e^{-\Theta(n^{1-5\varepsilon/6})} = o(2^{-n^{\varepsilon/3}}) = o\left(\frac{1}{p}\right).$$

(Here we use the upper bound assumption on ε.) Second,

$$\begin{aligned} C_0(10(k/|b|)^{1-\alpha})^k \cdot \frac{p}{M^{1/2}} &\leqslant C_0(10(n^{\varepsilon/3-(1-\varepsilon/2)})^{1-\alpha})^k \cdot p \\ &= (n^{-\Theta(1)})^{\Theta(n^{\varepsilon/3})} \cdot p \\ &= 2^{-\Theta(n^{\varepsilon/3}\log n)} \cdot \Theta(2^{n^{\varepsilon/3}}) \\ &= o(1). \end{aligned}$$

And last, we observe that, as $t \geqslant n$,

$$\frac{C_0 t}{M^{1/2}} \geqslant \frac{n}{\Theta(n^{\frac{1}{2}(1-5\varepsilon/6)})} = \omega(1).$$

Therefore the dominant term in the sum above is indeed $C_0 t/(pM^{1/2})$. Then, choosing the constant $C = C(\alpha, \varepsilon) > C_0$ sufficiently large, we obtain

$$\rho_{\mathbb{F}_p}(b) \leqslant \frac{Ct}{pM^{1/2}} \leqslant \frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}$$

as desired. (Note: in the last step we have incorporated the implicit constant in $M = \Theta(n^{1-5\varepsilon/6})$ into the constant C.) □

Lemma 2.3. *For every integer n and real number $t \geqslant n$,*

$$|\{a \in \mathbb{F}_p^n \colon |\mathrm{supp}(a)| \geqslant n^{1-\varepsilon/2} \text{ and } a \notin H_t\}| \leqslant 2^n \left(\frac{p}{\alpha t}\right)^n \cdot t^{n^{1-\varepsilon/2}}.$$

Proof. We may assume that $t \leqslant p$, as otherwise the left-hand side above is zero; see the comment below Definition 2.2. Let us now fix an $S \subseteq [n]$ with $|S| \geqslant n^{1-\varepsilon/2}$ and count only vectors a with $\mathrm{supp}(a) = S$. Since $a \notin H_t$, the restriction a_S of a to the set S must be contained in the set $B^{\alpha}_{k,n^{1-\varepsilon/2},\geqslant t}(|S|)$. Hence the number of choices for a_S is at most $|B^{\alpha}_{k,n^{1-\varepsilon/2},\geqslant t}(|S|)|$, which Theorem 2.2 bounds as follows:

$$|B^{\alpha}_{k,n^{1-\varepsilon/2},\geqslant t}(|S|)| \leqslant \left(\frac{n^{1-\varepsilon/2}}{|S|}\right)^{2k-1} (\alpha t)^{n^{1-\varepsilon/2}-|S|} p^{|S|}.$$

Since $|S| \geqslant n^{1-\varepsilon/2}$, we can simply write

$$\left(\frac{n^{1-\varepsilon/2}}{|S|}\right)^{2k-1} \leqslant 1.$$

For the remaining terms, we slightly rewrite it as

$$(\alpha t)^{n^{1-\varepsilon/2}-|S|} p^{|S|} = \left(\frac{p}{\alpha t}\right)^{|S|} (\alpha t)^{n^{1-\varepsilon/2}},$$

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

and then, since $|S| \leqslant n$ and $\alpha t \leqslant t \leqslant p$, we can bound this from above as

$$|B^{\alpha}_{k,n^{1-\varepsilon/2},\geqslant t}(|S|)| \leqslant \left(\frac{p}{\alpha t}\right)^n t^{n^{1-\varepsilon/2}}.$$

Now, since a_S completely determines a, we obtain the desired conclusion by summing the above bound over all sets S. □

3. Proof of Theorem 1.1

In this section we gradually construct the entire proof of Theorem 1.1.

For convenience, we introduce some notation to indicate the distance of two Rademacher matrices.

Definition 3.1. For two $n \times m$ matrices M, M' we let $d(M, M')$ denote the number of entries where M and M' differ.

With this definition in hand, Theorem 1.1 can be stated as follows.

Theorem 3.1. *For every $\varepsilon > 0$ and $m \geqslant n + n^{1-\varepsilon/6}$, a.a.s. we have rank$(M') = n$ for all $n \times m$, ± 1 matrices M' with $d(M_{n,m}, M') \leqslant (1-\varepsilon)m/2$.*

First we will prove Theorem 1.1 under the assumption that $m = \omega(n)$.

3.1 Proof of Theorem 1.1 under the assumption $m = \omega(n)$

Let $\varepsilon > 0$ be any fixed constant, and let $m \geqslant C(\varepsilon)n$, where $C(\varepsilon)$ is a sufficiently large constant. We wish to show that a.a.s. $M = M_{n,m}$ is such that every $n \times m$ matrix M' with $d(M, M') \leqslant (1-\varepsilon)m/2$ has rank n.

In order to do so, let us take (say) $p = 3$ and work over $\mathbb{F}_3$. Observe that if the above statement holds over $\mathbb{F}_3$ then it trivially holds over $\mathbb{Z}$.

Let $a \in \mathbb{F}_3^n \setminus \{\mathbf{0}\}$, and note that for a randomly chosen $x \in \{\pm 1\}^n$ we have

$$\mathbb{P}[a^T x = 0\,] \leqslant \frac{1}{2}.$$

Therefore, as the columns of M are independent, it follows that the random variable $X_a =$ 'the number of zeros in $a^T M$' is stochastically dominated by Bin$(m, 1/2)$. Hence, by Chernoff's bound, we obtain that

$$\mathbb{P}[X_a \geqslant (1+\varepsilon)m/2] \leqslant e^{-C_1 m}$$

for some C_1 that depends on ε. By applying the union bound over all $a \in \mathbb{F}_3^n \setminus \{0\}$, we obtain that

$$\mathbb{P}[\exists a \in \mathbb{F}_3^n \setminus \{0\} \text{ with } X_a \geqslant (1+\varepsilon)m/2] \leqslant 3^n e^{-C_1 m} = o(1),$$

where the last inequality follows from the fact that $m \geqslant C(\varepsilon)n$ and $C(\varepsilon)$ is sufficiently large.

Thus M is typically such that in *every* non-zero linear combination of its rows, there are less than $(1+\varepsilon)m/2$ zeros. In particular, since by changing at most $(1-\varepsilon)m/2$ entries one can affect at most $(1-\varepsilon)m/2$ columns, it follows that for all M' with $d(M, M') \leqslant (1-\varepsilon)m/2$, no non-trivial combination of the rows of M' is the 0 vector. In particular, every such M' is of rank n. This completes the proof for this case. □

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

3.2 Proof of Theorem 1.1 under the assumption $m = O(n)$

In what follows we always assume that $m = O(n)$. Therefore, whenever convenient, in appropriate asymptotic formulas we may switch between m and n without further explanation. This case is more involved than the case $m = \omega(n)$ and it will be further divided into a few subcases. From now on, we fix p to be some prime $p = \Theta(2^{n^{\varepsilon/3}})$, and concretely, we write $m \leqslant C(\varepsilon)n$ for some constant $C(\varepsilon)$.

Now, write $m' = m - \lfloor n^{1-\varepsilon/6} \rfloor$ (the width of the matrix under consideration minus the $\lfloor n^{1-\varepsilon/6} \rfloor$ 'extra' columns). In the following two subsections, we will show that with high probability, for every $a \in \mathbb{F}_p^n$, if $a^T M' = 0$ for some $n \times m'$ matrix M' with $d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2$, then a has 'many' non-zero entries, and is 'pseudorandom' in some sense (Lemma 3.1 and Lemma 3.2). From here, we can apply the Halász inequality (in the form of Lemma 2.2) almost directly, using the fact that there are $m - m' = \lfloor n^{1-\varepsilon/6} \rfloor$ extra columns, to conclude that for any such a, the probability that $a^T M_{n,m} = 0$ is small.

3.2.1 Eliminating small linear dependencies

First we wish to show that if $a^T M' = 0$ (over $\mathbb{F}_p$) for some M' with $d(M_{n,m'}, M') \leqslant (1-\varepsilon)m/2$, then a has 'many' non-zero entries (assuming $a \neq 0$ of course).

Lemma 3.1. *Let $\varepsilon > 0$, let $p = \Theta(2^{n^{\varepsilon/3}})$ be a prime and let $n + n^{1-\varepsilon/6} \leqslant m \leqslant C(\varepsilon)n$. Write $m' = m - \lfloor n^{1-\varepsilon/6} \rfloor$. Then, working in $\mathbb{F}_p$, the probability that there exists a matrix M' with $d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2$ and a non-zero vector $a \in \mathbb{F}_p^n$ with $|\mathrm{supp}(a)| \leqslant n^{1-\varepsilon/2}$ and with $a^T M' = 0$ is at most $2^{-\Theta(n)}$.*

Proof. Given a vector $a \in \mathbb{F}_p^n$, we let $\ell := |\mathrm{supp}(a)|$. Note that for any $a \neq 0$ and a uniformly chosen vector $x \in \{\pm 1\}^n$, we trivially have

$$\mathbb{P}[a^T x = 0] \leqslant \frac{1}{2}.$$

Moreover, as we are only allowed to change at most $(1-\varepsilon)m/2$ coordinates of $M_{n,m'}$, it follows that at most $(1-\varepsilon)m/2$ entries of $a^T M_{n,m'}$ can be altered. In particular, if there exists a vector a for which $a^T M' = 0$, where $d(M_{n,m'}, M') \leqslant (1-\varepsilon)m/2$, then this implies that $a^T M_{n,m'}$ already contained at least $m' - (1-\varepsilon)m/2 = (1 + \varepsilon - o(1))m/2$ zero entries.

Now, since the random variable counting the number of 0 entries is stochastically dominated by $\mathrm{Bin}(n, 1/2)$, by Chernoff's bound we obtain that for a given $a \neq 0$, the probability of having at least $(1 + \varepsilon - o(1))m/2$ zeros in $a^T M_{n,m'}$ is at most $2^{-c(\varepsilon)m}$, where $c(\varepsilon)$ is some constant depending only on ε. Thus the probability that for a given non-zero vector a there exists some M' with $d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2$ and $a^T M' = 0$ is at most $2^{-c(\varepsilon)m}$.

All in all, by applying the union bound over all $a \neq 0$ with $\ell \leqslant n^{1-\varepsilon/2}$ non-zero entries, the probability that we are seeking to bound is at most

$$\sum_{\ell=1}^{n^{1-\varepsilon/2}} \binom{n}{\ell} p^\ell 2^{-c(\varepsilon)m} \leqslant \sum_{\ell=1}^{n^{1-\varepsilon/2}} 2^{\ell \log n + \ell n^{\varepsilon/3} - c(\varepsilon)m} = 2^{-\Theta(n)},$$

where the last equality holds due to the assumption $\ell \leqslant n^{1-\varepsilon/2}$. □

3.2.2 Eliminating 'bad' vectors

We now show that, almost surely, any vector a with many non-zero entries and with $a^T M' = 0$ for some M' with $d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2$ will be 'good' or 'unstructured'.

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

Lemma 3.2. *Let $\varepsilon > 0$, let $p = \Theta(2^{n^{\varepsilon/3}})$ be a prime and let $n + n^{1-\varepsilon/6} \leqslant m \leqslant C(\varepsilon)n$. Write $m' = m - \lfloor n^{1-\varepsilon/6} \rfloor$. Then, working in $\mathbb{F}_p$, the probability that there exists a matrix M' with $d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2$ and a vector $a \in \mathbb{F}_p^n \setminus H_n$ with at least $n^{1-\varepsilon/2}$ non-zero entries such that $a^T M' = 0$ is at most $2^{-\Theta(n \log n)}$.*

Proof. Our first step is to take a union bound over choices of a; we wish to bound the quantity

$$\sum_{\substack{a \in \mathbb{F}_p^n \setminus H_n \\ |\mathrm{supp}(a)| \geqslant n^{1-\varepsilon/2}}} \mathbb{P}[\exists M' \text{ with } d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2 \text{ and } a^T M' = 0]. \tag{3.1}$$

Now we use the sets H_t to divide the vectors a into different classes. As observed after Definition 2.2, every $a \in \mathbb{F}_p^n$ with at least $n^{1-\varepsilon/2}$ non-zero entries is in H_t for some $t \leqslant p$. Moreover, notice that $H_t \subseteq H_{t+1}$ for any $t > 0$. So we can write $\mathbb{F}_p^n \setminus H_n$ as a union $\bigcup_{n+1 \leqslant t \leqslant p} H_t \setminus H_{t-1}$. Therefore, taking a union bound over integers $t > n$, the probability (3.1) that we are trying to bound is at most

$$\sum_{t=n+1}^{p} \left(\sum_{a \in H_t \setminus H_{t-1}} \mathbb{P}[\exists M' \text{ with } d(M', M_{n,m'}) \leqslant (1-\varepsilon)m/2 \text{ and } a^T M' = 0] \right).$$

Now we take another union bound, this time over the possible edits to the matrix; by changing at most $(1-\varepsilon)m/2$ entries, an adversary can form

$$\sum_{i=0}^{(1-\varepsilon)m/2} \binom{nm'}{i} \leqslant \left(\frac{2en}{1-\varepsilon} \right)^{(1-\varepsilon)m/2} = 2^{(1-\varepsilon+o(1))(m/2)\log n}$$

$n \times m'$ matrices. Thus (3.1) is at most

$$\sum_{t=n+1}^{p} \left(\sum_{a \in H_t \setminus H_{t-1}} 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot \mathbb{P}[a^T M_{n,m'} = 0] \right).$$

(Note: this is possible because by conditioning on the locations of the entries edited, each altered matrix M' is distributed identically to $M_{n,m'}$.)

We now wish to bound the probability that $a^T M_{n,m'} = 0$ for any fixed $a \in H_t \setminus H_{t-1}$. By Lemma 2.2 (as $a \in H_t$), and by the independence of the columns in $M_{n,m'}$, this probability is at most

$$\left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}} \right)^{m'}.$$

Therefore (3.1) is at most

$$\sum_{t=n+1}^{p} \left(\sum_{a \in H_t \setminus H_{t-1}} 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot \left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}} \right)^{m'} \right).$$

We now bound the number of vectors a in each $H_t \setminus H_{t-1}$. By definition, $H_t \setminus H_{t-1} \subset \mathbb{F}_p^n \setminus H_{t-1}$, and by Lemma 2.3, the size of $\mathbb{F}_p^n \setminus H_{t-1}$ is bounded above by

$$\left(\frac{2p}{\alpha t} \right)^n \cdot t^{n^{1-\varepsilon/2}},$$

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

where $\alpha \in (0,1)$ is any fixed constant (note that the constant C above depends on α). Thus (3.1) is bounded by the following explicit expression:

$$\sum_{t=n+1}^{p} \left(\frac{2p}{\alpha t}\right)^n \cdot t^{n^{1-\varepsilon/2}} \cdot 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot \left(\frac{Ct}{pn^{\frac{1}{2}(1-5\varepsilon/6)}}\right)^{m'}$$
$$= 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot n^{-(1-5\varepsilon/6)(m'/2)} \cdot \left(\frac{2}{\alpha}\right)^n \cdot C^{m'} \sum_{t=n+1}^{p} \left(\frac{t}{p}\right)^{m'-n} t^{n^{1-\varepsilon/2}}.$$

Now, bounding each piece separately, and recalling that $n \leqslant m' = O(n)$,

$$\left(\frac{2}{\alpha}\right)^n C^{m'} = 2^{O(n)},$$
$$\sum_{t=n+1}^{p} \left(\frac{t}{p}\right)^{m'-n} t^{n^{1-\varepsilon/2}} \leqslant p \cdot 1 \cdot p^{n^{1-\varepsilon/2}} = 2^{n^{\varepsilon/3}} \cdot 2^{n^{\varepsilon/3} \cdot n^{1-\varepsilon/2}} = 2^{o(n)},$$
$$2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot n^{-(1-5\varepsilon/6)(m'/2)} = 2^{-(1-o(1))(\varepsilon/12)\cdot m\log n},$$

where in the last equality we use the fact that $m' = m - \lfloor n^{1-\varepsilon/6} \rfloor = (1-o(1))m$. Thus, in total, (3.1) is at most

$$2^{(-\varepsilon/12+o(1))m\log n} = 2^{-\Theta(n\log n)}.$$

This completes the proof of the lemma. □

3.2.3 Completing the proof

Given the assumption $m \leqslant C(\varepsilon)n$, we will in fact prove something slightly stronger, namely that Theorem 1.1 holds over $\mathbb{F}_p$ for an appropriate choice of p. We wish to bound the probability that there exists some non-zero vector $a \in \mathbb{F}_p^n$ with $a^T M_{n,m} = 0$, even after at most $(1-\varepsilon)m/2$ edits. Let $p = \Theta(2^{n^{\varepsilon/3}})$ be prime. We begin by dividing into 'structured' and 'unstructured' vectors; for brevity, given a non-zero vector a and matrix M, we let $\mathcal{E}(a, M)$ denote the event that there exists a matrix M' with $d(M', M) \leqslant (1-\varepsilon)m/2$ and $a^T M' = 0$.

$$\mathbb{P}[\exists\, a \in \mathbb{F}_p^n \text{ with } \mathcal{E}(a, M_{n,m})]$$
$$\leqslant \mathbb{P}[\exists\, a \in H_n \text{ with } \mathcal{E}(a, M_{n,m})] \tag{3.2}$$
$$+ \mathbb{P}[\exists\, a \in \mathbb{F}_p^n \setminus H_n \text{ with } \mathcal{E}(a, M_{n,m})], \tag{3.3}$$

where H_n is the set of 'good' or 'unstructured' vectors defined in Section 2.3. The first summand (3.2) is bounded as follows. First take a union bound over possible edits to $M_{n,m}$. There are

$$\sum_{i=0}^{(1-\varepsilon)m/2} \binom{nm}{i} = 2^{(1-\varepsilon+o(1))(m/2)\log n}$$

possible choices for M'. Thus, for the first term (3.2), we obtain a bound of

$$2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot \mathbb{P}[\exists\, a \in H_n \text{ with } a^T M_{n,m} = 0].$$

(As in the proof of Lemma 3.2, this is possible because, by conditioning on the locations of the entries edited, each M' is distributed identically to $M_{n,m}$.) And for $a \in H_n$, and $x \in \{\pm 1\}^n$ chosen uniformly at random, Lemma 2.2 gives

$$\mathbb{P}[a^T x = 0] \leqslant \frac{Cn}{pn^{(1/2-5\varepsilon/12)}} < \frac{n}{p}.$$

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

So for $M_{n,m}$ with $m \geqslant n + n^{1-\varepsilon/6}$ columns, the probability of having $a^T M_{n,m} = 0$ is at most

$$\left(\frac{n}{p}\right)^{n+n^{1-\varepsilon/6}}.$$

Therefore, as there are at most p^n vectors $a \in H_n$, and as $m \leqslant C(\varepsilon) \cdot n$, the first summand (3.2) is bounded by

$$\begin{aligned} 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot p^n \left(\frac{n}{p}\right)^{n+n^{1-\varepsilon/6}} &= 2^{(1-\varepsilon+o(1))(m/2)\log n} \cdot p^{-n^{1-\varepsilon/6}} n^{n+n^{1-\varepsilon/6}} \\ &= 2^{O(n\log n)} \cdot 2^{-n^{\varepsilon/3} n^{1-\varepsilon/6}} \\ &= 2^{-\Theta(n^{1+\varepsilon/6})}. \end{aligned}$$

Now we bound the second summand (3.3). We begin by restricting to the first $m' = m - \lfloor n^{1-\varepsilon/6} \rfloor$ columns of $M_{n,m}$. This gives a strictly larger probability, as it is more likely that there is a linear dependence among the rows of a matrix when we restrict to only a subset of its columns. So (3.3) is bounded above by

$$\begin{aligned} &\mathbb{P}[\exists\, a \in \mathbb{F}_p^n \setminus H_n \text{ with } \mathcal{E}(a, M_{n,m'})] \\ &\quad \leqslant \mathbb{P}[\exists\, a \in \mathbb{F}_p^n \setminus H_n \text{ with } |\mathrm{supp}(a)| \geqslant n^{1-\varepsilon/2} \text{ and } \mathcal{E}(a, M_{n,m'})] \\ &\qquad + \mathbb{P}[\exists\, a \in \mathbb{F}_p^n \setminus H_n \text{ with } |\mathrm{supp}(a)| < n^{1-\varepsilon/2} \text{ and } \mathcal{E}(a, M_{n,m'})], \end{aligned}$$

and these are respectively the precise probabilities bounded in Lemmas 3.2 and 3.1. Therefore this is at most

$$2^{-\Theta(n\log n)} + 2^{-\Theta(n)}.$$

Thus, in total, the probability that there exists a non-zero vector $a \in \mathbb{F}_p^n$ with $a^T M_{n,m} = 0$, even after at most $(1-\varepsilon)m/2$ edits, is at most

$$2^{-\Theta(n^{1+\varepsilon/6})} + 2^{-\Theta(n\log n)} + 2^{-\Theta(n)} = 2^{-\Theta(n)}.$$

Acknowledgement

The authors would like to thank Wojciech Samotij for many fruitful discussions.

References

[1] Bourgain, J., Vu, V. H. and Wood, P. M. (2010) On the singularity probability of discrete random matrices. *J. Funct. Anal.* **258** 559–603.
[2] Campos, M., Mattos, L., Morris, R. and Morrison, N. (2019) On the singularity of random symmetric matrices. arXiv:1904.11478
[3] Ferber, A. and Jain, V. (2019) Singularity of random symmetric matrices–a combinatorial approach to improved bounds. *Forum of Mathematics, Sigma* **7** E22. doi:10.1017/fms.2019.21
[4] Ferber, A., Jain, V., Luh, K. and Samotij, W. (2019) On the counting problem in inverse Littlewood–Offord theory. arXiv:1904.10425
[5] Halász, G. (1977) Estimates for the concentration function of combinatorial number theory and probability. *Period. Math. Hungar.* **8** 197–211.
[6] Jain, V. (2019) The strong circular law: a combinatorial view. arXiv:1904.11108
[7] Jain, V. (2019) Approximate Spielman–Teng theorems for the least singular value of random combinatorial matrices. To appear in *Israel J. Math.* arXiv:1904.10592
[8] Kahn, J., Komlós, J. and Szemerédi, E. (1995) On the probability that a random ±1-matrix is singular. *J. Amer. Math. Soc.* **8** 223–240.

https://doi.org/10.1017/S0963548320000413 Published online by Cambridge University Press

[9] Komlós, J. (1967) On the determinant of (0, 1) matrices. *Studia Sci. Math. Hungar* **2** 7–21.
[10] Luh, K., Meehan, S. and Nguyen, H. H. (2019) Random matrices over finite fields: methods and results. arXiv:1907.02575
[11] Nguyen, H. H. (2013) On the singularity of random combinatorial matrices. *SIAM J. Discrete Math.* **27** 447–458.
[12] Rudelson, M. and Vershynin, R. (2008) The Littlewood–Offord problem and invertibility of random matrices. *Adv. Math.* **218** 600–633.
[13] Tao, T. and Vu, V. (2007) On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.* **20** 603–628.
[14] Tao, T. and Vu, V. H. (2009) Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. *Ann. of Math. (2)* **169** 595–632.
[15] Tikhomirov, K. (2020) Singularity of random Bernoulli matrices. *Annals of Mathematics* **191**, 593–634.
[16] Vershynin, R. (2014) Invertibility of symmetric random matrices. *Random Struct. Algorithms* **44** 135–182.
[17] Vu, V. (2008) Random discrete matrices. In *Horizons of Combinatorics*, Vol. 17 of Bolyai Society Mathematical Studies, pp. 257–280. Springer.

Cite this article: Ferber A, Luh K and McKinley G (2021). Resilience of the rank of random matrices. *Combinatorics, Probability and Computing* **30**, 163–174. https://doi.org/10.1017/S0963548320000413