Combinatorics, Probability and Computing (2022), **31**, pp. 21–28
doi:10.1017/S0963548321000146

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Singularity of sparse random matrices: simple proofs

Asaf Ferber[1,*†], Matthew Kwan[2,‡] and Lisa Sauermann[3,§]

[1]Department of Mathematics, University of California, Irvine, CA 92697, USA [2]Department of Mathematics, Stanford University, Stanford, CA 94305, USA and [3]School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA
*Corresponding author. Email: mattkwan@stanford.edu

(Received 2 November 2020; revised 20 April 2021; first published online 15 June 2021)

Abstract

Consider a random $n \times n$ zero-one matrix with 'sparsity' p, sampled according to one of the following two models: either every entry is independently taken to be one with probability p (the 'Bernoulli' model) or each row is independently uniformly sampled from the set of all length-n zero-one vectors with exactly pn ones (the 'combinatorial' model). We give simple proofs of the (essentially best-possible) fact that in both models, if $\min(p, 1-p) \geq (1+\varepsilon)\log n/n$ for any constant $\varepsilon > 0$, then our random matrix is nonsingular with probability $1 - o(1)$. In the Bernoulli model, this fact was already well known, but in the combinatorial model this resolves a conjecture of Aigner-Horev and Person.

2020 MSC Codes: 60B20

1. Introduction

Let M be an $n \times n$ random matrix with i.i.d. Bernoulli (p) entries (meaning that each entry M_{ij} satisfies $\mathbb{P}(M_{ij} = 1) = p$ and $\mathbb{P}(M_{ij} = 0) = 1 - p$). It is a famous theorem of Komlós [15, 16] that for $p = 1/2$ a random Bernoulli matrix is *asymptotically almost surely* nonsingular: that is, $\lim_{n\to\infty} \mathbb{P}(M \text{ is singular}) = 0$. Komlós' theorem can be generalised to sparse random Bernoulli matrices as follows.

Theorem 1.1. *Fix $\varepsilon > 0$, and let $p = p(n)$ be any function of n satisfying* $\min(p, 1-p) \geq (1+\varepsilon)\log n/n$. *Then for a random $n \times n$ random matrix M with i.i.d.* Bernoulli (p) *entries, we have*

$$\lim_{n\to\infty} \mathbb{P}(M \text{ is singular}) = 0.$$

Theorem 1.1 is best-possible, in the sense that if $\min(p, 1-p) \leq (1-\varepsilon)\log n/n$, then we actually have $\lim_{n\to\infty} \mathbb{P}(M \text{ is singular}) = 1$ (because, for instance, M is likely to have two identical columns). That is to say, $\log n/n$ is a *sharp threshold* for singularity. It is not clear when Theorem 1.1 first appeared in print, but strengthenings and variations on Theorem 1.1 have been proved by several different authors (see, for example, [1, 3, 5, 6]).

Next, let Q be an $n \times n$ random matrix with independent rows, where each row is sampled uniformly from the subset of vectors in $\{0, 1\}^n$ having exactly d ones (Q is said to be a random

†Research supported in part by NSF Awards DMS-1954395 and DMS-1953799.
‡Research supported by NSF Award DMS-1953990.
§Research supported by NSF Grant CCF-1900460 and NSF Award DMS-2100157.

combinatorial matrix). The study of such matrices was initiated by Nguyen [19], who proved that if $d = n/2$ then Q is asymptotically almost surely nonsingular (where $n \to \infty$ along the even integers). Strengthenings of Nguyen's theorem have been proved by several authors, see, for example, [2, 10, 12, 13, 23]. Recently, Aigner-Horev and Person [2] conjectured an analogue of Theorem 1.1 for sparse random combinatorial matrices, which we prove in this note.

Theorem 1.2. *Fix $\varepsilon > 0$, and let $d = d(n)$ be any function of n satisfying $\min(d, n-d) \geq (1+\varepsilon)\log n$. Then for an $n \times n$ random zero-one matrix Q with independent rows, where each row is chosen uniformly among the vectors with d ones, we have*

$$\lim_{n\to\infty} \mathbb{P}(Q \text{ is singular}) \to 0.$$

Just like Theorem 1.1, Theorem 1.2 is best-possible in the sense that if $\min(d, n-d) \leq (1-\varepsilon)\log n$, then $\lim_{n\to\infty} \mathbb{P}(M \text{ is singular}) = 1$. Theorem 1.2 improves on a result of Aigner-Horev and Person: they proved the same fact under the assumption that $\lim_{n\to\infty} d/(n^{1/2}\log^{3/2} n) = \infty$ (assuming that $d \leq n/2$).

The structure of this note is as follows. First, in Section 2 we prove a simple and general lemma (Lemma 2.1) which applies to any random matrix with i.i.d. rows. This lemma distills the essence of (a special case of) an argument due to Rudelson and Vershinyn [22]. Essentially, it shows that in order to prove Theorems 1.1 and 1.2, one just needs to prove some relatively crude estimates about the typical structure of the vectors in the left and right kernels of our random matrices.

Then, in Sections 3 and 4 we show how to use Lemma 2.1 to give simple proofs of Theorem 1.1 and Theorem 1.2. Of course, Theorem 1.1 is not new, but its proof is extremely simple and it serves as a warm-up for Theorem 1.2. It turns out that in order to analyse the typical structure of the vectors in the left and right kernel, we can work over $\mathbb{Z}_q$ for some small integer q (in fact, we can mostly work over $\mathbb{Z}_2$). This idea is not new (see, for example, [2, 4, 8, 9, 10, 11, 18, 20, 21]), but the details here are much simpler.

We remark that with a bit more work, the methods in our proofs can also likely be used to prove the conclusions of Theorems 1.1 and 1.2 under the weaker (and strictly best-possible) assumptions that $\lim_{n\to\infty}(\min(pn, n-pn) - \log n) = \infty$ and $\lim_{n\to\infty}(\min(d, n-d) - \log n) = \infty$. However, in this note we wish to emphasise the simple ideas in our proofs and do not pursue this direction.

Notation. All logarithms are to base e. We use common asymptotic notation, as follows. For real-valued functions $f(n)$ and $g(n)$, we write $f = O(g)$ to mean that there is some constant $C > 0$ such that $|f| \leq Cg$. If g is nonnegative, we write $f = \Omega(g)$ to mean that there is $c > 0$ such that $f \geq cg$ for sufficiently large n. We write $f = o(g)$ to mean that $f(n)/g(n) \to 0$ as $n \to \infty$.

2. A general lemma

In this section, we prove a (very simple) lemma which will give us a proof scheme for both Theorems 1.1 and 1.2. For a vector x, let $\operatorname{supp}(x)$ (the *support* of x) be the set of indices i such that $x_i \neq 0$.

Lemma 2.1. *Let $\mathbb{F}$ be a field, and let $A \in \mathbb{F}^{n\times n}$ be a random matrix with i.i.d. rows $R_1, \dots, R_n$. Let $\mathcal{P} \subseteq \mathbb{F}^n$ be any property of vectors in $\mathbb{F}^n$. Then for any $t \in \mathbb{R}$, the probability that A is singular is upper-bounded by*

$$\mathbb{P}(x^T A = 0 \text{ for some nonzero } x \in \mathbb{F}^n \text{ with } |\operatorname{supp}(x)| < t) \tag{1}$$

$$+ \frac{n}{t}\mathbb{P}(\text{there is nonzero } x \notin \mathcal{P} \text{ such that } x \cdot R_i = 0 \text{ for all } i = 1, \dots, n-1) \tag{2}$$

$$+ \frac{n}{t}\sup_{x\in\mathcal{P}} \mathbb{P}(x \cdot R_n = 0). \tag{3}$$

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

Proof. Note that A is singular if and only if there is a nonzero $x \in \mathbb{F}^n$ satisfying $x^T A = 0$. Let $\mathcal{E}_i$ be the event that $R_i \in \text{span}\{R_1, \dots, R_{i-1}, R_{i+1}, \dots, R_n\}$, and let X be the number of i for which $\mathcal{E}_i$ holds. Then by Markov's inequality and the assumption that the rows $R_1, \dots, R_n$ are i.i.d., we have

$$\mathbb{P}\Big(x^T M = 0 \text{ for some } x \text{ with } |\operatorname{supp}(x)| \geq t\Big) \leq \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t} = \frac{n}{t}\mathbb{P}(\mathcal{E}_n).$$

It now suffices to show that $\frac{n}{t}\mathbb{P}(\mathcal{E}_n)$ is upper-bounded by the sum of the terms (2) and (3). Note that we can always choose a nonzero vector $x \in \mathbb{F}^n$ with $x \cdot R_i = 0$ for $i = 1, \dots, n-1$. We interpret x as a random vector depending on $R_1, \dots, R_{n-1}$ (but not R_n). If the event $\mathcal{E}_n$ occurs, we must have $x \cdot R_n = 0$, so

$$\frac{n}{t}\mathbb{P}(\mathcal{E}_n) \leq \frac{n}{t}\mathbb{P}(x \notin \mathcal{P}) + \frac{n}{t}\mathbb{P}(x \cdot R_n = 0 \mid x \in \mathcal{P}).$$

Then $\frac{n}{t}\mathbb{P}(x \notin \mathcal{P})$ is upper-bounded by the expression in (2), and, since x and R_n are independent, $\frac{n}{t}\mathbb{P}(x \cdot R_n = 0 \mid x \in \mathcal{P})$ is upper-bounded by the expression in (3). □

3. Singularity of sparse Bernoulli matrices: a simple proof

Let us fix $0 < \varepsilon < 1$. We will take $t = cn$ for some small constant c (depending on ε), and let $\mathcal{P}$ be the property $\{x \in \mathbb{Q}^n : |\operatorname{supp}(x)| \geq t\}$. All we need to do is to show that the three terms (1), (2) and (3) in Lemma 2.1 are each of the form $o(1)$. The following lemma is the main part of the proof.

Lemma 3.1. *Let $R_1, \dots, R_{n-1}$ be the first $n-1$ rows of a random* Bernoulli(p) *matrix, with $\min(p, 1-p) \geq (1+\varepsilon)\log n/n$. There is $c > 0$ (depending only on ε) such that with probability $1 - o(1)$, no nonzero vector $x \in \mathbb{Q}^n$ with $|\operatorname{supp}(x)| < cn$ satisfies $R_i \cdot x = 0$ for all $i = 1, \dots, n-1$.*

Proof. If such a vector x were to exist, we would be able to multiply by an integer and then divide by a power of two to obtain a vector $v \in \mathbb{Z}^n$ with at least one odd entry also satisfying $|\operatorname{supp}(v)| < cn$ and $R_i \cdot v = 0$ for $i = 1, \dots, n-1$. Interpreting v as a vector in $\mathbb{Z}_2^n$, we would have $R_i \cdot v \equiv 0 \pmod 2$ for $i = 1, \dots, n-1$ and furthermore $v \in \mathbb{Z}_2^n$ would be a nonzero vector consisting of less than cn ones. We show that such a vector v is unlikely to exist (working over $\mathbb{Z}_2$ discretises the problem, so that we may use a union bound).

Let $p^* = \min(p, 1-p) \geq (1+\varepsilon)\log n/n$. Consider any $v \in \{0,1\}^n$ with $|\operatorname{supp}(v)| = s$. Then $R_i \cdot v$ for $i = 1, \dots, n-1$ are i.i.d. Binomial(s, p) random variables. Let $P_{s,p}$ denote the probability that a Binomial(s, p) random variable is even. We observe

$$\begin{aligned} P_{s,p} &= \frac{1}{2}\left(\sum_{i=0}^{s}\binom{s}{i}p^i(1-p)^{s-i} + \sum_{i=0}^{s}\binom{s}{i}(-1)^i p^i(1-p)^{s-i}\right) \\ &= \frac{1}{2} + \frac{(1-2p)^s}{2} \leq \frac{1}{2} + \frac{(1-2p^*)^s}{2}. \end{aligned}$$

Then, using the fact that $e^{-t} = 1 - t + O(t^2)$ for $t = o(1)$, we deduce

$$P_{s,p} \leq \begin{cases} e^{-(1+o(1))sp^*} & \text{if } sp^* = o(1), \\ e^{-\Omega(1)} & \text{if } sp^* = \Omega(1). \end{cases}$$

Taking $r = \delta/p^*$ for sufficiently small δ (relative to ε), and recalling that $p^* \geq (1+\varepsilon)\log n/n$, the probability that there exists nonzero $v \in \mathbb{Z}_2^n$ with $|\operatorname{supp}(v)| < cn$ and $R_i \cdot v \equiv 0 \pmod 2$ for all $i = 1, \dots, n-1$ is at most

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

$$\sum_{s=1}^{cn}\binom{n}{s}P_{s,p}^{n-1} \le \sum_{s=1}^{r} e^{s\log n-(1-\varepsilon/3)snp^*} + \sum_{s=r+1}^{cn} e^{s(\log(n/s)+1)-\Omega(n)}$$
$$\le \sum_{s=1}^{\infty} n^{-s\varepsilon/3} + \sum_{s=1}^{cn} e^{n\left((s/n)(\log(n/s)+1)-\Omega(1)\right)} = o(1),$$

provided c is sufficiently small (relative to δ). □

Taking c as in Lemma 3.1, we immediately see that the term (2) is of the form $o(1)$. Observing that the rows and columns of M have the same distribution, and that the event $x^TM = 0$ is simply the event that $x \cdot C_i = 0$ for each column C_i of M; it also follows from Lemma 3.1 that the term (1) is of the form $o(1)$. Finally, the following straightforward generalisation of the well-known Erdős–Littlewood–Offord theorem shows that the term (3) is of the form $o(1)$, which completes the proof of Theorem 1.1. This lemma is the only nontrivial ingredient in the proof of Theorem 1.1. It appears as [5, Lemma 8.2], but it can also be quite straightforwardly deduced from the Erdős–Littlewood–Offord theorem itself.

Lemma 3.2. *Consider a (non-random) vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, and let $\xi_1, \dots, \xi_n$ be i.i.d.* Bernoulli (p) *random variables, and let $p^* = \min(p, 1-p)$. Then*

$$\max_{a\in\mathbb{R}} \mathbb{P}(x_1\xi_1 + \dots + x_n\xi_n = a) = O\left(\frac{1}{\sqrt{|\operatorname{supp}(x)|p^*}}\right).$$

4. Singularity of sparse combinatorial matrices

Let us again fix $0 < \varepsilon < 1$. The proof of Theorem 1.2 proceeds in almost exactly the same way as the proof of Theorem 1.1, but there are three significant complications. First, since the entries are no longer independent, the calculations become somewhat more technical. Second, the rows and columns of Q have different distributions, so we need two versions of Lemma 3.1: one for vectors in the left kernel and one for vectors in the right kernel. Third, the fact that each row has exactly d ones means that we are not quite as free to do computations over $\mathbb{Z}_2$ (for example, if d is even and v is the all-ones vector, then we always have $Qv = 0$ over $\mathbb{Z}_2$). For certain parts of the argument, we will instead work over $\mathbb{Z}_{d-1}$.

Before we start the proof, the following lemma will allow us to restrict our attention to the case where $d \le n/2$, which will be convenient.

Lemma 4.1. *Let $Q \in \mathbb{R}^{n\times n}$ be a matrix whose every row has sum d, for some $d \notin \{0, n\}$. Let J be the $n \times n$ all-ones matrix. Then Q is singular if and only if $J - Q$ is singular.*

Proof. Note that the all-ones vector $\mathbf{1}$ is in the column space of Q (since the sum of all columns of Q equals $d\mathbf{1}$). Hence, every column of $J - Q$ is in the column space of Q. Therefore, if Q is singular, then $J - Q$ is singular as well. The opposite implication can be proved the same way. □

In the rest of the section, we prove Theorem 1.2 under the assumption that $(1+\varepsilon)\log n \le d \le n/2$ (note that if Q is a uniformly random zero-one matrix with every row having exactly d ones, then $J - Q$ is a uniformly random zero-one matrix with every row having exactly $n - d$ ones).

The first ingredient we will need is an analogue of Lemma 3.2 for 'combinatorial' random vectors. In addition to the notion of the support of a vector, we define a *fibre* of a vector to be a set of all indices whose entries are equal to a particular value.

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

Lemma 4.2. *Let $0 \le d \le n/2$, and consider a (non-random) vector $x \in \mathbb{R}^n$ whose largest fibre has size $n - s$, and let $\gamma \in \{0,1\}^n$ be a random zero-one vector with exactly d ones. Then*

$$\max_{a \in \mathbb{R}} \mathbb{P}(x \cdot \gamma = a) = O\Big(\sqrt{n/(sd)}\Big).$$

We deduce Lemma 4.2 from the $p = 1/2$ case of Lemma 3.2 (that is, from the Erdős–Littlewood–Offord theorem [7]).

Proof. The case $p = 1/2$ is treated in [17, Proposition 4.10]; this proof proceeds along similar lines. Let $p = d/n \le 1/2$. We realise the distribution of γ as follows. First choose $d = pn$ random disjoint pairs $(i_1, j_1), \dots, (i_{pn}, j_{pn}) \in \{1, \dots, n\}^2$ (each having distinct entries), and then determine the 1-entries in γ by randomly choosing one element from each pair.

We first claim that with probability $1 - e^{-\Omega(sp)}$, at least $\Omega(sp)$ of our pairs (i, j) have $x_i \neq x_j$ (we say such a pair is *good*). To see this, let I be a union of fibres of x, chosen such that $|I| \ge n/3$ and $n - |I| \ge s/3$ (if $s \le 2n/3$, we can simply take I to be the largest fibre of x, and otherwise we can greedily add fibres to I until $|I| \ge n/3$). To prove our claim, we will prove that in fact with the desired probability there are $\Omega(sp)$ different ℓ for which $i_\ell \notin I$ and $j_\ell \in I$.

Let $f = \lceil pn/6 \rceil$ and let S be the set of $\ell \le f$ for which $i_\ell \notin I$. So, $|S|$ has a hypergeometric distribution with mean $(n - |I|)f/n = \Omega(sp)$, and by a Chernoff bound (see, for example, [14, Theorem 2.10]), we have $|S| = \Omega(sp)$ with probability $1 - e^{-\Omega(sp)}$. Condition on such an outcome of $i_1, \dots, i_f$. Next, let T be the set of $\ell \in S$ for which $j_\ell \in I$. Then, conditionally, $|T|$ has a hypergeometric distribution with mean at least $(|I| - f)|S|/n = \Omega(sp)$, so again using a Chernoff bound we have $|T| = \Omega(sp)$ with probability $1 - e^{-\Omega(sp)}$, as claimed.

Now, condition on an outcome of our random pairs such that at least $\Omega(sp)$ of them are good. Let ξ_ℓ be the indicator random variable for the event that i_ℓ is chosen from the pair (i_ℓ, j_ℓ), so $\xi_1, \dots, \xi_{pn}$ are i.i.d. Bernoulli(1/2) random variables, and $x \cdot \gamma = a$ if and only if

$$(x_{i_1} - x_{j_1})\xi_1 + \cdots + (x_{i_{pn}} - x_{j_{pn}})\xi_1 = a - x_{j_1} - \cdots - x_{j_{pn}}.$$

Under our conditioning, $\Omega(sp)$ of the $x_{i_\ell} - x_{j_\ell}$ are nonzero, so by Lemma 3.2 with $p = 1/2$, conditionally we have $\mathbb{P}(x \cdot \gamma = a) \le O(1/\sqrt{sp})$. We deduce that unconditionally

$$\mathbb{P}(x \cdot \gamma = 0) \le e^{-\Omega(sp)} + O(1/\sqrt{sp}) = O(1/\sqrt{sp}) = O(\sqrt{n/(sd)}),$$

as desired. □

The proof of Theorem 1.2 then reduces to the following two lemmas. Indeed, for a constant $c > 0$ (depending on ε) satisfying the statements in Lemma 4.3 and 4.4, we can take $t = cn/\log d$, and

$$\mathcal{P} = \{x \in \mathbb{Q}^n : x \text{ has largest fibre of size at most } (1 - c/\log d)n\}.$$

We can then apply Lemma 2.1. By Lemma 4.3, the term (1) is bounded by $o(1)$, by Lemma 4.4 the term (2) is bounded by $(n/t) \cdot n^{-\Omega(1)} = (\log d/c) \cdot n^{-\Omega(1)} = o(1)$, and by Lemma 4.2 the term (3) is bounded by $(n/t) \cdot O\Big(\sqrt{n \log d/(cnd)}\Big) = O(\log^{3/2} d/\sqrt{d}) = o(1)$.

Lemma 4.3. *Let Q be a random combinatorial matrix (with d ones in each row), with $(1 + \varepsilon) \log n \le d \le n/2$. There is $c > 0$ (depending only on ε) such that with probability $1 - o(1)$, there is no nonzero vector $x \in \mathbb{Q}^n$ with $|\operatorname{supp}(x)| < cn/\log d$ and $x^T Q = 0$.*

Lemma 4.4. *Let $R_1, \dots, R_{n-1}$ be the first $n - 1$ rows of a random combinatorial matrix (with d ones in each row), with $(1 + \varepsilon) \log n \le d \le n/2$. There is $c > 0$ (depending only on ε) such that with probability $1 - n^{-\Omega(1)}$, every nonzero $x \in \mathbb{Q}^n$ satisfying $R_i \cdot x = 0$ for all $i = 1, \dots, n - 1$ has largest fibre of size at most $(1 - c/\log d)n$.*

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

Proof of Lemma 4.3. As in Lemma 3.1, it suffices to work over $\mathbb{Z}_2$. Let $C_1, \dots, C_n$ be the columns of Q, consider any $v \in \mathbb{Z}_2^n$ with $|\operatorname{supp}(v)| = s$, and let $\mathcal{E}_v$ be the event that $C_i \cdot v \equiv 0 \pmod 2$ for $i = 1, \dots, n$. Note that $\mathcal{E}_v$ only depends on the submatrix Q_v of Q containing only those rows j with $v_j = 1$ (and $\mathcal{E}_v$ is precisely the event that every column of Q_v has an even sum).

Let $p = d/n \le 1/2$, let M_v be a random $s \times n$ matrix with i.i.d. Bernoulli (p) entries and let $\mathcal{E}'_v$ be the event that every column in M_v has an even sum. Note that M_v is very similar to Q_v, so the probability of $\mathcal{E}_v$ is very similar to the probability of $\mathcal{E}'_v$. Indeed, writing $R_1, \dots, R_s$ and $R'_1, \dots, R'_s$ for the rows of Q_v and M_v, respectively, and writing $s_j = |\operatorname{supp}(R'_j)|$, for each j we have $s_j \sim \text{Binomial}(n, p)$, so an elementary computation using Stirling's formula shows that $\mathbb{P}(s_j = d) = \Omega(1/\sqrt{d}) = e^{-O(\log d)}$. Hence

$$\mathbb{P}(\mathcal{E}_v) = \mathbb{P}(\mathcal{E}'_v \mid s_j = d \text{ for all } j) \le \mathbb{P}(\mathcal{E}'_v)/\mathbb{P}(s_j = d \text{ for all } j) = e^{O(s \log d)}\mathbb{P}(\mathcal{E}'_v) = e^{O(s \log (pn))}\mathbb{P}(\mathcal{E}'_v).$$

Recalling the quantity $P_{s,p}$ from the proof of Lemma 3.1, we have

$$\mathbb{P}(\mathcal{E}'_v) = P_{s,p}^n = \begin{cases} e^{-(1+o(1))spn} & \text{if } sp = o(1), \\ e^{-\Omega(n)} & \text{if } sp = \Omega(1), \end{cases}$$

so if $s \le cn/\log d = cn/\log(pn)$ for small $c > 0$, then we also have

$$\mathbb{P}(\mathcal{E}_v) \le \begin{cases} e^{-(1+o(1))spn} & \text{if } sp = o(1), \\ e^{-\Omega(n)} & \text{if } sp = \Omega(1). \end{cases}$$

Let $P_s = \mathbb{P}(\mathcal{E}_v)$ (which only depends on s). We can now conclude the proof in exactly the same way as in Lemma 3.1. Taking $r = \delta/p$ for sufficiently small δ (relative to ε), the probability that there exists nonzero $v \in \mathbb{Z}_2^n$ with $|\operatorname{supp}(v)| < cn/\log d$ and $C_i \cdot v \equiv 0 \pmod 2$ for all $i = 1, \dots, n$ is at most

$$\begin{aligned} \sum_{s=1}^{cn/\log d} \binom{n}{s} P_s &\le \sum_{s=1}^{r} e^{s \log n - (1-\varepsilon/3)snp} + \sum_{s=r+1}^{cn/\log d} e^{s(\log(n/s)+1) - \Omega(n)} \\ &\le \sum_{s=1}^{\infty} n^{-s\varepsilon/3} + \sum_{s=1}^{cn/\log d} e^{n((s/n)(\log(n/s)+1) - \Omega(1))} = o(1), \end{aligned}$$

provided c is sufficiently small (relative to δ). □

We will deduce Lemma 4.4 from the following lemma.

Lemma 4.5. *Suppose $p \le 1/2$ and $pn \to \infty$, and let $\gamma \in \{0,1\}^n$ be a random vector with exactly pn ones. Let $q \ge 2$ be an integer and consider a (non-random) vector $v \in \mathbb{Z}_q^n$ whose largest fibre has size $n - s$. Then for any $a \in \mathbb{Z}_q$ we have $\mathbb{P}(v \cdot \gamma \equiv a \pmod q) \le P_{p,n,s}$ for some $P_{p,n,s}$ (only depending on p, n and s) satisfying*

$$P_{p,n,s} = \begin{cases} e^{-\Omega(1)} & \text{when } sp = \Omega(1), \\ e^{-(1-o(1))sp} & \text{when } sp = o(1) \end{cases}$$

Proof. As in the proof of Lemma 4.2, we realise the distribution of γ by first choosing pn random disjoint pairs $(i_1, j_1), \dots, (i_{pn}, j_{pn}) \in \{1, \dots, n\}^2$, and then randomly choosing one element from each pair to comprise the 1-entries of γ.

Let $\mathcal{E}$ be the event that $v_i \ne v_j$ for at least one of our random pairs (i, j). Then $\mathbb{P}(v \cdot \gamma \equiv a \pmod q \mid \mathcal{E}) \le 1/2$, and therefore $\mathbb{P}(v \cdot \gamma \equiv a \pmod q) \le 1 - \mathbb{P}(\mathcal{E})/2$. So, it actually suffices to prove that

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

$$\mathbb{P}(\mathcal{E}) \geq \begin{cases} \Omega(1) & \text{when } sp = \Omega(1), \\ (2 - o(1))sp & \text{when } sp = o(1). \end{cases}$$

If $s \geq n/3$ (this can only occur if $sp = \Omega(1)$), then we can choose $J \subseteq \{1, \dots, n\}$ to be a union of fibres of the vector $v \in \mathbb{Z}_q^n$ such that $n/3 \leq |J| \leq 2n/3$. In this case,

$$\mathbb{P}(\mathcal{E}) \geq \mathbb{P}(i_1 \in J,\ j_1 \notin J) = \Omega(1),$$

as desired. So, we assume $s < n/3$, and let $I \subseteq \{1, \dots, n\}$ be the set of indices in the largest fibre of v (so $|I| = n - s$). Note that $\mathcal{E}$ occurs whenever there is a pair $\{i_k, j_k\}$ with exactly one element in I.

Let $\mathcal{F}$ be the event that $i_k \in I$ for all $k = 1, \dots, pn$. We have

$$\mathbb{P}(\mathcal{E} \mid \mathcal{F}) \geq 1 - (1 - s/n)^{pn} = \begin{cases} \Omega(1) & \text{when } sp = \Omega(1), \\ (1 - o(1))sp & \text{when } sp = o(1), \end{cases}$$

and

$$\mathbb{P}(\mathcal{E} \mid \overline{\mathcal{F}}) \geq (n - s - pn)/(n - pn) = \begin{cases} \Omega(1) & \text{when } sp = \Omega(1), \\ 1 - o(1) & \text{when } sp = o(1). \end{cases}$$

This already implies that if $sp = \Omega(1)$, then $\mathbb{P}(\mathcal{E}) = \Omega(1)$ as desired. If $sp = o(1)$, then $\mathbb{P}(\mathcal{F}) \leq (1 - s/n)^{pn} = 1 - (1 + o(1))sp$, so

$$\mathbb{P}(\mathcal{E}) = \mathbb{P}(\mathcal{F})\mathbb{P}(\mathcal{E} \mid \mathcal{F}) + \mathbb{P}(\overline{\mathcal{F}})\mathbb{P}(\mathcal{E} \mid \overline{\mathcal{F}}) \geq (2 - o(1))sp,$$

as desired. □

Proof of Lemma 4.4. Let $q = d - 1$. It suffices to prove that with probability $1 - o(1)$ there is no nonconstant 'bad' vector $v \in \mathbb{Z}_q^n$ whose largest fibre has size at least $(1 - c/\log q)n$ and which satisfies $R_i \cdot v \equiv 0 \pmod q$ for all $i = 1, \dots, n - 1$. (Note that by the choice of q, if $v \in \mathbb{Z}_q^n$ is constant and nonzero, then it is impossible to have $v \cdot R_1 = 0$).

Let $p = d/n$, consider any $v \in \mathbb{Z}_q^n$ whose largest fibre has size $n - s$, and consider any $i \in \{1, \dots, n - 1\}$. Then $R_i \cdot v$ is of the form in Lemma 4.5, so taking $r = \delta/p$ for sufficiently small δ (relative to ε), the probability that such a bad vector exists is at most

$$\sum_{s=1}^{c'n/\log q} \binom{n}{s} q^{s+1} P_{p,n,s}^{n-1} \leq \sum_{s=1}^{r} e^{s\log n + (s+1)2\sqrt{pn} - (1-\varepsilon/3)spn} + \sum_{s=r+1}^{c'n/\log q} e^{s(\log(n/s)+1) + cn + 2\sqrt{pn} - \Omega(n)}$$

$$\leq \sum_{s=1}^{\infty} n^{-s\varepsilon/3} + \sum_{s=1}^{c'n/\log q} e^{n\left((s/n)(\log(n/s)+1) - \Omega(1)\right)} = n^{-\Omega(1)},$$

provided $c' > 0$ is sufficiently small (relative to δ) and n is sufficiently large. □

Acknowledgements

We would like to thank Elad Aigner-Horev, Yury Person and the anonymous referee for helpful comments and suggestions.

References

[1] Addario-Berry, L. and Eslava, L. (2014) Hitting time theorems for random matrices. *Combin. Probab. Comput.* **23**(5) 635–669.

[2] Aigner-Horev, E. and Person, Y. (2020) On sparse random combinatorial matrices. arXiv preprint arXiv:2010.07648.

https://doi.org/10.1017/S0963548321000146 Published online by Cambridge University Press

[3] Basak, A. and Rudelson, M. (2018) Sharp transition of the invertibility of the adjacency matrices of sparse random graphs. arXiv preprint arXiv:1809.08454.
[4] Campos, M., Mattos, L., Morris, R. and Morrison, N. On the singularity of random symmetric matrices. *Duke Math. J.* **170**(5) 881–907.
[5] Costello, K. P. and Vu, V. H. (2008) The rank of random graphs. *Random Struct. Alg.* **33**(3) 269–285.
[6] Costello, K. P. and Vu, V. (2010) On the rank of random sparse matrices. *Combin. Probab. Comput.* **19**(3) 321–342.
[7] Erdős, P. (1945) On a lemma of Littlewood and Offord. *Bull. Amer. Math. Soc.* **51** 898–902.
[8] Ferber, A. (2020) Singularity of random symmetric matrices – simple proof. arXiv preprint arXiv:2006.07439.
[9] Ferber, A. and Jain, V. (2019) Singularity of random symmetric matrices—a combinatorial approach to improved bounds. *Forum Math. Sigma* **7** e22.
[10] Ferber, A., Jain, V., Luh, K. and Samotij, W. (to appear) On the counting problem in inverse Littlewood–Offord theory. *J. Lond. Math. Soc.*.
[11] Huang, J. (2018) Invertibility of adjacency matrices for random d-regular graphs. arXiv preprint arXiv:1807.06465.
[12] Jain, V. (to appear) Approximate Spielman-Teng theorems for the least singular value of random combinatorial matrices *Israel J. Math.*
[13] Jain, V., Sah, A. and Sawhney, M. (2020) Singularity of discrete random matrices II. arXiv preprint arXiv:2010.06554.
[14] Janson, S., Łuczak, T. and Rucinski, A. (2000) *Random Graphs*, *Wiley-Interscience Series in Discrete Mathematics and Optimization*. Wiley-Interscience, New York.
[15] Komlós, J. (1967) *On the determinant of* $(0, 1)$ *matrices*. *Studia Sci. Math. Hungar.* **2** 7–21.
[16] Komlós, J. (1968) On the determinant of random matrices. *Studia Sci. Math. Hungar.* **3** 387–399.
[17] Litvak, A. E., Lytova, A., Tikhomirov, K., Tomczak-Jaegermann, N. & Youssef, P. (2017) Adjacency matrices of random digraphs: singularity and anti-concentration. *J. Math. Anal. Appl.* **445**(2) 1447–1491.
[18] Mészáros, A. (2020) The distribution of sandpile groups of random regular graphs. *Trans. Amer. Math. Soc.* **373**(9) 6529–6594.
[19] Nguyen, H. H. (2013) On the singularity of random combinatorial matrices. *SIAM J. Discrete Math.* **27**(1) 447–458.
[20] Nguyen, H. H. and Wood, M. M. (2018) Cokernels of adjacency matrices of random r-regular graphs. arXiv preprint arXiv:1806.10068.
[21] Nguyen, H. H. and Wood, M. M. (2018) Random integral matrices: universality of surjectivity and the cokernel. arXiv preprint arXiv:1806.00596.
[22] Rudelson, M. and Vershynin, R. (2008) The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.* **218**(2) 600–633.
[23] Tran, T. (2020) The smallest singular value of random combinatorial matrices. arXiv preprint arXiv:2007.06318.

Cite this article: Ferber A, Kwan M and Sauermann L (2022). Singularity of sparse random matrices: simple proofs. *Combinatorics, Probability and Computing* **31**, 21–28. https://doi.org/10.1017/S0963548321000146