On the permanent of random Bernoulli matrices

Terence Tao [a,1], Van Vu [b,*,2]

[a] *Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, United States*
[b] *Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019, United States*

Abstract

We show that the permanent of an $n \times n$ matrix with iid Bernoulli entries ± 1 is of magnitude $n^{(\frac{1}{2}+o(1))n}$ with probability $1 - o(1)$. In particular, it is almost surely non-zero.
© 2008 Elsevier Inc. All rights reserved.

MSC: 05D40; 15A15; 15A52

Keywords: Random matrices; Permanent

1. Introduction

Let M be an $n \times n$ matrix. Two basic parameters of M are its *determinant*

$$\mathrm{Det}(M) := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)},$$

and its *permanent*

$$\mathrm{Per}(M) := \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}.$$

* Corresponding author.
  E-mail addresses: tao@math.ucla.edu (T. Tao), vanvu@math.rutgers.edu (V. Vu).

Let M_n denote the random Bernoulli matrix of size n (the entries of M_n are iid random variables taking values ± 1 with probability $1/2$ each). For some time, it has been a central problem in probabilistic combinatorics to determine the asymptotic behavior of $\mathrm{Det}(M_n)$ and $\mathrm{Per}(M_n)$, as n tends to infinity (here and later we use the asymptotic notation under this assumption).

In the 1960s, Komlós [6,7] proved that asymptotically almost surely (i.e. with probability $1 - o(1)$), $\mathrm{Det}\, M_n \neq 0$. Since then, the problem of estimating the singular probability $\mathbf{P}(\mathrm{Det}\, M_n = 0)$ was studied in many papers [2,5,8,12]. It is easy to see that $\mathbf{P}(\mathrm{Det}(M_n) = 0) \geqslant (1/2 + o(1))^n$ and it has been conjectured that this lower bound is sharp. The most current upper bound is $\mathbf{P}(\mathrm{Det}(M_n) = 0) \leqslant (\frac{1}{\sqrt{2}} + o(1))^n$ [2].

The order of magnitude of $\mathrm{Det}\, M_n$ was computed recently. In [11], the authors showed that

$$\textit{Asymptotically almost surely, } \left|\mathrm{Det}(M_n)\right| = n^{(1/2 - o(1))n}. \tag{1}$$

On the other hand, little has been known about the permanent. Prior to this work, it was not known whether $\mathrm{Per}(M_n) \neq 0$ almost surely. It was observed by Alon (see also [13]) that if $n + 1$ is a power of 2, then any $n \times n$ ± 1 matrix has permanent equal $(n + 1)/2$ modulo $n + 1$ and thus is non-zero.

Similar to the situation with Det, the second moment of Per is easy to compute, using the definition of permanent and linearity of expectation

$$\mathbf{E}\left|\mathrm{Per}(M_n)\right|^2 = n!. \tag{2}$$

Few higher moments of $\mathrm{Per}(M_n)$ can also be computed (with some difficulty), but they do not appear to reveal much useful information.

The main goal of this paper is to establish an analogue of (1) for $\mathrm{Per}(M_n)$.

Theorem 1.1. *Asymptotically almost surely,*

$$\left|\mathrm{Per}(M_n)\right| = n^{(\frac{1}{2} + o(1))n}.$$

The upper bound follows from (2), Chebyshev's inequality and the fact that $n! = n^{(1 + o(1))n}$. The main task is to prove the lower bound and we are going to show

Theorem 1.2. *There is a positive constant c such that for every $\varepsilon > 0$ and n sufficiently large depending on ε, we have*

$$\mathbf{P}\left(\left|\mathrm{Per}(M_n)\right| \geqslant n^{(\frac{1}{2} - \varepsilon)n}\right) \geqslant 1 - n^{-c}.$$

Remark 1.3. The constant $c > 0$ in Theorem 1.2 can be made explicit (e.g. one can take $c = 1/10$) but we have not attempted to optimise it here. In any case, our method does not seem to lead to any value of c larger than $1/2$, due to its reliance on the Erdős–Littlewood–Offord inequality (Lemma 2.4) at the very last step (to get from $(n - 1) \times (n - 1)$-minors to the $n \times n$ matrix). In principle, one can obtain better results by using more advanced Littlewood–Offord inequalities, but it is not clear to the authors how to restructure the rest of the argument so that such inequalities can be exploited.

Remark 1.4. The lower bound $n^{(\frac{1}{2}-\varepsilon)n}$ can probably be sharpened to $n^{n/2}\exp(-\Theta(n))$, but we do not pursue this direction here.

Remark 1.5. Our proof also works (verbatim) for Det and thus we obtains a new proof for (1). The lower bound obtained for the determinant is however inferior to that in [11].

Remark 1.6. The Bernoulli distribution does not play a significant role. The theorem holds for virtually any (not too degenerate) discrete distribution. Also, it is not necessary to assume that the entries have identical distribution. The independence is, however, critical. In particular, our arguments do not seem to easily yield any non-trivial result for the permanent of a random *symmetric* Bernoulli matrix.

All previous proofs concerning Det(M_n) proceeded by geometric arguments (for instance, interpreting Det(M_n) = 0 as the event that the rows of M_n lie in a hyperplane). Such geometric arguments are unavailable for the permanent and thus one needs to find a new approach. In this paper, we proceed by a combinatorial method, studying the propagation of probabilistic lower bounds for the permanent from small minors to large ones. Roughly speaking, we are going to expose the rows of the matrix one at the time and try to show that, with high probability, the magnitude of the permanent of many (full-size) minors increases by a large factor (close to $\sqrt{n}$) at every step. This can be done in most of the process except the last few steps, where we simply keep the permanents from dropping.

In the next section, we present our probabilistic tools. The proof is outlined in Section 3, modulo many propositions. The rest of the paper is devoted to the verification of these propositions. As already mentioned, we are going to use the standard asymptotic notation (O, o, Ω, Θ) under the assumption that $n \to \infty$.

Remark 1.7. Random matrices in which the entries are bounded away from zero were studied in [9,10]. In this situation there is much less cancellation and a stronger result is known, namely a central limit theorem for the permanent. For random 0–1 matrices, the problem is closely related to that of counting perfect matchings in a random graph [4]. We also mention that some general results for the permanent rank of a matrix A (i.e. the size of the largest minor of A with non-vanishing permanent) were established in [14].

2. Probabilistic tools

We shall rely frequently on three standard tools from probability theory. The first one asserts that if there are a collection of events that are individually likely to be true, then it is likely that most of them are true at once, even if there are strong correlations between such events:

Lemma 2.1 (*First moment*). *Let $E_1, \ldots, E_m$ be arbitrary events (not necessarily independent) such that $\mathbf{P}(E_i) \geqslant 1 - \delta$ for all $1 \leqslant i \leqslant m$ and some $\delta > 0$, and let $0 < c < 1$. Then*

$$\mathbf{P}(\text{At most } cm \text{ of the } E_1, \ldots, E_m \text{ are false}) \geqslant 1 - \frac{\delta}{c}.$$

Proof. Let $I(E)$ be the indicator of an event E. From Markov's inequality we have

$$\mathbf{P}\left(\sum_{i=1}^{m} I(\overline{E_i}) \geqslant cm\right) \leqslant \frac{1}{cm}\mathbf{E}\sum_{i=1}^{m} I(\overline{E_i}),$$

and the claim follows from linearity of expectation. □

Our next tool is the following concentration result, a well-known consequence of Azuma's inequality [1].

Lemma 2.2. *Let $T > 0$, let $\xi_1, \ldots, \xi_n$ be iid Bernoulli variables, and let $Y = Y(\xi_1, \ldots, \xi_n)$ be a function such that $|Y(x) - Y(x')| \leqslant T$ for all pairs $x = (\xi_1, \ldots, \xi_n)$, $x' = (\xi_1', \ldots, \xi_n')$ of Hamming distance one. Then*

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geqslant S\big) \leqslant 2\exp\left(-\frac{S^2}{2nT^2}\right).$$

We also need the following (also standard) one-sided version of Azuma's inequality, which can be proved in the same way as Azuma's inequality itself.

Lemma 2.3. *Let $\mathbf{F}_0 \subset \mathbf{F}_2 \subset \cdots \subset \mathbf{F}_m$ be a sequence of nested σ-algebras in a probability space Ω and W_i, $0 \leqslant i \leqslant m$, be $\mathbf{F}_i$-measurable real functions obeying the submartingale-type property*

$$\mathbf{E}(W_i|\mathbf{F}_{i-1}) \leqslant W_{i-1}$$

for all $1 \leqslant i \leqslant m$. Assume also that $|W_i - W_{i-1}| \leqslant 1$ for all $1 \leqslant i \leqslant m$. Then for any $\lambda \geqslant 0$ we have

$$\mathbf{P}(W_m - W_0 \geqslant \lambda) \leqslant \exp\left(-\frac{\lambda^2}{2m}\right).$$

Finally, we need the classical Littlewood–Offord–Erdős inequality [3].

Lemma 2.4. *Let $\lambda > 0$ and $m, k \geqslant 1$, and let $v_1, \ldots, v_m$ be real numbers such that $|v_i| \geqslant \lambda$ for at least k values of i. Let $a_1, \ldots, a_m$ be iid signs drawn uniformly from $\{-1, +1\}$. Then we have*

$$\mathbf{P}\big(|a_1 v_1 + \cdots + a_m v_m| \leqslant x\lambda\big) = O\left(\frac{x}{\sqrt{k}}\right)$$

for all $x \geqslant 1$.

3. Preliminary reductions

Fix a small $\varepsilon_0 > 0$. Our goal is to show that

$$\mathbf{P}\big(|\mathrm{Per}(M_n)| \geqslant n^{(\frac{1}{2}-\varepsilon_0)n}\big) \geqslant 1 - O\big(n^{-\Omega(1)}\big), \tag{3}$$

as $n \to \infty$.

We shall do this by first establishing lower bounds on many minors of $\text{Per}(M_n)$, starting with 1×1 minors and increasing the size of the minors one at a time, until reaching the full $n \times n$ matrix M_n. The main point will be to ensure that lower bounds on $k \times k$ minors are passed on to many "children" $(k+1) \times (k+1)$ minors, and that the lower bounds improve by almost $n^{1/2}$ for the majority of k.

When we talk about a $k \times k$ minor (of M_n), we always understand that it is formed by some k columns and the *first k rows*. Thus, such a minor can be indexed by its k columns, which can be identified with an element of $\binom{[n]}{k} := \{A \subset [n] := \{1, \ldots, n\}: |A| = k\}$. We use M_A to denote the minor of M_n associated to such an element $A \in \binom{[n]}{k}$. We also use M_k to denote the $k \times n$ matrix formed by the first k rows of M_n, thus M_A is completely determined by M_k.

Let $1 \leqslant k \leqslant n$. For any $A \in \binom{[n]}{k}$ and $\lambda > 0$, we say that A is λ-*heavy* if $|\text{Per}(M_A)| \geqslant \lambda$. For any $N > 0$, let $E_{k,N,\lambda}$ denote the event that at least N elements of $\binom{[n]}{k}$ are λ-heavy. For instance, it is clear that

$$\mathbf{P}(E_{1,n,1}) = 1. \tag{4}$$

Our objective is to show that

$$\mathbf{P}\big(E_{n,1,n^{(\frac{1}{2}-\varepsilon)n}}\big) \geqslant 1 - O\big(n^{-\Omega(1)}\big). \tag{5}$$

Our strategy will be to move from the $k = 1$ bound (4) to the $k = n$ bound (5) by "growing" N and λ for many values of k.

For small values of k (e.g. $k \leqslant \varepsilon n$, for some small ε to be chosen later) we will just use a crude bound that does not grow N or λ, but has an exponentially high probability of success:

Proposition 3.1 (*Maintaining a single large minor*). *Let $1 \leqslant k < n$ and $\lambda > 0$. Then we have*

$$\mathbf{P}(E_{k+1,1,\lambda}|E_{k,1,\lambda}) \geqslant 1 - 2^{-(n-k)}.$$

This result is quite easy and is established in Section 4.

Proposition 3.1 does not grow N or λ. To handle the intermediate values of k (e.g. between εn and $(1-\varepsilon)n$) we will need more sophisticated estimates. We first need a variant of Proposition 3.1 in which the number N of minors can be large.

Proposition 3.2 (*Maintaining many large minors*). *Let $1 \leqslant k \leqslant (1 - \varepsilon)n$ for some $\varepsilon > 0$, let $N \geqslant 1$ and let $\lambda > 0$. Then we have*

$$\mathbf{P}(E_{k+1,\varepsilon N/6,\lambda}|E_{k,N,\lambda}) \geqslant 1 - \exp\big(-\Omega(\varepsilon n)\big).$$

We prove Proposition 3.2 in Section 4. This proposition has a very small failure rate, but does not improve either N or λ. To achieve such growth, we need a further proposition, which has much higher failure rate but has a good chance of increasing either N or λ significantly.

Proposition 3.3 *(Growing many large minors). Let $1 \leqslant k \leqslant (1 - \varepsilon)n$ for some $\varepsilon > 0$, let $N \geqslant 1$, let $1 > c > 0$, and let $\lambda > 0$. Then we can partition the event $E_{k,N,\lambda}$ as $E'_{k,N,\lambda,c} \vee E''_{k,N,\lambda,c}$, where the events $E'_{k,N,\lambda,c}, E''_{k,N,\lambda,c}$ depend only on M_k, and where*

$$\mathbf{P}\big(E_{k+1,n^c N,\lambda}\big|E'_{k,N,\lambda,c}\big) \geqslant 1/3 \tag{6}$$

and

$$\mathbf{P}\big(E_{k+1,\varepsilon N/4,n^{1/2-c}\lambda}\big|E''_{k,N,\lambda,c}\big) \geqslant 1 - n^{-c/4}. \tag{7}$$

This proposition will be proven in Section 5. Finally, to handle the last few values of k ($(1 - \varepsilon)n \leqslant k \leqslant n$) we need the following result.

Proposition 3.4 *(Endgame). Let $1 \leqslant k \leqslant (1 - \varepsilon)n$ for some $\varepsilon > 0$, and let $\lambda > 0$. Then*

$$\mathbf{P}(E_{n,1,n^{-\log n}\lambda}|E_{k,1,\lambda}) \geqslant 1 - n^{-\Omega(1)}$$

if n is sufficiently large depending on ε.

This proposition will be proven in Section 6.

In the rest of this section, we show how Propositions 3.1–3.4 imply the desired bound (5).

Recall that $\varepsilon_0 > 0$ is fixed. We choose a number $\varepsilon > 0$ sufficiently small compared to ε_0, and a number ε' sufficiently small compared to ε. Let $k_1 := \lfloor (1 - \varepsilon)n \rfloor$. In view of Proposition 3.4, it suffices to show that

$$\mathbf{P}(E_{k_1,1,n^{(1/2-\varepsilon_0/2)}n}) \geqslant 1 - n^{-\Omega(1)}. \tag{8}$$

Applying Proposition 3.1 repeatedly, combined with (4), we obtain

$$\mathbf{P}(E_{k_0,1,1}) \geqslant 1 - \exp\big(-\Omega(n)\big) \tag{9}$$

for $k_0 := \lfloor \varepsilon n \rfloor + 1$. (One can also use here Alon's observation from the introduction, replacing k_0 with $2^m - 1$ for some suitable m. However, this observation is specific to the permanent (as opposed to the determinant).)

To get from k_0 to k_1, we construct random variables N_k, λ_k and W_k for $k_0 \leqslant k \leqslant k_1$ by the following algorithm.

- Step 0. Initialise $k := k_0$. If $E_{k_0,1,1}$ holds, then set $N_{k_0} := 1$, $\lambda_{k_0} := 1$, $W_{k_0} = 0$. Otherwise, set $N_{k_0} := 0$, $\lambda_{k_0} := 1$, $W_{k_0} := 0$.
- Step 1. If $N_k = 0$ then set $N_{k+1} := 0$, $\lambda_{k+1} := \lambda_k$, $W_{k+1} := W_k$. Move to Step 5. Otherwise, move on to Step 2.
- Step 2. If $k = k_1$ then terminate the algorithm. Otherwise, move on to Step 3.
- Step 3. By Proposition 3.3, we are either in event $E'_{k,N_k,\lambda_k,\varepsilon}$ or $E''_{k,N_k,\lambda_k,\varepsilon}$. Expose the row $k + 1$.
- Step 4. Define N_{k+1} and λ_{k+1} by the following rule:
    (I) If $E'_{k,N_k,\lambda,\varepsilon} \wedge E_{k+1,n^\varepsilon N_k/4,\lambda_k}$ holds then we say that k is *Type I*.
        Set $N_{k+1} := n^\varepsilon N_k/4$ and $\lambda_{k+1} := \lambda_k$.

(II) If $E'_{k,N_k,\lambda,\varepsilon} \wedge \overline{E_{k+1,n^\varepsilon N_k/4,\lambda_k}} \wedge E_{k+1,\varepsilon N_k/6,\lambda_k}$ holds then we say that k is *Type II*. Set $N_{k+1} := \varepsilon' N_k$ and $\lambda_{k+1} := \lambda_k$. (Here we use the fact that $\varepsilon' \leqslant \varepsilon/6$.)

(III) If $E''_{k,N_k,\lambda,\varepsilon} \wedge E_{k+1,\varepsilon' N_k,n^{1/2-\varepsilon}\lambda_k}$ holds then we say that k is *Type III*. Set $N_{k+1} := \varepsilon' N_k$, $\lambda_{k+1} := n^{1/2-\varepsilon}\lambda_k$.

(IV) If $E''_{k,N_k,\lambda,\varepsilon} \wedge \overline{E_{k+1,\varepsilon' N_k,n^{1/2-\varepsilon}\lambda_k}} \wedge E_{k+1,\varepsilon N_k/6,\lambda_k}$ holds then we say that k is *Type IV*. Set $N_{k+1} := \varepsilon' N_k$ and $\lambda_{k+1} := \lambda_k$. (Here we use the fact that $\varepsilon' \leqslant \varepsilon/6$.)

(V) If none of the above holds then set $N_{k+1} := 0$, $\lambda_{k+1} := \lambda_k$.

Set $W_{k+1} := W_k + (1 - \varepsilon/2) - 3\mathbf{I}_{k\ type\ I} - \mathbf{I}_{k\ type\ III}$.

- Step 5. Increment k to $k + 1$, and then return to Step 1.

We say that the algorithm is *successful* if at the terminating time ($k = k_1$), $N_{k_1} \neq 0$ and $W_{k_1} \leqslant \varepsilon' n/2$. We first show

Proposition 3.5. *The probability that the algorithm is successful is* $1 - \exp(-\Omega(\varepsilon'n))$.

Proof. From (9), we know that the probability of failure at $k = k_0$ is $\exp(-\Omega(n))$. From Proposition 3.2, we know that the probability that $E_{k+1,\varepsilon N_k/6,\lambda_k}$ fails given E_{k,N_k,λ_k}, for any given $k_0 < k \leqslant k_1$, is $\exp(-\Omega(\varepsilon n)) \leqslant \exp(-\Omega(\varepsilon'n))$. The union bound then implies that the probability that $N_{k_1} = 0$ is $\exp(-\Omega(\varepsilon'n))$.

From Proposition 3.3 and the definition of W_{k+1}, we obtain the submartingale-type property

$$\mathbf{E}(W_{k+1}|M_k) \leqslant W_k.$$

Also we have $|W_{k+1} - W_k| = O(1)$. By Lemma 2.3 (with the σ-algebra generated by M_k playing the role of $\mathbf{F}_k$), we obtain

$$\mathbf{P}(W_{k_1} \geqslant \varepsilon'n/2) \leqslant \exp(-\Omega(\varepsilon'n)).$$

The claim follows. □

Next, we prove the following (deterministic) proposition, which, together with the previous proposition, imply (8).

Proposition 3.6. *If the algorithm is successful, then* $E_{k_1,1,n^{(1/2-\varepsilon_0/2)n}}$ *holds.*

Proof. Assume that the algorithm is successful. We have $W_{k_1} \leqslant \varepsilon'n/2$, which implies (via the definition of W_{k+1}) that

$$\sum_{i=k_0}^{k_1-1} 3\mathbf{I}_{k\ type\ I} + \mathbf{I}_{k\ type\ III} \geqslant (k_1 - k_0) - \varepsilon'n.$$

On the other hand, the number of steps of type I is only $o(n)$. Indeed, each such step increases N_k by a huge factor $n^\varepsilon/4$ while any other step decreases N_k by at most a constant factor. These combined with the fact that $N_k \leqslant 2^n$ for any k yield the desired bound. Thus, the number of steps of type III is at least

$$(k_1 - k_0) - (\varepsilon' + o(1))n \geqslant (1 - 2\varepsilon - O(1) - (\varepsilon' + o(1)))n \geqslant (1 - 3\varepsilon)n$$

thanks to the definition of k_0, k_1 and the fact that ε is larger than ε'. Since each type III step increases λ_k by $n^{1/2-\varepsilon}$, it follows that

$$\lambda_{k_1} \geqslant n^{(1/2-\varepsilon)(1-3\varepsilon)} \geqslant n^{(1-\varepsilon_0/2)n}$$

as we set ε much smaller than ε_0. The proof is complete. $\quad\square$

Remark 3.7. The above consideration in fact gives an exponentially small probability bound for (8). Unfortunately, the argument used to prove Proposition 3.4 only yields a polynomial bound, especially in the last step of the argument (dealing with the bottom row of M_n). This is why the final bound in Theorem 1.2 is only polynomial in nature.

It remains to prove Propositions 3.1–3.4. This will be the focus of the remaining sections.

4. Child and parent minors

To prove Propositions 3.1–3.4, it is important to understand the relationship between the permanent of a "parent" minor M_A and the permanent of a "child" minor $M_{A'}$. More precisely, we say that $M_{A'}$ is a *child* of M_A (or M_A is a *parent* of $M_{A'}$) if we have $A' = A \cup \{i\}$ for some $i \notin A$ (or equivalently if $A = A' \setminus \{i\}$ for some $i \in A'$).

Let $A \in \binom{[n]}{k+1}$ for some $1 \leqslant k < n$. From the definition of permanent we have the cofactor expansion

$$\mathrm{Per}(M_A) = \sum_{i \in A} a_{k+1,i} \, \mathrm{Per}(M_{A\setminus\{i\}}). \tag{10}$$

We can draw an easy consequence of this:

Lemma 4.1 (*Large parent often has large child*). *Let $A \in \binom{[n]}{k}$ for some $1 \leqslant k < n$, and let $i \notin A$. Assume that the submatrix M_k is fixed and we expose the (random) row $k+1$. Then*

$$\mathbf{P}\big(\big|\mathrm{Per}(M_{A\cup\{i\}})\big| \geqslant \big|\mathrm{Per}(M_A)\big|\big) \geqslant \frac{1}{2}.$$

In fact, this bound is still true if we condition on all the entries of the row $k+1$ except for $a_{k+1,i}$.

Proof. Let $M'_{A\cup\{i\}}$ denote the same minor as $M_{A\cup\{i\}}$ but with the sign $a_{k+1,i} \in \{-1,+1\}$ replaced by $-a_{k+1,i}$. From (10), we have

$$\big|\mathrm{Per}(M_{A'\cup\{i\}}) - \mathrm{Per}(M_{A\cup\{i\}})\big| = 2\big|\mathrm{Per}(M_A)\big|.$$

The claim follows. $\quad\square$

We can amplify this probability $\frac{1}{2}$ to an exponentially small probability by exploiting the fact that one parent has many "independent" children.

Lemma 4.2 *(Large parent often has many large children). Let $A \in \binom{[n]}{k}$ for some $1 \leqslant k < n$, and let $I \subset [n] \setminus A$. Assume that the submatrix M_k is fixed and we expose the (random) row $k + 1$. Then*

$$\mathbf{P}\big(\big|\text{Per}(M_{A \cup \{i\}})\big| \geqslant \big|\text{Per}(M_A)\big| \text{ for some } i \in I\big) \geqslant 1 - 2^{-|I|} \tag{11}$$

and

$$\mathbf{P}\big(\big|\text{Per}(M_{A \cup \{i\}})\big| \geqslant \big|\text{Per}(M_A)\big| \text{ for at least } |I|/3 \text{ values of } i \in I\big) \geqslant 1 - O\big(\exp\big(-\Omega\big(|I|\big)\big)\big). \tag{12}$$

Proof. We further condition on all entries of the $k + 1$ row except for $a_{k+1,i}$ where $i \in I$. The first claim follows from the previous lemma and independence. The second follows from Chernoff's bound. (One can, of course, use Azuma's inequality as well.) □

We can now immediately prove Proposition 3.1:

Proof of Proposition 3.1. Let us condition on the first k rows M_k, and assume that $E_{k,1,\lambda}$ holds, thus there exists a λ-heavy $A \in \binom{[n]}{k}$. Applying (11) with $I := [n] \setminus A$ we conclude that

$$\mathbf{P}\left(A' \text{ is } \lambda\text{-heavy for some } A' \in \binom{[n]}{k+1}\right) \geqslant 1 - 2^{-(n-k)}$$

and the claim follows. □

A slightly more sophisticated argument also gives Proposition 3.2.

Proof of Proposition 3.2. We may take N to be an integer. Let us condition on the first k rows M_k, and assume that $E_{k,N,\lambda}$ holds, thus there exist N λ-heavy minors $A_1, \ldots, A_N \in \binom{[n]}{k}$. Each A_j has at least εn children $A_j \cup \{i\}$. Let us call A_j *good* if it has at least $\varepsilon n/3$ λ-heavy children $A_j \cup \{i\}$. By (12), each j has a probability $1 - \exp(-\Omega(\varepsilon n))$ of being good. Applying Lemma 2.1 with $c := 1/2$, we conclude that with probability $1 - \exp(-\Omega(\varepsilon n))$, at least $N/2$ of the j are good.

Let us now suppose that at least $N/2$ of the j are good. By definition, each good A_j has at least $\varepsilon n/3$ λ-heavy children $A_j \cup \{i\}$. On the other hand, each child has at most n parents. By the usual double counting argument, this implies that at least $\varepsilon N/6$ elements in $A' \in \binom{[n]}{k+1}$ are λ-heavy, and the claim follows. □

5. Growing large minors

The purpose of this section is to prove Proposition 3.3. Fix $k, \varepsilon, N, c, \lambda$; we may take N to be an integer. We condition on M_k of M_n and assume that $E_{k,N,\lambda}$ holds. Thus we may find distinct λ-heavy $A_1, \ldots, A_N \in \binom{[n]}{k}$.

For each $l \geqslant 1$, let F_l denote the number of $A' \in \binom{[n]}{k+1}$ which have exactly l parents in the set $\{A_1, \ldots, A_N\}$. Since each A_j has at least εn children $A_j \cup \{i\}$, a double counting argument shows

$$\sum_{l=1}^{n} l F_l \geqslant \varepsilon n N.$$

Now set $K := \lfloor \frac{\varepsilon}{8} n^{1-c} \rfloor$. Since

$$\sum_{l=1}^{n} l F_l \leqslant K(F_1 + \cdots + F_K) + n(F_{K+1} + \cdots + F_n)$$

we see that either

$$F_1 + \cdots + F_K \geqslant \frac{\varepsilon n N}{2K} \tag{13}$$

or

$$F_{K+1} + \cdots + F_n \geqslant \frac{\varepsilon N}{2}. \tag{14}$$

We let $E'_{k,N,\lambda,c}$ be the event that (13) (and $E_{k,N,\lambda}$, of course) holds, and $E''_{k,N,\lambda,c}$ be the event that (13) fails but (14) (and $E_{k,N,\lambda}$) holds.

Suppose first that $E'_{k,N,\lambda,c}$ holds. Then by (13), we can find at least $\frac{\varepsilon n N}{2K}$ elements A' in $\binom{[n]}{k+1}$, each of which has at least one parent in $\{A_1, \ldots, A_N\}$. By Lemma 4.1, each such A' is λ-heavy with probability at least $1/2$. Applying Lemma 2.1, we conclude that with probability at least $1/3$, at least $\frac{\varepsilon n N}{8K}$ of these A' will be λ-heavy. The claim (6) now follows from the choice of K.

Now suppose instead that $E''_{k,N,\lambda,c}$ holds. Then by (14), we can find at least $\frac{\varepsilon N}{2}$ elements A' in $\binom{[n]}{k+1}$, each one of which has at least K parents in $\{A_1, \ldots, A_N\}$. By (10) and Lemma 2.4, we see that each of these A' is $n^{1/2-c}\lambda$-heavy with probability $1 - O(n^{1/2-c}/K^{1/2})$. Applying Lemma 2.1, we see that with probability $1 - O(n^{1/2-c}/K^{1/2})$, at least $\varepsilon N/4$ of the A' will be $n^{1/2-c}\lambda$-heavy. The claim (7) now follows from the choice of K (and the assumption that n is large). This concludes the proof of Proposition 3.3.

6. The endgame

The purpose of this section is to prove Proposition 3.4. Fix k, n, λ. We condition on M_k and assume that $E_{k,1,\lambda}$ holds, thus one of the elements of $\binom{[n]}{k}$ is λ-heavy. By symmetry we may assume without loss of generality that $[k]$ is λ-heavy. Our task is to show that $[n]$ is $n^{-\log n}\lambda$-heavy with probability $1 - n^{-\Omega(1)}$.

Set $L := \frac{1}{100} \lfloor \log n \rfloor$. We first show that there are plenty of heavy minors in $\binom{[n]}{n-L}$.

Lemma 6.1 (*Many heavy minors of order $n - L$*). *Let $B \subset \binom{[n] \setminus [k]}{2L}$. Then with probability $1 - \exp(-\Omega(L))$, there exists a λ-heavy minor $A \in \binom{[n]}{n-L}$ which contains $[n] \setminus B$.*

Proof. We construct $A_j \in \binom{[n]}{j}$ for $k \leqslant j \leqslant n - L$ by the following algorithm.

- Step 0. Initialise $j := k$ and $A_j := [k]$.
- Step 1. If there exists $i \in [n] \setminus (B \cup A_j)$ such that $A_j \cup \{i\}$ is λ-heavy, then choose one of these i arbitrarily, set $A_{j+1} := A_j \cup \{i\}$, and go onto Step 4. Otherwise, go to Step 2.
- Step 2. If there exists $i \in B \setminus A_j$ such that $A_j \cup \{i\}$ is λ-heavy, then choose one of these i arbitrarily, $A_{j+1} := A_j \cup \{i\}$, and go onto Step 4. Otherwise, go to Step 3.

- Step 3. Choose $i \in [n] \setminus A_j$ arbitrarily, and set $A_{j+1} := A_j \cup \{i\}$.
- Step 4. If $j = n - L - 1$ then STOP. Otherwise increment j to $j + 1$ and return to Step 1.

Applying (11) we see that if A_j is λ-heavy for some $k \leqslant j < n - L$, then with probability at least $1 - 2^{-(n-j)}$ $A_j \cup \{i\}$ is λ-heavy for at least one $i \in [n] \setminus A_j$. By construction, this implies that A_{j+1} is λ-heavy with probability at least $1 - 2^{-(n-j)}$. By the union bound (and the fact that A_k is λ-heavy), we thus conclude that with probability $1 - O(2^{-L})$, A_j is λ-heavy for all $k \leqslant j \leqslant n - L$.

Let $W_j := |[n] \setminus (B \cup A_j)|$, thus $W_k = n - k - 2L$ and $\min(W_j - 1, 0) \leqslant W_{j+1} \leqslant W_j$ for all $k \leqslant j < n - L$. By (11), we see that if A_j is λ-heavy, and $W_j > 0$ then $W_{j+1} = W_j - 1$ with probability at least $1 - 2^{-W_j}$. By the union bound, we conclude that $W_{n-\lfloor 2.01L \rfloor} = \lfloor 2.01L \rfloor - 2L$ with probability $1 - \exp(-\Omega(L))$. We condition on this event.

For any $n - \lfloor 2.01L \rfloor \leqslant j < n - L$, we see from the previous discussion that if $W_j > 0$, then $W_{j+1} = W_j - 1$ with probability at least 0.4 (say), and $W_{j-1} = W_j$ otherwise. From this we see that

$$\mathbf{E}\big(2^{W_{j+1}} - 1 | W_{n-\lfloor 2.01L \rfloor} = \lfloor 2.01L \rfloor - 2L\big) \leqslant \frac{1}{\sqrt{2}} \mathbf{E}\big(2^{W_j} - 1 | W_{n-\lfloor 2.01L \rfloor} = \lfloor 2.01L \rfloor - 2L\big)$$

(say) for all $n - \lfloor 2.1L \rfloor \leqslant j < n - L$. Since

$$\mathbf{E}\big(2^{W_{n-\lfloor 2.01L \rfloor}} - 1 | W_{n-\lfloor 2.01L \rfloor} = \lfloor 2.01L \rfloor - 2L\big) \leqslant 2^{0.01L},$$

we conclude by iteration that

$$\mathbf{E}\big(2^{W_{n-L}} - 1 | W_{n-\lfloor 2.01L \rfloor} = \lfloor 2.01L \rfloor - 2L\big) \leqslant \exp\big(-\Omega(L)\big)$$

and thus $W_{n-L} = 0$ with probability $1 - \exp(-\Omega(L))$. Since A_{n-L} is also λ-heavy with probability $1 - \exp(-\Omega(L))$, the claim follows. $\quad\square$

For any integer $N \geqslant 1$, any $1 \leqslant j \leqslant L$, and any $\lambda' > 0$, let $F_{j,N,\lambda'}$ denote the event that there exists N λ'-heavy sets (minors) $A_1, \ldots, A_N \in \binom{[n]}{n-j}$ whose complements $[n] \setminus A_1, \ldots, [n] \setminus A_N$ are disjoint.

Corollary 6.2 *(Many complement-disjoint heavy minors of order $n - L$). We have* $\mathbf{P}(F_{L,\lfloor \varepsilon n/10L \rfloor,\lambda}) = 1 - \exp(-\Omega(L)) = 1 - n^{-\Omega(1)}$.

Proof. Choose $\lfloor \varepsilon n/4L \rfloor$ disjoint sets $B_1, \ldots, B_{\lfloor \varepsilon n/4L \rfloor} \in \binom{[n] \setminus [k]}{2L}$ arbitrarily. For each of these B_i, Lemma 6.1 shows that with probability $1 - \exp(-\Omega(L))$, there exists a heavy $A_i \in \binom{[n]}{n-L}$ with $[n] \setminus A_i \subset B_i$ (in particular, the sets $[n] \setminus A_i$ are disjoint). The claim now follows from Lemma 2.1. $\quad\square$

We now propagate the events $F_{j,N}$ downward from $j = L$ to $j = 1$ (accepting some loss in the weight threshold λ' and in the population N of heavy minors when doing so) by means of the following lemma.

Lemma 6.3 *(Many heavy minors of order $n - j$ imply many heavy minors of order $n - j + 1$).*
Let $1 < j \leqslant L$, $N \geqslant n^{0.5}$ (say), and $\lambda' > 0$. Then

$$\mathbf{P}(F_{j-1,\lfloor N/10 \rfloor, \lambda'/n} | F_{j,N,\lambda}) \geqslant 1 - n^{-\Omega(1)}. \tag{15}$$

Proof. Fix j, N. We condition on M_{n-j} so that $F_{j,N}$ hold. Thus we can find λ'-heavy sets $A_1, \ldots, A_N \in \binom{[n]}{n-j}$ with disjoint complements, which we now fix. For each A_i, we arbitrarily choose a child $B_i = A_i \cup \{h_i\} \in \binom{[n]}{n-j+1}$. By construction, the $B_1, \ldots, B_N$ also have disjoint complements and the h_i are different.

Let $T := \lfloor n^{0.1} \rfloor$. Call a child B_i *good* if it has at least T λ'/n-heavy parents (of which A_i will be one of them), and *bad* otherwise. There are two cases.

Case 1: at least half of the B_i are good. By (10) and Lemma 2.4, each B_i has a probability $1 - O(T^{-1/2})$ of being λ'/n-heavy. The claim now follows from Lemma 2.1.

Case 2: at least half of the B_i are bad. Let I be the set of all i such that B_i is bad and H be the set of h_i, $i \in I$. Draw a bipartite graph G between I and H by connection i to $h_{i'}$ if $B_i \setminus \{h_{i'}\}$ is λ'/n-heavy. As the B_i are bad, each $i \in I$ has degree at most T. By double counting the edges in this graph, we have

$$\sum_{h \in H} \deg_h \leqslant T|I| = T|H|$$

where $\deg_h$ denotes the degree of h.

Again by a double counting argument, one can easily shows that the set $I' := \{i \mid \deg_{h_i} \leqslant 2T\}$ is at least $|I|/2 \geqslant N/4$. We condition on the entries of the $n - j + 1$ row not in the columns determined by I'. For each $i \in I'$, let

$$Y_i := \min\left(\frac{|\operatorname{Per} M_{B_i}|}{\lambda'}, 1 \right)$$

and $Y := \sum_{i \in I'} Y_i$. By Lemma 4.1, $\mathbf{E}(Y_i) \geqslant 1/2$ since each B_i has a λ'-heavy parent. Thus, by linearity of expectation, $\mathbf{E}(Y) \geqslant |I'|/2 \geqslant |N/8|$.

Now we estimate the effect of each random entry $a_{n-j+1,h}$ on Y. If $h \notin B_i$, then flipping $a_{n-j+1,h}$ does not change Y_i. If $h \in B_i$ and the $(n - j) \times (n - j)$ minor corresponding to $a_{n-j+1,h}$ is not λ'/n-heavy, then flipping $a_{n-j+1,h}$ changes Y_i by at most $2/n$. Finally, if $h \in B_i$ and the $(n - j) \times (n - j)$ minor corresponding to $a_{n-j+1,h}$ is λ'/n-heavy, then flipping $a_{n-j+1,h}$ changes Y_i by at most 1. On the other hand, the number of such i is at most $2T$ by the definition of I'. Thus, flipping $a_{n-j+1,h}$ changes Y by at most $2T + 2 \leqslant 3T$.

By Lemma 2.2 and the definitions of N and T

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geqslant |I'|/100\big) \leqslant 2\exp\left(-\Omega\left(\frac{|I'|^2}{T^2|I'|} \right) \right) = \exp\left(-\Omega\left(\frac{N}{T^2} \right) \right) = n^{-\Omega(1)}.$$

Since $\mathbf{E}(Y) \geqslant |I'|/2 \geqslant N/8$, it follows that $Y \geqslant N/9$ with probability $1 - n^{-\Omega(1)}$. Finally, notice that if $Y \geqslant N/9$, then the definition of Y and Y_i implies (with room to spare) that for at least $N/10$ indices i, $\frac{|\operatorname{Per} M_{B_i}|}{\lambda'} \geqslant \frac{1}{n}$. This concludes the proof. $\square$

Iterating Lemma 6.3 $L \leqslant \frac{\log n}{100}$ times starting with Corollary 6.2, we conclude that

$$\mathbf{P}(F_{1,\lfloor n^{0.5}\rfloor, n^{-\log n}\lambda}) \geqslant 1 - n^{-\Omega(1)}.$$

Now suppose that $F_{1,\lfloor n^{0.5}\rfloor, n^{-\log n}\lambda}$ holds, thus there are at least $\lfloor n^{0.5}\rfloor$ $n^{-\log n}\lambda$-heavy minors in $\binom{[n]}{n-1}$. Applying (10) and Lemma 2.4 we conclude that $[n]$ is $n^{-\log n}\lambda$-heavy with probability at least $1 - O(1/\sqrt{\lfloor n^{0.5}\rfloor}) = 1 - n^{-\Omega(1)}$. This completes the proof of Theorem 1.2.

References

[1] N. Alon, J. Spencer, The Probabilistic Method, second ed., Wiley, 2000.
[2] J. Bourgain, V. Vu, P. Wood, On the singularity probability of random discrete matrices, preprint.
[3] P. Erdős, On a lemma of Littlewood and Offord, Bull. Amer. Math. Soc. 51 (1945) 898–902.
[4] S. Janson, The number of spanning trees, Hamilton cycles, and perfect matchings in a random graph, Combin. Probab. Comput. 3 (1) (1994) 97–126.
[5] J. Kahn, J. Komlós, E. Szemerédi, On the probability that a random ±1 matrix is singular, J. Amer. Math. Soc. 8 (1995) 223–240.
[6] J. Komlós, On the determinant of (0, 1) matrices, Studia Sci. Math. Hungar. 2 (1967) 7–22.
[7] J. Komlós, On the determinant of random matrices, Studia Sci. Math. Hungar. 3 (1968) 387–399.
[8] J. Komlós, Circulated note, reproduced as Theorem 14.11 in: B. Bollobás (Ed.), Random Graphs, second ed., Cambridge Univ. Press, 2001.
[9] G. Rempala, A. Gupta, Some extensions of Girko's limit theorems for permanents of random matrices, Random Oper. Stochastic Equations 8 (4) (2000) 305–318.
[10] G. Rempala, J. Wesołowski, Asymptotic behavior of random permanents, Statist. Probab. Lett. 45 (1999) 149–158.
[11] T. Tao, V. Vu, On random ±1 matrices: Singularity and Determinant, Random Structures Algorithms 28 (1) (2006) 1–23.
[12] T. Tao, V. Vu, On the singularity probability of random Bernoulli matrices, J. Amer. Math. Soc. 20 (3) (2007) 603–628.
[13] I.M. Wanless, Permanents of matrices of signed ones, Linear Multilinear Algebra 53 (2005) 427–433.
[14] Y. Yu, The permanent rank of a matrix, J. Combin. Theory Ser. A 85 (1999) 237–242.