EXPONENTIAL ANTICONCENTRATION OF THE PERMANENT

ZACH HUNTER, MATTHEW KWAN, AND LISA SAUERMANN

Abstract. Let $A \in \mathbb{R}^{n\times n}$ be a random matrix with independent entries, and suppose that the entries are "uniformly anticoncentrated" in the sense that there is a constant $\varepsilon > 0$ such that each entry a_{ij} satisfies $\sup_z \mathbb{P}[a_{ij} = z] \le 1 - \varepsilon$ (for example, A could be a uniformly random $n \times n$ matrix with ± 1 entries). Significantly improving previous bounds of Tao and Vu, we prove that the permanent of A is exponentially anticoncentrated: there is $c_\varepsilon > 0$ such that $\sup_z \mathbb{P}[\mathrm{per}(A) = z] \le \exp(-c_\varepsilon n)$. Our proof also works for the determinant, giving an alternative proof of a classical theorem of Kahn, Komlós and Szemerédi. As a consequence, we see that there are at least exponentially many different permanents of $n \times n$ matrices with ± 1 entries, resolving a problem of Ingram and Razborov.

1. Introduction

Two important matrix parameters are the *determinant* and *permanent*: for an $n\times n$ matrix $A = (a_{ij})$, they are defined as

$$\det(A) = \sum_{\pi\in S_n} \mathrm{sign}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} \qquad \text{and} \qquad \mathrm{per}(A) = \sum_{\pi\in S_n} \prod_{i=1}^{n} a_{i\pi(i)}. \tag{1.1}$$

Determinants are omnipresent throughout mathematics and the sciences. Permanents are of fundamental importance in combinatorics and computational complexity theory, and also play an important role in quantum physics and statistics (see for example the books and surveys [2, 16, 17, 20, 27]).

In combinatorics, computer science and random matrix theory, it is of great interest to understand statistical properties of determinants and permanents, among various classes of discrete matrices. In particular, one of the most intensive directions of research is the study of the singularity probability of random sign matrices. Let A_n be a uniformly random $n\times n$ matrix with ± 1 entries; what is the probability that $\det(A_n) = 0$? Perhaps the foundational theorem in combinatorial random matrix theory, due to Komlós [12], is that $\mathbb{P}[\det(A_n) = 0] = o(1)$ as $n \to \infty$. A subsequent breakthrough by Kahn, Komlós and Szemerédi [10] (simplified by Tao and Vu [22]) showed that the singularity probability is in fact *exponentially small*: that is to say, $\mathbb{P}[\det(A_n) = 0] = \exp(-cn)$ for some constant $c > 0$. Quantitative aspects were later improved by Tao and Vu [23] and by Bourgain, Vu and Wood [3], culminating in a celebrated result of Tikhomirov [26] proving the essentially optimal bound $\mathbb{P}[\det(A_n) = 0] \le (1/2+o(1))^n$.

It is reasonable to expect similar results for the permanent. In fact one may expect that *even stronger* bounds should hold; for example, Vu (see [29, Conjecture 6.12] and [28]) conjectured that $\mathbb{P}[\mathrm{per}(A_n) = 0]$ decays faster than exponentially. However, the permanent is much more difficult to study than the determinant, and much less is known: until now, the only theorem in this direction is due to Tao and Vu [25], who proved that

$$\mathbb{P}[\mathrm{per}(A_n) = 0] \le n^{-c}$$

for some constant $c > 0$ (see also [24] for an earlier, simpler version of their argument that gives a bound of shape $\exp(-c\sqrt{\log n})$). Although Tao and Vu did not attempt to optimise the constant c, they note that their argument cannot hope to obtain a stronger bound than about $1/\sqrt{n}$, owing to their use of the so-called *Erdős–Littlewood–Offord inequality* (they write that "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").

In this paper, we overcome this barrier, obtaining the first exponential bound.

Theorem 1.1. *There is an absolute constant $c > 0$ such that the following holds. Let A_n be a uniformly random $n \times n$ matrix with ± 1 entries. Then, for any $x \in \mathbb{R}$,*

$$\mathbb{P}[\mathrm{per}(A_n) = x] \le \exp(-cn).$$

Zach Hunter was supported by SNSF grant 200021-228014. Matthew Kwan was supported by ERC Starting Grant "RANDSTRUCT" No. 101076777. Lisa Sauermann was supported by the DFG Heisenberg Program.

Here, we stated a bound for the probability that $\mathrm{per}(A_n) = x$ for general x, not just for $x = 0$, but this should not be viewed as a new feature of our work: it is easy to make the same generalisation with all previous work on determinants and permanents of discrete random matrices.

It is worth remarking that our proof works in exactly the same way for the determinant instead of the permanent, so as a byproduct this also gives a new proof of the Kahn–Komlós–Szemerédi theorem. All previous work on determinants of random matrices fundamentally used linear-algebraic properties of the determinant; since such properties are unavailable for the permanent, our proof is necessarily very different. In fact, our proof is also very different to the previous approach of Tao and Vu in [25] (which was until now the only known approach to study permanents of random matrices).

1.1. **General entry distributions.** Some of the previous work on determinants and permanents of discrete random matrices readily generalises to other entry distributions. For example, shortly after Komlós' foundational paper showing that $\mathbb{P}[\det(A_n) = 0] = o(1)$, in a second paper [13] he proved that the same result holds for a random matrix with independent μ-distributed entries, for any non-degenerate distribution μ (fixing μ and sending $n \to \infty$).

Bourgain, Vu and Wood [3] observed that in fact one does not need to fix the entry distribution; it suffices for the entries to be independent and "uniformly anticoncentrated" in the sense that there is a constant $\varepsilon > 0$ such that each entry a_{ij} satisfies $\sup_z \mathbb{P}[a_{ij} = z] \leq 1 - \varepsilon$. Specifically, they observed (see [3, Theorem 1.4]) that some of their ideas could be combined with an approach of Tao and Vu [22], to prove that $\mathbb{P}[\det(A_n) = 0]$ decays exponentially in this very general setting. Here, we prove the same statement for the permanent.

For a random variable X, we write $Q[X] = \sup_x \mathbb{P}[X = x]$ for the maximum point probability of X.

Theorem 1.2. *For any $0 < p < 1$ there is $c_p > 0$ such that the following holds. Let $A_n \in \mathbb{R}^{n\times n}$ be a random matrix with independent entries, such that every entry ξ of A_n satisfies $Q[\xi] \leq p$. Then*

$$Q[\mathrm{per}(A_n)] \leq \exp(-c_p n).$$

Note that Theorem 1.1 is an immediate corollary of Theorem 1.2 for $p = 1/2$. We remark that while Theorem 1.1 might be improvable to a super-exponential bound (as conjectured by Vu), in the more general setting of Theorem 1.2 an exponential bound is best-possible. Indeed, if the entries in the first row of A_n are each equal to zero with probability p, then with probability p^n the entire first row is zero, in which case $\mathrm{per}(A_n) = 0$.

1.2. **The range of the permanent.** For any set $S \subseteq \mathbb{R}$ and any $n \in \mathbb{N}$, we write $\Phi_S(n)$ for the set of all permanents of $n \times n$ matrices with entries in S.

For any discrete random variable X, it is easy to see that X must be supported on at least $1/Q[X]$ different values. So, by applying Theorem 1.2 to random matrices whose entries are uniformly sampled from any pair $\{a, b\} \subseteq \mathbb{R}$, we obtain an exponential lower bound on $|\Phi_{\{a,b\}}(n)|$.

Corollary 1.3. *For any set $S \subseteq \mathbb{R}$ of size at least 2, taking $c_{1/2}$ as in Theorem 1.2, we have*

$$|\Phi_S(n)| \geq \exp(c_{1/2} n),$$

This result is not new when $S = \{0, 1\}$. Indeed, in this case a classical theorem of Brualdi and Newman [4] says that $\Phi_{\{0,1\}}(n) \supseteq \{0, \ldots, 2^{n-1}\}$ (see also the improvements in [7]). However, Corollary 1.3 significantly improves on previous results when $S = \{-1, 1\}$. Indeed, in 1993 Kräuter [14] showed that $|\Phi_{\{-1,1\}}(n)| \geq n + 1$. This has remained essentially the state of the art[1] until very recently, when Ingram and Razborov [9] showed how to use ideas from additive combinatorics and Diophantine geometry to prove that $|\Phi_{\{-1,1\}}(n)| \geq \exp(c(\log n)^2/\log\log n)$ for some constant $c > 0$. Actually, they explicitly raised the problem (see [9, Problem 1]) of proving that $|\Phi_{\{-1,1\}}(n)|$ grows exponentially in n; this problem is solved by Corollary 1.3.

1.3. **Proof ideas.** The full proof of Theorem 1.2 is not very long, but it nonetheless seems worthwhile to give a brief glimpse of the ideas. The starting point is a "relative anticoncentration inequality" (Theorem 3.1), proved using a theorem of Kesten [11] (which appears as Theorem 3.2, and is the only aspect of our proof which is not self-contained). Very roughly speaking, Theorem 3.1 says that, under certain assumptions, sums of independent random variables are more anticoncentrated than "thinned-out versions of themselves" obtained by taking a randomly subsampled subset of the summands. This inequality is related to certain inequalities of Halász [8] that played a crucial role in the work of Kahn, Komlós and

[1] One can slightly improve Kräuter's bound using Tao and Vu's results on permanents of random matrices, but it is unclear how to get more than $o(n)$ additional permanents this way.

Szemerédi [10] on the singularity probability (mentioned earlier in this introduction). However, the way we use Theorem 3.1 is rather different to all previous work in this area: we use it to prove a *recursive* anticoncentration bound, taking advantage of the recursive properties of the formula in (1.1).

In a bit more detail: in the setting of Theorem 1.2, it is not hard to see that $\mathrm{per}(A_n)$ can be interpreted as a weighted sum $a_1\xi_1 + \cdots + a_n\xi_n$, where $\xi_1, \dots, \xi_n$ are the entries of the last row, and $a_1, \dots, a_n$ are permanents of $(n-1) \times (n-1)$ submatrices of the first $n-1$ rows. If we condition on the first $n-1$ rows of A_n (so the only remaining randomness is in the last row) then this is a sum of independent random variables, to which we can apply our relative anticoncentration theorem. It turns out that if we average the resulting bound over all possible outcomes of the first $n-1$ rows, we get a bound on the anticoncentration of $\mathrm{per}(A_n)$, in terms of the anticoncentration of the permanents of certain "smaller" random matrices.

To see why this is useful, suppose that we were able to prove a bound of the form $Q[\mathrm{per}(A_n)] \leq \alpha_p Q[\mathrm{per}(A_{n-1})]$, for some $\alpha_p < 1$. We would then be able to iterate this bound to prove that $Q[\mathrm{per}(A_n)] \leq \alpha_p^n$; i.e., $\mathrm{per}(A_n)$ is exponentially anticoncentrated, as desired. Unfortunately, the situation is not quite this simple: it turns out that for our smaller random matrices, the point probabilities of the entries are not bounded by p anymore, and also there are "error terms" corresponding to certain assumptions we need to make to effectively apply Theorem 3.1. Nonetheless, with various additional ideas, we can prove a sufficiently strong recursive anticoncentration bound (see Theorem 2.2) for a certain more general class of random matrices (which are allowed to contain some fixed deterministic rows in addition to their random rows).

It is worth remarking that our proof is "purely combinatorial"; in particular, it features no geometry, no linear algebra and no Fourier analysis.

1.4. **Further directions.** The most obvious direction for further research is to improve the bound in Theorem 1.1, towards the super-exponential bound conjectured by Vu.

Also, it would be very interesting to generalise these ideas to *symmetric* random matrices. For a random symmetric $n \times n$ matrix S_n with ± 1 entries, the second and third author [15] proved that $\mathbb{P}[\mathrm{per}(S_n) = 0] \leq n^{-c}$ for some constant c, but this method cannot prove a bound stronger than $1/\sqrt{n}$. Even for the determinant, it is much more difficult to prove exponential bounds in the symmetric setting: this was accomplished only recently by Campos, Jenssen, Michelen and Sahasrabudhe [5].

Another direction is to consider the "small-ball" setting: namely, instead of just studying the probability that $\mathrm{per}(A_n)$ is equal to a given point, we could study the probability that it lies in an interval of given length. Actually, Tao and Vu [24] already considered this small-ball setting: their proof incorporates a "growth" step, which allows them to prove that $\mathbb{P}[|\mathrm{per}(A_n) - x| \leq n^{(1/2-\varepsilon)n}] \leq n^{-1/10}$ for any $x \in \mathbb{R}$ and any constant $\varepsilon > 0$ (assuming n is sufficiently large). By incorporating this growth step into our own proof scheme, one can prove the same result with an exponential probability bound. (This introduces some additional technicalities, so in the interest of keeping the arguments simple, in this paper we decided to restrict our attention to the point-probability setting.)

In the latter direction it is worth mentioning the *permanent anticoncentration conjecture* of Aaronson and Arkhipov [1] (a key ingredient in the analysis of the hardness of the *BosonSampling* problem in quantum computing). This is a precise conjecture about the typical size of the permanent of a Gaussian random matrix, seemingly well outside the reach of current techniques. See [1] for some discussion of the relevant obstacles; some of these obstacles can be overcome by ideas in this paper, but the critical obstacle is that existing techniques are not sufficiently sensitive to small differences in the permanent.

1.5. **Notation.** Let $\mathbb{N} = \{0, 1, 2, \dots\}$ be the set of non-negative integers. We use asymptotic notation throughout: for functions $f = f(n)$ and $g = g(n)$, we write $f = O(g)$ or $g = \Omega(f)$ or $f \lesssim g$ to mean that there is a constant C such that $|f(n)| \leq C|g(n)|$ for sufficiently large n, and we write $f = o(g)$ to mean that $f/g \to 0$ as $n \to \infty$. Subscripts on asymptotic notation indicate quantities that should be treated as constants. Also, all random variables throughout this paper are assumed to be real-valued.

2. A recursive bound

We will actually prove a slight strengthening of Theorem 1.2, where we consider matrices which have a random part and a fixed part. This is necessary for our recursive approach.

Definition 2.1. For $0 < p < 1$ and $n \in \mathbb{N}$, let $f_p(n)$ be the supremum of $Q[\mathrm{per}(A)]$, over all random matrices A of the following form.

- For some $T \in \mathbb{N}$, let $A_{\text{fix}} \in \mathbb{R}^{T\times(n+T)}$ be a fixed $T \times (n+T)$ matrix containing at least one $T \times T$ submatrix with nonzero permanent.
- Let $A_{\text{rand}} \in \mathbb{R}^{n\times(n+T)}$ be a random matrix with independent entries, such that every entry ξ of A_{rand} satisfies $Q[\xi] \le p$.
- Let $A \in \mathbb{R}^{(n+T)\times(n+T)}$ be the $(n+T) \times (n+T)$ matrix obtained by concatenating A_{fix} and A_{rand} (i.e., the first T rows are the rows of A_{fix}, and the last n rows are the rows of A_{rand}).

In Definition 2.1, we think of n as being the "size" of A (the random part A_{rand} has n rows, and in the proof of Theorem 1.2, we will only really use the randomness of n of its columns). We will prove a recursive bound for $f_p(n)$, see Theorem 2.2 below. To state this bound, for any $0 < p < 1$, let us define

$$\tau_p = 1 - \prod_{s=1}^{\infty}(1-p^s), \tag{2.1}$$

and observe that $\tau_p < 1$, no matter how close p is to 1. To see this, let C_p be sufficiently large in terms of p such that $\exp(-C_p p) \le 1-p$, so for every $0 \le x \le p$ we have $1-x \ge \exp(-C_p x)$, and hence $1-\tau_p = \prod_{s=1}^{\infty}(1-p^s) \ge \prod_{s=1}^{\infty}\exp(-C_p p^s) = \exp\big(-C_p\sum_{s=1}^{\infty}p^s\big) = \exp(-C_p p/(1-p)) > 0$.

Theorem 2.2. *For any $0 < p < 1$ and $t \in \mathbb{N}$ and $\varepsilon > 0$, there is $\delta > 0$ and $N \in \mathbb{N}$ such that for all $n \ge N$ we have*

$$f_p(n) \le \exp(-\delta n) \;+\; \varepsilon\sum_{s=1}^{t}p^s f_p(n-s) \;+\; (1+\varepsilon)\tau_p\cdot f_p(n-t) \;+\; (1+\varepsilon)\sum_{s=t+1}^{n}p^s f_p(n-s).$$

In the course of proving Theorem 2.2, we will also prove the following easier bound (which shall serve as a base case when proving Theorem 1.2 by induction).

Proposition 2.3. *For any $p < 1$ and $n \in \mathbb{N}$, we have*

$$f_p(n) \le 1 - \prod_{s=1}^{n}(1-p^s) \le \tau_p < 1.$$

Proof of Theorem 1.2 given Theorem 2.2 and Proposition 2.3. As in the theorem statement, fix $0 < p < 1$. First, note that it suffices to prove that $f_p(n) \le \exp(-c_p n)$ for some $c_p > 0$. Indeed, any random matrix A_n as in Theorem 1.2 trivially satisfies the conditions in Definition 2.1 for $T = 0$, so $Q[\text{per}(A_n)] \le f_p(n)$.

Recalling that $\tau_p < 1$, choose $\mu > 0$ such that $\tau_p < 1-5\mu$. Then, choose $\varepsilon > 0$ such that $\varepsilon < (1-\sqrt{p})\mu$ and $(1+\varepsilon)\tau_p < (1+\varepsilon)(1-5\mu) < 1-4\mu$. Also, choose $t \in \mathbb{N}$ such that $p^{t/2} < (1-\sqrt{p})\mu$. Now, let $\delta > 0$ and $N \in \mathbb{N}$ be as in Theorem 2.2. By making N larger if needed, we may assume that $\exp(-\delta N/2) \le \mu$.

By Proposition 2.3 we have $f_p(n) < 1$ for all $n \in \mathbb{N}$. So, we can choose some $c_p > 0$ such that $f_p(n) < \exp(-c_p n)$ for $n = 0, 1, \ldots, N$, and also $c_p < \delta/2$ and $\exp(c_p) < p^{-1/2}$ and $(1-4\mu)\exp(c_p t) < 1-3\mu$.

Let us now prove by induction that $f_p(n) \le \exp(-c_p n)$ for all $n \in \mathbb{N}$. For $n \le N$ this is automatically guaranteed by the choice of c_p. So let us now consider $n > N$, and assume that we have already shown $f_p(n') \le \exp(-c_p n')$ for all $n' < n$. Then, using Theorem 2.2, we have

$$\begin{aligned}
f_p(n) &\le \exp(-\delta n) + \varepsilon\sum_{s=1}^{t}p^s f_p(n-s) + (1+\varepsilon)\tau_p\cdot f_p(n-t) + (1+\varepsilon)\sum_{s=t+1}^{n}p^s f_p(n-s)\\
&\le \exp(-\delta n) + \varepsilon\sum_{s=1}^{t}p^s\exp(-c_p(n-s)) + (1-4\mu)\exp(-c_p(n-t)) + (1+\varepsilon)\sum_{s=t+1}^{n}p^s\exp(-c_p(n-s)).
\end{aligned}$$

This yields

$$\begin{aligned}
f_p(n)\cdot\exp(c_p n) &\le \exp(-(\delta-c_p)n) + \varepsilon\sum_{s=1}^{t}p^s\exp(c_p s) + (1-4\mu)\exp(c_p t) + (1+\varepsilon)\sum_{s=t+1}^{n}p^s\exp(c_p s)\\
&\le \exp(-\delta n/2) + \varepsilon\sum_{s=1}^{t}p^s p^{-s/2} + (1-3\mu) + (1+\varepsilon)\sum_{s=t+1}^{n}p^s p^{-s/2}\\
&\le (1-3\mu) + \exp(-\delta N/2) + \varepsilon\sum_{s=1}^{n}p^{s/2} + \sum_{s=t+1}^{n}p^{s/2}\\
&\le (1-3\mu) + \exp(-\delta N/2) + \frac{\varepsilon}{1-\sqrt{p}} + \frac{p^{t/2}}{1-\sqrt{p}} \le (1-3\mu)+\mu+\mu+\mu = 1,
\end{aligned}$$

as desired. □

3. A relative anticoncentration inequality

A key ingredient in the proof of Theorem 2.2 is the following "relative anticoncentration inequality", stating that under certain assumptions, a sum of independent random variables is more anticoncentrated than a "randomly thinned-out version of itself". This is similar to [21, Lemma 7.14] and related inequalities of Halász [8], used in previous work on the singularity probability of random matrices.

Theorem 3.1. *Consider independent random variables $X_1, \dots, X_n$, and for each $i = 1, \dots, n$, let X_i' be an independent copy of X_i. For some integer $1 \le k < n/4$, let $K \subseteq \{1, \dots, n\}$ be a uniformly random subset of exactly k indices (independent from the variables $X_1, \dots, X_n, X_1', \dots, X_n'$). Now, define $X = X_1 + \cdots + X_n$ and $X_K^* = \sum_{i \in K}(X_i - X_i')$, and assume that $\mathbb{P}[X_K^* = 0] < 1 - 4k/n$. Then we have*

$$Q[X] \lesssim \sqrt{\frac{k}{n}} \cdot \frac{\mathbb{P}[X_K^* = 0]}{1 - \mathbb{P}[X_K^* = 0]}.$$

The main ingredient to prove Theorem 3.1 is the following (special case of a) theorem of Kesten [11].

Theorem 3.2 ([11, Theorem 2])**.** *Consider independent random variables $X_1, \dots, X_m$ and $0 < p < 1$, such that $Q[X_i] \le p$ holds for each $i = 1, \dots, m$. Then*

$$Q[X_1 + \cdots + X_m] \lesssim \frac{p}{\sqrt{(1-p)m}}.$$

We note that Theorem 3.2 is the only part of this paper that is not self-contained. It was proved with a simple (but quite ingenious!) combinatorial argument, via an anticoncentration inequality of Rogozin [19] (which, in this context, is essentially equivalent to the so-called *Erdős–Littlewood–Offord theorem* [6]). See also [18] for an alternative Fourier-analytic proof.

In our proof of Theorem 3.1, we will actually need the following slight generalisation of Theorem 3.2, which compares $Q[X_1 + \cdots + X_m]$ to probabilities that the random variables X_i are equal to independent copies of themselves.

Corollary 3.3. *Consider independent random variables $X_1, \dots, X_n$ and $0 < p < 1$. For each $i = 1, \dots, n$, let X_i' be an independent copy of X_i. Suppose that we have $\mathbb{P}[X_i = X_i'] \le p$ for at least $m \ge 2$ different indices $i \in \{1, \dots, n\}$. Then*

$$Q[X_1 + \cdots + X_n] \lesssim \frac{p}{\sqrt{(1-p)m}}.$$

To deduce Corollary 3.3 from Theorem 3.2 we will need a few more general inequalities. First, we record the (near-trivial) fact that a sum of independent random variables is at least as anticoncentrated as its summands.

Fact 3.4. *For independent random variables X, Y we have $Q[X + Y] \le Q[X]$.*

Proof. For any $z \in \mathbb{R}$ we have

$$\mathbb{P}[X + Y = z] = \mathbb{P}[X = z - Y] = \mathbb{E}\big[\mathbb{P}[X = z - Y \mid Y]\big] \le \mathbb{E}\big[Q[X]\big] = Q[X].$$ □

Next, we record a simple "duplication" inequality for a sum of independent random variables.

Lemma 3.5. *Consider independent discrete random variables X_1 and X_2. Letting X_1' and X_2' be independent copies of X_1 and X_2, we have*

$$Q[X_1 + X_2] \le \mathbb{P}[X_1 = X_1']^{1/2} \cdot \mathbb{P}[X_2 = X_2']^{1/2}.$$

In particular (applying the above inequality to the case where X_2 has the distribution of $-X_1$), we have

$$Q[X_1 - X_1'] \le \mathbb{P}[X_1 = X_1'].$$

Proof. We have

$$Q[X_1 + X_2] = \sup_x \sum_z \mathbb{P}[X_1 = z]\mathbb{P}[z + X_2 = x] \le \sup_x \left(\sum_z \mathbb{P}[X_1 = z]^2\right)^{1/2} \left(\sum_z \mathbb{P}[z + X_2 = x]^2\right)^{1/2}$$

by the Cauchy–Schwarz inequality. But note that (for any $x \in \mathbb{R}$)

$$\sum_z \mathbb{P}[X_1 = z]^2 = \sum_z \mathbb{P}[X_1 = z]\mathbb{P}[X_1' = z] = \mathbb{P}[X_1 = X_1'],$$

$$\sum_z \mathbb{P}[z + X_2 = x]^2 = \sum_z \mathbb{P}[X_2 = x - z]\mathbb{P}[X_2' = x - z] = \mathbb{P}[X_2 = X_2'],$$

and the desired result follows. □

Now, we are ready to prove Corollary 3.3.

Proof of Corollary 3.3. We may assume that $\mathbb{P}[X_i = X_i'] \le p$ for $i = 1, \dots, m$. By Fact 3.4 we may furthermore assume that $n = m$ and that $m \ge 2$ is even. Also, note that we may assume that $X_1, \dots, X_m$ are discrete random variables[2].

By the first part of Lemma 3.5, we see that $Q[X_1 + \cdots + X_m]$ is at most

$$\mathbb{P}[(X_1 + \cdots + X_{m/2}) = (X_1' + \cdots + X_{m/2}')]^{1/2} \cdot \mathbb{P}[(X_{m/2+1} + \cdots + X_m) = (X_{m/2+1}' + \cdots + X_m')]^{1/2}.$$

Then, note that

$$\begin{aligned}\mathbb{P}[(X_1 + \cdots + X_{m/2}) = (X_1' + \cdots + X_{m/2}')] &= \mathbb{P}[(X_1 - X_1') + \cdots + (X_{m/2} - X_{m/2}') = 0] \\ &\le Q[(X_1 - X_1') + \cdots + (X_{m/2} - X_{m/2}')].\end{aligned}$$

By the second part of Lemma 3.5, for each $i = 1, \dots, m$, we have $Q[X_i - X_i'] \le \mathbb{P}[X_i = X_i'] \le p$, so Theorem 3.2 yields

$$\mathbb{P}[(X_1 + \cdots + X_{m/2}) = (X_1' + \cdots + X_{m/2}')] \lesssim \frac{p}{\sqrt{(1-p)(m/2)}} \lesssim \frac{p}{\sqrt{(1-p)m}}.$$

By exactly the same argument, we obtain

$$\mathbb{P}[(X_{m/2+1} + \cdots + X_m) = (X_{m/2+1}' + \cdots + X_m')] \lesssim \frac{p}{\sqrt{(1-p)m}},$$

and the desired result follows. □

Finally, we use Corollary 3.3 to prove Theorem 3.1.

Proof of Theorem 3.1. Let $h = \lfloor n/k \rfloor$ and let $I_1, \dots, I_h \in \{1, \dots, n\}$ be uniformly random disjoint subsets of exactly k indices (independent from all other random objects in the statement of Theorem 3.1). For each $j = 1, \dots, h$, define

$$X_j^* = \sum_{i \in I_j} (X_i - X_i'), \quad \rho_j = \mathbb{P}[X_j^* = 0 \,|\, I_j] = \mathbb{P}\Big[\sum_{i \in I_j} X_i = \sum_{i \in I_j} X_i' \,\Big|\, I_j\Big],$$

and let $\lambda = 1 - \mathbb{P}[X_K^* = 0] > 4k/n > 2/h$. Note that for each $j = 1, \dots, h$, the random variable X_j^* on its own has exactly the same distribution as X_K^*. Thus, we have $\mathbb{E}[\rho_1 + \cdots + \rho_h] = h \cdot \mathbb{P}[X_K^* = 0] = (1-\lambda)h$. So, we can fix an outcome of $I_1, \dots, I_h$ satisfying $\rho_1 + \cdots + \rho_h \le (1-\lambda)h$.

Now, we can have at most $(1 - \lambda/2)h$ indices $j \in \{1, \dots, h\}$ with $\rho_j \ge (1-\lambda)/(1-\lambda/2)$. Thus, there are at least $(\lambda/2)h$ indices $j \in \{1, \dots, h\}$ satisfying $\rho_j < (1-\lambda)/(1-\lambda/2)$. Then, Fact 3.4 and Corollary 3.3 (applied with $m = \lceil (\lambda/2)h \rceil \ge 2$ and $p = (1-\lambda)/(1-\lambda/2) < 1 - \lambda/2$, with the random variables $\sum_{i \in I_j} X_i$ and their independent copies $\sum_{i \in I_j} X_i'$ for $j = 1, \dots, h$) give

$$Q[X] \le Q\Big[\sum_{i \in I_1} X_i + \cdots + \sum_{i \in I_h} X_i\Big] \lesssim \frac{(1-\lambda)/(1-\lambda/2)}{\sqrt{(\lambda/2) \cdot (\lambda/2)h}} \lesssim \frac{1}{\sqrt{h}} \cdot \frac{1-\lambda}{\lambda} \lesssim \sqrt{\frac{k}{n}} \cdot \frac{\mathbb{P}[X_K^* = 0]}{1 - \mathbb{P}[X_K^* = 0]}. \quad \square$$

Note that Theorem 3.1 only applies under the assumption that $\mathbb{P}[X_K^* = 0]$ is sufficiently far away from 1. We record a corollary with an additional condition that allows us to easily verify this assumption.

Corollary 3.6. *Consider $\gamma > 0$ and $0 < p < 1$, and in the setting of Theorem 3.1, assume that $2/\gamma \le k < (1-p)n/8$ and that we have $Q[X_i] \le p$ for at least γn indices $i \in \{1, \dots, n\}$. Then $\mathbb{P}[X_K^* = 0] \le (1+p)/2 < 1 - 4k/n$ and consequently*

$$Q[X] \lesssim \sqrt{\frac{k}{n}} \cdot \frac{\mathbb{P}[X_K^* = 0]}{1 - \mathbb{P}[X_K^* = 0]} \lesssim_p \sqrt{\frac{k}{n}} \cdot \mathbb{P}[X_K^* = 0].$$

[2]Formally: for each i, write $A_i = \{x : \mathbb{P}[X_i = x] > 0\}$ as the "atomic part" of X_i, and note that for any z we have $\mathbb{P}[X_1 + \cdots + X_n = z$ and $X_i \notin A_i$ for some $i] = 0$. So, if we alter the behaviour of X_i when $X_i \notin A_i$ (to make X_i discrete), this cannot decrease $Q[X_1 + \cdots + X_m]$. We can make these alterations in a way that has an arbitrarily small impact on probabilities of the form $\mathbb{P}[X_i = X_i']$.

Proof. Let I be a set of $\lceil \gamma n \rceil$ indices $i \in \{1, \dots, n\}$ satisfying $Q[X_i] \le p$. Then $|K \cap I|$ has a hypergeometric distribution, with mean $\mathbb{E}[|K \cap I|] = |I| \cdot k/n \ge \gamma k$ and variance $\mathrm{Var}[|K \cap I|] \le |I| \cdot k/n = \mathbb{E}[|K \cap I|]$. So by Chebyshev's inequality, we have $\mathbb{P}[K \cap I = \emptyset] \le 1/\mathbb{E}[|K \cap I|] \le 1/(\gamma k) \le 1/2$. Thus, we have $K \cap I \ne \emptyset$ with probability at least $1/2$. For any such outcome of K we have $\mathbb{P}[X^*_K = 0 \,|\, K] \le p$ by Fact 3.4. So, we can conclude that $\mathbb{P}[X^*_K = 0] \le (1+p)/2$. Using $k < (1-p)n/8$, it is easy to check that $(1+p)/2 < 1 - 4k/n$, and the rest of the statement is obtained by applying Theorem 3.1. □

4. Proof of the recursive bound

In this section we prove Theorem 2.2 (and along the way, Proposition 2.3). So, let $0 < p < 1$, and let A be as in Definition 2.1. Throughout, we will assume without loss of generality that the last T columns of A_{fix} induce a $T \times T$ submatrix with nonzero permanent. We will also use the following notation.

Definition 4.1. In the setting of Definition 2.1:

- For a set $I \subseteq \{1, \dots, n+T\}$, let $-I = \{1, \dots, n+T\} \setminus I$ be the complement of I.
- For $i \in I$, let $I - i = I \setminus \{i\}$, and for $i \notin I$, let $I + i = I \cup \{i\}$.
- For a set $J \subseteq \{1, \dots, n+T\}$, let $A[J]$ be the $|J| \times |J|$ submatrix of A indexed by the first $|J|$ rows and the columns indexed by J.
- For $s \in \{1, \dots, n\}$, let $A^{\uparrow s}$ be the $(n+T-s) \times (n+T)$ submatrix of A consisting of the first $n+T-s$ rows.
- For $s \in \{1, \dots, n\}$, $z \in \mathbb{R}$ and $\alpha > 0$, let $\mathcal{E}_z(s, \alpha)$ be the event that for more than $\alpha\binom{n}{s}$ of the $\binom{n}{s}$ size-s subsets $I \subseteq \{1, \dots, n\}$, we have $\mathrm{per}(A[-I]) \ne z$. We also use the shorthand $\mathcal{E}(s, \alpha) = \mathcal{E}_0(s, \alpha)$ and $\mathcal{E}_z(s) = \mathcal{E}_z(s, 0)$ and $\mathcal{E}(s) = \mathcal{E}_0(s, 0)$.

Note that the event $\mathcal{E}(s)$ precisely means that the $(n+T-s) \times (n+T)$ matrix $A^{\uparrow s}$ contains some $(n+T-s) \times (n+T-s)$ submatrix with non-zero permanent. In particular, $\mathcal{E}(n)$ does not depend on the randomness of A_{rand} (since $A^{\uparrow n} = A_{\text{fix}}$), and holds deterministically by the conditions in Definition 2.1.

The permanent enjoys an analogue of the minor expansion formula for the determinant, which we record below. This fact allows us to study permanents recursively.

Fact 4.2. *Fix a set $J \subseteq \{1, \dots, n\}$ of size s. We have*

$$\mathrm{per}(A[-J]) = \sum_{\substack{i \in \{1, \dots, n+T\} \\ i \notin J}} \xi_i \, \mathrm{per}(A[-(J+i)]),$$

where $\xi_1, \dots, \xi_{n+T}$ are the entries of the $(n+T-s)$-th row of A.

We collect various estimates on probabilities involving the events $\mathcal{E}_z(s, \alpha)$. First, the following simple lemma gives a bound on the probability that $\mathrm{per}(A[-J]) = z$ for all size-$(s-1)$ subsets $J \subseteq \{1, \dots, n\}$, given that there is at least one size-s subset $I \subseteq \{1, \dots, n\}$ with $\mathrm{per}(A[-I]) \ne 0$.

Lemma 4.3. *For any $s \in \{1, \dots, n\}$, any $z \in \mathbb{R}$ and any outcome of $A^{\uparrow s}$ satisfying $\mathcal{E}(s)$, we have*

$$\mathbb{P}\left[\overline{\mathcal{E}_z(s-1)} \,\middle|\, A^{\uparrow s}\right] \le p^s.$$

Proof. Conditioning on the given outcome of $A^{\uparrow s}$, there is some size-s subset $I \subseteq \{1, \dots, n\}$ satisfying $\mathrm{per}(A[-I]) \ne 0$. Write $\xi_1, \dots, \xi_{n+T}$ for the entries of the $(n-s+1)$-th row of A_{rand}, and reveal ξ_i for all $i \in \{1, \dots, n+T\} \setminus I$.

Now, consider all the s permanents of the form $\mathrm{per}(A[-(I-i)])$ for $i \in I$. It suffices to show that (conditioned on the revealed information, using only the randomness of ξ_i for $i \in I$), the probability that all these permanents are equal to z is at most p^s. This is a fairly immediate consequence of Fact 4.2, using that $Q[\xi_i] \le p$ for all $i \in I$. Indeed, in our conditional probability space, for each $i \in I$ the random variable $\mathrm{per}(A[-(I-i)])$ only depends on the randomness of ξ_i, and at most one of the outcomes of ξ_i will cause $\mathrm{per}(A[-(I-i)]) \ne z$ (since $\mathrm{per}(A[-I]) \ne 0$). □

Note that Lemma 4.3 in particular implies $\mathbb{P}[\overline{\mathcal{E}_z(s-1)} \,|\, \mathcal{E}(s)] \le p^s$ for any $s \in \{1, \dots, n\}$ and any $z \in \mathbb{R}$. This is already enough to prove Proposition 2.3, as follows.

Proof of Proposition 2.3. Consider any $z \in \mathbb{R}$. Recalling that the event $\mathcal{E}(n)$ always holds, we have

$$\mathbb{P}[\mathrm{per}(A) \ne z] = \mathbb{P}[\mathcal{E}_z(0)] \ge \mathbb{P}\left[\mathcal{E}_z(0) \cap \bigcap_{s=1}^{n} \mathcal{E}(s)\right] = \mathbb{P}[\mathcal{E}_z(0) \,|\, \mathcal{E}(1)] \prod_{s=2}^{n} \mathbb{P}[\mathcal{E}(s-1) \,|\, \mathcal{E}(s)] \ge \prod_{s=1}^{n} (1 - p^s),$$

by Lemma 4.3. The desired result follows. □

The next lemma is a simple application of Markov's inequality.

Lemma 4.4. *For any $s \in \{1, \dots, n\}$ and $0 < \alpha < 1/2$, we have*

$$\mathbb{P}\Big[\overline{\mathcal{E}(s,\alpha)}\Big] \le \frac{f_p(n-s)}{1-\alpha} \le (1+2\alpha) f_p(n-s).$$

Proof. If $\overline{\mathcal{E}(s,\alpha)}$ holds, there must be at least $(1-\alpha)\binom{n}{s}$ size-s subsets $I \subseteq \{1, \dots, n\}$ with $\mathrm{per}(A[-I]) = 0$. For each size-s subset $I \subseteq \{1, \dots, n\}$, the matrix $A[-I]$ has the form in Definition 2.1 (with $n-s$ random rows), so the expected number of such subsets $I \subseteq \{1, \dots, n\}$ satisfying $\mathrm{per}(A[-I]) = 0$ is at most $f_p(n-s)\binom{n}{s}$. The desired result then follows from Markov's inequality. □

Finally, the next lemma is where most of the difficulty lies. It combines the ideas from Lemmas 4.3 and 4.4, together with Theorem 3.1 (our relative anticoncentration inequality).

Lemma 4.5. *Fix $\alpha > 0$. For large n, any $s \in \{1, \dots, n\}$, and any $z \in \mathbb{R}$, we have*

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1)} \cap \mathcal{E}(s,\alpha)\Big] \le \mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \cap \mathcal{E}(s,\alpha)\Big] \le \alpha p^s f_p(n-s) + \exp(-\Omega_{\alpha,p}(n)).$$

In our proof of Lemma 4.5 we will use the following simple fact about dense set families.

Lemma 4.6. *For $n \in \mathbb{N}$, $s \in \{1, \dots, n\}$ and $\alpha > 0$, let $\mathcal{F}$ be a collection of at least $\alpha\binom{n}{s}$ size-s subsets of $\{1, \dots, n\}$. Then there is a collection $\mathcal{H}$ of $|\mathcal{H}| \ge (\alpha/2)s\binom{n}{s}$ pairs of the form (F, G) with $F \in \mathcal{F}$ and $G \subseteq F$ of size $|G| = s-1$, such that for each $(F, G) \in \mathcal{H}$, there are at least $(\alpha/2)(n-s+1)$ sets $F' \in \mathcal{F}$ satisfying $(F', G) \in \mathcal{H}$.*

Proof. Consider the bipartite graph whose vertices on the left side correspond to the sets $F \in \mathcal{F}$ and whose vertices on the right side correspond to the $\binom{n}{s-1}$ different size-$(s-1)$ subsets $G \subseteq \{1, \dots, n\}$. We draw an edge between a vertex F on the left and a vertex G on the right if and only if $G \subseteq F$. Then the graph has $s \cdot |\mathcal{F}| \ge \alpha s\binom{n}{s}$ edges. There can be at most $\binom{n}{s-1} \cdot (\alpha/2)(n-s+1) = (\alpha/2)s\binom{n}{s}$ edges (F, G) with $\deg(G) \le (\alpha/2)(n-s+1)$. Thus, the graph has at least $(\alpha/2)s\binom{n}{s}$ edges (F, G) with $\deg(G) > (\alpha/2)(n-s+1)$. We can now take $\mathcal{H}$ to be the set of these edges. □

Now we prove Lemma 4.5.

Proof of Lemma 4.5. Let C_p be a constant depending only on p such that in the conclusion of Corollary 3.6 we always have $Q[X] \le C_p \cdot \sqrt{k/n} \cdot \mathbb{P}[X_K^* = 0]$. Let us fix some small $\nu > 0$ depending on α and p such that $\nu < (1-p)/8$ and $\sqrt{\nu} \le (\alpha^2/4) \cdot (p/C_p)$.

The first inequality in the statement of Lemma 4.5 follows from the fact that the event $\mathcal{E}_z(s-1) = \mathcal{E}_z(s-1, 0)$ contains the event $\mathcal{E}_z(s-1, \alpha/(4s))$. So it suffices to prove the second inequality. Due to the asymptotic term $\exp(-\Omega_{\alpha,p}(n))$ in the statement, we may assume that n is large with respect to p, α, ν.

Let $\xi_1, \dots, \xi_{n+T}$ be the entries of the $(n+T-s+1)$-th row of A. Our strategy is to use Corollary 3.6 to estimate the probability of the event $\overline{\mathcal{E}_z(s-1, \alpha/(4s))}$, conditioned on an outcome of $A^{\uparrow s}$ such that $\mathcal{E}(s, \alpha)$ holds (using the randomness of $\xi_1, \dots, \xi_{n+T}$). Then, we will sum this estimate over all outcomes of $A^{\uparrow s}$ satisfying $\mathcal{E}(s, \alpha)$ to obtain a bound for the probability of $\overline{\mathcal{E}_z(s-1, \alpha/(4s))} \cap \mathcal{E}(s, \alpha)$. In order to bound the probability of the event $\overline{\mathcal{E}_z(s-1, \alpha/(4s))}$ conditioned on an outcome of $A^{\uparrow s}$, we will consider the events $\mathrm{per}(A[-I]) \ne 0$ and $\mathrm{per}(A[-J]) \ne 0$ for various choices of nested subsets $J \subseteq I \subseteq \{1, \dots, n\}$ of sizes $|I| = s$ and $|J| = s-1$. Let (I^*, J^*) be a uniformly random pair of subsets $J^* \subseteq I^* \subseteq \{1, \dots, n\}$ with $|I^*| = s$ and $|J^*| = s-1$, independent from the random matrix A (note that the total number of such subset pairs is exactly $s\binom{n}{s}$).

Step 1: Reducing to a linear anticoncentration problem. Consider an outcome of $A^{\uparrow s}$ such that $\mathcal{E}(s, \alpha)$ holds. Let $\mathcal{F}$ be the family of all size-s subsets $I \subseteq \{1, \dots, n\}$ with $\mathrm{per}(A[-I]) \ne 0$ (note that this family $\mathcal{F}$ is determined by the outcome of $A^{\uparrow s}$). By the definition of the event $\mathcal{E}(s, \alpha) = \mathcal{E}_0(s, \alpha)$, we have $|\mathcal{F}| > \alpha\binom{n}{s}$. Thus, by Lemma 4.6 we can find a collection $\mathcal{H}$ of at least $(\alpha/2)s\binom{n}{s}$ pairs (I, J) with $I \in \mathcal{F}$ and $J \subseteq I$ of size $|J| = s-1$, such that for each $(I, J) \in \mathcal{H}$, there are at least $(\alpha/2)(n-s+1)$ sets $I' \in \mathcal{F}$ satisfying $(I', J) \in \mathcal{H}$.

Let us say that a set $I \in \mathcal{F}$ is *parental* if it has a size-$(s-1)$ subset $I' \subseteq I$ with $\mathrm{per}(A[-I']) \ne z$ (we can think of I' as a *child* of I). Note that for each size-$(s-1)$ subset $I' \subseteq \{1, \dots, n\}$ there can be at most $n-s+1$ different sets $I \in \mathcal{F}$ with $I' \subseteq I$ (so every child has at most $n-s+1$ parents). In particular, the number of size-$(s-1)$ subsets $I' \subseteq \{1, \dots, n\}$ with $\mathrm{per}(A[-I']) \ne z$ is at least

$$\frac{|\{I \in \mathcal{F} : I \text{ parental}\}|}{n-s+1} \ge \frac{|\{(I, J) \in \mathcal{H} : I \text{ parental}\}|}{s(n-s+1)},$$

observing that for every $I \in \mathcal{F}$ there can be at most s different sets J with $(I, J) \in \mathcal{H}$.

Thus, if for more than half of the pairs $(I, J) \in \mathcal{H}$ the set I is parental, we have more than

$$\frac{|\mathcal{H}|/2}{s(n-s+1)} \geq \frac{(\alpha/4)s\binom{n}{s}}{s(n-s+1)} = \frac{(\alpha/4)(n-s+1)\binom{n}{s-1}}{s(n-s+1)} = \frac{\alpha}{4s} \cdot \binom{n}{s-1}$$

different size-$(s-1)$ subsets $I' \subseteq \{1, \dots, n\}$ with $\mathrm{per}(A[-I']) \neq z$, so the event $\mathcal{E}_z(s-1, \alpha/(4s))$ holds.

Therefore, we can conclude

$$\begin{aligned}
\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \,\Big|\, A^{\uparrow s}\Big] &\leq \mathbb{P}\Big[I \text{ is not parental for at least half of the pairs } (I,J) \in \mathcal{H} \,\Big|\, A^{\uparrow s}\Big] \\
&\leq \frac{\mathbb{E}\big[|\{(I,J) \in \mathcal{H} : I \text{ not parental}\}| \,\big|\, A^{\uparrow s}\big]}{|\mathcal{H}|/2} \\
&= \frac{2}{|\mathcal{H}|} \sum_{(I,J)\in\mathcal{H}} \mathbb{P}\Big[I \text{ not parental} \,\Big|\, A^{\uparrow s}\Big] \\
&= 2 \cdot \mathbb{P}\Big[I^* \text{ not parental} \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big], \qquad (4.1)
\end{aligned}$$

where $\mathcal{E}_{\mathcal{H}}$ denotes the event that $(I^*, J^*) \in \mathcal{H}$ (noting that conditional on this event, the pair (I^*, J^*) is a unformly random pair in $\mathcal{H}$). Now, note that

$$\mathbb{P}\Big[I^* \text{ not parental} \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big] \leq \mathbb{P}\Big[\mathrm{per}(A[-J^*]) = z \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big] \cdot p^{s-1}. \qquad (4.2)$$

The justification of (4.2) is very similar to the proof of Lemma 4.3 (recalling Fact 4.2). If we reveal $A^{\uparrow s}$ and an outcome $(I^*, J^*) \in \mathcal{H}$, and we reveal ξ_i for $i \in \{1, \dots, n+T\} \setminus J^*$, then this already determines $\mathrm{per}(A[-J^*])$. If $\mathrm{per}(A[-J^*]) \neq z$ then I^* is automatically parental. Furthermore, in the resulting conditional probability space given this revealed information, for each $i \in J^*$ the random variable $\mathrm{per}(A[-(I^* - i)])$ only depends on the randomness of ξ_i, and at most one of the possible outcomes of ξ_i will cause $\mathrm{per}(A[-(I^* - i)]) = z$ (here we are using that $\mathrm{per}(A[-I^*]) \neq 0$ as $I^* \in \mathcal{F}$). We need $\mathrm{per}(A[-(I^* - i)]) = z$ for all $i \in J^*$ for I^* not to be parental, and we recall our assumption $Q[\xi_i] \leq p$.

Combining (4.1) and (4.2) yields

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \,\Big|\, A^{\uparrow s}\Big] \leq 2p^{s-1} \cdot \mathbb{P}\Big[\mathrm{per}(A[-J^*]) = z \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big] \qquad (4.3)$$

for any outcome of $A^{\uparrow s}$ satisfying $\mathcal{E}(s, \alpha)$.

Note that this in particular implies $\mathbb{P}[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \,|\, A^{\uparrow s}] \leq 2p^{s-1}$. If $s \geq n/2$, then averaging this over the possible outcomes of $A^{\uparrow s}$ satisfying $\mathcal{E}(s, \alpha)$ already yields the bound

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \cap \mathcal{E}(s,\alpha)\Big] \leq \mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \,\Big|\, \mathcal{E}(s,\alpha)\Big] \leq 2p^{s-1} \leq 2p^{n/2-1} = \exp(-\Omega_p(n)).$$

Therefore for the rest of the proof we may assume $s < n/2$.

Now, note that if we reveal $A^{\uparrow s}$ and (I^*, J^*), then in the resulting conditional probability space, $\mathrm{per}(A[-J^*])$ is a linear function of $\xi_1, \dots, \xi_{n+T}$ (by Fact 4.2). For every $i \in \{1, \dots, n+T\} \setminus J^*$, the coefficient of ξ_i in this linear function is precisely $\mathrm{per}(A[-(J^* + i)])$. We will apply our relative anticoncentration result in Corollary 3.6 to this linear function (or, more precisely, to the linear function obtained by omitting the last T terms).

Step 2: Applying the relative anticoncentration inequality. Let $k = \lfloor \nu(n-s+1) \rfloor$, and let A^* be a random matrix obtained from $A[-J^*]$ as follows. First, for every entry ξ_i in the $(n+T-s+1)$-th row of $A[-J^*]$, replace this entry with $\xi_i - \xi_i'$, where ξ_i' is an independent copy of ξ_i. Then, choose a uniformly random subset $K \subseteq \{1, \dots, n\} \setminus J^*$ of exactly k of the first $n-s+1$ column indices of $A[-J^*]$, and set all the entries in the $(n-s+1)$-th row to zero except those indexed by K.

Claim 4.7. *For any fixed outcomes of $A^{\uparrow s}$ and $(I^*, J^*) \in \mathcal{H}$, we have*

$$\mathbb{P}\Big[\mathrm{per}(A[-J^*]) = z \,\Big|\, A^{\uparrow s}, (I^*, J^*)\Big] \leq C_p \cdot \sqrt{\nu} \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}, (I^*, J^*)\Big].$$

Proof. As in the claim statement, let us fix outcomes of $A^{\uparrow s}$ and $(I^*, J^*) \in \mathcal{H}$, and consider the conditional probability space obtained by conditioning on these outcomes. For a random variable Y, we write $Q[Y \,|\, A^{\uparrow s}, (I^*, J^*)] = \sup_y \mathbb{P}[Y = y \,|\, A^{\uparrow s}, (I^*, J^*)]$.

As $(I^*, J^*) \in \mathcal{H}$, there are at least $(\alpha/2)(n-s+1)$ different sets $I' \in \mathcal{F}$ with $(I', J^*) \in \mathcal{H}$. For all of these sets I' we have $J^* \subseteq I'$ and $\mathrm{per}(A[-I']) \neq 0$. In other words, there are at least $(\alpha/2)(n-s+1)$ indices $i \in \{1, \dots, n\} \setminus J^*$ such that $\mathrm{per}(A[-(J^* + i)]) \neq 0$.

Now, let us define $X_i = \mathrm{per}(A[-(J^*+i)]) \cdot \xi_i$ for all $i \in \{1,\dots,n+T\} \setminus J^*$. Here, the coefficient $\mathrm{per}(A[-(J^*+j)])$ is already determined by $A^{\uparrow s}$ and (I^*, J^*), so after revealing $A^{\uparrow s}$ and (I^*, J^*), the random variable X_i is a rescaling of ξ_i. Thus, the random variables X_i for $i \in \{1,\dots,n+T\} \setminus J^*$ are independent in the conditional conditional probability space obtained by revealing $A^{\uparrow s}$ and (I^*, J^*). Furthermore, by Fact 4.2 we have $\mathrm{per}(A[-J^*]) = \sum_{i\in\{1,\dots,n+T\}\setminus J^*} X_i$. We can now apply Corollary 3.6 with $\gamma = \alpha/2$ and the random variables X_i for $i \in \{1,\dots,n\} \setminus J^*$. At least $(\alpha/2)(n-s+1)$ of these random variables satisfy $Q[X_i \mid A^{\uparrow s}, (I^*, J^*)] \le p$, since this holds for all indices $i \in \{1,\dots,n\} \setminus J^*$ such that $\mathrm{per}(A[-(J^*+i)]) \neq 0$ (also noting that $k = \lfloor \nu(n-s+1) \rfloor \ge \lfloor \nu n/2 \rfloor \ge 4/\alpha$ if n is sufficiently large in terms of p, α and ν). This yields

$$\begin{aligned} Q\Big[\mathrm{per}(A[-J^*]) \,\Big|\, A^{\uparrow s}, (I^*, J^*)\Big] &= Q\Bigg[\sum_{i\in\{1,\dots,n+T\}\setminus J^*} X_i \,\Bigg|\, A^{\uparrow s}, (I^*, J^*)\Bigg] \\ &\le Q\Bigg[\sum_{i\in\{1,\dots,n\}\setminus J^*} X_i \,\Bigg|\, A^{\uparrow s}, (I^*, J^*)\Bigg] \\ &\le C_p \cdot \sqrt{\frac{k}{n-s+1}} \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}, (I^*, J^*)\Big]. \end{aligned}$$

Here, for the first inequality we used Fact 3.4, and for the second inequality we used Corollary 3.6, noting that $\mathrm{per}(A^*) = \sum_{i\in K} \mathrm{per}(A[-(J^*+i)])(\xi_i - \xi_i') = \sum_{i\in K}(X_i - X_i')$ if we define $X_i' = \mathrm{per}(A[-(J^*+i)]) \cdot \xi_i'$ for $i \in \{1,\dots,n\} \setminus J^*$ (then X_i' is an independent copy of X_i in this conditional probability space). Recalling $k \le \nu(n-s+1)$, we obtain the desired inequality. □

Averaging the inequality in Claim 4.7 over all possible outcomes of $(I^*, J^*) \in \mathcal{H}$ yields

$$\mathbb{P}\Big[\mathrm{per}(A[-J^*]) = z \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big] \le C_p \cdot \sqrt{\nu} \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big]. \tag{4.4}$$

Now, for every fixed outcome of $A^{\uparrow s}$ satisfying $\mathcal{E}(s,\alpha)$, recall that $|\mathcal{H}| \ge (\alpha/2)s\binom{n}{s}$, and that (I^*, J^*) is independent from $A^{\uparrow s}$. This implies

$$\mathbb{P}\Big[\mathcal{E}_{\mathcal{H}} \,\Big|\, A^{\uparrow s}\Big] = \mathbb{P}\Big[(I^*, J^*) \in \mathcal{H} \,\Big|\, A^{\uparrow s}\Big] = \frac{|\mathcal{H}|}{s\binom{n}{s}} \ge \alpha/2.$$

Therefore we obtain

$$\mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}\Big] \ge \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \text{ and } \mathcal{E}_{\mathcal{H}} \,\Big|\, A^{\uparrow s}\Big] \ge \frac{\alpha}{2} \cdot \mathbb{P}\Big[\mathrm{per}(A[-J^*]) = z \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big].$$

Combining this with (4.3) and (4.4) yields

$$\begin{aligned} \mathbb{P}\Big[\overline{\mathcal{E}_z(s-1, \alpha/(4s))} \,\Big|\, A^{\uparrow s}\Big] &\le 2C_p\sqrt{\nu}p^{s-1} \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}, \mathcal{E}_{\mathcal{H}}\Big] \\ &\le \frac{4C_p\sqrt{\nu}}{\alpha} p^{s-1} \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}\Big] \le \alpha p^s \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \,\Big|\, A^{\uparrow s}\Big] \end{aligned}$$

for every outcome of $A^{\uparrow s}$ satisfying $\mathcal{E}(s,\alpha)$, where for the last inequality we used $\sqrt{\nu} \le (\alpha^2/4) \cdot (p/C_p)$.

Summing over the possible outcomes of $A^{\uparrow s}$ satisfying $\mathcal{E}(s,\alpha)$, we deduce that

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1, \alpha/(4s))} \cap \mathcal{E}(s,\alpha)\Big] \le \alpha p^s \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0 \cap \mathcal{E}(s,\alpha)\Big] \le \alpha p^s \cdot \mathbb{P}\Big[\mathrm{per}(A^*) = 0\Big]. \tag{4.5}$$

Step 3: Recursion. Now, note that the permanent is unaffected by rearranging rows. Let A' be obtained from A^* by moving row $(n+T-s+1)$ (the row we modified to get A^* from $A[-J^*]$) directly after row T, to become the new $(T+1)$-th row. Note that the last T entries of the $(n+T-s+1)$-th row of A^* are always zero, hence the last T entries of the $(T+1)$-th row of A' are always zero (and also note that the $T \times T$ submatrix of A' in the first T rows and last T columns has a non-zero permanent, since this is precisely the submatrix in the last T columns of A_{fix}).

Suppose we reveal an outcome of J^* and an outcome of the $(T+1)$-th row of A' which is not all zero. Then, the first $T+1$ rows of A' contain a $(T+1) \times (T+1)$ submatrix with nonzero permanent (in the last T columns, together with any column for which there is a nonzero entry in the $(T+1)$-th row). So, conditional on revealed information, A' has the form in Definition 2.1, with $n-s$ random rows (and $T+1$ fixed rows), so the conditional probability that $\mathrm{per}(A') = 0$ is at most $f_p(n-s)$. This shows

$$\mathbb{P}\big[\mathrm{per}(A^*) = 0\big] = \mathbb{P}\big[\mathrm{per}(A') = 0\big] \le \mathbb{P}\big[\text{the } (T+1)\text{-th row of } A' \text{ is all-zero}\big] + f_p(n-s).$$

Now, note that the $(T+1)$-th row of A' is all-zero if and only if all entries of the $(n+T-s+1)$-th row of A^* are zero. The probability of this is at most $p^k \le p^{\lfloor \nu n/2 \rfloor} = \exp(-\Omega_{p,\alpha}(n))$ (since $Q[\xi_i - \xi_i'] \le Q[\xi_i] \le p$ for every $i \in K$). Thus, we obtain

$$\mathbb{P}\big[\operatorname{per}(A^*) = 0\big] \le f_p(n-s) + \exp(-\Omega_{p,\alpha}(n)),$$

and together with (4.5) we can conclude

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(s-1,\alpha/(4s))} \cap \mathcal{E}(s,\alpha)\Big] \le \alpha p^s \cdot \big(f_p(n-s) + \exp(-\Omega_{\alpha,p}(n))\big). \qquad \square$$

Now, we deduce Theorem 2.2.

Proof of Theorem 2.2. Recalling that the event $\mathcal{E}(n)$ holds deterministically, we can observe that

$$\mathbb{P}[\operatorname{per}(A) = z] = \mathbb{P}\Big[\overline{\mathcal{E}_z(0)}\Big] = \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(n)\Big] \le \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(t)\Big] + \sum_{s=t+1}^{n} \mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap \mathcal{E}(s)\Big]. \tag{4.6}$$

We can further break down the above terms, as follows. Let $\alpha = \alpha_t = \varepsilon/3$, and for all $1 \le s \le t$ let $\alpha_{s-1} = \alpha_s/(4s)$. Then for $s = t+1, \dots, n$, we have

$$\mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap \mathcal{E}(s)\Big] = \mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap (\mathcal{E}(s) \setminus \mathcal{E}(s,\alpha))\Big] + \mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap \mathcal{E}(s,\alpha)\Big], \tag{4.7}$$

and furthermore

$$\begin{aligned}\mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(t)\Big] &\le \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap (\mathcal{E}(t) \setminus \mathcal{E}(t,\alpha))\Big] + \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(t,\alpha)\Big] \\ &\le \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap (\mathcal{E}(t) \setminus \mathcal{E}(t,\alpha))\Big] + \mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(1,\alpha_1)\Big] + \sum_{s=2}^{t} \mathbb{P}\Big[\overline{\mathcal{E}(s-1,\alpha_{s-1})} \cap \mathcal{E}(s,\alpha_s)\Big] \end{aligned} \tag{4.8}$$

holds. We can directly substitute Lemma 4.5 into some of the terms in (4.7) and (4.8). To bound the other terms, let us write $\sup_{\mathcal{E}} \mathbb{P}[\mathcal{F} \mid A^{\uparrow s}]$ and $\inf_{\mathcal{E}} \mathbb{P}[\mathcal{F} \mid A^{\uparrow s}]$ to indicate a supremum or infimum over all outcomes of $A^{\uparrow s}$ satisfying an event $\mathcal{E}$. Then, by Lemmas 4.3 and 4.4, for $s = t+1, \dots, n$ we obtain

$$\begin{aligned}\mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap (\mathcal{E}(s) \setminus \mathcal{E}(s,\alpha))\Big] &\le \mathbb{P}\Big[\overline{\mathcal{E}(s,\alpha)}\Big] \cdot \mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \,\Big|\, \mathcal{E}(s) \setminus \mathcal{E}(s,\alpha)\Big] \\ &\le \mathbb{P}\Big[\overline{\mathcal{E}(s,\alpha)}\Big] \cdot \sup_{\mathcal{E}(s)} \mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \,\Big|\, A^{\uparrow s}\Big] \le (1+2\alpha) f_p(n-s) \cdot p^s, \end{aligned} \tag{4.9}$$

as well as (recalling the definition of τ_p in (2.1))

$$\begin{aligned}\mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap (\mathcal{E}(t) \setminus \mathcal{E}(t,\alpha))\Big] &\le \mathbb{P}\Big[\overline{\mathcal{E}(t,\alpha)}\Big] \cdot \left(1 - \mathbb{P}\Big[\mathcal{E}_z(0) \,\Big|\, \mathcal{E}(t) \setminus \mathcal{E}(t,\alpha)\Big]\right) \\ &\le \mathbb{P}\Big[\overline{\mathcal{E}(t,\alpha)}\Big] \cdot \left(1 - \inf_{\mathcal{E}(1)} \mathbb{P}\Big[\mathcal{E}_z(0) \,\Big|\, A^{\uparrow s}\Big] \cdot \prod_{s=2}^{t} \inf_{\mathcal{E}(s)} \mathbb{P}\Big[\mathcal{E}(s-1) \,\Big|\, A^{\uparrow s}\Big]\right) \\ &\le (1+2\alpha) f_p(n-t) \cdot \left(1 - \prod_{s=1}^{t} (1-p^s)\right) \le (1+2\alpha)\tau_p \cdot f_p(n-t). \end{aligned} \tag{4.10}$$

Substituting Lemma 4.5 and (4.9) into (4.7), for $s = t+1, \dots, n$, we obtain

$$\mathbb{P}\Big[\overline{\mathcal{E}(s-1)} \cap \mathcal{E}(s)\Big] \le (1+3\alpha) p^s f_p(n-s) + \exp(-\Omega_{\varepsilon,p}(n)),$$

and substituting Lemma 4.5 and (4.10) into (4.8), we obtain

$$\mathbb{P}\Big[\overline{\mathcal{E}_z(0)} \cap \mathcal{E}(t)\Big] \le (1+2\alpha)\tau_p \cdot f_p(n-t) + \sum_{s=1}^{t} \alpha p^s f_p(n-s) + \exp(-\Omega_{\varepsilon,p,t}(n)).$$

Substituting the above two estimates into (4.6) yields

$$\mathbb{P}[\operatorname{per}(A) = z] \le (1+2\alpha)\tau_p \cdot f_p(n-t) + \alpha \sum_{s=1}^{t} p^s f_p(n-s) + (1+3\alpha) \sum_{s=t+1}^{n} p^s f_p(n-s) + \exp(-\Omega_{\varepsilon,p,t}(n)),$$

as desired (recalling $\alpha = \varepsilon/3$). $\square$

References

[1] S. Aaronson and A. Arkhipov, *The computational complexity of linear optics*, Theory Comput. **9** (2013), 143–252.

[2] S. Arora and B. Barak, *Computational complexity: A modern approach*, Cambridge University Press, April 2009.

[3] J. Bourgain, V. H. Vu, and P. M. Wood, *On the singularity probability of discrete random matrices*, J. Funct. Anal. **258** (2010), no. 2, 559–603.

[4] R. A. Brualdi and M. Newman, *Some theorems on the permanent*, J. Res. Nat. Bur. Standards Sect. B **69B** (1965), 159–163.

[5] M. Campos, M. Jenssen, M. Michelen, and J. Sahasrabudhe, *The singularity probability of a random symmetric matrix is exponentially small*, J. Amer. Math. Soc. **38** (2025), no. 1, 179–224.

[6] P. Erdös, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51** (1945), 898–902.

[7] A. E. Guterman and K. A. Taranin, *On the values of the permanent of* $(0,1)$*-matrices*, Linear Algebra Appl. **552** (2018), 256–276.

[8] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Period. Math. Hungar. **8** (1977), no. 3-4, 197–211.

[9] D. Ingram and A. Razborov, *On the range of the permanent of* (± 1)*-matrices*, arXiv preprint arXiv:2507.09433.

[10] J. Kahn, J. Komlós, and E. Szemerédi, *On the probability that a random* ± 1*-matrix is singular*, J. Amer. Math. Soc. **8** (1995), no. 1, 223–240.

[11] H. Kesten, *A sharper form of the Doeblin–Lévy–Kolmogorov–Rogozin inequality for concentration functions*, Math. Scand. **25** (1969), 133–144.

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

[13] J. Komlós, *On the determinant of random matrices*, Studia Sci. Math. Hungar. **3** (1968), 387–399.

[14] A. R. Kräuter, *Permanents of* $(1,-1)$*-matrices*, Proceedings of the Topology and Geometry Research Center, vol. 4, Kyungpook National University, 1993, pp. 151–204.

[15] M. Kwan and L. Sauermann, *On the permanent of a random symmetric matrix*, Selecta Math. (N.S.) **28** (2022), no. 1, Paper No. 15, 29 pp.

[16] M. B. Marcus and J. Rosen, *Asymptotic properties of permanental sequences—related to birth and death processes and autoregressive Gaussian sequences*, SpringerBriefs in Probability and Mathematical Statistics, Springer, Cham, 2021.

[17] H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, vol. 6, Addison-Wesley Publishing Co., Reading, MA, 1978, With a foreword by Marvin Marcus.

[18] L. P. Postnikova and A. A. Judin, *An analytic method for estimates of the concentration function*, Trudy Mat. Inst. Steklov. **143** (1977), 143–151, 210.

[19] B. A. Rogozin, *On the increase of dispersion of sums of independent random variables*, Teor. Verojatnost. i Primenen. **6** (1961), 106–108.

[20] S. Scheel, *Permanents in linear optical networks*, arXiv preprint arXiv:quant-ph/0406127.

[21] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006.

[22] T. Tao and V. Vu, *On random* ± 1 *matrices: singularity and determinant*, Random Structures Algorithms **28** (2006), no. 1, 1–23.

[23] T. Tao and V. Vu, *On the singularity probability of random Bernoulli matrices*, J. Amer. Math. Soc. **20** (2007), no. 3, 603–628.

[24] T. Tao and V. Vu, *On the permanent of a random Bernoulli matrix*, Unpublished note, appearing on the blog of Terence Tao https://terrytao.wordpress.com/2008/04/16/on-the-permanent-of-a-random-bernoulli-matrix/. Available directly at https://terrytao.wordpress.com/wp-content/uploads/2008/04/permanent.pdf, 2008.

[25] T. Tao and V. Vu, *On the permanent of random Bernoulli matrices*, Adv. Math. **220** (2009), no. 3, 657–669.

[26] K. Tikhomirov, *Singularity of random Bernoulli matrices*, Ann. of Math. (2) **191** (2020), no. 2, 593–634.

[27] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, second ed., Cambridge University Press, Cambridge, 2001.

[28] V. Vu, *Random matrices: a survey*, slides for a presentation at an IPAM workshop, http://campuspress.yale.edu/vanvu/files/2013/01/ipam09.pdf, 2009.

[29] V. H. Vu, *Recent progress in combinatorial random matrix theory*, Probab. Surv. **18** (2021), 179–200.

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

Institute of Science and Technology Austria (ISTA).
Email address: matthew.kwan@ist.ac.at

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