On the Rank of Random Matrices

C. Cooper*
School of Mathematical Sciences, University of North London, London N7 8DB, UK

Received 11 March 1998; accepted 25 June 1999

ABSTRACT: Let $M = (m_{ij})$ be a random $n \times n$ matrix over $GF(2)$. Each matrix entry m_{ij} is independently and identically distributed, with $\Pr(m_{ij} = 0) = 1 - p(n)$ and $\Pr(m_{ij} = 1) = p(n)$. The probability that the matrix M is nonsingular tends to $c_2 \approx 0.28879$ provided $\min(p, 1 - p) \geq (\log n + d(n))/n$ for any $d(n) \to \infty$. Sharp thresholds are also obtained for constant $d(n)$. This answers a question posed in a paper by J. Blömer, R. Karp, and E. Welzl (Random Struct Alg, 10(4) (1997)). © 2000 John Wiley & Sons, Inc. Random Struct. Alg., 16, 209–232, 2000

1. INTRODUCTION

We consider the following model of random $(n \times n)$ matrices over $GF(t)$. Let $M = (m_{ij})$ be a $(n \times n)$ matrix with entries, r, in $GF(t)$, independently and identically distributed as

$$\Pr(m_{ij} = r) = \begin{cases} 1 - p, & r = 0, \\ \dfrac{p}{t-1}, & r \neq 0. \end{cases}$$

We denote the space of these matrices by $\mathbf{M}(n, p; t)$

* Research supported by the STORM Research Group

In their paper, Blömer, Karp, and Welzl, [1] posed the following question about $\mathbf{M}(n, p; t)$ at the end of Section 6: Is there a function $p(n)$ that tends to 0 as n goes to infinity and a constant $c > 0$ such that a random $(n \times n)$ matrix over $GF(2)$, where each matrix entry is 0 with probability $1 - p(n)$ and 1 with probability $p(n)$, is nonsingular with probability at least c?

In fact, provided $p(n)$ does not tend to either 0 or 1 too rapidly, the asymptotic probability that such a matrix is nonsingular is given by $c \approx 0.28879$. This value of c is the limit of the right-hand side of (1) below, in the case where $t = 2$.

We will require $p(n)$ to satisfy $\min(p(n), 1 - p(n)) \geq (\log n + d(n))/n$, where $d(n) \to \infty$ arbitrarily slowly. For, if $p(n) \geq (\log n + d(n))/n$, where $d(n) \to \infty$, with probability tending to 1, there are no rows or columns of the matrix that are identically 0. Thus any linear dependencies are nontrivial. However, if $p \to 0$ in such a way that $d(n)$ is constant or $d(n) \to -\infty$, then the matrix may have rows or columns that are identically 0. Similarly if $p \to 1$ too rapidly, then the matrix may have two rows or columns consisting entirely of 1s. The value c given above extends to the case where $d(n) \to -\infty$ slowly, provided we restrict our attention to the subset of random matrices that avoid these two types of linear dependency.

Consider first the special case in $GF(\mathrm{t})$ when $p = (t - 1)/t$, and thus all values and vectors are equiprobable. If the first s columns of the matrix M are linearly independent, they span a vector space of dimension s and size t^s. The probability that the next column avoids this space is $(1 - t^s/t^n)$, and thus

$$\Pr(M \text{ is nonsingular}) = \prod_{s=1}^{n} \left(1 - \frac{1}{t^s}\right). \tag{1}$$

This result is a special case of a theorem (given, e.g., in [KLS] p 33) for the limiting probability of rank of a square matrix in the equiprobable model.

Theorem 1. *Let $p = (t - 1)/t$ and let M be a random $(n \times n)$ matrix with entries in $GF(t)$. Let $p_n(k, t)$ be the probability that* $\operatorname{rank}(M) = n - k$. *Then*

$$\lim_{n\to\infty} p_n(k, t) = \pi(k, t) = \begin{cases} \displaystyle\prod_{j=1}^{\infty} \left(1 - \left(\frac{1}{t}\right)^j\right), & k = 0, \\[2ex] \displaystyle\frac{\prod_{j=k+1}^{\infty}(1 - (1/t)^j)}{\prod_{j=1}^{k}(1 - (1/t)^j)} \frac{1}{(t)}^{k^2}, & k \geq 1. \end{cases} \tag{2}$$

The probabilities $\pi(k, t)$ are $\Theta(t^{-k^2})$ and tend to 0 very rapidly. Some values are given in Table 1. Any value not tabulated is less than 1×10^{-10}.

The precise number, $\eta_n(k, t)$, of $(n \times n)$ matrices of rank $(n - k)$, is a known quantity, and is given by (15) of Section 5. Thus $\pi(k, t)$ can also be obtained as $\pi(k, t) = \lim_{n\to\infty} \eta_n(k, t)/t^{n^2}$.

For $GF(2)$, $c_2 = \pi(0, 2)$ is the limiting probability that M is nonsingular for a wide range of p, as the following theorem shows.

TABLE 1

k	$\pi(k,2)$	$\pi(k,3)$	$\pi(k,5)$	$\pi(k,7)$
0	0.2887880951	0.5601260779	0.7603327959	0.8367954017
1	0.5775761902	0.4200945584	0.2376039987	0.1627102180
2	0.1283502645	0.0196919324	0.0020625347	0.0004943453
3	0.0052387863	0.0000873902	0.0000006707	0.0000000296
4	0.0000465670	0.0000000410		
5	0.0000000969			

Theorem 2. *Let $M \in \mathbf{M}(n, p;\, 2)$ be a random matrix over $GF(2)$.*

(i) *If $p = (\log n + d(n))/n \le 1/2$, then*

$$\lim_{n\to\infty} \Pr(M \text{ is nonsingular}) = \begin{cases} 0, & d(n) \to -\infty, \\ c_2 \exp(-2e^{-d}), & d(n) \to d \text{ constant}, \\ c_2, & d(n) \to \infty. \end{cases} \tag{3}$$

(ii) *If $p = 1 - (\log n + d(n))/n \ge 1/2$, then*

$$\lim_{n\to\infty} \Pr(M \text{ is nonsingular}) = \begin{cases} 0, & d(n) \to -\infty, \\ c_2 \exp(-2e^{-d})(1+e^{-d})^2, & d(n) \to d \text{ constant}, \\ c_2, & d(n) \to \infty. \end{cases} \tag{4}$$

(iii) *Let a unity row of a matrix, be a row consisting entirely of 1s. Let $\mathcal{F}$ denote the event that M has no zero rows or columns and at most one unity row and column. If $(\log n - \omega)/n \le p \le 1 - (\log n - \omega)/n$, where $\omega = o(\log\log n)$, then for any finite nonnegative integer k,*

$$\lim_{n\to\infty} \Pr(M \text{ has rank } n-k \mid \mathcal{F}) = \pi(k,2) \tag{5}$$

and, in particular,

$$\lim_{n\to\infty} \Pr(M \text{ is nonsingular} \mid \mathcal{F}) = c_2. \tag{6}$$

The case in Theorem 2(iii) given by (6) extends the results of Theorem 2(i) and (ii) slightly.

Our approach is to count the number, W, of sequences of columns $(M_{i_1}, \ldots, M_{i_s})$ of M such that $i_1 < i_2 < \cdots < i_s$ and

$$c_1 M_{i_1} + \cdots + c_s M_{i_s} = 0,$$

where c_i, $i = 1, \ldots, s$, are nonzero elements of $GF(t)$. In a vector space over $GF(t)$ a linear dependency $c_1\mathbf{a} + c_2\mathbf{b} = 0$ has $(t-2)$ *associated* dependencies $\alpha(c_1\mathbf{a} + c_2\mathbf{b}) = 0$ for $\alpha \in GF(t) \setminus \{0,1\}$. In our estimate $\mathbf{E}W$ of the expected number of linear dependencies in Section 3, we divide our result by $1/(t-1)$ to count only nonassociated dependencies.

Theorem 3. *Let* $t \geq 3$ *and* $t = o(\sqrt{\log n})$ *be a prime power. Let* $M \in \mathbf{M}(n, p;\ t)$ *be a random matrix over* $GF(t)$. *Let* $p = (\log n + \log(t-1) + \omega)/n$, *where* $\omega \to \infty$. *Then the expected number of nonassociated linear dependencies in the columns of* M *tends to* $1/(t-1)$ *as* $n \to \infty$.

A subsequence $A = (M_{i_1}, M_{i_2}, \ldots, M_{i_s})$ of the columns of M is uniquely described by its index set $I_A = \{i_1, i_2, \ldots, i_s\}$. We will write $A = (a_1, a_2, \ldots, a_s)$, where $a_j = M_{i_j}$. If there is no ambiguity, we do not distinguish between A and I_A and shall speak of A as a *column set*. We reserve the notation $A_1, A_2, \ldots, A_j, \ldots$ to refer to sets $1, 2, \ldots, j, \ldots$ of columns.

Suppose the rank of M is $n - k$, so that the *defect* is k (following the notation of [1]). We can partition the columns of M into $(M(n-k), M(k))$. Here $M(n-k)$ is an $(n \times (n-k))$ matrix of full column rank, $(n-k)$, and the columns of $M(k)$ lie in the column space of $M(n-k)$. Thus there are $t^k - 1$ (not necessarily distinct) linear dependencies in the columns of M induced by nontrivial linear combinations of the columns of $M(k)$.

We can also regard M as a sequence of columns $(M_1, \ldots, M_n)$ which we reveal one at a time. Suppose that the defective columns have indices $i_1, \ldots, i_k$. We can define unique *smallest* linearly dependent column sets $A_1, \ldots, A_k$, where the final column of A_j is i_j. If the columns of A_j are $(M_{s_1}, \ldots, M_{i_j})$, then $(i_j - s_1)$ is the smallest of all index differences of linearly dependent sequences of columns with final column M_{i_j}.

Let $A \,\Delta\, B$ be the symmetric set difference $A \cup B - A \cap B$ of the sets A and B. In the case of column sets A, B we shall write $A \,\Delta\, B$ to mean the set of columns with index set $I_A \,\Delta\, I_B$. If the defect of M is k and the field is $GF(2)$, the smallest column sets $A_1, \ldots, A_k$ generate $2^k - (k+1)$ other linearly dependent sets $A_1 \,\Delta\, A_2$, $A_1 \,\Delta\, A_3, \ldots, A_1 \,\Delta\, A_2 \,\Delta\, \cdots \,\Delta\, A_k$.

For reasons given below, we definitely do not want to count all $2^k - 1$ dependent sets arising from a defect of k. However, we cannot quite count just the smallest dependencies $A_1, \ldots, A_k$ either. We settle for a compromise and count *simple* sequences of linearly dependent column sets $\mathbf{B} = (B_1, \ldots, B_k)$. (Kolchin [4] also recognized the value of simple sequences. He called such sequences *independent*, perhaps a better notation. We retain the notation *simple* to make a distinction from general discussions of linear independence.) A k-tuple of sets $\mathbf{B} = (B_1, \ldots, B_k)$ is simple, if no set B_i in $\mathbf{B}$ is a set difference of other sets B_j in $\mathbf{B}$. In other words, events

$$B_{j_1} \,\Delta\, B_{j_2} \,\Delta\, \cdots \,\Delta\, B_{j_l} = \varnothing \qquad (j_1 < j_2 < \cdots < j_l; 1 \leq l \leq k)$$

do not occur.

Let $V(M) = \{\varnothing\} \cup \{B : B \text{ is zero sum in } M\}$, then $(V(M), \Delta)$ is a vector space over $GF(2)$ under the convention that $0 \cdot B = \varnothing$, $1 \cdot B = B$. In $V(M)$ a simple sequence $(B_1, \ldots, B_k)$ is an ordered basis of dimension k.

If W is the number of linear dependencies in M, we calculate the expected number of simple l-sequences, for $l \geq 1$. Denote this by $\mathbf{E}(W, l; \text{simple})$. We will show that in $GF(2)$, $\mathbf{E}(W, l; \text{simple}) \sim 1$. Of course the restricted expectation $\mathbf{E}(W, l; \text{simple})$ is much less than the usual (unrestricted) lth factorial moment $\mathbf{E}(W)_l$ of W, but that is just as well. As proved in [6, pp. 34–35], the limiting probabilities $\{\pi(k,t)\}$ do not satisfy the Carleman condition (see, e.g. [3]) to recover the distribution $\{\pi(k,t)\}$ from the moments $\mathbf{E}W^l$, where W is the number of linear dependencies in M.

For simplicity of notation, we denote $\mathbf{E}(W, l; \text{simple})$ by $\mathbf{E}(W)_l$ throughout this paper.

Let $A = (a_1, \ldots, a_m)$ be a subsequence of the columns $(M_1, \ldots, M_n)$ of M and let $\mathbf{c} = (c_1, \ldots, c_m)$ be a sequence of *nonzero* coefficients from $GF(t)$. The matrix M is singular if and only if there exist $\mathbf{c}$ and A such that

$$c_1 a_1 + c_2 a_2 + \cdots + c_m a_m = 0, \qquad c_i \neq 0, i = 1, \ldots, m. \tag{7}$$

Given a subsequence A of M, we count $\sum_{\mathbf{c}} 1_{(\mathbf{c}A=0)}$, where 1_X is the indicator for the event X and where $\mathbf{c} \in (GF(t) \setminus \{0\})^m$. Because $GF(t)$ is a field, and by definition any nonzero element has probability $p/(t-1)$, for *fixed* nonzero c_i, $i = 1, \ldots, m$, the event given in (7) has the same probability as the event, $a_1 + \cdots + a_m = 0$, that the set of columns A is *zero sum*.

The probability $\rho_m(r)$ that row j of A has sum $a_{j1} + \cdots + a_{jm} = r$ $(r \in GF(t))$ is

$$\rho_m(r) = \begin{cases} \dfrac{1}{t}\left(1 + (t-1)\left(1 - \dfrac{t}{t-1}p\right)^m\right), & r = 0, \\[2ex] \dfrac{1}{t}\left(1 + (-1)\left(1 - \dfrac{t}{t-1}p\right)^m\right), & r \neq 0. \end{cases} \tag{8}$$

The results (8) are derived by recurrence relations (see, e.g., [1]).

Let $\rho_m(0) = \rho_m$. The probability that a subsequence $A = (a_1, \ldots, a_m)$ of the columns of M satisfies $a_1 + \cdots + a_m = 0$ is ρ_m^n. The number W_m of nonassociated linearly dependent m subsets has expectation

$$\mathbf{E}W_m = \frac{1}{t-1}\binom{n}{m}(t-1)^m \rho_m^n.$$

2. THE STRUCTURE OF THE PROOF OF THEOREM 2

We prove Theorem 2(i) and (ii) in three parts corresponding to the three cases listed in (3) or (4). The case where $d(n) \to -\infty$ is more or less trivial as zero rows (resp. pairs of unity rows) occur with probability tending to 1 as $p \to 0$ (resp. $p \to 1$), and this case is not treated in any detail. The case treated in Sections 3 and 4 is for the range $\min(p, 1-p) \geq (\log n + d(n))/n$, where $d(n) \to \infty$. The case of Theorem 2(i) and (ii), where $d(n) = d$, constant (the sharp threshold), follows directly from the proof of Theorem 2(iii) given in Sections 6 and 7.

In Section 3 we calculate the expected number of linear dependencies $\mathbf{E}W$ in the columns of M over $GF(t)$. We show that, provided $d(n) \to \infty$, whp (with high probability: with probability tending to 1 as $n \to \infty$) any linear dependencies in the columns of M must be of size about $n(t-1)/t$, and we condition on this event in all subsequent calculations.

In Section 4 we focus on $GF(2)$. We calculate the expected number $\mathbf{E}(W)_k = \mathbf{E}(W, k; \text{simple})$ of *simple* k-tuples of linearly dependent column sets $(A_1, A_2, \ldots,$

A_k), where each $|A_i| \sim n/2$. We show that $\mathbf{E}(W)_k \sim 1$. This result holds for $k \geq 1$ and provided k does not tend to infinity too quickly. The result is formally stated in Lemma 6, but the entire section is devoted to the proof.

The proof that $\mathbf{E}(W)_k \sim 1$ requires a technical lemma, Lemma 7, to bound the error terms. The most important part of Lemma 7 is part (iii), which deals with the cancellation of these error terms. To maintain continuity of exposition, the proof of Lemma 7 is given in the Appendix.

In Section 5 we obtain the limiting probability distribution $(\pi(j,2), j \geq 0)$ of the defect of M from the moments $\mathbf{E}(W)_k$. The values of $\pi(j,2)$ are given by Theorem 1. The preliminary discussion is true for a general finite field $GF(t)$. This discussion includes a proof of Theorem 1. We then restrict our attention to $GF(2)$ to complete the proof of Theorem 2. The only detail required for convergence to the distribution $(\pi(j,2), j \geq 0)$ is that $\mathbf{E}(W)_k \sim 1$, so that proving this result in Sections 4, 6, and 7 establishes the required convergence for the various cases of Theorem 2.

In working with $\mathbf{E}(W)_k$, the lattice we count over is subspaces of a vector space, rather than subsets of a set. Thus the usual formula for recovering the probability distribution from the factorial moments is incorrect here. The main function of Section 5 is to draw together various standard results on Möbius inversion on a lattice of vector spaces in an asymptotic context.

Section 6 considers the proof of Theorem 2(i), (ii), and (iii) for $GF(2)$ in the case where $p = (\log n + d(n))/n$ and $|d(n)| = o(\log\log n)$ (the sharp threshold). When $|d|$ is constant or tends to infinity very slowly, the number of zero rows and columns occurring in M is asymptotically Poisson with parameter $2e^{-d}$. Fortunately, the correlation between any zero rows or columns and large ($m \sim n/2$) linear dependent sets of columns is asymptotically zero. Conditioning on the event that there are no zero rows or columns, we can show $\mathbf{E}(W)_k \sim 1$, and the results of Sections 5 hold.

Section 7 considers the sharp threshold for $GF(2)$ in the case where $1 - p = (\log n + d(n))/n$ and $|d(n)| = o(\log\log n)$. Because (8) is symmetric in p and $1 - p$ when $t = 2$ and m is even (see (23)), we can adapt the proofs of the previous section.

3. THE EXPECTED NUMBER OF LINEAR DEPENDENCIES

We calculate the expected number of linear dependencies $\mathbf{E}W$ in the columns of M, where

$$\mathbf{E}W = \sum_{m=1}^{n} \mathbf{E}W_m.$$

This calculation is for any $GF(t)$, $t = o(\sqrt{\log n})$ under the assumption that $p \geq p_0$, where $p_0 = (\log n + \log(t-1) + d(n))/n$. In the case of $GF(2)$ we further require that $p \leq 1 - p_0$. Provided $d(n) \to \infty$, the main contribution to $\mathbf{E}W$ is from $\mathbf{E}W_m$, where $m \sim n(t-1)/t$. Specifically, $\sum_{m \sim n(t-1)/t} \mathbf{E}W_m \to 1/(t-1)$ and $\sum_{m \not\sim n(t-1)/t} \mathbf{E}W_m \to 0$. The case $m \sim n(t-1)/t$ is straightforward. The "other m" analysis requires more care. To preserve continuity of exposition, the calculations are given in the Appendix.

Lemma 4. *Let the field be $GF(t)$, where $t = o(\sqrt{\log n})$. Let $p = c/n$, where $c = \log n + \log(t-1) + \omega$, and $\omega \to \infty$:*

$$\mathbf{E}W = \frac{1}{t-1}\big(1 + O(e^{-\omega})\big).$$

Lemma 5. *Let the field be $GF(t)$, where $t = o(\sqrt{\log n})$. Let $p = c/n$, where $c = \log n + \log(t-1) + d(n)$, where $d(n) > -o(\log\log n)$.*

- **(i)** *With probability $1 - o(1/n)$ no linearly dependent set, A, of columns of M has $\log n \le |A| \le m_2 = n(t-1)/t(1 - 4\sqrt{(\log n)/n})$ or $|A| \ge m_3 = n(t-1)/t(1 + 4\sqrt{(\log n)/n})$.*
- **(ii)** *If $d(n) = \omega \to \infty$, then with probability $1 - O(e^{-\omega})$ there are no linearly dependent sets A of size $1 \le |A| \le \log n$.*

4. FACTORIAL MOMENTS OF W IN $GF(2)$

We now restrict our attention to $GF(2)$ and $d(n) \to \infty$. We calculate $\mathbf{E}(W)_k$ over all simple k-tuples $(A_1, \dots, A_k)$ of zero-sum sets. Let $(W)_k = \sum 1_{(A_1,\dots,A_k)}$, where $1_{(A_1,\dots,A_k)}$ is the indicator for the event that the column sets $(A_j : j = 1, \dots, k)$ of M are zero sum.

Say A is *small* if $|A| < m_2$, where $m_2 = n(t-1)/t(1 - 4\sqrt{(\log n)/n})$. Let $\mathcal{N}$ be the property that no zero-sum set is small. By Lemma 5, $\mathcal{N}$ occurs with probability $1 - O(e^{-\omega})$. In this section we prove the following result.

Lemma 6. $\mathbf{E}(W)_k = 1 + (2^k - 1)e^{-\omega}(1 + o(1))$. *The counting is over all simple k-tuples and restricted to $\mathcal{N}$.*

This result is proved for $k = 1$ in (31) of the Appendix.

To calculate the kth factorial moment of W, we dissect any k tuple of zero-sum sets $A_1, \dots, A_k$ into disjoint sets I_j, $j = 1 \cdots L$, and express each A_i as the union of a suitable subset of the I_j, indexed by a set $C(A_i) \subset \{1, \dots, L\}$. The parity of A_i depends on the parity of the selected subsets I_j. The parities of the subsets, I_j which allow $A_1, \dots, A_k$ to be simultaneously zero sum, is described by an equation $\mathbf{Ay} = \mathbf{0}$.

As a motivating example, consider two zero-sum sets A_1 and A_2. This defines three (index) sets $I_1 = A_1 \cap \overline{A}_2$, $I_2 = A_1 \cap A_2$, and $I_3 = \overline{A}_1 \cap A_2$. The parity restrictions on the subsets, I_j, given by $\mathbf{Ay} = \mathbf{0}$, have vector $\mathbf{y} = (y_1, y_2, y_3)$ for the parities of (I_1, I_2, I_3) and matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Thus, in *each* row of M, the only possible simultaneous parities of the sum of the entries in these subsets are

$A_1 \cap \overline{A}_2$	$A_1 \cap A_2$	$\overline{A}_1 \cap A_2$
0	0	0
1	1	1

Suppose, now, we have zero-sum column sets $A_1, A_2, \ldots, A_k$. To calculate factorial moments, we need to consider the $L = 2^k - 1$ subsets arising from the intersection of the sets A_j. The general intersection is $I = X_1 X_2 \cdots X_k$, where $X_j \in \{A_j, \overline{A}_j\}$ and where XY denotes $X \cap Y$. We write these intersections (in some order) as $I_1, \ldots, I_L$ and let $I_{L+1} = \overline{A}_1 \overline{A}_2 \cdots \overline{A}_k$.

We fix our attention on row s of M. Let

$$f(I) = \sum_{i \in I} m_{si} \pmod 2$$

be the total (mod 2) in row s of entries m_{si} of M with columns idexed by the subset I. For $A_1, \ldots, A_k$ to be simultaneously zero sum in row s of M, we have k equations $\mathrm{EQ}_1(A_1), \ldots, \mathrm{EQ}_k(A_k)$ restricting the parity (in row s) of the sums $f(I)$ on the sets $I_1, \ldots, I_L$. Thus, for example, A_1 is dissected into 2^{k-1} subsets $A_1 X_2 \cdots X_k$ and $\mathrm{EQ}_1(A_1)$ is

$$f(A_1 \overline{A}_2 \cdots \overline{A}_k) + f(A_1 A_2 \overline{A}_3 \cdots \overline{A}_k) + \cdots + f(A_1 A_2 \cdots A_k) = 0 \pmod 2.$$

In general, EQ_i is $\sum_{I_j \subset A_i} f(I_j) = 0 \pmod 2$. We will record which sets A_i $(i = 1, \ldots, k)$ occur uncomplemented in the subsets I_j $(j = 1, \ldots, L)$ as a $(k \times L)$ matrix $\mathbf{A} = (a_{ij})$. The ith row of $\mathbf{A}$ corresponds to A_i, and

$$a_{ij} = \begin{cases} 0, & X_i = \overline{A}_i \text{ in } I_j, \\ 1, & X_i = A_i \text{ in } I_j. \end{cases}$$

The terms $y = f(I)$ of EQ_i are evaluated on the nonzero entries a_{ij} in row i of $\mathbf{A}$. We write the equations $\mathrm{EQ}_1, \ldots, \mathrm{EQ}_k$ as $\mathbf{Ay} = \mathbf{0}$. Here $\mathbf{y} = (y_1, \ldots, y_L)$ is a column vector over $GF(2)$. We give the proof of the following lemma in the Appendix.

Lemma 7. *Let $(A_1, A_2, \ldots, A_k)$ be a simple k-tuple of zero-sum sets.*

(i) *At most $L - k$ sets I can be empty.*

(ii) *Let $T = \{\mathbf{y} : \mathbf{Ay} = \mathbf{0}\}$. If $0 \le i \le L - k$ sets I are empty, then $|T| = 2^{L-k-i}$.*

(iii) *Let $J = \{j_1, \ldots, j_s\}$ index a subset of the columns of $\mathbf{A}$. Let*

$$\lambda(J) = \sum_{\mathbf{y} \in T} (-1)^{y_{j_1} + \cdots + y_{j_s}}.$$

There is a set S consisting of $|S| = (2^k - 1)$ index sets $C \subset \{1, \ldots, L\}$ such that

$$\lambda(J) = \begin{cases} |T|, & \text{if } J = C \in S, \\ 0, & \text{otherwise}. \end{cases}$$

(iv) *For $C \in S$ let the sets I_j indexed by $C = \{j_1, \ldots, j_N\}$ satisfy*

$$I_{j_1} \cup \cdots \cup I_{j_N} = B_C.$$

There are sets $A_{i_1}, \ldots, A_{i_l}$ indexed by $\{i_1, \ldots, i_l\} \subset \{1, \ldots, k\}$ such that $C = C(i_1, \ldots, i_l)$ and

$$B_C = A_{i_1} \Delta \cdots \Delta A_{i_l},$$

and thus the set B_C is zero sum.

Basically, this lemma says the following: Let $\mathbf{e} = (e_j : j = 1, \ldots, L)$ be a vector and let $J = \{j : e_j = 1\}$. The only J with nonzero $\lambda(J)$ are those with $\mathbf{e}(J)$ in the row space of $\mathbf{A}$. This is because such an $\mathbf{e}(J)$ describes a linear combination of the zero-sum sets $A_1, \ldots, A_k$ so that $y_{j_1} + \cdots + y_{j_s} \equiv 0 \pmod 2$.

Let x_j denote the number of columns of M in I_j. Consider the sth row of M. Provided no $I = \varnothing$, from Eq. (8) we find

$$\begin{aligned}\pi(x_1, \ldots, x_L) &= \Pr\big(A_1, A_2, \ldots, A_k \text{ are zero sum in the } s\text{th row of } M\big)\\ &= \sum_{(y_1,\ldots,y_L)\in T} \prod_{j=1}^{L} \left\{ \tfrac{1}{2}\left(1 + (-1)^{y_j}(1-2p)^{x_j}\right)\right\}\\ &= \frac{1}{2^L}\left(|T| + \sum_{i=1}^{L} \sum_{J=\{j_1,\ldots,j_i\}} (1-2p)^{x_{j_1}+\cdots+x_{j_i}} \sum_{\mathbf{y}\in T} (-1)^{y_{j_1}+\cdots+y_{j_i}}\right),\end{aligned}$$

where $|T| = 2^{L-k}$. Thus, by Lemma 7(iii),

$$\pi = \frac{1}{2^k}\left(1 + \sum_{C\in S} (1-2p)^{x_{j_1}+\cdots+x_{j_N}}\right). \tag{9}$$

Let $B = I_{j_1} \cup \cdots \cup I_{j_N}$. If $|B| \geq m_2$ then $x_{j_1} + \cdots + x_{j_N} \geq n/2(1 - 4\sqrt{\log n/n})$ and

$$(1-2p)^{x_{j_1}+\cdots+x_{j_N}} = \frac{e^{-\omega}}{n}(1 + o(1)). \tag{10}$$

However, by Lemma 7(iv), if $C = \{j_1, \ldots, j_N\} \in S$, then $B = B_C$ is zero sum, and hence $|B| \geq m_2$ as we have conditioned on $\mathcal{N}$. Thus as $|S| = 2^k - 1$ and $L + 1 = 2^k$,

$$\begin{aligned}\mathbf{E}(W)_k &= \sum \binom{n}{x_1, \ldots, x_{L+1}} \big(\pi(x_1, \ldots, x_L)\big)^n\\ &= \big(1 + (2^k - 1)e^{-\omega}(1 + o(1))\big) \sum \binom{n}{x_1, \ldots, x_{L+1}} \frac{1}{2^{kn}}\\ &= 1 + (2^k - 1)e^{-\omega}(1 + o(1))\end{aligned}$$

as required. The multinomial sum was approximated in a manner similar to (29).

If i of the sets I are empty, then

$$\pi(x_1, \ldots, x_{L-i}) = \frac{1}{2^{L-i}}\left(|T| + \sum_{s=1}^{L-i} \sum_{J=\{j_1,\ldots,j_s\}} (1-2p)^{x_{j_1}+\cdots+x_{j_s}} \sum_{\mathbf{y}\in T} (-1)^{y_{j_1}+\cdots+y_{j_s}}\right),$$

where all the sets indexed by J are nonempty. However, now $|T| = 2^{L-k-i}$ so that

$$\pi = \frac{1}{2^k}\left(1 + \sum_{C\in S}(1-2p)^{x_{j_1}+\cdots+x_{j_N}}\right)$$

as before. Possibly some of the $x_j = 0$, but $x_{j_1} + \cdots + x_{j_N} \geq m_2$ and $|S| = 2^k - 1$. Thus

$$\mathbf{E}(W)_k \sim \left(\frac{2^k - i}{2^k}\right)^n = o(1).$$

5. LIMITING PROBABILITY DISTRIBUTION OF MATRIX DEFECT

The initial discussion in this section is true for general $GF(t)$ and concerns the distribution $(\pi(r,t),\ r \geq 0)$ given in (2). The t-nomial theorem (see [2, p. 125], [5, p. 78], or [7, p. 291] for more details) states that, for $r \geq 1$,

$$(1+x)(1+tx)\cdots(1+t^{r-1}x) = \sum_{k=0}^{r}\begin{bmatrix} r \\ k \end{bmatrix}_t t^{\binom{k}{2}}x^k, \tag{11}$$

where the Gaussian coefficients are defined, for $t > 0$ by $\begin{bmatrix} r \\ 0 \end{bmatrix}_t = 1$ and

$$\begin{bmatrix} r \\ k \end{bmatrix}_t = \frac{(t^r-1)(t^{r-1}-1)\cdots(t^{r-k+1}-1)}{(t^k-1)(t^{k-1}-1)\cdots(t-1)}.$$

If we define

$$\begin{bmatrix} \infty \\ k \end{bmatrix}_z = \frac{1}{(1-z^k)\cdots(1-z)}, \qquad |z| < 1, k \geq 1,$$

then

$$\prod_{r=0}^{\infty}(1+z^r x) = \sum_{k=0}^{\infty}\begin{bmatrix} \infty \\ k \end{bmatrix}_z z^{\binom{k}{2}}x^k. \tag{12}$$

We note that if $t > 1$, then

$$\frac{1}{\prod_{j=1}^{k}(t^j-1)} = \begin{bmatrix} \infty \\ k \end{bmatrix}_{1/t}\left(\frac{1}{t}\right)^{\binom{k}{2}}\left(\frac{1}{t}\right)^k.$$

Using the convention that $\prod_{j=1}^{0}(t^j-1) = 1$, let us define $a(r)$, $b(r,l)$, and $c(r,l)$ by

$$c(r,l) = (-1)^l\frac{t^{-\binom{r}{2}}}{\prod_{j=1}^{r}(t^j-1)}\frac{t^{-rl}}{\prod_{j=1}^{l}(t^j-1)} = (-1)^l a(r)b(r,l). \tag{13}$$

Then from (2) and (12),

$$\begin{aligned}\sum_{l\geq 0} c(r,l) &= \frac{t^{-\binom{r}{2}}}{\prod_{j=1}^{r}(t^j-1)}\sum_{l\geq 0}\begin{bmatrix}\infty\\ l\end{bmatrix}_{(1/t)}\left(\frac{1}{t}\right)^{\binom{l}{2}}\left(\frac{-1}{t^{r+1}}\right)^{l}\\ &= \frac{t^{-\binom{r}{2}}}{\prod_{j=1}^{r}(t^j-1)}\prod_{l\geq 0}\left(1+\left(\frac{1}{t}\right)^{l}\left(\frac{-1}{t^{r+1}}\right)\right)\\ &= \pi(r,t).\end{aligned}\tag{14}$$

To obtain the values $c(r,l)$ we used the result (see [7, p. 303]) that the number of $(n\times n)$ matrices over $GF(t)$ with defect r is

$$\eta_n(r,t) = \begin{bmatrix}n\\ r\end{bmatrix}_t\sum_{l=0}^{n-r}(-1)^l\begin{bmatrix}n-r\\ l\end{bmatrix}_t t^{n(n-(r+l))+\binom{l}{2}}.\tag{15}$$

When $p=(t-1)/t$ and all matrices of the space $\mathbf{M}(n,p;t)$ are equiprobable, we have $\Pr(\text{defect}=r)=\eta_n(r,t)/t^{n^2}$. It can be easily shown that

$$\lim_{n\to\infty}\frac{\eta_n(r,t)}{t^{n^2}} = \sum_{l\geq 0}c(r,l),$$

which we have already proved is $\pi(r,t)$. We now prove the following lemma:

Lemma 8. *Let the field be $GF(2)$ and let $\pi(r,2)$ be given by Theorem 1. Let $\{M\}$ be a space of random $(n\times n)$ matrices for which $\mathbf{E}(W)_k = 1+\epsilon(k)$. Let r be constant. If $\sum_{l\geq 0}c(r,l)\epsilon(r+l)\to 0$, then*

$$\lim_{n\to\infty}\Pr\left(M \text{ has defect } r\right) = \pi(r,2).$$

Proof. Let $t=2$. Let $(A_1,\dots,A_k)$ be a k-tuple of simple zero-sum sets of columns A_j of M. Let $N_{k,r}$ be the number of simple k-tuples if the defect of M is r. Simple k-tuples are ordered k-sequences of linearly independent vectors in the vector space over $GF(2)$ generated by the set differences $A\,\Delta\,B$ of zero-sum columns of M. Thus $N_{k,r}$ is the number of ordered k-bases of an r-dimensional vector space (see [4] for an alternative discussion of this), and for $1\leq k\leq r$,

$$N_{k,r} = (t^r-1)(t^r-t)\cdots(t^r-t^{k-1}).$$

If the probability of defect k is $p(k,t)=p_k$, we can write

$$\mathbf{E}(W)_k = \sum_{r\geq k}N_{k,r}p_r.$$

Thus we have the following system of equations which we wish to solve for p_k:

$$\begin{aligned}
1 &= p_0 + p_1 + p_2 + \cdots + p_k + \cdots,\\
\mathbf{E}W &= N_{1,1}p_1 + N_{1,2}p_2 + \cdots + N_{1,k}p_k + \cdots,\\
\mathbf{E}(W)_2 &= N_{2,2}p_2 + \cdots + N_{2,k}p_k + \cdots,\\
&\vdots \qquad \cdots \qquad \vdots\\
\mathbf{E}(W)_k &= N_{k,k}p_k + \cdots.
\end{aligned}$$

The tail of the distribution $\{p(k,t)\}$ is small even for constant k, as

$$\sum_{j\ge k} p_j \le \frac{\mathbf{E}(W)_k}{N_{k,k}} < t^{-\binom{k}{2}}.$$

We wish to prove $p_r = p(r,t) \to \pi(r,t)$ as given in (2). To extract $p_r = p(r,t)$, we will multiply $\mathbf{E}(W)_{r+l}$ for $l \ge 0$ by $c(r,l)$, given by (13), and add. The right-hand side gives p_r exactly. For, if $j = r+s$, $s \ge 0$, then the coefficient of p_j is

$$a(r)\big(N_{r,j}b(r,0) - N_{r+1,j}b(r,1) + \cdots + (-1)^l N_{r+l,j}b(r,l) \\ + \cdots + (-1)^{j-r}N_{j,j}b(r,j-r)\big).$$

However,

$$\begin{aligned}
b(r,l)N_{r+l,j} &= (t^j-1)\cdots(t^j-t^{r+l-1})\frac{t^{-rl}}{\prod_{i=1}^{l}(t^i-1)}\\
&= (t^j-1)\cdots(t^j-t^{r-1})\begin{bmatrix} j-r \\ l \end{bmatrix}_t t^{\binom{l}{2}}.
\end{aligned}$$

Thus the coefficient of p_j is

$$a(r)(t^j-1)\cdots(t^j-t^{r-1})\sum_{l=0}^{j-r}\begin{bmatrix} j-r \\ l \end{bmatrix}_t t^{\binom{l}{2}}(-1)^l,$$

so that for $j > r$, this coefficient is identically zero, from (11), with $x = (-1)$.

As $\mathbf{E}(W)_k \sim 1$, the left-hand side will tend to $\pi(r,t)$ by the argument of (14) above. The least accurate estimate of $\mathbf{E}(W)_k$ occurs in Section 4, so we consider this case in the most detail. Specifically, from Lemma 6 let $\mathbf{E}(W)_k = 1 + (2^k - 1)e^{-\omega}(1+o(1))$. Then

$$\begin{aligned}
\sum_{l\ge 0} c(r,l)\mathbf{E}(W)_{r+l} &= \pi(r,2) + \sum_{l\ge 0}\Big((2^{r+l}-1)e^{-\omega}\Big)(1+o(1))(-1)^l a(r)\frac{2^{-rl}}{\prod_{j=1}^{l}(2^j-1)}\\
&= \pi(r,2) + 2^r e^{-\omega}O(1)\\
&\sim \pi(r,2),
\end{aligned}$$

provided $2^r e^{-\omega} \to 0$, which it does as r is constant and $\omega \to \infty$. ∎

6. THE SHARP THRESHOLD AS $p \to 0$

Let $np = \log n + d(n)$, where $|d(n)| = o(\log\log n)$. To be specific, we will choose $|d| \le \log(\frac{1}{2}\log\log n)$ so that $1/(\log n) \le \exp(-2e^{-d}) \le 1$. We wish to condition on the event $\mathcal{F}$ that M has no zero rows or columns, and at most one unity row and column. The following remarks constitute the proof of Theorem 2(iii) and of Theorem 2(i) and (ii) when d is constant (the sharp threshold) and $p \to 0$.

(R1) *The Probability the Matrix Has No Zero Rows and Columns*: We show that

$$\Pr(\text{no zero rows}) = (1-(1-p)^n)^{2n}(1+o(1)) \tag{16}$$

$$\sim e^{-2e^{-d}}. \tag{17}$$

In fact, for (R2), we will consider a matrix M on m rows and n columns, with a fixed set R of r distinguished columns. The number of zero rows in M is $B(m,(1-p)^n)$, and $\Pr(\text{no zero rows}) = (1-(1-p)^n)^m$. Let the values m, r, and k satisfy $|n-m|$, r, $k = O(\log n)$. Let X be the number of zero columns in $M - R$ and let $(X)_k$ denote the (ordinary) factorial moment of X. For fixed $k \le \log n$,

$$\begin{aligned}\mathbf{E}((X)_k \mid \text{no zero rows in } M) &= (n-r)_k(1-p)^{km}\frac{\left(1-(1-p)^{n-k}\right)^m}{(1-(1-p)^n)^m}\\ &= (n-r)_k(1-p)^{mk}\left(1-kpe^{-d}(1+o(1))\right).\end{aligned}$$

Thus by (standard) inclusion–exclusion, and using the Bonferroni inequalities to halt the summation at $k = \log n$ (see, e.g., [8, p. 141]),

$$\begin{aligned}&\Pr\big(\text{no zero columns in } M-R \mid \text{no zero rows in } M\big)\\ &\quad = (1-(1-p)^m)^{n-r}\left(1+o\left(\frac{\log^7 n}{n}\right)\right)\end{aligned} \tag{18}$$

and (16) follows when $r = 0$ and $m = n$.

(R2) *Lemma 5 Is Still True*: In particular, conditional on "no zero rows or columns," with probability $1 - o(1)$ there are no small ($< m_2$) zero-sum subsets of columns.

Proof. We note that

$$\begin{aligned}\mathbf{E}(W \mid \text{ no zero rows or columns }) &\le (\mathbf{E}W)/(\Pr(\text{no zero rows or columns}))\\ &\le 2\mathbf{E}We^{2e^{-d}},\end{aligned}$$

where $e^{2e^{-d}} \le \log n$. We can use the estimates of $\sum \mathbf{E}W_m$ of cases $\log n \le m \le m_2$ and $m_3 \le m \le n$ from Section 3, Lemma 5, as these are $o(1/n)$. We now consider the case of $2 \le m \le \log n$.

We first prove that whp a set of m columns of M has at most e^2mnp nonzero entries. Let Z be a binomial random variable with mean μ and let $\alpha \geq 1$. Then by the Chernoff inequality,

$$\Pr(Z \geq \alpha\mu) \leq \left(\frac{e}{\alpha}\right)^{\alpha\mu} e^{-\mu}.$$

The number of entries in the columns is $B(nm, p)$ with mean $\mu = mnp$. Let $\mathscr{R}$ be the event that there exists a set of columns with at least $e^2\mu$ entries. Then

$$\begin{aligned}\Pr(\mathscr{R}) &\leq \binom{n}{m} e^{-(e^2+1)\mu} \\ &< \exp -em\log n. \end{aligned} \tag{19}$$

Let N be the $(n \times m)$ matrix consisting of the selected columns. Each column of N must have at least one nonzero entry, as no column of M is zero. Let $n - s$ rows of N be zero. The remaining s rows must each have at least two nonzero entries. Thus N has at least $\theta = \max(m, 2s)$ nonzero entries. The expected number of (m, s) pairs satisfying this condition is at most

$$\nu(m, s) = \binom{n}{m}\binom{n}{s}(1 - p)^{(n-s)m} \Pr(X \geq \theta),$$

where $X \sim B(sm, p)$. Thus by the Chernoff inequality,

$$\begin{aligned}\Pr(X \geq m) &\leq (esp)^m e^{-smp} \\ \Pr(X \geq s) &\leq (\tfrac{1}{2}emp)^{2s} e^{-smp}.\end{aligned}$$

Case $s \leq \lfloor m/2 \rfloor$:

$$\begin{aligned}\nu(m, s) &\leq \frac{n^m}{m!}\frac{n^s}{s!}(\exp -npm)(esp)^m \\ &\leq O(1)\left(e^{-d+2}(\log n)\right)^m e^s \left(\frac{s}{n}\right)^{m-s} \\ &\leq (\log n)^{2m}\left(\frac{m}{n}\right)^{m-\lfloor m/2 \rfloor}.\end{aligned}$$

Case $s > \lfloor m/2 \rfloor$:

$$\begin{aligned}\nu(m, s) &\leq \frac{n^m}{m!}\frac{n^s}{s!}(\exp -npm)(\tfrac{1}{2}emp)^{2s} \\ &\leq \left(\frac{e^{-d+1}}{m}\right)^m \left(\frac{e^3 m(\log n)^2}{n}\right)^s.\end{aligned}$$

Thus

$$\sum_s \nu(m, s) \leq O\left(\left(\frac{\log n^3}{n}\right)^{m/2}\right),$$

so that $\sum_{m,s} \nu(m, s) = O(\log^4 n/n)$.

Let N have s nonzero rows, S. By (19), $|S| = s \le e^2 mnp$. We can rearrange M into

$$M = \begin{pmatrix} S & S_1 & O \\ O & R & M_1 \end{pmatrix}.$$

The matrix S_1 is $s \times r$ and each column contains nonzero entries, arising from the rows of S in $M - N$. By (19), S_1 almost always has $r \le e^2 snp$ columns. The matrix M_1 is $(n-s) \times (n-m-r)$. By the condition that M has no zero rows or columns, the matrix $(R \;\; M_1)$ has no zero rows, and the columns of M_1 are nonzero. By (18) of (R1) the probability of this event is asymptotically equal to the conditioning probability (16) of the event that M has no zero rows or columns. The result follows. ∎

(R3) We now prove $\mathbf{E}((W)_k \mid \text{ no zero rows or columns }) = 1 + o(\log^4 n/\sqrt{n})$.

(R3a) We modify (9), so that π becomes $\phi = \pi - (1-p)^n$, for the event that $A_1, \ldots, A_k$ are zero sum in the given row, but the row is not identically zero:

$$\begin{aligned}\phi &= \Pr\big(((A_1, \ldots, A_k) \text{ are zero sum in row } j\big) \text{ and } (\text{row } j \text{ is not zero})\,\big)\\ &= \frac{1}{2^k}\Big(1 + \sum_{C \in S}(1-2p)^{x_{j_1}+\cdots+x_{j_N}} - 2^k(1-p)^n\Big)\\ &= \frac{1}{2^k}\left(1 - \frac{e^{-d}}{n}\left(1 + o\left(\frac{\log^2 n}{\sqrt{n}}\right)\right)\right), \qquad (20)\end{aligned}$$

because, by (R2), $x_{j_1} + \cdots + x_{j_N} \ge m_2 = n/2(1 - 4\sqrt{(\log n)/n})$ and $|S| = 2^k - 1$, so

$$\sum_{C \in S}(1-2p)^{x_{j_1}+\cdots+x_{j_N}} = (2^k - 1)\frac{e^{-d}}{n}\left(1 + o\left(\frac{\log^2 n}{\sqrt{n}}\right)\right).$$

Thus

$$\begin{aligned}&\Pr\big(((A_1, \ldots, A_k) \text{ are zero sum}) \text{ and } (\text{no zero rows})\big)\\ &\quad = \frac{1}{2^{kn}} e^{-e^{-d}}\left(1 + o\left(\frac{\log^3 n}{\sqrt{n}}\right)\right). \qquad (21)\end{aligned}$$

(R3b) We next prove

$$\begin{aligned}&\Pr\big(\text{no zero columns} \mid ((A_1, \ldots, A_k) \text{ are zero-sum}), \text{ and } (\text{no zero rows})\big)\\ &\quad \sim e^{-e^{-d}}. \qquad (22)\end{aligned}$$

Proof. Let $X(k) = \sum_{s=1}^{n} X_s(k)$, where

$$X_s(k) = \begin{cases} 1, & \text{if } (M_s \text{ is a zero column}), ((A_1, \ldots, A_k) \text{ are zero sum}), \text{ and} \\ & (\text{no zero rows}), \\ 0, & \text{otherwise.} \end{cases}$$

Let $m = |\bigcup_{i=1\cdots k} A_i|$. Considering l-tuples of zero columns, we find for fixed $l \leq \log n$,

$$\begin{aligned}\frac{\mathbf{E}(X(k))_l}{l!} &= \sum_{j=0}^{l} \binom{n-m}{l-j}\binom{m}{j}(1-p)^{nl} \\ &\quad \times \left(\frac{1}{2^k}\left(1 + \sum_{C \in S}(1-2p)^{\mu(j,C)} - 2^k(1-p)^{n-l}\right)\right)^n \\ &= \left[\frac{e^{-e^{-d}}}{2^{nk}}\left(1 + o\left(\frac{\log^3 n}{\sqrt{n}}\right)\right)\right]\binom{n}{l}(1-p)^{nl}.\end{aligned}$$

The superscript $\mu(j, C) = \sum_{s \in C} x_s - j_C$, where j_C is the number of zero columns in the subset of $\bigcup A_i$ indexed by C. The final line follows from the argument used on ϕ in (R3a). We then use (standard) inclusion–exclusion as in (R1). ■

(R3c) Thus from (16), (21), and (22),

$$\begin{aligned}&\Pr\big((A_1, \ldots, A_k) \text{ is zero sum } \mid \text{ no zero rows or columns }\big) \\ &\quad = \frac{1}{2^{nk}}\left(1 + o\left(\frac{\log^4 n}{\sqrt{n}}\right)\right).\end{aligned}$$

7. THE SHARP THRESHOLD AS $p \to 1$

Let $p = 1 - (\log n + d(n))/n$ and $|d(n)| = o(\log\log n)$. Substitution of $q = 1 - p$ into (8) gives

$$\begin{aligned}\rho_m(0) &= \tfrac{1}{2}(1 + (-1)^m(1-2q)^m), \\ \rho_m(1) &= \tfrac{1}{2}(1 + (-1)^{m+1}(1-2q)^m),\end{aligned} \tag{23}$$

which is the same as (8) provided m is even. Thus, we propose to imitate the previous arguments, substituting q for p.

(P1) A column of M is *unity* if it consists entirely of 1s. Trivial linear dependencies occur in the rows of M if two or more rows are unity, and similarly for the columns.

The number of unity columns is $B(n, p^n) = B(n, (1-q)^n)$. The total number of unity rows is asymptotically $Po(e^{-d})$ for constant d, by (R1) of the previous section. Let $\mathcal{U}$ be the event that at most one row and at most one column are unity. By an application of the techniques of (R1),

$$\Pr(\mathcal{U}) \sim \left(e^{-e^{-d}}(1 + e^{-d})\right)^2.$$

(P2) Let $\mathcal{V}$ be the event that there are no unity rows or columns. Let $\mathbf{c} \in \{0, 1\}^n$ be a fixed vector, where $\mathbf{c}$ is $\mathbf{0}$, $\mathbf{1}$, or has $O(\log n)$ zeroes. Let $W(\mathbf{c})$ be the number of sets of columns of M adding to $\mathbf{c}$. We now show that the calculations of (R2) and (R3) are essentially unaltered and $\mathbf{E}\left((W(\mathbf{c}))_k \mid \mathcal{V}\right) \sim 1$.

Case of Small Sets, Size $1 \le m \le \log n$. If $\mathbf{c} = \mathbf{0}, \mathbf{1}$, the evaluation of $\mathbf{E}W(\mathbf{c})$ leads to an analysis identical to that of $\mathbf{E}W_m$ in (R2), except we now condition on the number of rows, s, with at least two zeroes.

Let $\mathbf{c}$ have $k \ge 1$ fixed zeroes. Let m be odd, $m \ge 1$. Let $\theta = \max(k, m)$. The probability of the event that there exist m columns adding to $\mathbf{c}$ is at most

$$o(1) + \binom{n}{m} m^\theta q^\theta (1-q)^{(n-\Theta(m\log n))m} = o(1).$$

The first $o(1)$ is from (19), and the second term conditions on the event that the number of zero entries is at most $e^2 mnp$.

Case of Sets Size $m \sim n/2$. Let A be a column set of size m. Let $\psi(i)$ be the probability that A is i sum in row s and that row s is not unity. Then

$$\psi(0) = \tfrac{1}{2}\left(1 + (-1)^m(1-2q)^m\right) - (1-q)^n 1 \ (m \text{ even}),$$
$$\psi(1) = \tfrac{1}{2}\left(1 - (-1)^m(1-2q)^m\right) - (1-q)^n 1 \ (m \text{ odd}).$$

In either case and for any $m \sim n/2$,

$$\psi(i) = \tfrac{1}{2}\left(1 - \frac{e^{-d}}{n}(1 + o(1))\right),$$

so that $\psi(i)$ has the same value as ϕ in (20) of (R3a), with $k = 1$. Thus $\mathbf{E}(W(\mathbf{c}) \mid \mathcal{V}) \sim 1$ irrespective of the value of $\mathbf{c}$. The generalization of these calculations to $\mathbf{E}(W)_k$ in (R3) follows naturally.

(P3) Suppose we condition on $\mathcal{U}$. Either $\mathcal{V}$ holds, or there is at least one unity row or column. If one of each (Case (C1)), let the row be $\mathbf{u}$ and the column be $\mathbf{v}$. Else (Case (C2)) let it be a unity row, $\mathbf{u}$, and choose a column $\mathbf{v}$ at random. Let M_1 be the $(n-1) \times (n-1)$ matrix obtained from M by deleting $\mathbf{u}$, $\mathbf{v}$.

In Case (C1), the matrix M_1 satisfies the property $\mathcal{V}$. In Case (C2), what is the probability that we have made another unity row by deleting $\mathbf{v}$? It is at most ϕ, the expected number of rows where the column $\mathbf{v}$ had the only zero entry in that row. However,

$$\phi = n(1-p)p^{n-1} = O\left(\frac{\log^2 n}{n}\right).$$

Thus with probability $1 - o(1)$ either the matrix M or the matrix M_1 has property $\mathcal{V}$. If M has $\mathcal{V}$, we have proved in (P2) that $\mathbf{E}((W)_k \mid \mathcal{V}) \sim 1$.

(P4) Suppose M_1 has $\mathcal{V}$. Let $\mathbf{v}_1$ be the restriction of $\mathbf{v}$ to the rows of M_1. Events in M_1 leading to a *possible* linear dependence in M are that a set of columns of M_1 adds to $\mathbf{0}$ or to $\mathbf{v}_1$. Such events are counted by the random variables $W(\mathbf{0})$ and $W(\mathbf{v}_1)$, evaluated over the columns of M_1.

The effect on $W(\mathbf{0})$ of adding back the row $\mathbf{u}$ is to restrict the evaluation of $\mathbf{E}W(\mathbf{0})$ to *even size* column sets $m = 2l$, as only even sets will add to zero in the row $\mathbf{u}$. Thus $\mathbf{E}(W(\mathbf{0}) \mid \mathscr{V}) \sim \frac{1}{2}$. Similarly, the effect on $W(\mathbf{v}_1)$ of adding back the row $\mathbf{u}$ is to restrict the evaluation of $W(\mathbf{v}_1)$ to odd size column sets $m = 2l + 1$ (because the entry of $\mathbf{v}$ in the row $\mathbf{u}$ is 1).

Let $W^*(\mathbf{0})$ be the evaluation of $W(\mathbf{0})$ on even column sets of M_1 and let $W^*(\mathbf{v}_1)$ be the evaluation of $W(\mathbf{v}_1)$ over odd size column sets of M_1. Let $Z = W^*(\mathbf{0}) + W^*(\mathbf{v}_1)$. Then Z is the number of linear dependencies in the columns of M, and

$$\mathbf{E}(Z \mid M_1 \text{ has } \mathscr{V}) \sim \tfrac{1}{2} + \tfrac{1}{2} = 1.$$

(P5) To extend this argument to higher moments, we retain the definition of simple k-tuples of sets $(C_1, \dots, C_k)$ except now, some sets are zero sum (denoted by A_j) and some are $\mathbf{v}_1$ sum (denoted by B_j). Set differences $A_1 \Delta A_2$, $A \Delta B$, or $B_1 \Delta B_2$ on these sets preserves the joint property "$\mathbf{v}_1$ sum or zero sum."

Consider row s of M_1. $(C_1, \dots, C_k)$ defines a vector $\mathbf{c}(s) = (c_1, \dots, c_k)$, where $c_i = 0$ if C_i is zero sum and $c_i = v_{1s}$ if C_i is $\mathbf{v}_1$ sum. In the notation of Section 4 we now require that the intersection structure of $(C_1, \dots, C_k)$ gives solutions satisfying $\mathbf{Ay} = \mathbf{c}(s)$ in each row s of M_1. The extension of Lemma 7, given by Lemma A6 of the Appendix, ensures that the proofs of Section 4 and their extension to (R3) are intact.

Let $(C_1, \dots, C_k) = (A_1, \dots, A_j, B_1, \dots, B_{k-j})$ and let $\mathbf{E}(Z(j))_k$ be the expectation of Z on simple k-tuples with exactly j zero-sum sets, and conditional on $\mathscr{V}$. Thus

$$\mathbf{E}(Z(j))_k \sim \left(\tfrac{1}{2}\right)^k$$

and

$$\mathbf{E}(Z)_k = \sum_{j=0}^{k} \binom{k}{j} \mathbf{E}(Z(j))_k \sim 1,$$

as before.

APPENDIX

Expected Number of Linear Dependencies

The proof of Lemma 5(i) follows from (25), (34), (28), and (30), and the proof of Lemma 5(ii) follows from (26). The proof of Lemma 4 then follows from (31).

Let the field be $GF(t)$, where $t = o(\sqrt{\log n})$. Let $p = c/n$, $p \leq (t-1)/t$, where $c = \log n + \log(t-1) + d(n)$. We will prove here that for $d(n) \geq -o(\log\log n)$ there are whp no linear dependencies in the range $m(L) = \{\log n, \dots, ((t-1)/t)n\left(1 - 4\sqrt{\log n/n}\right)\}$ or $m(U) = \{((t-1)/t)n \quad (1 + 4\sqrt{\log n/n}), n\}$. Furthermore, for $d(n) \to \infty$ there are no dependencies in the range $\{1, \dots, \log n\}$. Let $\omega_1 \to \infty$ arbitrarily slowly. Let $m_0 = n/c^2$ and $m_i =$

$n(t-1)/t(1-\epsilon_i)$ for $i = 1, 2, 3$. Specifically let $m_1 = n(t-1)/t(1-1/\omega_1 \log n)$, let $m_2 = n(t-1)/t(1-4\sqrt{\log n/n})$, and let $m_3 = n(t-1)/t(1+4\sqrt{\log n/n})$.

We now investigate the behavior of

$$\mathbf{E}W_m = \frac{1}{t-1}\binom{n}{m}(t-1)^m \rho_m^n.$$

Case of $1 \le m \le n/c^2 = m_0$. The absolute value of the terms of $(1 - tp/(t-1))^m$ tends steadily to zero for any value of m in this interval, so

$$\left(1 - \frac{t}{t-1}p\right)^m \le 1 - \frac{t}{t-1}mp + \left(\frac{t}{t-1}\right)^2 \frac{m^2 p^2}{2}$$

by a standard property of alternating series. Thus $\rho_m \le (1 - mp(1-mp))$ and

$$\begin{aligned}\mathbf{E}W_m &\le \frac{1}{t-1}\frac{(n(t-1))^m}{m!}\exp -nmp(1-mp)\\ &= \frac{1}{t-1}\frac{(e^{-d(n)+1})^m}{m!}.\end{aligned} \tag{24}$$

Thus, when $d(n) \ge -o(\log\log n)$,

$$\sum_{m=\log n}^{m_0} \mathbf{E}W_m = o(1/n^2), \tag{25}$$

and when $d(n) = \omega$,

$$\sum_{1}^{m_0} \mathbf{E}W_m = O\left(e^{-\omega}\right). \tag{26}$$

Case of $m_1 \le m \le n$. We note that $f(m,p) = (1 + (t-1)(1 - ((t-1)/t)p)^m)^n$ tends rapidly to 1 as $d(n) \to \infty$. If $d(n) \ge 3\log n$, then $1 \le f(m,p) \le 1 + o(1/n)$ so we consider the case $-o(\log\log n) < d(n) < 3\log n$ in most detail.

Write $m = n(1-\epsilon)(t-1)/t$,

$$\begin{aligned}\rho_m(0) &= \frac{1}{t}\left(1 + (t-1)\left(1 - \frac{t}{t-1}p\right)^m\right)\\ &= \frac{1}{t}\left(1 + (t-1)\exp -c(1-\epsilon) + O(np^2)\right)\\ &= \frac{1}{t}\left(1 + \frac{1}{n}\exp(-d(n) + c\epsilon(1+o(1)))\right).\end{aligned} \tag{27}$$

Thus,

$$\rho_m^n = \frac{1}{t^n}\left(\exp(e^{-d+c\epsilon(1+o(1))})\right)$$

and

$$\sum_{m_1}^{m_2} \mathbf{E}W_m = \frac{1}{t-1} \sum_{m=m_1}^{m_2} \binom{n}{m} \frac{(t-1)^m}{t^n} (\exp(e^{-d+c\epsilon(1+o(1))}))$$
$$\leq \frac{1}{t-1} (\exp(e^{-d+c\epsilon_1(1+o(1))})) \exp\left(-(\epsilon_2)^2 \frac{t-1}{3t} n\right)$$
$$= o\left(\frac{1}{n^{4/3}}\right). \tag{28}$$

This will follow because the binomial random variable $X \sim B(n, (t-1)/t)$ is sharply concentrated around the mean $\mu = n(t-1)/t$. Thus,

$$\Pr(|X - \mu| > \delta\mu) \leq 2 \exp - \frac{\delta^2 \mu}{3}, \tag{29}$$

by the Hoeffding inequality (see, e.g., [8, p. 136]). Let $\delta = \epsilon_2$. It is also an immediate consequence that

$$\sum_{m_3}^{n} \mathbf{E}W_m = o(1/n^2). \tag{30}$$

Finally

$$\sum_{m_2}^{m_3} \mathbf{E}W_m = \begin{cases} \dfrac{1}{t-1} \exp\left(e^{-d} + o\left(\dfrac{\log^{5/2} n}{\sqrt{n}}\right)\right), & d > -o(\log\log n), \\ \dfrac{1}{t-1}(1 + e^{-\omega})(1 + o(1)), & d(n) = \omega \to \infty, \\ \dfrac{1}{t-1}\left(1 + o\left(\dfrac{1}{n}\right)\right), & d(n) \geq 3\log n. \end{cases} \tag{31}$$

Case of $m_0 \leq m \leq m_1$. We will show that $\mathbf{E}W_m$ is a decreasing function of m up to $m = \frac{1}{2} n(\log c/c)(t-1)/t$ and an increasing function of m for $2n(\log c/c)(t-1)/t \leq m \leq m_1$. Thus

$$\sum_{m_0}^{m_1} \mathbf{E}W_m \leq n \max\{\mathbf{E}W_{m_0}, \mathbf{E}W_{m_1}, \mathbf{E}W_{m^*}\}, \tag{32}$$

where $m^* \in \{bn(\log c/c)(t-1)/t$ and $\frac{1}{2} \leq b \leq 2\}$.

Let $\Delta = t/(t-1)p$. If $R = (\mathbf{E}W_{m+1})/(\mathbf{E}W_m)$, we see that

$$R = \frac{n-m}{m+1}(t-1)\left(1 - \frac{(t-1)\Delta(1-\Delta)^m}{1+(t-1)(1-\Delta)^m}\right)^n$$
$$\leq \frac{n-m}{m+1} t \exp -(t-1)n\Delta\left(\frac{\exp -m\Delta/(1-\Delta)}{1+(t-1)\exp -m\Delta}\right)$$
$$\leq \frac{n-m}{m+1} t \exp -ntp\left(\frac{\exp -(t/(t-1)mp(1+O(p)))}{1+(t-1)\exp -(t/(t-1)mp)}\right)$$

by repeated application of $\exp -x/(1-x) \leq 1-x \leq \exp -x$.

Let $m = bn(\log c/c)(t-1)/t$, where $1/(c\log c) < b < c/(\log c)$. Then $m\Delta = b\log c$ and

$$R \leq \frac{n-m}{m+1} t \exp -\left(\frac{tc^{1-b(1+O(p))}}{1+(t-1)c^{-b}}\right).$$

Thus provided $b \leq 1/2$, then $R < 1$ and $\mathbf{E}W_m$ is decreasing.

We now consider $b \geq 2$:

$$\begin{aligned} R &\geq \frac{n-m}{m+1}(t-1)\exp -\left(\frac{ntpe^{-m\Delta/(1-\Delta)}}{1+(t-1)e^{-m\Delta} - tpe^{-m\Delta/(1-\Delta)}}\right) \\ &\geq \frac{n-m}{m+1}(t-1)\exp\left(-\frac{t}{c^{b-1}}(1-o(1))\right). \end{aligned}$$

Thus, $R \geq 1$ for $2n(\log c/c)(t-1)/t \leq m \leq m_1$ and $\mathbf{E}W_m$ is increasing in this range.

To evaluate $\mathbf{E}W_{m^*}$, we note that $\binom{n}{m} \leq (ne/m)^m$, so

$$\mathbf{E}W_m \leq \left(\frac{ne(t-1)}{m}\right)^m \frac{1}{t^n} \exp\left(n(t-1)e^{-m\Delta}\right).$$

When $m = m^* = nb(\log c/c)(t-1)/t$,

$$(\mathbf{E}W_m)^{1/n} \leq \left(\frac{tec}{b\log c}\right)^{(t-1)/t(b\log c/c)} \left(\frac{1}{t}\right) \exp\left((t-1)c^{-b}\right).$$

Taking logarithms, the right-hand side is

$$-\log t + O\left(b\frac{\log^2 c}{c} + \frac{t}{c^b}\right) \leq -\frac{1}{2}\log t,$$

because, as $t = o(\sqrt{\log n})$, $-\log t$ is the dominant term.

We now consider

$$\mathbf{E}W_{m_1} \leq (1+(t-1)(1-\Delta)^{m_1})^n \Pr(X \leq m_1),$$

where $X \sim B(n,(t-1)/t)$ is a binomial random variable. Now

$$\begin{aligned} (1+(t-1)(1-\Delta)^{m_1})^n &\leq \exp\left((t-1)n\exp -m_1\Delta\right) \\ &\leq \exp\left(2(t-1)n^{1/\omega_1 \log n}e^{-d(n)}\right). \end{aligned}$$

So from (29) with $d(n) \geq -o(\log\log n)$,

$$\mathbf{E}W_{m_1} \leq 2\exp\left(o(t\log n) - \frac{n}{6(\omega_1\log n)^2}\right). \tag{33}$$

Finally

$$\sum_{m_0}^{m_1} \mathbf{E}W_m = o(1/n^2) \tag{34}$$

from (32) using (24) and (33), and because $\mathbf{E}W_{m^*} = O(t^{-n/2})$.

Proof of Lemma 7

If column j of row i of $\mathbf{A}$ has $a_{ij} = 1$, this means the set A_i is uncomplemented in the intersection I_j. The columns of $\mathbf{A}$ are exactly the $(2^k - 1)$ distinct nonzero vectors of length k. Thus the rank of $\mathbf{A}$ is k.

Let $\mathscr{L}$ be the k-dimensional vector space generated by the rows $(\mathbf{r}_1, \ldots, \mathbf{r}_k)$ of $\mathbf{A}$. An element $\mathbf{e}$ of $\mathscr{L}$ can be written as $\mathbf{e} = c_1\mathbf{r}_1 + \cdots + c_k\mathbf{r}_k$, where $c_i \in \{0, 1\}$, or as $\mathbf{e} = (e_j, j = 1, \ldots, L)$. Associated with $\mathbf{e}$ are two sets of indices. The set $R(\mathbf{e}) = \{i : c_i = 1\}$, the indices of the selected sets A_i (rows of $\mathbf{A}$), and $C(\mathbf{e}) = \{j : e_j \equiv 1 (\bmod 2)\}$, the indices of the subsets I_j (columns of $\mathbf{A}$) with nonzero entries in $\mathbf{e}$.

Lemma A1. *Let $S = \{C(\mathbf{e}) : \mathbf{e} \in \mathscr{L}\}$.*

(i) $|S| = 2^k - 1$.
(ii) *Let $C(\mathbf{e}) \in S$. The vector $\mathbf{e} = c_1\mathbf{r}_1 + \cdots + c_k\mathbf{r}_k$, which has $R(\mathbf{e}) = \{i_1, \ldots, i_s\}$, corresponds to the zero-sum column set*

$$B = A_{i_1} \,\Delta\, A_{i_2} \,\Delta\, \cdots \,\Delta\, A_{i_s}.$$

Proof. **(i)** There are $2^k - 1$ distinct nonzero linear combinations $\mathbf{e}$ of the rows of $\mathbf{A}$, giving $2^k - 1$ distinct sets of column indices $C(\mathbf{e})$.

(ii) This is because

$$I_j \subset B \iff \sum_{i \in R(\mathbf{e})} a_{ij} \equiv 1 \ (\bmod 2) \iff e_j = 1.$$

Thus $B = \bigcup_{j \in C(\mathbf{e})} I_j$ and B is zero sum. ■

This completes the proof of Lemma 7(iv).

Let $T = \{\mathbf{y} : \mathbf{A}\mathbf{y} = \mathbf{0}\}$. Let $J = \{j_1, \ldots, j_s\}$. Say a solution $\mathbf{y} \in T$ is even on J if $y_{j_1} + \cdots + y_{j_s} = 0$ and odd (on J) otherwise. Let $S = S(J) = \{\mathbf{y} \in T : \mathbf{y}$ is even on $J\}$.

Lemma A2. *Either $S = T$ or $|S| = \frac{1}{2}|T|$.*

Proof. We note that T is a group under vector addition. Now $\mathbf{0} \in S$ and S is closed under addition, so S is a subgroup of T. If $\mathbf{y}, \mathbf{y}' \in T$, the cosets $\mathbf{y} + S$, $\mathbf{y}' + S$ are equal if and only if $(\mathbf{y} - \mathbf{y}') \in S$. This occurs if $\mathbf{y}$ and $\mathbf{y}'$ are both even on J or both odd on J. If $\mathbf{y}$ is any odd solution, $\mathbf{y} + S$ is the unique odd coset so that $|S| = |\mathbf{y} + S|$ and $S \cup (\mathbf{y} + S) = T$. ■

Lemma A3. *If $S = T$, then*

$$\sum_{\mathbf{y} \in T} (-1)^{y_{j_1} + \cdots + y_{j_s}} = |T|.$$

If $S \neq T$, then

$$\sum_{\mathbf{y} \in T} (-1)^{y_{j_1} + \cdots + y_{j_s}} = \tfrac{1}{2}|T| - \tfrac{1}{2}|T| = 0.$$

We now consider the existence of odd solutions. For now, let us assume that no set I is empty.

Lemma A4. *Let $T = \{\mathbf{y} : \mathbf{A}\mathbf{y} = \mathbf{0}\}$. Let $J = \{j_1, \dots, j_s\}$. T has odd solutions on J if and only if $J \neq C(\mathbf{e})$ for some $\mathbf{e} \in \mathcal{L}$ the row space of $\mathbf{A}$.*

Proof. Let $J = \{j_1, \dots, j_s\}$ and let $\alpha = (\alpha_j : j = 1, \dots, L)$, where

$$\alpha_j = \begin{cases} 1, & j \in J, \\ 0, & j \in L \setminus J. \end{cases}$$

Then $\mathbf{y}$ is an even solution on J iff

$$\mathbf{a}\mathbf{y} = 0 \quad \Longleftrightarrow \quad y_{j_1} + \cdots + y_{j_s} = 0.$$

Let $\mathbf{A}^* = \left[\begin{smallmatrix} \mathbf{A} \\ \boldsymbol{\alpha} \end{smallmatrix}\right]$ and $R = \{\mathbf{w} : \mathbf{A}^*\mathbf{w} = \mathbf{0}\}$, so that $R \subseteq T$. Using the rank and nullity theorem,

$$\dim(T) = L - \operatorname{rank}(\mathbf{A}) = L - k,$$
$$\dim(R) = L - \operatorname{rank}(\mathbf{A}^*).$$

If $\alpha = \mathbf{e}$ for some $\mathbf{e} \in \mathcal{L}$ the row space of $\mathbf{A}$, then $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^*)$ and $R = T$.
If $\alpha \notin \mathcal{L}$, then $\operatorname{rank}(\mathbf{A}^*) = \operatorname{rank}(\mathbf{A}) + 1$ and $|R| < |T|$. Any $\mathbf{y} \in T \setminus R$ satisfies $\alpha\mathbf{y} = 1$, an odd solution. ■

What if some sets $I = \varnothing$?

Lemma A5. *If $0 \le i \le L - k$ sets I are empty,*

(i) $|T| = 2^{L-k-i}$.
(ii) *Let $S = \{C(\mathbf{e}) : \mathbf{e} \in \mathcal{L}\}$. Then $|S| = 2^k - 1$.*

Proof. Let i sets be empty. We delete the corresponding columns of $\mathbf{A}$ so that $\mathbf{A}$ is now a $k \times (L - i)$ matrix. We claim that the row space of $\mathbf{A}$ still has dimension k. Recall that $(A_1, \dots, A_k)$ is simple iff

$$A_{i_1} \Delta \cdots \Delta A_{i_l} \neq \varnothing \qquad (1 \le l \le k).$$

Let $\mathbf{e} = \mathbf{r}_{i_1} + \cdots + \mathbf{r}_{i_l}$. As $A_{i_1} \Delta \cdots \Delta A_{i_l} \neq \varnothing$, then $C(\mathbf{e}) \neq \varnothing$ so that $\mathbf{e} \neq \mathbf{0}$. Thus the rows are linearly independent. The solution space now has dimension $L - k - i$, where $0 \le i \le L - k$. ■

This completes the proof of Lemma 7(i), (ii), and (iii).
To prove Theorem 2(ii) as $p \to 1$, we need to consider a more general form of Lemma 7(iii), which we now state.

Lemma A6. *Let* $\mathbf{c} \in \{0, 1\}^k$. *Let* $T(\mathbf{c}) = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{c}\}$. *Let*

$$\lambda(J, \mathbf{c}) = \sum_{\mathbf{x} \in T(\mathbf{c})} (-1)^{\sum_{j \in J} x_j}.$$

Then

$$\lambda(J, \mathbf{c}) = \begin{cases} 0, & J \notin S, \\ (-1)^{\xi} |T|, & J \in S, \end{cases}$$

where $\xi \in \{0, 1\}$.

Proof. Let $\mathbf{z} = (z_i : i = 1, \ldots, k)$ be a specific solution of $\mathbf{A}\mathbf{x} = \mathbf{c}$ so that $T(\mathbf{c}) = T + \mathbf{z}$. Then

$$\begin{aligned} \lambda(J, \mathbf{c}) &= \sum_{\mathbf{x} \in T(\mathbf{c})} (-1)^{\sum_{j \in J} x_j} \\ &= \sum_{\mathbf{y} \in T} (-1)^{\sum_{j \in J} (y_j + z_j)} \\ &= (-1)^{\sum_{j \in J} z_j} \sum_{\mathbf{y} \in T} (-1)^{\sum_{j \in J} y_j}. \end{aligned}$$

The result now follows from the standard form of Lemma 7.

REFERENCES

[1] J. Blömer, R. Karp, and E. Welzl, The rank of sparse random matrices over finite fields, Random Struct Alg, 10(4) (1997), 407–419.
[2] P.J. Cameron, Combinatorics: Topics, Techniques, Algorithms, Cambridge Univ. Press, Cambridge, UK, 1994.
[3] R. Durrett, Probability: Theory and Examples, Wadsworth, Belmont, CA, 1991.
[4] V. Kolchin, Random graphs and systems of linear equations in finite fields, Random Struct Alg, 5(1) (1994), 135–146.
[5] D.E. Knuth, The Art of Computer Programming, Addison-Wesley, Reading, MA, 1973.
[6] I.N. Kovalenko, A.A. Levitskya, and M.N. Savchuk, Selected Problems in Probabilistic Combinatorics, Naukova Dumka, Kiev, 1986 (in Russian).
[7] J.H. van Lint and R. M. Wilson, A Course in Combinatorics, Cambridge Univ. Press, Cambridge, UK, 1992.
[8] E.M. Palmer, Graphical Evolution, Wiley-Interscience, New York, 1985.