Journal of Combinatorics
Volume 13, Number 3, 397–436, 2022

Rank of near uniform matrices*

Jake Koenig and Hoi Nguyen

A central question in random matrix theory is universality. When an emergent phenomena is observed from a large collection of chosen random variables it is natural to ask if this behavior is specific to the chosen random variable or if the behavior occurs for a larger class of random variables.

The rank statistics of random matrices chosen uniformly from $\mathrm{Mat}(\mathbf{F}_q)$ over a finite field are well understood. The universality properties of these statistics are not yet fully understood however. Recently Wood [40] and Maples [26] considered a natural requirement where the random variables are not allowed to be too close to constant and they showed that the rank statistics match with the uniform model up to an error of type e^{-cn}. In this paper we explore a condition called near uniform, under which we are able to prove tighter bounds q^{-cn} on the asymptotic convergence of the rank statistics.

Our method is completely elementary, and allows for a small number of the entries to be deterministic, and for the entries to not be identically distributed so long as they are independent. More importantly, the method also extends to near uniform symmetric, alternating matrices.

Our method also applies to two models of perturbations of random matrices sampled uniformly over $\mathrm{GL}_n(\mathbf{F}_q)$: subtracting the identity or taking a minor of a uniformly sampled invertible matrix.

AMS 2000 subject classifications: 60B20.

1. Introduction

A central question in random matrix theory is that of universality. If instead of taking a Haar-uniform random matrix, we sample by some other methods, under what conditions will we observe similar statistics in some large scale matrix phenomena? While there have been many results addressing universality of random matrices in characteristic zero to study the spectral

arXiv: 2101.00107
*The authors are supported by NSF grant DMS-1752345.

behavior of various models of random matrices, we have not seen much in the literature addressing universality behavior in the finite fields setting. In fact, to the best of our knowledge, although there had been results for matrices over finite fields such as [1, 3, 6, 9, 14, 12, 20, 21, 19, 35], universality results only appeared very recently in [15, 23, 26, 27, 31, 32, 33, 40, 41].

The rank statistics of a uniform random matrix over a finite field $\mathbf{F}_q$ are well studied. An exact formula for the rank of a uniform matrix in $\mathbf{F}_q$ are folklore and can be found in for instance [2]. A formula for the rank of a symmetric matrix was found by MacWilliams [24] and for an alternating matrix by MacWilliams and Sloane [25]. We will recover these formulas throughout our presentation under a more probabilistic viewpoint.

Consider random matrices with entries being iid copies of ξ satisfying a so called min-entropy condition that $\mathbf{P}(\xi = k) \leq 1 - \alpha$ for all $k \in \mathbf{F}_q$. The main result of [26] by Maples (see also [31] for corrections) and a special case of the recent result [41] by Wood showed that the rank distribution of such matrices is asymptotically equal to the distribution of the Haar distributed random matrices. Maples' approach relies on a very fine machinery (swapping method) originating from [5, 11, 37], which achieves an exponential error bound $e^{-c_\alpha n}$, while Wood's approach relies on counting surjections (moment method) which yields weaker error but the method has wider applications.

In this paper we introduce a condition on the entries which can be seen as an interpolating between uniform and constant min-entropy. Let $\xi \in \mathbf{F}_q$ be a random variable satisfying,

$$c_k = \mathbf{P}(\xi = k) \leq C/q, \forall k \in \mathbf{F}_q, \tag{1}$$

where $C \geq 1$ is a given constant. We refer to ξ satisfying equation 1 as near uniform. In a sense, it resembles distributions with bounded density in the characteristic zero case.

Under this stronger hypothesis we show that the error bound can be improved to q^{-cn}, by a completely elementary approach. Our other main contributions are: (1) we do not require the entries to be identically distributed, only that they each satisfy the near uniform condition, and we do not even need all the entries to be near uniform as we can allow a small linear number of entries in each column to be arbitrary (see Theorem 1.4 and Theorem 1.6); (2) our method can be extended to symmetric and alternating matrices (see Theorem 3.3 and Theorem 3.7) where we also allow many entries to be arbitrary, (3) furthermore, our method also yields interesting applications towards the uniform model itself, such as perturbation of

matrices sampled uniformly over $\mathrm{GL}_n(\mathbf{F}_q)$ (see Theorem 4.1 and Theorem 4.2), or the evolution of rank of a matrix evolved from an arbitrary matrix (see Theorem 3.5). All of these results seem to be new.

In the following subsection we detail our main result for the independent model. The statements of our results in the other models are deferred to their corresponding section.

1.1. Random matrices of independent entries

In this section, if not stated otherwise, we assume that $X_1, \dots, X_n$ are independent random vectors in $\mathbf{F}_q^n$ with independent and near uniform entries. Let M_n be the random $n \times n$ matrix with the X_i as column vectors. Denote the corank of the matrix M_n by

$$Q(M_n) := n - \mathrm{rank}(M_n).$$

It is known (see for instance [2]) that for the uniform model $M_{\mathrm{uniform},n}$, for any $0 \le k \le n$ we have

$$\mathbf{P}(Q(M_{\mathrm{uniform},n}) = k) = \frac{1}{q^{k^2}} \frac{\prod_{i=1}^{n}(1-1/q^i)\prod_{i=k+1}^{n}(1-1/q^i)}{\prod_{i=1}^{n-k}(1-1/q^i)\prod_{i=1}^{k}(1-1/q^i)}.$$

Let Q_∞ be the "limiting" random variable with

$$\mathbf{P}(Q_\infty = k) = \frac{1}{q^{k^2}} \frac{\prod_{i=k+1}^{\infty}(1-1/q^i)}{\prod_{i=1}^{k}(1-1/q^i)}.$$

The following was shown by Fulman and Goldstein [17]

$$\frac{1}{8q^{n+1}} \le \|Q(M_{\mathrm{uniform},n}) - Q_\infty\|_{TV} \le \frac{3}{q^{n+1}}. \tag{2}$$

This is a striking result, but as their method compares $Q(M_{\mathrm{uniform},n})$ and Q_∞ directly, it does not seem to extend beyond the uniform model. One of the main results of this note is the following.

Theorem 1.2 (rank distribution of independent entry square matrices)**.** *There exist constants c, C' depending on C such that*

$$\|Q(M_n) - Q_\infty\|_{TV} \le \left(\frac{C'}{q}\right)^{cn}.$$

Next, motivated by problems from algebraic combinatorics to enumerate invertible matrices with given constraints of zero entries (see for instance [22] and the references therein), we consider perturbations in the following way. For each $1 \leq i \leq n$, let $F_i \subset [n]$ be the set of coordinates of X_i which are allowed to be any random variable (independent of the others). These sets F_i satisfy the following.

Condition 1. *Let α be a sufficiently small constant.*

- $|F_i| < \alpha n$.
- *Every index $i \in [n]$ is contained in at most $(1 - 12\alpha)n$ sets F_j*[1].

These conditions are necessary as our result does not hold if the matrix has (many) fixed rows or columns. One typical example is $F_i = \{i, \dots, i+\alpha n\}$ (which corresponds to the case that the entries of distance αn from the main diagonals are allowed to be deterministic, for instance all of them can be zero).

Definition 1.3. We say that X_i is a near uniform vector of *type* F_i if the entries of X_i are independent and near uniform, except those with indices in F_i, which may be arbitrary.

Theorem 1.4. *Let $F_1, \dots, F_n$ be index sets satisfying Condition 1. Assume that $X_1, \dots, X_n$ are independent random vectors, where X_i is of type F_i. Then there exist constants C', c depending on C and α such that*

$$\|Q(M_n) - Q_\infty\|_{TV} \leq \Big(\frac{C'}{q}\Big)^{cn}.$$

Note that Theorem 1.2 is a corollary of Theorem 1.4 where one can simply take the sets F_i to be empty.

Corollary 1.5. *The number of matrices $M_n \in \mathrm{Mat}_n(\mathbf{F}_q)$ of rank r, where $r \geq (1-\alpha)n$, and where the entries in $F_1, \dots, F_n$ are all zero, is asymptotically*

$$N_{F_1,\dots,F_n} = q^{n^2 - \sum_i |F_i|}\Big(Q_\infty(n-r) + O((C'/q)^{cn})\Big).$$

We remark that invertible matrices with vanishing diagonal are studied in [22]. A precise formula for the number of $n \times (m+n)$ matrices of rank n with vanishing diagonal is given in [22, Proposition 2.2]. Relatedly, we

[1]There is nothing special about the constant 12, we use it for convenience, and without intention to optimize.

show that the above estimate also holds for more general rectangular matrices (stated here without perturbations for convenience). Let $X_1, \dots, X_{n+m}$ be random independent vectors with near uniform entries in $\mathbf{F}_q^n$, and let $M_{n\times(n+m)}$ be the random $n \times (n+m)$ matrix spanned by the column vectors. It is known (again from [2]) that for the uniform model $M_{\text{uniform},n,m}$, for any $0 \le k \le n$ we have

$$\mathbf{P}(Q(M_{\text{uniform},n,m}) = k) = \frac{1}{q^{k(m+k)}} \frac{\prod_{i=1}^{n+m}(1-1/q^i)\prod_{i=k+1}^{n}(1-1/q^i)}{\prod_{i=1}^{n-k}(1-1/q^i)\prod_{i=1}^{m+k}(1-1/q^i)}.$$

Let $Q_{m,\infty}$ be "limiting" (as $n \to \infty$) random variable with

$$\mathbf{P}(Q_{m,\infty} = k) = \frac{1}{q^{k(m+k)}} \frac{\prod_{i=k+1}^{\infty}(1-1/q^i)}{\prod_{i=1}^{m+k}(1-1/q^i)}.$$

Note that [17] Fulman and Goldstein showed that the two distributions above are very close,

$$\frac{1}{8q^{n+m+1}} \le \|Q(M_{\text{uniform},n,m}) - Q_\infty\|_{TV} \le \frac{3}{q^{n+m+1}}. \tag{3}$$

Here we show

Theorem 1.6 (rank distribution of independent entry rectangular matrices)**.** *For the near uniform model $M_{n\times(n+m)}$ there exist constants c, C' depending on C such that*

$$\|Q(M_{n\times(n+m)}) - Q_{m,\infty}\|_{TV} \le \left(\frac{C'}{q}\right)^{c(m+n)}.$$

1.7. Organization of paper and proof method

All our results are obtained through a column or column and row exposure process. See Figures 1 and 2 as well as Claim 4.3.

This exposure process has two regimes. For the first linear stretch of the process we have with high probability full or nearly full rank. See Lemma 2.2, Proposition 3.10 and Lemma 4.4 for the precise statements in the different contexts. All of these results are similar to Odlyzko's lemma and obtained through the same method.

For the remaining columns we show with high probability our matrix so far has no structured normal vectors. See Proposition 2.6 and Proposition 3.12. Our notion of structure is related to the concentration probability

in [30]. See their Equation (10). Our goal with the structure is similar but less technical and easier. They contain their structured set in a GAP and we simply avoid a bad set $\mathcal{B}$. See Lemma 2.9 which is used in both the independent and symmetric cases.

This structure on the normal vectors to the column space implies a column is contained in the column space with probability close to what one would predict from the uniform model. This is shown in Lemma 2.11.

For the independent and alternating models these observations suffice. But for the symmetric model the rank sometimes increases by one or two so we need to understand quadratic forms. This analysis uses the decoupling lemma and is done in Lemma 3.18 of Section 3.

In Section 2 we show our results for the independent model. In Section 3 we show our results for the symmetric and alternating models. In Section 4 we give the perturbed and corner model of a random matrix and give our results for that model.

1.8. Notation

We use $M_n = (M_n(ij))_{1\le i,j\le n}$ to denote a square matrix of size n with entries $M_n(ij)$ from $\mathbf{F}_q$. This matrix will be either non-symmetric, symmetric, or alternating of near uniform entries or a perturbed invertible matrix depending on the section.

We denote by $r_i(M_n), c_j(M_n)$ the i^{th} row and j^{th} column of M_n respectively. When M_n is one of our models of random matrix and N is a fixed $k \times k$ matrix we let $M_n(N)$ denote a matrix with upper left corner N and other entries distributed like those of M_n.

We denote by $W_k(M_n)$ the column span of the first k columns of M_n. We write simply W_k when the matrix is clear.

Let $\mathbf{e}_i$ denote the i-th basis vector $(0,\dots,0,1,0,\dots,0)$.

We write $\mathbf{P}$ for probability and $\mathbf{E}$ for expected value. Sometimes we write $\mathbf{P}_X(.)$ to emphasize that the probability is taken with respect to X. For an event $\mathcal{E}$, we write $\bar{\mathcal{E}}$ for its complement.

For a given index set $J \subset [n]$ and a vector $X = (x_1,\dots,x_n)$, we write $X|_J$ or sometimes X_J to be the subvector of X of components indexed from J. Similarly, if H is a subspace then $H|_J$ or H_J is the subspace spanned by $X|_J$ for $X \in H$.

For $W \subset V$ a subspace we write $\overline{W}$ for the complement of W in V and $W^\perp$ for the orthogonal complement.

Finally we let

$$e_q(t) := \exp(2\pi i \operatorname{tr}(t)/p)$$

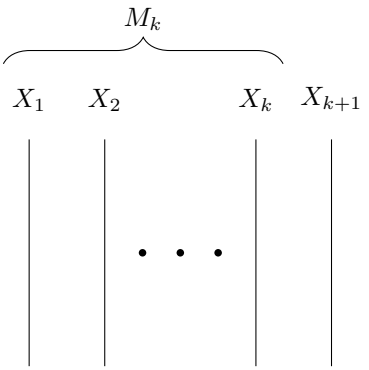

Figure 1: column exposure process.

where p is the characteristic of $\mathbf{F}_q$. Note that the trace is equally distributed in $\mathbf{F}_p$ so for every $a \in \mathbf{F}_q$ we have,

$$1_{a=0} = \frac{1}{q} \sum_{t \in \mathbf{F}_q} e_q(at). \tag{4}$$

2. Independent entry random matrices: proof of Theorem 1.4 and Theorem 1.6

We first note that throughout the paper the constants c, C' in $(C'/q)^{cn}$ might differ from case to case.

Let W_k be the subspace generated by $X_1, \ldots, X_k$ and M_k be the matrix with $X_1, \ldots X_k$ as columns. The approach is a column exposure process illustrated by Figure 1.

The idea is to show that the rank statistics evolve similarly to how they evolve in the uniform case when we expose the columns $X_1, \ldots, X_n$ one by one. Note that in the uniform case we have

$$\begin{aligned} \mathbf{P}_{X_{k+1}}\Big(\operatorname{rank}(M_{k+1}) = \operatorname{rank}(M_k) | W_k \wedge \operatorname{rank}(M_k) = l\Big) \\ = \mathbf{P}(X_{k+1} \in W_k | W_k \wedge \operatorname{rank}(M_k) = l) \\ = \frac{|W_k|}{q^n} = \frac{1}{q^{n-l}}. \end{aligned} \tag{5}$$

To show that the rank for the near uniform model evolves asymptotically as in (5), there are two regimes to understand. We first show that for $k \leq$

$(1-\varepsilon)n$, M_k is full rank with high probability. This is Lemma 2.3. For $k \geq (1-\varepsilon)n$, M_k is no longer full rank with high probability. We show the probability an additional column increases the rank matches with the uniform model by combining Proposition 2.6 and Lemma 2.11. These pieces are combined to prove the main theorems at the end of the section.

2.1. Full rank for thin matrices and non-sparsity of normal vectors

We first show that if $k < (1-6\alpha)n$ then with high probability our vectors X_i are independent.

Lemma 2.2 (Odlyzko)**.** *Let V be a subspace of codimension d in $\mathbf{F}_q^n$ and X be a random vector with independent near uniform entries except for up to k entries which may have arbitrary distribution. Then,*

$$\mathbf{P}(X \in V) \leq \left(\frac{C}{q}\right)^{d-k}.$$

Proof. Let $V = \langle v_1, \dots, v_{n-d}\rangle$ and consider the $n \times (n-d)$ matrix with the v_i as columns. Without loss of generality suppose the first $n-d$ rows are linearly independent. Restricting to the first $n-d$ rows we have $X|_{[n-d]} = c_1 v_1|_{[n-d]} + \cdots + c_{n-d} v_{n-d}|_{[n-d]}$ for unique c_i.. If $X \in V$ then $X = c_1 v_1 + \cdots + c_{n-d} v_{n-d}$. This implies we have equality in the last d coordinates of X. By assumption at least $d-k$ of the last d entries of X are independent and equal a fixed value with probability at most C/q. The result follows. □

Lemma 2.3. *Let α be as in Condition 1. With probability at least $1 - n(C/q)^{5\alpha n}$, the matrix generated by $X_1, \dots, X_{\lfloor(1-6\alpha)n\rfloor}$ has full rank.*

Proof. Recall that X_i has type F_i and $|F_i| < \alpha n$. If the vectors $X_1, \dots, X_{\lfloor(1-6\alpha)n\rfloor}$ do not have full rank then for some i we have,

$$X_i \in \langle X_1 \dots X_{i-1}\rangle \text{ and } X_1 \dots X_{i-1} \text{ are linearly independent.}$$

By Lemma 2.2 we can bound the probability of this event by,

$$\begin{aligned}\mathbf{P}(X_i \in \langle X_1 \dots X_{i-1}\rangle | X_1 \dots X_{i-1} \text{ are lin. ind.}) &\leq (C/q)^{n-i+1-n\alpha} \\ &\leq (C/q)^{n-(\lfloor(1-6\alpha)n\rfloor - n\alpha} \\ &\leq (C/q)^{5\alpha n}.\end{aligned}$$

Taking the union bound over all $(1-6\alpha)n$ places where the rank may drop we get the desired bound. □

Therefore, with a loss of at most $n(C/q)^{5\alpha n}$ in probability it suffices to assume that $W_{(1-6\alpha)n}$ has full rank.

Next we show that W_k does not have a "sparse" normal vector for $k \geq (1-6\alpha)n$ with high probability. This is important to the argument because we allow some entries of our matrix to be deterministic. If we had a normal vector $\mathbf{a}$ with support contained in one of our bad sets $|F_i|$ we wouldn't be able to estimate the probability $X_{k+1} \cdot \mathbf{a} = 0$. Roughly speaking, these events are needed towards establishing any asymptotic form of (5) as the event $X_{k+1} \in W_k$ is equivalent with $X_{k+1} \cdot \mathbf{a} = 0$ for all normal vectors $\mathbf{a}$ of W_k.

Lemma 2.4. *Let α be as in Condition 1. Then there exist constants C', c such that the following hold. For any $(1-6\alpha)n \leq k \leq n$, with probability at least $1-(C'/q)^{cn}$ the following holds: any normal vector $\mathbf{a}$ of W_k has at least αn non-zero coordinates.*

Proof. Let $\mathbf{a}$ be a prospective normal vector with $S = \mathbf{supp}(\mathbf{a})$, and $s = |S| \leq \alpha n$. There are less than $2^n q^{\alpha n}$ such vectors. For each $1 \leq i \leq k$, if we have that $\bar{F}_i \cap S \neq \emptyset$ (that is $S \not\subset F_i$), then by (1), it is clear that

$$\mathbf{P}(X_i \cdot \mathbf{a} = 0) \leq C/q.$$

Letting k' denote the number of such indices, by Condition 1 we see that

$$k' \geq k - (1-12\alpha)n \geq 6\alpha n.$$

Hence the probability that there exists a sparse $\mathbf{a}$ serving as the normal vector to W_k is bounded by

$$\sum_{\mathbf{a}, |\mathbf{supp}(\mathbf{a})| \leq \alpha n} \mathbf{P}(X_i \cdot \mathbf{a} = 0, 1 \leq i \leq \lfloor 6\alpha n \rfloor) \leq 2^n q^{\alpha n} (C/q)^{\lfloor 6\alpha n \rfloor} \leq (C'/q)^{cn}.$$

□

In the next step we show that with high probability the normal vectors are not only sparse but do not have structures (see Proposition 2.6). Thanks to the near uniform assumption, our method is much simpler compared to those of [23, 26, 31, 32, 41].

2.5. Anti-concentration probability

We first introduce concentration inequalities for vectors with randomness from (1). Let $X = (x_1, \dots, x_n)$ be a random vector of type F. Let $c_k^i = \mathbf{P}(x_i = k)$, where by definition we know that $c_k^i \leq C/q$.

Let X be a near uniform random vector of type F. By the Fourier transform for any $r \in \mathbf{F}_q$ and fixed $\mathbf{a} \in F_q^n$ we have

$$\mathbf{P}(X \cdot \mathbf{a} = r) = \mathbf{E}(1_{X \cdot \mathbf{a} = r}) = \frac{1}{q} \sum_{t \in \mathbf{F}_q} \prod_{i \notin F} \mathbf{E} e_p(\mathrm{tr}(x_i a_i t)) e_p(-\mathrm{tr}(r' t))$$

where r' depends on r and other deterministic entries of X. By the triangle inequality

$$\begin{aligned} |\mathbf{P}(X \cdot \mathbf{a} = r) - \frac{1}{q}| &\leq \frac{1}{q} \sum_{t \in \mathbf{F}_q, t \neq 0} \prod_{i \notin F} |\mathbf{E} e_p(\mathrm{tr}(x_i a_i t))| \\ &= \frac{1}{q} \sum_{t \in \mathbf{F}_q, t \neq 0} \prod_{i \notin F} \Big| \sum_{k \in \mathbf{F}_q} c_k^i e_p(\mathrm{tr}(k a_i t)) \Big| \\ &=: \frac{1}{q} \sum_{t \neq 0} \prod_{i \notin F} f_i(t a_i). \end{aligned} \tag{6}$$

Where $f_i(t) = |\sum_{k \in \mathbf{F}_q} c_k^i e_p(\mathrm{tr}(kt))|$. Motivated by this, for convenience we define the following,

$$\rho_F(\mathbf{a}) := \frac{1}{q} \sum_{t \neq 0} \prod_{i \notin F} f_i(t a_i).$$

Conceptually $\rho_F(\mathbf{a})$ measures the "structure" of $\mathbf{a}$. Vectors with smaller ρ_F are "less structured". The concept is a waypoint in the proof. At a high level our approach is to show that normal vectors are likely unstructured and unstructured vectors are orthogonal to near uniform vectors with probability close to uniform.

The goal of this subsection is to show the following proposition which states that with high probability the column space of W_k has no structured normal vectors.

Proposition 2.6. *Given C and α where α is sufficiently small. There exist constants C', c such that the following holds for $m \geq (1 - 6\alpha)n$: with probability at least $1 - (C'/q)^{cn}$ with respect to $X_1, \ldots, X_m$ any nonzero vector $\mathbf{a} = (a_1, \ldots, a_n)$ orthogonal to each X_i has,*

$$\rho_{F_{m+1}}(\mathbf{a}) \leq (C'/q)^{cn}.$$

To prove the above result, observe that if $a \neq 0$ then, noting that $\sum_{t\in\mathbf{F}_q} e_p(\mathrm{tr}(at)) = q1_{a=0}$, we have

$$\begin{aligned}\sum_{t\in\mathbf{F}_q} f_i(ta)^2 &= \sum_{t\in F_q}\left|\sum_k c_k^i e_p(\mathrm{tr}(kta))\right|^2\\ &= \sum_{t\in F_q}\left[\sum_k (c_k^i)^2 + \sum_{k\neq k'} c_k^i c_{k'}^i e_p(\mathrm{tr}((k-k')ta))\right]\\ &= q\sum_k (c_k^i)^2 \leq q\max_k c_k^i \sum_k c_k^i \leq C.\end{aligned}$$

As such, for any $K > 0$, let $T_i \subset \mathbf{F}_q$ be the set of a where $|f_i(a)| \geq Kq^{-1/2}$ then

$$|T_i| \leq Cq/K^2. \tag{7}$$

This implies the following claim.

Claim 2.7. *For any* M *if* $\mathbf{a} = (a_1, \ldots, a_n)$ *is such that for any* $t \neq 0$*, it has at least* M *indices* $i \in [n]\backslash F$ *such that* $ta_i \notin T_i$*. Then we have*

$$\rho_F(\mathbf{a}) = \frac{1}{q}\sum_t \prod_{i\notin F} f_i(ta_i) \leq (Kq^{-1/2})^M. \tag{8}$$

Thus towards Proposition 2.6, it suffices to show the following and then choose suitable M and K.

Lemma 2.8. *With probability at least* $1 - q^{6\alpha n}(2C)^n(C/K^2)^{n-M-|F|}$ *with respect to* $X_1, \ldots, X_m$*, any normal vector* $\mathbf{a}$ *of* W_m *satisfies the condition in Claim 2.7.*

To prove Lemma 2.8 we bound the number of vectors not satisfying equation (8). Then we bound the probability each of those vectors is a normal vector to our matrix. Counting the bad set is also important in the symmetric section so we split it into the following lemma.

Lemma 2.9 (Counting lemma)**.** *Let* $\mathcal{B}$ *be the set of vectors in* $\mathbf{F}_p^n$ *where*

$$\frac{1}{q}\sum_t \prod_{i\notin F} f_i(ta_i) \geq (Kq^{-1/2})^M.$$

Then

$$|\mathcal{B}| \leq q^{n+1}2^n(C/K^2)^{n-M-|F|}.$$

Proof. By Claim 2.7 for any $\mathbf{a} \in \mathcal{B}$ there exists t_0 such that $t_0 a_i \in T_i$ for at least $n - M - |F|$ indices i and $T_i| \leq (Cq/K^2)$ by 7. The other $M + |F|$ entries of $\mathbf{a}$ can be any element of $\mathbf{F}_q$. Thus

$$\begin{aligned}|\mathcal{B}| &\leq q2^n(Cq/K^2)^{n-M-|F|}q^{M+|F|} \\ &\leq q^{n+1}2^n(C/K^2)^{n-M-|F|}.\end{aligned}$$

□

Proof of Lemma 2.8. We estimate the probability of the complement event. For a given $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{B}, \mathbf{a} \neq 0$, we estimate the probability that it is a normal vector to W_m. By Lemma 2.4, we just need to focus on the event that $\mathbf{a}$ has at least αn non-zero entries. As such, by Condition 1, for each $1 \leq i \leq m$ we have

$$\mathbf{P}(X_i \cdot \mathbf{a} = 0) \leq C/q.$$

So the probability that $\mathbf{a}$ is a normal vector to W_m is bounded by

$$\prod_{i=1}^{m} \mathbf{P}(\mathbf{a} \cdot \mathbf{X}_i = 0) \leq (C/q)^m.$$

Taking a union bound over $\mathcal{B}$, we obtain an upper bound for the probability of the existence of structured normal vectors,

$$q^{n+1}2^n(C/K^2)^{n-M-|F|} \times (C/q)^{n-6\alpha n} \leq q^{6\alpha n}(2C)^n(C/K^2)^{n-M-|F|}. \tag{9}$$

□

Proof of Proposition 2.6. Choose $M = \beta n, K = q^\beta$ with small $\beta < 1/2$ and with even smaller α. We obtain a bound $(C'/q)^{-cn}$ for the bound on probability in Equation (9), and for the bound on structure in Claim 2.7. □

2.10. The rank statistics in the exposure process

The motivation to consider structure is if all vectors orthogonal to a given subspace have small ρ then the probability it contains a near uniform random vector behaves like the probability it contains a uniform random vector.

Lemma 2.11. *Let $H \subset \mathbf{F}_q^n$ be a fixed subspace of codimension d such that for every nonzero $w \in H^\perp$ $|\mathbf{P}(X \cdot w = 0) - 1/q| < \delta$ and let X be a near uniform random vector of type F as above. Then*

$$|\mathbf{P}(X \in H) - 1/q^d| \leq 2\delta.$$

Note when applying the lemma it suffices to bound $\rho_F(w)$ because $|\mathbf{P}(X \cdot w = 0) - 1/q| \leq \rho_F(w)$.

We also refer the reader to [23] for a variant for fields of prime order.

Proof. We have the following identity,

$$1_{a_1=0 \wedge \cdots \wedge a_d=0} = \frac{1}{q^d} \sum_{t_1,\ldots,t_d \in \mathbf{F}_q} e_q(a_1 t_1 + \cdots + a_d t_d).$$

Therefore, letting $\mathbf{v}_1, \ldots, \mathbf{v}_d$ be a basis for $H^\perp$,

$$\begin{aligned}
\mathbf{P}(X \in H) &= \mathbf{P}(X \cdot \mathbf{v}_1 \wedge \cdots \wedge X \cdot \mathbf{v}_d) \\
&= \mathbf{E}\left[\frac{1}{q^d} \sum_{t_1,\ldots,t_d \in \mathbf{F}_q} e_q(t_1[X \cdot \mathbf{v}_1] + \cdots + t_d[X \cdot \mathbf{v}_d])\right] \\
&= \mathbf{E}\left[\frac{1}{q^d} + \frac{1}{q^d} \sum_{\substack{t_1,\ldots,t_d \in \mathbf{F}_q \\ \text{not all } t_i \text{ are zero}}} e_q(t_1[X \cdot \mathbf{v}_1] + \cdots + t_d[X \cdot \mathbf{v}_d])\right] \\
&= \frac{1}{q^d} + \frac{1}{q^d} \sum_{\substack{t_1,\ldots,t_d \in \mathbf{F}_q \\ \text{not all } t_i \text{ are zero}}} \mathbf{E} e_q(t_1[X \cdot \mathbf{v}_1] + \cdots + t_d[X \cdot \mathbf{v}_d]).
\end{aligned}$$

We now split the sum into projective equivalence classes. Let $\sim$ be the equivalence relation given by $(t_1, \ldots, d_d) \sim (t'_1, \ldots, t'_d)$ if there exists $t \neq 0$ such that $(t_1, \ldots, t_d) = (t \cdot t'_1, \ldots, t \cdot t'_d)$. Not worrying about our choice of representative on account of our inner sum we can continue,

$$\begin{aligned}
&= \frac{1}{q^d} + \frac{1}{q^d} \sum_{(t_1,\ldots,t_d) \in (\mathbf{F}_q^n)^\times/\sim} \left[\sum_{t \in \mathbf{F}_q^\times} e_q(t(t_1[X \cdot \mathbf{v}_1] + \cdots + t_d[X \cdot \mathbf{v}_d]))\right] \\
&= \frac{1}{q^d} + \frac{1}{q^d} \sum_{(t_1,\ldots,t_d) \in (\mathbf{F}_q^n)^\times/\sim} q\left[\frac{1}{q} \sum_{t \in \mathbf{F}_q^\times} e_q\left(t\left(X \cdot \sum_{i=1}^n t_i \mathbf{v}_i\right)\right)\right].
\end{aligned}$$

Now observe,

$$\left[\frac{1}{q}\sum_{t\in\mathbf{F}_q^\times} e_q\left(t\left(X\cdot\sum_{i=1}^n t_i\mathbf{v}_i\right)\right)\right] = \mathbf{P}\left(X\cdot\sum_{i=1}^n t_i\mathbf{v}_i = 0\right) - 1/q.$$

The $\mathbf{v}_i$ are a basis and the t_i are not all zero so $\sum_{i=1}^n t_i\mathbf{v}_i$ is a nonzero element of $H^\perp$. By assumption then this is bounded by δ. There are $\frac{q^d-1}{q-1}$ elements in $(t_1,\dots,t_d)\in(\mathbf{F}_q^n)^\times/\sim$. We have,

$$\frac{q(q^d-1)}{q^d(q-1)} \le 2.$$

So by the triangle inequality we obtain,

$$|\mathbf{P}(X\in H) - 1/q^d| \le 2\delta.$$

□

Let $\mathcal{E}_m$ be the event considered in Proposition 2.6, namely the event that all normal vectors $\mathbf{a}$ have $\rho(\mathbf{a}) < (C'/q)^{cn}$. By the proposition we know that

$$\mathbf{P}(\mathcal{E}_m) \ge 1 - (C'/q)^{cn}.$$

Proposition 2.12. *Assume that $(1-6\alpha)n \le m \le n$ and the non-symmetric matrix M_n is as in Theorem 1.2. For each $l \le m$ we have*

$$\left|\mathbf{P}(X_{m+1}\in W_m|\mathrm{rank}(W_m) = l\wedge\mathcal{E}_m) - (1/q)^{n-l}\right| \le (C'/q)^{cn}.$$

In other words this result confirm that our rank evolution matches with Equation (5) of the uniform model.

Proof. By Proposition 2.6 on $\mathcal{E}_m$, for any $\mathbf{a}\perp W_n$, we have $\rho_{F_{m+1}}(\mathbf{a}) \le (1/q)^{cn}$. Therefore by Lemma 2.11,

$$|\mathbf{P}(X_{m+1}\in W_m|\mathrm{rank}(W_m) = l\wedge\mathcal{E}_m) - (1/q)^{n-l}| \le (C'/q)^{cn}.$$

□

Now we prove our main results for non-symmetric matrices.

Proof of Theorem 1.4. We first condition on the events in Lemma 2.3, Lemma 2.4, and on $\wedge_{m\ge\lfloor(1-6\alpha)n\rfloor}\mathcal{E}_m$ from Proposition 2.6. For each $k \ge \lfloor(1-6\alpha)n\rfloor =: m_0$, the event $Q(M_n) = n-k$ can be written as follows using the

column exposure process

$$\begin{aligned}&\mathbf{P}(\operatorname{rank}(M_n) = k)\\ &= \sum_{0<i_1<\cdots<i_{k-m_0}\le n-m_0} \mathbf{P}\big(\wedge_j X_{m_0+i_j} \notin W_{m_0+i_j-1} \wedge \text{no other rank increase}\big).\end{aligned}$$

Now using Proposition 2.12 we can estimate the RHS above as

$$\begin{aligned}&= \sum_{0<i_1<\cdots<i_{k-m_0}<n-m_0} \mathbf{P}(\text{a uniform matrix of size } n-m_0 \text{ drops rank}\\ &\qquad\qquad \text{at } i_1,\dots,i_{k-m_0}) + O(n(C'/q)^{cn})\\ &= \mathbf{P}(\text{a uniform matrix of size } n-m_0 \text{ has rank } k) + O(n(C'/q)^{cn})\\ &= Q_\infty(n-k) + O\Big(1/q^{n-m_0} + n(C'/q)^{cn}\Big),\end{aligned}$$

where we used (2) in the last estimate. □

Proof of Theorem 1.6. For this we observe that the column rank is equal to the row rank and consider the $(n+m)\times n$ transpose of M. Our proof of Theorem 1.4 shows that as we expose the columns the rank statistics match up to an error of type $(C'/q)^{c(m+n)}$ at each step. □

3. Random symmetric and alternating matrices

3.1. Random symmetric matrices

Now we discuss our result for symmetric matrices. Let the entries $m_{ij}, i \ge j$ of M_n be independent near uniform random variables. It is known (from [7, 24]) that for the uniform model M_{uniform}, for any $0 \le k \le n$ we have

$$\mathbf{P}(Q(M_{\text{uniform},n}) = k) = \frac{1}{q^{\binom{n+1}{2}}} \prod_{i=1}^{\lfloor (n-k)/2\rfloor} \frac{q^{2i}}{q^{2i-1}} \prod_{i=0}^{n-k-1} (q^{n-i}-1).$$

Let $Q_{sym,\infty}$ be the random variable with

$$\mathbf{P}(Q_{sym,\infty} = k) = \frac{\prod_{i=0}^{\infty}(1-q^{-2i-1})}{\prod_{l=1}^{k}(q^i-1)}.$$

In the uniform model of symmetric matrices Fulman and Goldstein [17] showed

$$\begin{aligned} \frac{0.18}{q^{n+1}} &\leq \|Q(M_{\text{uniform},n}) - Q_\infty\|_{TV} \leq \frac{2.25}{q^{n+1}} \text{ for } n \text{ even} \\ \frac{0.18}{q^{n+2}} &\leq \|Q(M_{\text{uniform},n}) - Q_\infty\|_{TV} \leq \frac{2}{q^{n+2}} \text{ for } n \text{ odd.} \end{aligned} \tag{10}$$

We show that

Theorem 3.2. *For the near uniform symmetric model M_n*

$$\|Q(M_n) - Q_{sym,\infty}\|_{TV} \leq \Big(\frac{C'}{q}\Big)^{cn}.$$

Next, as in the iid case, we can extend the result to perturbations. More precisely we will sample M_n, a random symmetric matrix of size n in the following way. All the entries on and above its diagonal are near uniform and independent except for those falling in index sets $F_i \subseteq [n]$ which satisfy the following condition:

Condition 2. *Let α be a sufficiently small constant.*

- $|F_i| < \alpha n$.
- *For symmetry, we also assume that for every $i, j \in [n]$, F_j contains i if and only if F_i contains j. In particular each index is contained in at most αn sets.*

Theorem 3.3 (rank distribution for random symmetric matrices)**.** *With M_n as above, there exist constants C', c such that*

$$\|Q(M_n) - Q_{sym,\infty}\|_{TV} \leq \Big(\frac{C'}{q}\Big)^{cn}.$$

Corollary 3.4. *The number of matrices $M_n \in \text{Mat}_n(\mathbf{F}_q)$ of rank r, where $r \geq (1-\alpha)n$, and where the entries in $F_1, \dots, F_n$ are all zero, is asymptotically*

$$N_{F_1,\dots,F_n} = q^{\binom{n+1}{2} - \sum_i |F_i|}\Big(Q_{sym,\infty}(n-r) + O((C'/q)^{cn})\Big).$$

In particular, when $F_1 = \{1\}, \dots, F_n = \{n\}$ then the number of full rank symmetric matrices with zero entries on the diagonal is $q^{\binom{n}{2}}\big(Q_{sym,\infty}(0) +$

$O(1/q^{cn})\big)$, *which is asymptotically the number of full rank symmetric matrices of size* $n-1$. *(In fact these two quantities are the same, see [22, Theorem 3.3].) Similarly one can also deduce an asymptotic formula for the number of invertible matrices of the first* k *entries zero (and so on), see also [22, Theorem 4.25].*

For all these results we depend on the following special case of Theorem 3.3 and Theorem 3.7. We state it as a theorem because it may be of independent interest.

Theorem 3.5. *Let* M_{m_0} *be an arbitrary matrix of size* m_0 *which is symmetric (or alternating). Let* $M_n^u(M_{m_0})$ *be the* $n \times n$ *random matrix with* M_{m_0} *as its upper left hand corner and the remaining entries above the diagonal sampled uniformly and those below chosen to make the matrix symmetric (or alternating). Then,*

$$\|Q(M_n^u(M_{m_0})) - Q_\bullet\|_{TV} \leq \frac{3^{n/2}}{q^{n/2-m_0}}.$$

Where $Q_\bullet$ *is either* $Q_{sym,\infty}$, $Q_{alt,e}$ *or* $Q_{alt,o}$ *depending on whether* M_n *is symmetric or alternating and the parity of* n.

We remark that the range of m_0 above is nearly optimal as the result no longer holds for $m_0 \geq (1 + o(1))n/2$. To see how this result is used to deduce Theorem 3.3 and Theorem 3.7 see Proposition 3.21.

3.6. Random alternating matrices (skew symmetric matrices)

Here we assume q is odd. For alternating matrices $M_n^T = -M_n$. It is well-known that the rank is even and is implied for example by Proposition 3.17. It is known (from [8]) that for the uniform model $M_{\text{uniform},n}$, for any $0 \leq k \leq n$ of the same parity with n we have

$$\mathbf{P}(Q(M_{\text{uniform},n}) = k) = \frac{1}{q^{\binom{n}{2}}} \prod_{i=1}^{(n-k)/2} \frac{q^{2i-2}}{q^{2i}-1} \prod_{i=0}^{n-k-1} (q^{n-i}-1).$$

Let $Q_{alt,e}$ and $Q_{alt,o}$ be the limiting random variables given by

$$\mathbf{P}(Q_{alt,e} = k) := \begin{cases} \prod_{i=0}^{\infty}(1-q^{-2i-1})\frac{q^k}{\prod_{i=1}^{k}(q^i-1)}, & k \text{ is even} \\ 0, \ k \text{ is odd} \end{cases}$$

and

$$\mathbf{P}(Q_{alt,o} = k) := \begin{cases} \prod_{i=0}^{\infty}(1-q^{-2i-1})\frac{q^k}{\prod_{i=1}^{k}(q^i-1)}, & k \text{ is odd} \\ 0, \ k \text{ is even} \end{cases}$$

Let M_n be a random matrix with independent near uniform entries above the diagonal except for entries in index sets F_i satisfying 2. The entries on the diagonal are zero and the entries below the diagonal are the negative of the entries above the diagonal.

Theorem 3.7 (Rank distribution for random alternating matrices)**.** *There exist constants c, C' depending on C such that*

$$\|(Q(M_n) - Q_{alt,e}\|_{TV} \leq \left(\frac{C'}{q}\right)^{cn} \text{ if } n \text{ is even}$$

and

$$\|(Q(M_n) - Q_{alt,o}\|_{TV} \leq \left(\frac{C'}{q}\right)^{cn} \text{ if } n \text{ is odd.}$$

Note that for the uniform model of random alternating matrices Fulman and Goldstein [17] showed

$$\begin{aligned} \frac{0.18}{q^{n+1}} \leq \|Q(M_{\text{uniform},n}) - Q_\infty\|_{TV} \leq \frac{1.5}{q^{n+1}} \text{ for } n \text{ even.} \\ \frac{0.37}{q^{n+1}} \leq \|Q(M_{\text{uniform},n}) - Q_\infty\|_{TV} \leq \frac{2.2}{q^{n+1}} \text{ for } n \text{ odd.} \end{aligned} \tag{11}$$

3.8. Corner exposure process

In this section, if we don't specify otherwise, the arguments will work for both M_n symmetric or alternating simultaneously. We let M_m be the upper left minor of size m of M_n. (Hence the notation is different from the non-symmetric case.)

Some of our ideas here are motivated by [10, 24, 28, 29, 38], and in particular [28][2]. The basic idea of our approach in the symmetric and alternating cases is similar to our approach in the independent case. But instead of considering a column exposure process, to preserve the independence between what we have and haven't seen, we expose one row and column at a time. In other words at the m^{th} step we will have observed the upper left M_m block and we'll be asking for the effect on the corank of expanding this block by one in both directions.

[2]The main result of this paper has been improved substantially in [33].

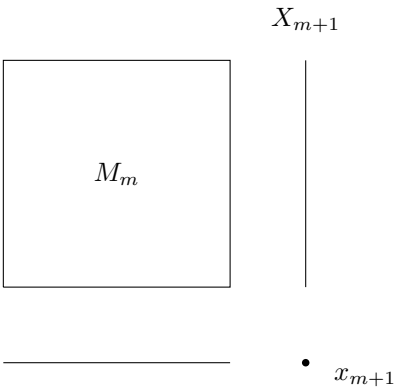

Figure 2: corner exposure process.

Throughout this section we will use $X = X_{m+1}$ to refer to the top m cells of column $m+1$ when we are at step m, that is $X = \mathbf{c}_{m+1}(M)|_{[m]}$. We also let x_{m+1} be the entry $M_n(m+1, m+1)$, and H_m be the column span of the exposed block at step m. This is summarized by Figure 2.

The principle difference between the models is that if $X \in H_m$ then in the alternating model M_{m+1} has the same rank as M_m. This is an elementary fact shown in Proposition 3.17. In the symmetric model when $X \in H_m$ the rank typically increases by 1, but it depends on x_{m+1}. In Lemma 3.18 we show the probability the rank increases by 1 is what we'd predict from the uniform model.

3.9. Almost full rank for M_m and non-sparsity of generalized normal vectors

This subsection is similar to Subsection 2.1 where we show the following Odlyzko's lemma for random symmetric and alternating matrices, here we assume

$$\sqrt{\alpha}n \le m \le n$$

where we recall that α is sufficiently small.

Lemma 3.10. *Let ε be positive constant that is small but larger than $\sqrt{\alpha}$. We have*

1. $\operatorname{rank}(M_m) \ge (1-\varepsilon)m$ *with probability* $1 - m(C/q)^{(\varepsilon-\sqrt{\alpha})m}$.

2. *(non-sparsity of solution vectors) Let Y_0 be a fixed vector in $\mathbf{F}_q^m$. The following holds with probability at least $1-(C'/q)^{(1-3\varepsilon-\sqrt{\alpha})m}$: any nonzero vector $\mathbf{v}$ for which $M_m\mathbf{v}$ agrees with Y_0 in at least $(1-\varepsilon)m$ coordinates must have at least εm non-zero components.*

We will denote the intersection of these events by $\mathcal{F}_m$. Also, for convenience, if $\mathbf{v}$ satisfies (2) then we say that $\mathbf{v}$ is a *generalized normal vector* of M_m (with parameter ε).

Proof. (1) Let $k \le m$ be an intermediate step. If $\operatorname{corank}(M_k) \ge r$ then by Lemma 2.2 (where we recall that the number of deterministic components is at most $\alpha n \le \sqrt{\alpha}m$) with probability at least $1-(C/q)^{(r-\sqrt{\alpha}m)}$ we have X_{k+1} is not in the column span of M_k which implies $\operatorname{rank}(M_{k+1}) = \operatorname{rank}(M_k)+2$.
If M_m has corank greater than εm there must be an intermediate step M_k where the corank is greater than or equal to εm but the rank doesn't increase by 2. Taking a union bound over the possible indices we get,

$$\mathbf{P}(\operatorname{corank}(M_m) \ge \varepsilon m) \le m\Big(\frac{C}{q}\Big)^{(\varepsilon-\sqrt{\alpha})m}.$$

(2) The number of candidate sparse vectors is,

$$\binom{m}{\varepsilon m} q^{\varepsilon m}.$$

Note that the probability such a vector dotted with row i is equal to any fixed value is bounded above by $\frac{C}{q}$ provided that F_i does not contain the support of the vector. By Condition 2 at most $\sqrt{\alpha}m$ rows do. We don't have independence among all $\mathbf{v}\cdot X_i$ but we do have independence between $\mathbf{v}\cdot X_i$ and the other columns if i is not in the support of $\mathbf{v}$. So because the support is assumed to be at most εm the probability of equality is bounded above by,

$$\binom{m}{\varepsilon m} q^{\varepsilon m}\Big(\frac{C}{q}\Big)^{(1-2\varepsilon-\sqrt{\alpha})m} \le \Big(\frac{C'}{q}\Big)^{(1-3\varepsilon-\sqrt{\alpha})m}.$$

Provided ε, α are sufficiently small. □

3.11. Anti-concentration probability

Here and later we continue to assume that $\sqrt{\alpha}n \le m \le n$. Our next step is similar to Subsection 2.5 where we will show that generalized normal vectors

do not have structures with very high probability. We next recall from Equation (6) that $\rho_F(\mathbf{a}) = \frac{1}{q}\sum_{t\neq 0}\prod_{i\notin F} f(ta_i)$, where $f(y) = |\sum_k c_k e_q(ky)|$, and

$$\left|\sup_a \mathbf{P}(X\cdot\mathbf{a} = a) - 1/q\right| \leq \rho_F(\mathbf{a}).$$

There is the following analog of Proposition 2.6 for generalized normal vectors. Loosely, it states that large symmetric near uniform matrices are unlikely to have structured vectors orthogonal to their row span. By symmetry the same result applies to the column span.

Proposition 3.12. *Assume that α is sufficiently small. There exist constants c_1, c_2, C' such that with probability $1-(C'/q)^{c_1 m}$ every nonzero generalized normal vector $\mathbf{a}$ of M_m of parameter $\varepsilon = 2\sqrt{\alpha}$ has*

$$\rho_F(\mathbf{a}) \leq (C'/q)^{c_2 m}.$$

Proof. Let $c_2 = (1/2-\beta)\beta$ where β is the constant in the proof of Proposition 2.6 (chosen appropriately, for instance $\beta = 8\sqrt{\alpha}$ would work). Let $\mathcal{B}$ denote the set of $\mathbf{v}\in\mathbf{F}_q^m$ with $\rho_F(\mathbf{v}) > (C'/q)^{c_2 m}$. From Lemma 2.9 we know,

$$|\mathcal{B}| < (2C')^m q^{m-\beta m/2}.$$

Without loss of generality we assume that the first $(1-\varepsilon)m$ entries of $M_m\mathbf{a}$ are zero, and our aim is to bound this event,

$$\sum_{\mathbf{a}\in\mathcal{B}\setminus\{0\}} \mathbf{P}\Big((M_m\mathbf{a})_i = 0, 1\leq i\leq (1-\varepsilon)m\Big).$$

For all $\mathbf{a}\in\mathcal{B}\setminus\{0\}$ we have the simple bound,

$$\mathbf{P}\Big((M_m\mathbf{a})_i = 0, 1\leq i\leq (1-\varepsilon)m\Big) \leq (C/q)^{(1-\varepsilon)m-1-\sqrt{\alpha}m}. \tag{12}$$

Because $\mathbf{a}\neq 0$ it has some entry $a_i\neq 0$. Conditioning on all the entries of M_m outside of the i^{th} column and row the probability that the j^{th} row of M_m has $r_j(M_m)\cdot Z = 0$, provided $j\notin F_i$, is bounded above by C/q because of the near uniformity of the entry $M_m(j,i)$. The $M_m(j,i)$ for j fixed and not equal to i are independent so we get inequality (12).

So we obtain,

$$\sum_{\mathbf{a}\in\mathcal{B}\setminus\{0\}} \mathbf{P}(\mathbf{a} \text{ is a generalized normal vector of } M_m)$$
$$\leq \binom{m}{\lfloor(1-\varepsilon)m\rfloor}(2C')^m q^{m-\beta m/2}(C/q)^{(1-\varepsilon)m-1-\sqrt{\alpha}m}$$
$$\leq (C''/q)^{(\beta/2-3\sqrt{\alpha})m}.$$

□

3.13. The rank statistics in the exposure process

The analysis to understand the rank evolution in the corner exposure process is more involved than in the independent case. The technique hinges on the following so called Decoupling Lemma. The technique was first used to analyze symmetric random matrices by Costello, Tao and Vu [10] though the idea is old and common in number theory.

Lemma 3.14 (Decoupling)**.** *Assume that $a_{ij} \in \mathbf{F}_q$ and $a_{ij} = a_{ji}$ and $b_i \in \mathbf{F}_q$. Assume that x_i are independent random variables. Then for any index set $I \subset [m]$ we have*

$$\sup_r \left|\mathbf{P}\left(\sum_{ij} a_{ij}x_ix_j + \sum_i b_ix_i = r\right) - 1/q\right|^4 \leq \left|\mathbf{P}\left(\sum_{i\in I, j\in I^c} a_{ij}y_iy_j = 0\right) - 1/q\right|,$$

where $y_i \equiv x_i - x_i'$ with x_i' an iid copy of x_i.

Proof. For short we write $f(X) = \sum_{ij} a_{ij}x_ix_j + \sum_i b_ix_i$. We write

$$\left|\mathbf{P}(f(X) = 0) - \frac{1}{q}\right| = \left|\frac{1}{q}\sum_{t\neq 0} \mathbf{E}e_q(-tf(X))\right|.$$

We then use Cauchy-Schwarz to complete squares,

$$LHS^2 \leq \frac{q-1}{q^2}\sum_{t\neq 0} |\mathbf{E}e_q(-tf(X))|^2 \tag{13}$$
$$\leq \frac{q-1}{q^2}\sum_{t\neq 0} \mathbf{E}_{X_I}|\mathbf{E}_{X_{I^c}}e_q(-tf(X_I, X_{I^c}))|^2$$
$$= \frac{q-1}{q^2}\sum_{t\neq 0} \mathbf{E}_{X_I}\mathbf{E}_{X_{I^c}, X'_{I^c}}e_q(-t[f(X_I, X_{I^c}) - f(X_I, X'_{I^c})])$$

$$= \frac{q-1}{q^2} \sum_{t \neq 0} \mathbf{E}_{X_{I^c}, X'_{I^c}} \mathbf{E}_{X_I} e_q(-t[f(X_I, X_{I^c}) - f(X_I, X'_{I^c})]).$$

Using Cauchy-Schwarz once more, we obtain

$$\begin{aligned} LHS^4 &\leq \left(\frac{q-1}{q^2}\right)^2 (q-1) \sum_{t \neq 0} \mathbf{E}_{X_{I^c}, X'_{I^c}} \mathbf{E}_{X_I, X'_I} e_q(-t[f(X_I, X_{I^c}) - f(X_I, X'_{I^c}) \\ &\qquad - f(X'_I, X_{I^c}) + f(X'_I, X'_{I^c})]) \\ &= \left(\frac{q-1}{q^2}\right)^2 q(q-1) \frac{1}{q} \sum_{t \neq 0} \mathbf{E}_{Y_I, Y_{I_c}} e_q\left(-t \sum_{i \in I, j \in I^c} a_{ij} y_i y_j\right) \\ &= \left(\frac{q-1}{q}\right)^3 \left(\mathbf{P}\left(\sum_{i \in I, j \in I^c} a_{ij} x_i y_j = 0\right) - 1/q\right). \end{aligned}$$

□

By Lemma 3.10, with probability at least $1 - (C'/q)^{cm}$ we can assume that M_m has rank at least $(1-\varepsilon)m$. Let us consider the event that M_m has rank $m-k$ (where $k \leq \varepsilon m$).

Claim 3.15. *Assume that G_m has rank $m-k$, then there is a set $I \subset [m], |I| = m-k$ such that the principle minor matrix $G_{I \times I}$ has full rank $m-k$.*

Proof. It suffices to consider the symmetric case. Let $M_{m-k \times m}$ denote the top $m-k$ rows of M_m and assume without loss of generality that $M_{m-k \times m}$ is full rank. Then by symmetry the leftmost $m-k$ columns $M_{m \times m-k}$ are also full rank. But we know the first $m-k$ rows of $M_{m \times m-k}$ span the row space by our assumption that $M_{m-k \times m}$ is full rank. Therefore the minor $M_{m-k \times m-k}$ is full rank. □

Therefore in what follows, we may assume without loss of generality that $\mathbf{r}_1(M_m), \ldots, \mathbf{r}_{m-k}(M_m)$ span the row space of M_m.

Probability for $\operatorname{rank}(M_{m+1}) \leq \operatorname{rank}(M_m) + 1$: When we add the new column $X_{m+1} = (x_{1(m+1)}, \ldots, x_{(m+1)(m+1)}) := (x_1, \ldots, x_{m+1})$ and its transpose (or negative transpose in the alternating case), the event $\operatorname{rank}(M_{m+1}) < \operatorname{rank}(M_m) + 2$ is equivalent with the event that the extended row vector $\mathbf{r}_1(M_{m+1}), \ldots, \mathbf{r}_{m-k}(M_{m+1})$ still generate the space spanned by the vectors $\mathbf{r}_1(M_{m+1}), \ldots, \mathbf{r}_m(M_{m+1})$. In particular, this hold if for $m-k+1 \leq i \leq m$

$$x_i = \sum_{j=1}^{n-k} a_{ij} x_j, \tag{14}$$

where a_{ij} are determined from M_m via

$$\mathbf{r}_i(M_m) = \sum_{j=1}^{m-k} a_{ij}\mathbf{r}_j(M_m).$$

In other words, equation (14) says that the vector $(x_1, \dots, x_m)$ is orthogonal to the vectors $(a_{i1}, \dots, a_{i(m-k)}, 0, \dots, -1, 0, \dots, 0)$, or equivalently it belongs to the hyperplane H_m generated by the column vectors of M_m. We next pause to record the evolution of the uniform model (where $(x_1, \dots, x_m)$ is chosen uniformly from $\mathbf{F}_q^m$): assume that $k \geq 1$. Then (see also [24, Lemma 4])

$$\mathbf{P}_{X_{m+1}}\big(\mathrm{rank}(M_{m+1}) \leq \mathrm{rank}(M_m) + 1 | \mathrm{rank}(M_m) = m - k\big) = \frac{1}{q^k}. \tag{15}$$

Now, let $\mathcal{E}_m$ be the event from Proposition 3.12. Then we have $\mathbf{P}(\mathcal{E}_m) = 1 - (C'/q)^{-cm}$. By this proposition, and by Lemma 2.11, we have

Lemma 3.16. *Assume that $k \geq 1$. Then*

$$\begin{aligned}&\Big|\mathbf{P}_{X_{m+1}}\big(\mathrm{rank}(M_{m+1}) \leq \mathrm{rank}(M_m) + 1 | \mathrm{rank}(M_m) = m - k \wedge \mathcal{E}_m\big) - \frac{1}{q^k}\Big| \\ &\quad = \Big|\mathbf{P}\big(X_{m+1} \in H_m | \mathrm{rank}(M_m) = m - k \wedge \mathcal{E}_m\big) - \frac{1}{q^k}\Big| \leq (C'/q)^{cn}.\end{aligned}$$

Probability for $\mathrm{rank}(M_{m+1}) = \mathrm{rank}(M_m)$: For alternating matrices we have $X_{m+1} \in H_m$ implies $Q(M_{m+1}) = Q(M_m) + 1$ by the following elementary proposition.

Proposition 3.17. *For A an alternating $m \times m$ matrix and $\mathbf{x}$ a column vector of length m the following block matrix has either corank $Q(A) - 1$ or $Q(A) + 1$,*

$$\begin{pmatrix} A & \mathbf{x} \\ -\mathbf{x}^T & 0 \end{pmatrix}.$$

Proof. If $\mathbf{x}$ is not in the column span of A then $[-\mathbf{x}^T\ 0]$ is not in the row span of $[A\ \mathbf{x}]$ so the rank increases by 2.

For the other direction, if $\mathbf{x}$ is in the column span of A, so $A\mathbf{a} = \mathbf{x}$ for some $\mathbf{a}$ we need to show that $-\mathbf{x}^T\mathbf{a} = 0$ so that the addition of the row will not increase the rank. Because A is alternating we have $A = -A^T$. Therefore $(\mathbf{a}^T A\mathbf{a})^T = \mathbf{a}^T A^T\mathbf{a} = -\mathbf{a}^T A\mathbf{a}$. This implies $\mathbf{a}^T A\mathbf{a} = 0$ and $\mathbf{a}^T\mathbf{x} = 0$. □

For symmetric matrices, $\operatorname{rank}(M_{m+1}) \leq \operatorname{rank}(M_m) + 1$ does not automatically imply that $\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m)$, so we have to consider the events $\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m)$ and $\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m) + 1$ separately. Let's focus on the event that

$$\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m).$$

Here besides the event considered in Lemma 3.16, $\mathbf{r}_{m+1}(M_{m+1})$ also belongs to the subspace generated by $\mathbf{r}_1(M_{m+1}), \ldots, \mathbf{r}_m(M_{m+1})$. Knowing that M_m has rank $m-k$, by Claim 3.15 we can assume that $M_{I\times I}$ is a submatrix of full rank $m-k$ in M_m, for some $I \subset [n]$ and $|I| = m-k$. Let $B = (b_{ij})$ be the inverse of $M_{I\times I}$ in $\mathbf{F}_q$. The exposure of $X = (x_1, \ldots, x_{m-k}, x_{m+1})$ would then increase the rank of M_{m+1} by one, except when

$$\sum_{ij} b_{ij} x_i x_j + x_{m+1} = 0. \tag{16}$$

For short we write $f(X_{m+1}) := \sum_{1\leq i,j\leq m-k} b_{ij}x_ix_j + x_{m+1}$. We next pause to record the evolution of the uniform model (where $(x_1, \ldots, x_{m+1})$ is chosen uniformly from $\mathbf{F}_q^m$). The probability $\mathbf{P}(X \in H_m \wedge f(X_{m+1}) = 0)$ can be simply reduced to $\mathbf{P}(X \in H_m) \times \mathbf{P}_{x_{m+1}}(f(X_{m+1} = 0|x_1, \ldots, x_m) = \frac{1}{q^k} \times \frac{1}{q}$ because x_{m+1} is uniform and independent from the other entries. Hence we have in the uniform case (see also [24, Lemma 4])

$$\mathbf{P}_{X_{m+1}}\Big(\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m)|\operatorname{rank}(M_m) = m-k\Big) = \frac{1}{q^{k+1}}. \tag{17}$$

For our case, we cannot rely on x_{m+1} because this entry can be deterministic. Furthermore, the random entries $x_1, \ldots, x_m$ are not uniform either. Nevertheless, we show that the evolution is still asymptotically the same as in the uniform case, under the events $\mathcal{E}_m \wedge \mathcal{F}_m$ from Lemma 3.10 and Proposition 3.12.

Lemma 3.18. *Let $k \geq 0$. We have*

$$\begin{aligned}&\Big|\mathbf{P}_{X_{m+1}}\big(\operatorname{rank}(M_{m+1}) = \operatorname{rank}(M_m)|\operatorname{rank}(M_m) = m-k \wedge \mathcal{E}_m \wedge \mathcal{F}_m\big) - q^{-k-1}\Big| \\ &\quad = \Big|\mathbf{P}_{X_{m+1}}\big(X \in H_m \wedge f(X_{m+1}) = 0|\mathcal{E}_m \wedge \mathcal{F}_m\big) - q^{-k-1}\Big| \leq (C'/q)^{cm},\end{aligned}$$

where $B = \{b_{ij}\}_{1\leq i,j,\leq m-k}$ is the inverse to the full rank $I \times I$ minor of M_m.

Proof. The probability can be rewritten in the following way.

$$\begin{aligned}\mathbf{P}(X \in H_m \wedge f(X) = 0) &= q^{-k-1} \sum_{\xi \in H_n^\perp, t \in \mathbf{F}_q} \mathbf{E} e_q(X \cdot \xi + f(X)t) \\ &= q^{-k-1} + q^{-k-1} \sum_{\xi \in H_m^\perp, t \in \mathbf{F}_q, t \neq 0} \mathbf{E} e_q(X \cdot \xi + f(X)t) \\ &\quad + q^{-k-1} \sum_{\xi \neq 0, \xi \in H_m^\perp} \mathbf{E} e_q(X \cdot \xi).\end{aligned}$$

Note that the third sum is

$$\begin{aligned}q^{-k-1} \sum_{\xi \neq 0, \xi \in H_m^\perp} \mathbf{E} e_q(X \cdot \xi) &= q^{-k-1} \Big(\sum_{\xi \in H_m^\perp} \mathbf{E} e_q(X \cdot \xi) - 1 \Big) \\ &= \frac{1}{q} \mathbf{P}(X \in H_m) - q^{-k-1}.\end{aligned}$$

Hence the third sum is bound by $(C'/q)^{cm}$ using Lemma 3.16.

For the second summand because multiplication by $t \neq 0$ is a bijection on $H_m^\perp$ we have the following equality

$$\begin{aligned}&\Big| q^{-k-1} \sum_{\xi \in H_m^\perp, t \in \mathbf{F}_q, t \neq 0} \mathbf{E} e_p(X \cdot \xi + f(X)t) \Big| \\ &\quad = \Big| q^{-k} \sum_{\xi \in H_m^\perp} q^{-1} \sum_{t \in \mathbf{F}_q, t \neq 0} \mathbf{E} e_p((X \cdot \xi + f(X))t) \Big|.\end{aligned}$$

The inner sum can be recognized as,

$$\begin{aligned}&\Big| q^{-1} \sum_{t \in \mathbf{F}_q, t \neq 0} \mathbf{E} e_p((X \cdot \xi + f(X))t) \Big| \\ &\quad = \mathbf{P}(X \cdot \xi + f(X) = 0) \\ &\quad = \mathbf{P}\Big(\sum_{1 \leq i,j \leq m-k} b_{ij} x_i x_j + \sum_{i=1}^{n} \xi_i x_i = -x_{m+1} \Big).\end{aligned}$$

Now applying Lemma 3.14 with ξ, b_i, X, x_{m+1} in the context of this proof serving the role of $b_i, a_i, x_i, -r$ in the context of the lemma respectively, we

obtain,

$$\left|q^{-k-1}\sum_{\xi\in H_m^{\perp},t\in\mathbf{F}_q,t\neq 0}\mathbf{E}e_q(X\cdot\xi+f(X)t)\right|^{1/4}\leq\left|\mathbf{P}(Y_{I_1}\cdot BY_{I_2}=0)-\frac{1}{q}\right|.$$

Where $I_1\cup I_2=I$ and Y_{I_1},Y_{I_2} are random vectors in $\mathbf{F}_q^{I_1},\mathbf{F}_q^{I_2}$ with entries y_i distributed by x_i-x_i'. Let Y'_{I_1},Y'_{I_2} denote random vectors in $\mathbf{F}_q^m$ which restrict to Y_{I_1},Y_{I_2} on I_1,I_2 and are zero elsewhere. Similarly let B' be the $m\times m$ matrix with $I\times I$ minor equal to B and zeros elsewhere. Then $Y_{I_1}\cdot BY_{I_2}=0$ is equivalent to

$$Y'_{I_1}\cdot B'Y'_{I_2}=0.$$

Now choose $|I_1|=\lfloor(1-\varepsilon)m\rfloor$ and $I_2=I\backslash I_1$, where $\varepsilon=2\sqrt{\alpha}$. Therefore the vector Y'_{I_2} is non-zero with probability at least $1-(C'/q)^{\varepsilon m}$ and

$$M_m(B'Y'_{I_2})=(M_mB')Y'_{I_2}=Y'_{I_2}.$$

It follows by Lemma 3.12 that, as $B'Y'_{I_2}$ is a generalized normal vector of M_m on $\mathcal{E}_m$, we have

$$\rho(B'Y'_{I_2})\leq(C'/q)^{cm}.$$

Noting $\rho(B'Y'_{I_2})=\rho((B'Y'_{I_2})_{I_1})$, it follows by the definition of ρ that

$$\left|\mathbf{P}_{Y'_{I_1}}(Y'_{I_1}\cdot BY'_{I_2}=0)-\frac{1}{q}\right|=\left|\mathbf{P}_{Y'_{I_1}}(Y'_{I_1}\cdot B'Y'_{I_2}=0)-\frac{1}{q}\right|\leq(C'/q)^{cm}.$$

□

Putting together by changing the constants, we obtain

Proposition 3.19 (Rank relations for symmetric matrices)**.** *Assume that $\sqrt{\alpha}n\leq m\leq n$ and the symmetric matrix M_n is as in Theorem 3.3. Then there exist positive constants c,C' such that*

- *Assume that $1\leq k\leq cm$, then*

$$\left|\mathbf{P}\big(\mathrm{rank}(M_{m+1})\leq\mathrm{rank}(M_m)+1|\mathrm{rank}(M_m)=m-k\wedge\mathcal{E}_m\wedge\mathcal{F}_m\big)-\frac{1}{q^k}\right|\leq\Big(\frac{C'}{q}\Big)^{cm}.$$

- *Assume that* $0 \le k \le cm$, *then*

$$\Big|\mathbf{P}\big(\mathrm{rank}(M_{m+1}) = \mathrm{rank}(M_m) | \mathrm{rank}(M_m) = m - k \wedge \mathcal{E}_m \wedge \mathcal{F}_m\big) - \frac{1}{q^{k+1}}\Big| \le \Big(\frac{C'}{q}\Big)^{cm}.$$

Proposition 3.20 (Rank relations for alternating matrices)**.** *Assume that* $\sqrt{\alpha}n \le m \le n$ *and the alternating matrix* M_m *is as in Theorem 3.7. Then there exist positive constants* c, C' *such that the following holds for any* $0 \le k \le cm$[3]

$$\Big|\mathbf{P}\big(\mathrm{rank}(A_{m+1}) = \mathrm{rank}(A_m) | \mathrm{rank}(A_m) = m - k \wedge \mathcal{E}_m \wedge \mathcal{F}_m\big) - \frac{1}{q^k}\Big| \le \Big(\frac{C'}{q}\Big)^{cm}.$$

Now Theorems 3.3 and 3.7 are almost in reach. We recall that in the independent case we started from a full rank matrix of size $n \times m_0$ (with $m_0 = \lfloor (1-6\alpha)n \rfloor$) and expose the remaining columns one by one. Aside from the negligible events, we showed there that the rank statistics is asymptotically that of the uniform matrix of size $n - m_0$. The situation in the symmetric and alternating cases is different: instead of a column exposure process we have a corner exposure process, and because of this the situation is more complicated. To motivate the reader, let's say we use Lemma 3.10 to evolve from $m_0 = \lfloor \sqrt{\alpha}n \rfloor$ with a matrix M_{m_0} whose corank is at most εm_0 (but we will not use this information). Assume for now that at each step we append a new *uniform* vector to move from a square matrix of size m to a square matrix of size $m+1$, what is the rank statistics of the final matrix? In this case we establish the following, which is Theorem 3.5 restated here for convenience.

Proposition 3.21. *Assume that* M_{m_0} *is an arbitrary matrix of size* m_0*. Let* $M_n^u(M_0)$ *be the* $n \times n$ *random matrix with* M_0 *as its upper right hand corner and the remaining entries above the diagonal sampled uniformly and those below chosen to make the matrix alternating or symmetric. Then,*

$$\|Q(M_n^u(M_0)) - Q_\bullet\|_{TV} \le \frac{3^{n/2}}{q^{n/2-m_0}}.$$

Where $Q_\bullet$ *is either* $Q_{sym,\infty}$, $Q_{alt,e}$ *or* $Q_{alt,o}$ *depending on whether* M_n *is symmetric or alternating and the parity of* n*.*

[3]Strictly speaking, Lemma 3.16 just gave $k \ge 1$, but for $k = 0$ the bound automatically holds with probability one.

We recall that the random matrix model above is a special case of Theorem 3.3 and Theorem 3.7, and hence inevitable. In order to prove Proposition 3.21, a key problem is that at any given step the corank isn't likely zero. In fact it is likely similar to the final distribution but not within the error tolerance we are aiming for. Our solution to this problem is to show that at *some* step of the corner exposure process with high probability the corank is zero. As a consequence, because at that step the corank matches with that of the uniform model at step zeroth, we can evolve as in the uniform model from there on. We will leave a linear length stretch to get our desired final distribution with our desired error bound.

Lemma 3.22. *Assume that $\varepsilon > \sqrt{\alpha}$. Let M_0 be a symmetric or alternating matrix of corank $x_0 \leq m_0$. If we add rows and columns according to the uniform model and let x_i denote the corank after adding the i^{th} row and column then*

$$\mathbf{P}(\exists i \leq \varepsilon n, x_i = 0) \geq 1 - \frac{3^{\varepsilon n}}{q^{\varepsilon n - m_0}}.$$

Proof of Lemma 3.22. When a row and column are added there are at most three possibilities for the corank: it could increase by one, decrease by one or stay the same. Therefore there are at most $3^{\varepsilon n}$ possible paths. At most because paths which have negative corank at some step are impossible and because the corank flatting is not possible in the alternating model. We compute the probability of the most likely path which has corank strictly greater than zero at every step then take a union bound.

Claim 3.23. *Among all sequences of coranks strictly greater than 0 the one with highest probability is the one which decreases to 1 and then alternates between 1 and 2. The probability of this path is less than,*

$$\left(\frac{1}{q}\right)^{\varepsilon n - m_0}.$$

Proof of Claim 3.23. We prove the claim for the symmetric rank relations and then describe the necessary modifications for the alternating model. We first recall the following transition probabilities for the uniform symmetric model

$$\mathbf{P}(x_{m+1} = x_m + 1) = \frac{1}{p^{x_m+1}}$$

and

$$\mathbf{P}(x_{m+1} = x_m - 1) = 1 - \frac{1}{p^{x_m}}$$

as well as

$$\mathbf{P}(x_{m+1} = x_m) = \frac{1}{p^{x_m}} - \frac{1}{p^{x_m+1}} = \frac{1}{p^{x_m}}(1 - 1/p).$$

Denote by $+, -, 0$ if the corank increases, decreases, and levels respectively, and by a string of those characters a sequence of those moves for our row-column appending process. We compare the probabilities of the following pairs of transitions at step m:

$$\mathbf{P}(+\ 0) = \frac{1}{q^{x_m+1}}\frac{1}{q^{x_m+1}}(1 - 1/q) < \frac{1}{q^{x_m}}(1 - 1/q)\frac{1}{q^{x_m+1}} = \mathbf{P}(0\ +).$$

Also

$$\mathbf{P}(-\ 0) = \Big(1 - \frac{1}{q^{x_m}}\Big)\frac{1}{q^{x_m-1}}(1 - 1/q) > \frac{1}{q^{x_m}}(1 - 1/q)\Big(1 - \frac{1}{q^{x_m}}\Big) = \mathbf{P}(0\ -)$$

and

$$\mathbf{P}(-\ +) = \Big(1 - \frac{1}{q^{x_m}}\Big)\frac{1}{q^{x_m}} > \frac{1}{q^{x_m+1}}\Big(1 - \frac{1}{q^{x_m+1}}\Big) = \mathbf{P}(+\ -).$$

So we have

$$\mathbf{P}(-... - +... + 0) \le \mathbf{P}(-... - +...0+) \le \mathbf{P}(-... - 0 + ...+).$$

Furthermore,

$$\mathbf{P}(-... - +... + -) \le \mathbf{P}(-... - +... - +) \le \mathbf{P}(-... - + - +...+).$$

Repeating the above comparison we can then arrive at sequences of the form

$$-... - 0...0 + -0...0 + -0...0 + -0...0 + ... + .$$

Observe we can replace the trailing $+...+$ by increasing the number of zeros, and we can move the zeros to the right side without changing the probability, so the maximum sequences might have the form

$$-... - + - + - ... + -0...0.$$

Lastly, we have

$$P(00) = \Big(\frac{1}{q^{x_m}}(1 - 1/q)\Big)^2 \le P(+\ -) = \frac{1}{q^{x_m+1}}\Big(1 - \frac{1}{q^{x_m}}\Big)$$

even when $x_m = 1$. So the most likely sequence is $-... - + - + - ... + -$ as claimed.

To bound the probability of this sequence observe there are at least $(\varepsilon n - (m_0 - x_0))/2$ pluses since the initial rank is at most $m_0 - x_0$, that is the maximum length of the leading string of $-$'s. Each plus occurs with probability at most $1/q^2$ so we get the desired bound.

For the alternating model observe the corank either increases or decreases by one at each step. Staying the same isn't an option. The probability of the corank decreasing in the alternating model is equal to the sum of the probability of the rank staying the same and decreasing in the symmetric model. Therefore the alternating model has the same most probable corank sequence with the same bounding probability. □

From the claim we have,

$$\mathbf{P}(\forall i \leq \varepsilon n(x_i > 0)) \leq 3^{\varepsilon n} q^{-(\varepsilon n - m_0)},$$

completing the proof. □

Proof of Proposition 3.21. By the previous lemma with probability $1 - 3^{\varepsilon n}/q^{\varepsilon n - m_0}$ for some index $i \leq \varepsilon n$, $Q(M_{m_0+i}) = 0$. The corank evolution only depends on the corank in the previous step so the distribution at the n^{th} step is the same as the distribution of the corank of an $(n-i) \times (n-i)$ matrix. By (10) and (11) the total variation distance between this and $Q_\bullet$ is less than $3^{\epsilon n}/q^{n - \varepsilon n - m_0}$. Choose $\varepsilon n = n/2$ we obtain as claimed. □

Proof of Theorem 3.3 and 3.7. Let M_0 be the upper left $m_0 \times m_0$ corner of our near uniform random matrix M_n. By our rank relations 3.19 and 3.20 at each of the remaining $n - m_0$ steps the evolution of the uniform and near uniform models differ by an error of at most $(C'/q)^{cm_0}$. Therefore summing over all possible indices of rank drops we have,

$$\|Q(M_n^u(M_0)) - Q(M_n)\|_{TV} \leq 2^n \left(\frac{C'}{q}\right)^{cm_0}.$$

By the previous Corollary we have,

$$\|Q(M_n^u(M_0) - Q_\bullet\|_{TV} \leq \frac{3^{n/2}}{q^{n/2 - m_0}}.$$

Therefore by the triangle inequality we have,

$$\|Q(M_n) - Q_\bullet\|_{TV} \leq 2^n \left(\frac{C'}{q}\right)^{cm_0} + \frac{3^{n/2}}{q^{n/2 - m_0}}.$$

□

4. Two models of perturbed GL_n matrices

Our method is quite general and allows us to give the rank statistics for two other models of random matrix. Rather than sampling entries independently or mostly independently we start with an invertible matrix and perturb it's diagonal or take a minor. Though these models have no independent entries at all our method is still able to handle them.

Theorem 4.1 (Perturbation of GL_n)**.** *Let A_n be chosen uniformly from $\mathrm{GL}_n(\mathbf{F}_q)$ and let $M_n = A_n - I$, then there exist constants C, c such that*

$$\|Q(M_n) - Q_\infty\|_{TV} \leq \left(\frac{C}{q}\right)^{cn}.$$

This model was first considered in [39] by Washington in his study of the Cohen-Lenstra heuristic, where it was shown that $\lim_{n\to\infty} \mathbf{P}(Q(M_n) = r) = Q_\infty(r)$. Here we show that the rate of convergence is extremely fast. We also remark that one might be able to study this model using the cycle index technique by Kung and Stong (see for instance [16]), however it is not clear how far one can quantify the speed of convergence.

Finally, our method also gives the following rank statistics of corners of $\mathrm{GL}_n(\mathbf{F}_q)$, a model considered recently by Van Peski [36] in his study of the singular numbers of products.

Theorem 4.2 (Corner of GL_n)**.** *Let $\varepsilon > 0$ be given and let A_n be chosen uniformly from $\mathrm{GL}_n(\mathbf{F}_q)$. Let $M_{n'}$ be the top left corner of size n' of A_n, where $n' \leq (1-\varepsilon)n$.*

$$\|Q(M_{n'}) - Q_\infty\|_{TV} \leq \frac{3}{q^{n'}} + \frac{2^{\min(n', \varepsilon n)+1}}{q^{\varepsilon n}}.$$

Note there's nothing special about the upper left corner so the same result holds for any fixed $n' \times n'$ minor.

Let $A_n \in \mathrm{GL}_n(\mathbf{F}_q)$ be chosen uniformly. Let $X_1, \ldots, X_n$ denote the columns of A_n. We start with two simple results which will be important for both theorems: a uniform element of $\mathrm{GL}_n(\mathbf{F}_q)$ can be sampled using the column exposure process and a bound on the probability a rectangular submatrix of A_n is full rank.

Claim 4.3. *Sample vectors $X_1, \ldots, X_n$ one at a time as follows: for $0 \leq k \leq n-1$ sample X_{k+1} uniformly from $\mathbf{F}_q^n \backslash \langle X_1, \ldots, X_k \rangle$. Let A be the matrix*

with X_i as columns. Then for each $B \in \mathrm{GL}_n(\mathbf{F}_q)$

$$\mathbf{P}(A = B) = \frac{1}{|\mathrm{GL}_n(\mathbf{F}_q)|}.$$

Proof. Let $\mathbf{c}_i$ be the i-th column vector of B. By definition $\mathbf{P}(X_{k+1} = \mathbf{c}_{k+1}(B)|X_1 = \mathbf{c}_1(B), \ldots, X_k = \mathbf{c}_k(B)) = \frac{1}{q^n - q^k}$, and so $\mathbf{P}(A = B) = \prod_{k=0}^{n-1} \frac{1}{q^n - q^k}$. □

Lemma 4.4. *For $1 \leq k \leq l \leq n$, let $X'_1, \ldots, X'_k$ be the vectors obtained from $X_1, \ldots, X_k$ by restricting to the first (or last) l coordinates. We have*

$$\mathbf{P}(X'_1, \ldots, X'_k \text{ are linearly independent}) \geq 1 - 2/q^{l-k}.$$

By symmetry for $X''_1, \ldots, X''_l$ the restrictions of $X_1, \ldots X_l$ to the first (or last) k coordinates we have,

$$\mathbf{P}(X''_1, \ldots, X''_l \text{ span } \mathbf{F}_q^k) \geq 1 - 2/q^{l-k}.$$

Proof. We show that the event that the X'_i are dependent is bounded by $2/q^{l-k}$. If the X'_i are dependent there exists $t < k$ such that $X'_1, \ldots, X'_t$ are linearly independent and $X'_{t+1} \in \langle X'_1, \ldots, X'_t \rangle$.

We have that $X'_{t+1} \in \langle X'_1, \ldots, X'_t \rangle$ if X'_{t+1} belongs to the set S_t, where S_t is the set whose projection onto the first l coordinates is a subset of $\langle X'_1, \ldots, X'_t \rangle$. If $X'_1, \ldots X'_t$ are linearly independent then S_t has size $q^t \times q^{n-l} = q^{n+t-l}$. Therefore summing over the indices $t < k$ we obtain,

$$\begin{aligned}
\mathbf{P}(X'_1, \ldots, X'_k \text{ are linearly dependent}) &\leq \frac{S_0}{|\overline{W}_0|} + \cdots + \frac{S_{k-1}}{|\overline{W}_{k-1}|} \\
&\leq \frac{q^{n-l}}{q^n - 1} + \cdots + \frac{q^{n+(k-1)-l}}{q^n - q^{k-1}} \\
&\leq 2(q^{-l} + \cdots + q^{k-l-1}) \\
&\leq 2q^{k-l}.
\end{aligned}$$

For the symmetric statement note that $X''_1, \ldots, X''_l$ spanning $\mathbf{F}_q^k$ is equivalent to the first k rows' restriction to the first l coordinates being linearly independent. □

4.5. Proof of Theorem 4.1

Let $M_n = A_n - I$. The method is very similar to the one in Section 2. Claim 4.3 justifies our use of a column exposure process. From Lemma 4.4 it

follows $X_1 - \mathbf{e}_1, \dots, X_k - \mathbf{e}_k$ are linearly independent for $k < (1/2 - c)n$ with probability $1 - q^{-cn}$. By Proposition 4.6 the normal vectors to W_k are likely unstructured. The proof concludes similarly to the proof of Theorem 1.4.

From Lemma 4.4 it follows that $W_k(M_n)$ has dimension k with probability at least $1 - q^{-cn}$ for $k \le 1/2(1-c)n$ so we may start analyzing the rank evolution process there. For $k \ge (1/2+\varepsilon)n$, let $\mathcal{F}_{GL,k}$ be the event that $X'_1, \dots, X'_k$ span $\mathbf{F}_q^{n-k}$ where $X'_1, \dots, X'_k$ are the restrictions of $X_1, \dots, X'_k$ to the last $n-k$ coordinates. For other k let $\mathcal{F}_{GL,k}$ be empty. By Lemma 4.4, $\mathbf{P}(\mathcal{F}_{GL,k}) > 1 - q^{n-2k}$. We now study how the rank evolves after exposing $X = X_{k+1}$.

Proposition 4.6. *Let $k \ge (1/2 - \varepsilon)n$. For any nonzero normal vector $\mathbf{w}$ of $\langle X_1 - \mathbf{e}_1, \dots, X_k - \mathbf{e}_k \rangle$ we have*

$$\left| \mathbf{P}((X_{k+1} - \mathbf{e}_{k+1}) \cdot \mathbf{w} = 0 | \mathcal{F}_{GL,k}) - 1/q \right| \le q^{(-1/2+\varepsilon)n}.$$

Proof. Let $Y_1, \dots, Y_{n-k}$ be an orthonormal basis for $\overline{W_k}$. We can view X_{k+1} as a random vector $\sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{n-k} \alpha_i Y_i$ where $(\alpha_1, \dots, \alpha_{n-k}, \beta_1, \dots, \beta_k)$ is chosen uniformly from the set of vectors where at least one of α_i is nonzero. In particular each β_i is independent of $\alpha_1, \dots, \alpha_{n-k}, \beta_1, \dots, \hat{\beta}_i, \dots, \beta_k$.

Case 1: $k \ge (1/2 + \varepsilon)n$.

Because we conditioned on $X'_1, \dots, X'_k$ spanning $\mathbf{F}_q^{n-k}$, $\mathbf{w}$ is not orthogonal to one of $\mathbf{e}_1, \dots, \mathbf{e}_k$. Without loss of generality assume that $\mathbf{w}$ is not orthogonal to $\mathbf{e}_1$. Now as $\mathbf{w}$ is orthogonal to $X_1 - \mathbf{e}_1, \dots, X_k - \mathbf{e}_k$, we have

$$\begin{aligned}
\mathbf{P}_{\beta_1}((X_{k+1} - \mathbf{e}_{k+1}) \cdot \mathbf{w} = 0) &= \mathbf{P}_{\beta_1}\left(\left(\sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{n-k} \alpha_i Y_i - \mathbf{e}_{k+1} \right) \cdot \mathbf{w} = 0 \right) \\
&= \mathbf{P}_{\beta_1}\left(\beta_1 X_1 \cdot \mathbf{w} = a \right) \\
&= \mathbf{P}_{\beta_1}(\beta_1 \mathbf{e}_1 \cdot \mathbf{w} = a) \\
&= 1/q.
\end{aligned}$$

Note $a = -\sum_{i=2}^{k} \beta_i X_i - \sum_{i=1}^{n-k} + \mathbf{e}_{k+1}$ is independent of β_1 and $\mathbf{e}_1 \cdot \mathbf{w} = X_1 \cdot \mathbf{w} \ne 0$.

Case 2: $(1/2 - \varepsilon)n \le k \le (1/2 + \varepsilon)n$.

If $\mathbf{w}$ is not orthogonal to some $\mathbf{e}_i, 1 \le i \le k$ we can use the argument from Case 1. So assume $\mathbf{e}_i \cdot \mathbf{w} = 0$ for $i \le k$. Because $\mathbf{w} \cdot (X_i - \mathbf{e}_i) = 0$ for $i \le k$ this also implies $X_i \cdot \mathbf{w} = 0$ for $i \le k$. Write $\mathbf{w} = (w_1, \dots, w_n) = (0, \dots, 0, w_{k+1}, \dots, w_n)$ with respect to the basis $\{X_1, \dots, X_k, Y_1, \dots, Y_{n-k}\}$.

The tuple $(\alpha_1, \ldots, \alpha_{n-k})$ is chosen uniformly from $\mathbf{F}_q^{n-k}\backslash\{\mathbf{0}\}$, and $(w_{k+1}, \ldots, w_n)$ is a fixed vector in $\mathbf{F}_q^{n-k}\backslash\{\mathbf{0}\}$. If X was chosen uniformly from $\mathbf{F}_q^{n-k}$ then $\mathbf{P}((X - \mathbf{e}_{k+1}) \cdot \mathbf{w} = 0) = 1/q$. Accounting for whether $X = 0$ would or would not have given an equality we get,

$$\left|\mathbf{P}\Big((X - \mathbf{e}_{k+1}) \cdot \mathbf{w} = 0\Big) - 1/q\right| \leq \max\left\{\frac{q^{n-k-1}}{q^{n-k}-1} - 1/q, 1/q - \frac{q^{n-k-1}-1}{q^{n-k}-1}\right\} \leq \frac{1}{q^{n-k-1}}.$$

□

Proof of Theorem 4.1. Combining Proposition 4.6 and Lemma 2.11 we then obtain that for any $k \geq (1/2 - \varepsilon)n$, with X_{k+1} being chosen uniformly from $\mathbf{F}_q^n \backslash W_k(A_n)$ we have

$$\left|\mathbf{P}_{X_{k+1}}(X_{k+1} - \mathbf{e}_{k+1} \in W_k(M_n) | \operatorname{codim}(W_k(M_n)) = l) - \frac{1}{q^l}\right| \leq q^{(-1/2+\varepsilon)n}.$$

Using this result together with the fact that $\operatorname{codim}(W_{\lfloor(1/2-\varepsilon)n\rfloor}(M_k)) = 0$ with probability $1 - p^{-\varepsilon n}$, we can complete the proof of Theorem 4.1 the same way we concluded Theorem 1.4. Namely there are 2^n ways the rank could evolve during the column exposure process. Choosing $\varepsilon = 1/4$ each occurs with the probability one would predict from the uniform model except for an error bounded by,

$$2q^{-n/2} + (1 - 1/4)nq^{-n/4}.$$

Where the errors come from the probability $W_{\lfloor n/4\rfloor}(M_n)$ is not full rank and the error we accumulate in the remaining $\lfloor(1 - 1/4)n\rfloor$ steps. So for $q > 2^4$ we can sum our errors and get our desired bound. □

4.7. Proof of Theorem 4.2

In this subsection, let $M_{n'}$ denote the upper left $n' \times n'$ corner of A_n where $n' \leq (1 - \varepsilon)n$ and let $X'_1, \ldots, X'_{n'}$ denote its columns.

Just as in the independent and perturbed models the goal is to apply Proposition 2.11 to determine the transition probabilities. With that in mind we show the following.

Proposition 4.8. *For all nonzero normal vectors* $\mathbf{w}$ *to* $W_k(M_{n'})$ *we have*

$$|\mathbf{P}(\mathbf{w} \cdot X'_{k+1} = 0) - 1/q| < \frac{1}{q^{n-k-1}}.$$

Proof. Let $Y_1, \dots, Y_{n-k}$ be a basis for $W_k(A_n)$. As in Proposition 4.6 $X_{k+1} = \sum_{i=1}^{k} \beta_i X_i + \sum_{i=1}^{n-k} \alpha_i Y_i$ where $(\alpha_1, \dots, \alpha_{n-k}, \beta_1, \dots, \beta_k)$ is chosen uniformly from the set of vectors where at least one of α_i is nonzero. Now consider the restriction of this equation to the first n' entries,

$$X'_{k+1} = \sum_{i=1}^{k} \beta_i X'_i + \sum_{i=1}^{n-k} \alpha_i Y'_i.$$

Taking the dot product with $\mathbf{w}$ we get,

$$\begin{aligned} \mathbf{w} \cdot X'_{k+1} &= \sum_{i=1}^{k} \beta_i X'_i \cdot \mathbf{w} + \sum_{i=1}^{n-k} \alpha_i Y'_i \cdot \mathbf{w} \\ &= \sum_{i=1}^{n-k} \alpha_i Y'_i \cdot \mathbf{w} \end{aligned}$$

At least one Y_i must have $Y_i \cdot \mathbf{w} \neq 0$. Therefore as in Proposition 4.6 we get,

$$\begin{aligned} \left| \mathbf{P}\Big((X'_{k+1} \cdot \mathbf{w} = 0\Big) - 1/q \right| &\leq \max\left\{ \frac{q^{n-k-1}}{q^{n-k}-1} - 1/q, 1/q - \frac{q^{n-k-1}-1}{q^{n-k}-1} \right\} \\ &\leq \frac{1}{q^{n-k-1}}. \end{aligned}$$

□

Proof of Theorem 4.2. Combining Proposition 4.8 and Proposition 2.11 we obtain,

$$\begin{aligned} |\mathbf{P}(X'_{k+1} \in W_k(M_{n'}) | \operatorname{codim}(W_k(M_{n'})) = d) - 1/p^d| &\leq \frac{1}{q^{n-k-1}} \\ &\leq q^{-\varepsilon n}. \end{aligned}$$

As a consequence, by directly comparing with the rank evolution of the uniform model over size n' (via (5)) and by taking union bound, we obtain a total variation distance of $2^{n'} q^{-\varepsilon n}$ for the rank distribution of this model and the uniform $n' \times n'$ model. This bound can be improved if n' is large, for instance when $n' \geq \varepsilon n$. Indeed, in this case, by Lemma 4.4 $W_{n'-\lfloor \varepsilon n \rfloor}(M_{n'})$ has full dimension $n' - \lfloor \varepsilon n \rfloor$ with probability at least $1 - 2q^{-\lfloor \varepsilon n \rfloor}$. Starting at this step leaves $2^{\lfloor \varepsilon n \rfloor}$ ways for the rank to evolve. As such we obtain a total variation distance of $2q^{-\lfloor \varepsilon n \rfloor} + 2^{\lfloor \varepsilon n \rfloor} q^{-\varepsilon n}$ between the rank distribution of this

model and the uniform $n' \times n'$ model. Finally, using Fulman and Goldstein's result we add another term of $\frac{3}{q^{n'}}$ for the total variation distance to the limiting distribution Q_∞. □

References

[1] G. V. Balakin, The distribution of the rank of random matrices over a finite field (Russian, with English summary), Teor. Veroyatn. Primen. 13 (1968), 631–641. MR0243571

[2] E. D. Belsley, Rates of convergence of Markov chains related to association schemes. Thesis (Ph.D.)–Harvard University, 1993. 116 pp. MR2689583

[3] J. Blomer, R. Karp, and E. Welzl, The rank of sparse random matrices over finite fields, Random Structures Algorithms 10 (1997), no. 4, 407–419. MR1608234

[4] A. M. Borodin, The law of large numbers and the central limit theorem for the Jordan normal form of large triangular matrices over a finite field, Jour. Math. Sci. (New York) 96 (1999), no. 5, 3455–3471. MR1691635

[5] J. Bourgain, V. Vu, and P. M. Wood, On the singularity probability of discrete random matrices, Journal of Functional Analysis 258 (2010), no. 2, 559–603. MR2557947

[6] R. P. Brent and B. D. McKay, Determinants and ranks of random matrices over $\mathbf{Z}_m$, Discrete Math. 66 (1987), no. 1-2, 35–49. MR0900928

[7] L. Carlitz, Representations by quadratic forms in a finite field, Duke Math. J. 21 (1954), 123–137. MR0059952

[8] L. Carlitz, Representations by skew forms in a finite field, Arch. Math. (Basel) 5 (1954), 19–31. MR0061122

[9] C. Cooper, On the rank of random matrices, Random Structures Algorithms 16 (2000), no. 2, 209–232. MR1742352

[10] K. Costello, T. Tao, and V. Vu, Random symmetric matrices are almost surely non-singular, Duke Math. J. 135 (2006), no. 2, 395–413. MR2267289

[11] J. Kahn, J. Komlós, and E. Szemerédi, On the probability that a random ± 1 matrix is singular, J. Amer. Math. Soc. 8 (1995), 223–240. MR1260107

[12] J. Kahn and J. Komlós, Singularity probabilities for random matrices over finite fields, Combin. Probab. Comput. 10 (2001), no. 2, 137–157. MR1833067

[13] M. V. Kozlov, On the rank of matrices with random Boolean elements, Sov. Math. Dokl. 7 (1966), 1048–1051. MR0224119

[14] L. S. Charlap, H. D. Rees, and D. P. Robbins, The asymptotic probability that a random biased matrix is invertible, Discrete Math. 82 (1990), no. 2, 153–163. MR1057484

[15] S. Eberhard, The characteristic polynomial of a random matrix, preprint, https://arxiv.org/abs/2008.01223.

[16] J. Fulman, Random matrix theory over finite fields, Bull. Amer. Math. Soc. 39 (2002), 51–85. MR1864086

[17] J. Fulman and L. Goldstein, Stein's method and the rank distribution of random matrices over finite fields, Ann. Probab. 43 (2015), no. 3, 1274–1314. MR3342663

[18] J. Fulman and N. Kaplan, Random Partitions and Cohen–Lenstra Heuristics, Ann. Comb. 23 (2019), 295–315. https://doi.org/10.1007/s00026-019-00425-y. MR3962859

[19] S. Koplewitz, The corank of a rectangular random integer matrix, Linear Algebra and its Applications 591 (2020), 160–168. MR4052583

[20] I. N. Kovalenko and A. A. Levitskaya, Limiting behavior of the number of solutions of a system of random linear equations over a finite field and a finite ring (Russian), Dokl. Akad. Nauk SSSR 221 (1975), no. 4, 778–781. MR0380957

[21] I. N. Kovalenko, A. A. Levitskaya, and M. N. Savchuk, Izbrannye zadachi veroyatnostnoi kombinatoriki (Russian), Naukova Dumka, Kiev, 1986.

[22] J. B. Lewis, R. Liu, G. Panova, A. H. Morales, S. V. Sam, Y. X. Zhang, Matrices with restricted entries and q-analogues of permutations, J. Comb. 2 (2012), no. 3, 355–396. MR2913199

[23] K. Luh, S. Meehan and H. Nguyen, Some new results in random matrices over finite fields, to appear in J. London Math. Soc.

[24] J. MacWilliams, Orthogonal Matrices Over Finite Fields, The American Mathematical Monthly 76 (1969), no. 2, 152–164. https://doi.org/10.2307/2317262. MR0238870

[25] J. MacWilliams and N. Sloane, The Theory of Error-Correcting Codes, 3rd ed. North-Holland, Amsterdam, 1997. MR0465510

[26] K. Maples, Singularity of Random Matrices over Finite Fields, preprint, https://arxiv.org/abs/1012.2372.

[27] K. Maples, Cokernels of random matrices satisfy the Cohen-Lenstra heuristics, preprint, https://arxiv.org/abs/1301.1239.

[28] K. Maples, Announcement: Symmetric random matrices over finite fields, preprint, http://user.math.uzh.ch/maples/maples.symma.pdf.

[29] H. Nguyen, Inverse Littlewood-Offord problems and the singularity of random symmetric matrices, Duke Math. J. 161 (2012), no. 4, 545–586. MR2891529

[30] H. Nguyen and Van Vu, Optimal inverse Littlewood-Offord theorems, Advances in Mathematics 226 (2011), no. 4, 5298–5319. MR2775902

[31] H. Nguyen and E. Paquette, Surjectivity of near square matrices, Combinatorics, Probability & Computing 29 (2020), no. 2, 267–292. MR4079637

[32] H. Nguyen and M. M. Wood, Random integral matrices: universality of surjectivity and the cokernel, submitted.

[33] H. Nguyen and M. M. Wood, in preparation.

[34] R. Stanley, Enumerative Combinatorics, vol. 1. Cambridge University Press, 1997. MR1442260

[35] R. Stanley and Y. Wang, The Smith normal form distribution of a random integer matrix, SIAM J. Discrete Math. 31 (2017), no. 3, 2247–2268. MR3706911

[36] R. V. Peski, Limits and fluctuations of p-adic random matrix products, preprint, https://arxiv.org/abs/2011.09356.

[37] T. Tao and V. Vu, On the singularity probability of random Bernoulli matrices, Journal of the A. M. S 20 (2007), 603–673. MR2291914

[38] R. Vershynin, Invertibility of symmetric random matrices, Random Structures & Algorithms 44 (2014), no. 2, 135–182. MR3158627

[39] L. Washington, Some remarks on Cohen-Lenstra heuristics, Math. Comp. 47 (1986), 741–747. MR0856717

[40] M. M. Wood, The distribution of sandpile groups of random graphs, Journal of the A. M. S 30 (2017), 915–958. MR3671933

[41] M. M. Wood, Random integral matrices and the Cohen-Lenstra Heuristics, American Journal of Mathematics, preprint, https://arxiv.org/abs/1504.04391. MR3928040

Jake Koenig
Ohio State University
231 West 18th Avenue
Columbus, Ohio 43210
USA
E-mail address: koenig.427@osu.edu

Hoi Nguyen
Ohio State University
231 West 18th Avenue
Columbus, Ohio 43210
USA
E-mail address: nguyen.1261@osu.edu

Received January 11, 2021