ERROR TERM IN THE COHEN-LENSTRA HEURISTIC VIA RANDOM MATRIX APPROACH

YUE XU AND XIUWU ZHU

ABSTRACT. The Cohen-Lenstra heuristic predicts the distribution of ideal class groups over number fields. Random matrix models provide a natural framework for explaining this heuristic, and recent results demonstrate the effectiveness of these tools. In this paper, we extend the analysis of the random matrix model to examine the error term in the Cohen-Lenstra heuristic. Additionally, we derive the asymptotic distribution of the corank of random matrices over finite fields, which can be modeled as a special class of Markov chains.

1. INTRODUCTION

1.1. **Cohen-Lenstra heuristic.** The Cohen-Lenstra-Martinet heuristics [4, 28] predict that for a family of number field extensions over a fixed base field, the distribution of ideal class groups is inversely proportional to the complexity of the algebraic structures of these groups, particularly the size of their automorphism groups. For example, $\mathbb{Z}/9\mathbb{Z}$ is expected to occur more frequently as a class group than $(\mathbb{Z}/3\mathbb{Z})^2$.

In this paper, we focus on quadratic fields, following Cohen and Lenstra's original formulation [3]. Let D be a fundamental discriminant and $\mathrm{Cl}(D)$ the ideal class group of $\mathbb{Q}(\sqrt{D})$. For any odd prime p and finite abelian p-group G, they conjectured:

$$\lim_{X \to \infty} \frac{\#\{0 < \pm D < X : \mathrm{Cl}(D)[p^\infty] \simeq G\}}{\#\{0 < \pm D < X\}} = \frac{\eta_\infty(p)/\eta_{u_\pm}(p)}{|G|^{u_\pm}|\mathrm{Aut}(G)|},$$

where $u_+ = 1$, $u_- = 0$, and $\eta_i(p) = \prod_{j=1}^{i}(1 - p^{-j})$ for $i = 0, 1, \ldots, \infty$. As a corollary,

$$\sum_{0 < \pm D < X} |\mathrm{Cl}(D)[p]| \sim C_\pm \sum_{1 < \pm D < X} 1 \sim C'_\pm X \quad \text{as } X \to \infty$$

for constants $C_\pm$ and $C'_\pm$. Davenport and Heilbronn [5] established the $p = 3$ case in 1971 with $C_+ = 4/3$, $C_- = 2$. For general p, recent work [18, equation (1.14)] shows that for any $\epsilon > 0$,

$$\sum_{0 < \pm D < X} |\mathrm{Cl}(D)[p]| \ll_{p,\epsilon} X^{\frac{3}{2} - \frac{1}{p+1} + \epsilon},$$

which remains far from the conjectured result.

For the 2-part of $\mathrm{Cl}(D)$, Gauss's genus theory and the Hardy-Ramanujan theorem [14] imply that $\dim_{\mathbb{F}_2} \mathrm{Cl}(D)[2]$ grows like $\log\log|D|$. Consequently, the 2-torsion subgroup is of density zero as $|D|$

increases. Gerth [12] extended the conjecture to finite abelian 2-groups G:

$$\lim_{X \to \infty} \frac{\#\{0 < \pm D < X : 2\mathrm{Cl}(D)[2^\infty] \simeq G\}}{\#\{0 < \pm D < X\}} = \frac{\eta_\infty(2)/\eta_{u_\pm}(2)}{|G|^{u_\pm}|\mathrm{Aut}(G)|}.$$

Smith [24] proved this for imaginary quadratic fields in 2017, and recently extended these results to ℓ^∞-class groups of cyclic ℓ-extensions over general base fields excluding 2ℓ-th roots of unity [26, 25].

We now examine the error term in the Cohen-Lenstra heuristic.

1.2. **Error term and random matrix model.** The counting of fundamental discriminants is well-understood (for example, see [6, equation (16)]):

$$\#\{0 < \pm D < X\} = \frac{3}{\pi^2}X + O(X^{1/2}).$$

For any finite abelian p-group G, define the error term:

$$E_{\pm,p}(G, X) := \#\{0 < \pm D < X : \mathrm{Cl}(D)[p^\infty] \simeq G\} - \frac{\eta_\infty(p)/\eta_{u_\pm}(p)}{|G|^{u_\pm}|\mathrm{Aut}(G)|} \cdot \frac{3}{\pi^2}X.$$

Smith's work [26, 25] established the bound:

$$E_{\pm,2}(G, X) \ll X \exp\left(-c \cdot (\log \log \log X)^{1/2}\right),$$

which naturally raises several questions about the error term's behavior:

Question 1. Does $E_{\pm,p}(G, X)$ admit a power-saving bound (i.e., $O(X^\theta)$ for some $\theta < 1$)? If so, does θ depend on G or p; moreover, can we determine an explicit main term for $E_{\pm,p}(G, X)$ as $X \to \infty$?

For a function f defined on all finite abelian p-groups, define the f-average error:

$$E_{\pm,p}(f, X) := \sum_G f(G) \cdot E_{\pm,p}(G, X).$$

Question 2. Do the error terms $E_{\pm,p}(f, X)$ share the same properties as in Question 1?

For the case $p = 3$ and $f(G) = |G[3]|$, Bhargava, Taniguchi, and Thorne [2] refined the Davenport-Heilbronn results, proving the existence of constants $B_\pm$ such that for any $\epsilon > 0$:

$$E_{\pm,3}(|G[3]|, X) = B_\pm X^{5/6} + O(X^{2/3+\epsilon}).$$

For general p, taking $f = \mathbf{1}_{\{G: \ G \text{ nontrivial}\}}$ (the indicator function for nontrivial groups), based on numerical experiments, Lewis and Williams [20] conjectured that

$$E_{+,p}(\mathbf{1}_{\{G: \ G \text{ nontrivial}\}}, X) \sim C_p X^{s_p},$$

where C_p depends on p, and s_p (potentially consistent across odd primes) appears to lie between 0.7 and 0.8.

In subfigures (a)-(d) of Figure 1, the prime p is set to 3, 5, 7, and 11, and the elementary divisors of the p-group G are $[1]$, $[p]$, $[p^2]$, and $[p, p]$, which correspond to $G \simeq 0$, $\mathbb{Z}/p\mathbb{Z}$, $\mathbb{Z}/p^2\mathbb{Z}$, and $(\mathbb{Z}/p\mathbb{Z})^2$, respectively. We plot the ratios $\log |E_{-,p}(G, X)|/\log X$ as X (the bound on the absolute discriminant of imaginary quadratic fields) increases.
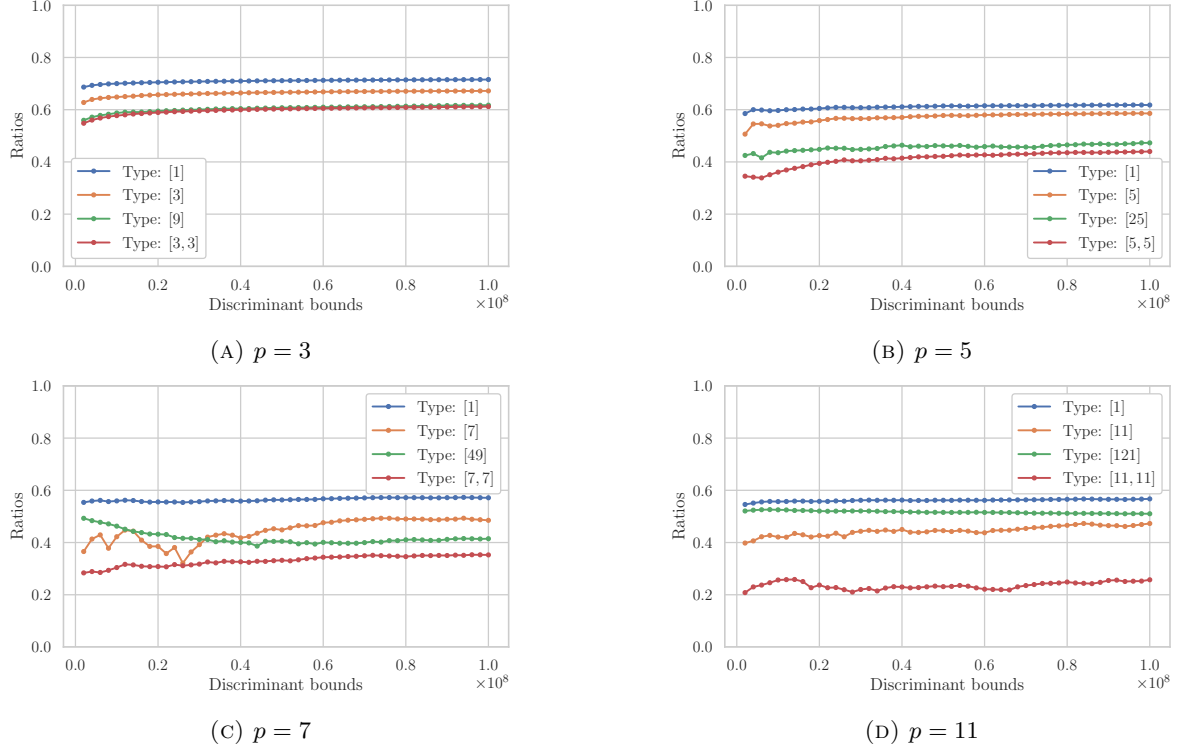
FIGURE 1. Plots of $\log|E_{-,p}(G,X)|/\log X$ for varying negative discriminant bounds and $p = 3, 5, 7, 11$

Our numerical experiments for imaginary quadratic fields with $p = 3, 5, 7, 11$ and $|D| < 10^8$ reveal that as X grows:

$$\text{the ratio } \frac{\log|E_{-,p}(G,X)|}{\log X} \text{ exhibits clear convergence, with a limit greater than } 1/4.$$

Remark 3. While the numerical experiments provide limited evidence for error term predictions in 2^∞-class groups, the Cohen-Lenstra heuristic for Selmer groups of quadratic twists of an elliptic curve E offers an illuminating parallel. As noted in [26, Remark 1.3], in this analogous setting, the error term corresponds to twisted curves E^d of Mordell-Weil rank greater than 1 and is conjectured to be $O(X^{3/4+\epsilon})$ for any $\epsilon > 0$ when $|d| \le X$.

We now consider the interpretation of the Cohen-Lenstra heuristic through random matrix models, which will allow us to analyze the error term more precisely.

The connection to random matrices was first established by Friedman and Washington [7] for function fields. Following [29, 27], we review how the p-class group arises as the cokernel of a random matrix. Let S be a finite set of primes in $\mathbb{Q}(\sqrt{D})$ generating $\mathrm{Cl}(D)$, with $O_S^\times$ the S-unit group and I_S the group of fractional ideals generated by S. This gives the exact sequence:

$$O_S^\times \otimes \mathbb{Z}_p \to I_S \otimes \mathbb{Z}_p \to \mathrm{Cl}(D)[p^\infty] \to 0.$$

Here, $I_S \otimes \mathbb{Z}_p$ (respectively $O_S^\times \otimes \mathbb{Z}_p$) is a free $\mathbb{Z}_p$-module of rank $n := |S|$ (respectively $n + u_\pm$), allowing us to express $\mathrm{Cl}(D)[p^\infty]$ as coker M_D for some matrix $M_D \in \mathrm{Mat}_{n \times (n+u_\pm)}(\mathbb{Z}_p)$.

Crucially, the matrix size $n \geq \dim_{\mathbb{F}_2} \mathrm{Cl}(D)[2]$ grows asymptotically as $\log \log |D|$, and thus increases with $|D|$ and X. If we model M_D as random in this limit, we obtain the correspondence:

$$(*) \qquad \frac{\#\{0 < \pm D < X : \mathrm{Cl}(D)[p^\infty] \simeq G\}}{\#\{0 < \pm D < X\}} \rightsquigarrow \mu\left(\{M \in \mathrm{Mat}_{n \times (n+u_\pm)}(\mathbb{Z}_p) : \mathrm{coker}\ M \simeq G\}\right),$$

where μ is the normalized Haar measure on $\mathrm{Mat}_{n \times (n+u_\pm)}(\mathbb{Z}_p)$. This leads to the Cohen-Lenstra distribution through the key result [7, 29]:

$$\lim_{n \to \infty} \mu\left(\{M \in \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p) : \mathrm{coker}\ M \simeq G\}\right) = \frac{\eta_\infty(p)/\eta_m(p)}{|G|^m |\mathrm{Aut}(G)|}.$$

In this paper, we establish the following refined version of the random matrix model distribution:

Theorem 4. *For a prime p and integer $m \geq 0$, consider the normalized Haar measure μ on $\mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)$. Then for any finite abelian p-group G,*

$$\mu\left(\{M \in \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p) : \mathrm{coker}\ M \simeq G\}\right) = w_m(G) + \lambda_m(G)p^{-n} + O(p^{-2n}),$$

where

$$w_m(G) = \frac{\eta_\infty(p)/\eta_m(p)}{|G|^m |\mathrm{Aut}(G)|}, \quad \lambda_m(G) = \frac{w_m(G)(1 + p^{-m} - p^{\mathrm{rk}_p(G)})}{p - 1}.$$

The implicit constant is at most $\left(\eta_m(p)^2/\eta_\infty(p)^2 - 1\right)^{1/2}$.

Remark 5. The proof technique actually yields higher-order expansions when needed.

For further discussion on the application of this random matrix model, we need to fix a method for choosing S, that is, choosing the primes that generate $\mathrm{Cl}(D)$. Let $T(D)$ be the smallest value such that the collection of all prime ideals with norm $\leq T(D)$ generates $\mathrm{Cl}(D)$. We take S to be this specific set of prime ideals. By the prime number theorem for number fields, we have the asymptotic relation:

$$n = |S| \asymp \frac{T(D)}{\log T(D)}.$$

Consider the correspondence $(*)$. Multiplying both sides by $\#\{0 < \pm D < X\}$ and subtracting the main term $w_{u_\pm}(G) \cdot \frac{3}{\pi^2} X$, we find that $E_{\pm,p}(G, X)$ corresponds to:

$$\mu\left(\{M \in \mathrm{Mat}_{n \times (n+u_\pm)}(\mathbb{Z}_p) : \mathrm{coker}\ M \simeq G\}\right) \cdot \#\{0 < \pm D < X\} - w_{u_\pm}(G) \cdot \frac{3}{\pi^2} X$$

$$= w_m(G)O(X^{1/4+\epsilon}) + \lambda_m(G)p^{-n}\frac{3}{\pi^2}X + \lambda_m(G)O(p^{-n}X^{1/4+\epsilon}) + O(p^{-2n}X).$$

Here we use the conjectural error bound $\#\{0 < \pm D < X\} = \frac{3}{\pi^2}X + O(X^{1/4+\epsilon})$ from [22, Remark 1.1]. The dominant contribution to $E_{\pm,p}(G, X)$ comes from comparing two terms: the error term $O(w_{u_\pm}(G)X^{1/4+\epsilon})$ from discriminant counting, and the secondary term $\lambda_{u_\pm}(G)p^{-n} \cdot \frac{3}{\pi^2}X$ from the matrix model. The larger of these two terms will dominate.

To analyze the behavior of X/p^n, we must consider the growth of $T(D)$. Current results in [15] show $T(D) \ll (\log D)^2$. Furthermore, [1] suggests that on average, $T(D)$ may grow more slowly: "It

even looks plausible that the average value of $T(D)$ as D increases is $O((\log D)^{1+\epsilon})$ for any $\epsilon > 0$". This implies:

$$\frac{1}{\#\{0 < \pm D < X\}} \sum_{0 < \pm D < X} T(D) \ll (\log X)^{1+\epsilon}.$$

Thus the average value of $n \asymp \frac{T(D)}{\log T(D)}$ is likely of size $o((\log X)^{1+\epsilon})$ for any $\epsilon > 0$, suggesting that X/p^n behaves essentially like a power-saving term.

Combining these analytic and heuristic considerations, we arrive at the following refined conjecture: For every prime p and any finite abelian p-group G, there exist explicit constants $B_{\pm,p}(G)$ and exponents $\theta_{\pm,p}(G) > \frac{1}{4}$ such that the error term satisfies:

$$E_{\pm,p}(G, X) \sim B_{\pm,p}(G) X^{\theta_{\pm,p}(G)} \quad \text{as } X \to \infty.$$

This conjecture naturally combines predictions from random matrix theory with established number-theoretic and numerical evidence.

1.3. **Random matrices over finite fields.** The cokernel distribution of p-adic matrices is closely related to the corank distribution of random matrices over finite fields. These distributions play a significant role in number theory, particularly in the Cohen-Lenstra conjecture for ideal class groups and Selmer groups [24, 17], with additional applications in coding theory (cf. [9]).

In this paper, we investigate several arithmetically significant random matrix models over finite fields, including uniform, symmetric, and skew-symmetric cases. These corank distributions share a key feature: they form reversible Markov chains with compactness properties. Consequently, analyzing their asymptotic behavior reduces to studying the convergence rates of the associated Markov chains.

For such chains, exponential convergence occurs precisely when certain drift conditions are satisfied [10], with the convergence rate determined by the largest absolute value of non-one eigenvalues in the transition matrix [21]. Our approach treats these transition matrices as compact operators on separable Hilbert spaces. Using q-series techniques, we determine their complete spectra and apply the spectral theorem to obtain detailed asymptotic behavior at all orders.

Let P be the transition matrix of an irreducible, aperiodic Markov chain on a countable set I, reversible with respect to π (i.e., $\pi(i)P(i,j) = \pi(j)P(j,i)$ for all $i, j \in I$). We work in the Hilbert space $\ell^2(\pi)$ of complex-valued sequences $\mu = (\mu(i))_{i \in I}$, equipped with the inner product and norm:

$$\langle \mu, \nu \rangle_\pi := \sum_{i \in I} \frac{\mu(i)\overline{\nu(i)}}{\pi(i)}, \quad \|\mu\|_\pi := \langle \mu, \mu \rangle_\pi^{1/2}.$$

In this framework, P acts as a bounded, self-adjoint linear operator on $\ell^2(\pi)$ via $P \cdot \mu := \mu P$.

Let q be a prime power and $m \geq 0$ a non-negative integer. We consider the random variables $\{X_{n,m}\}_{n \geq 1} : \mathrm{Mat}_{n \times (n+m)}(\mathbb{F}_q) \to \mathbb{Z}_{\geq 0}$ defined by $M \mapsto \mathrm{corank}\, M := n - \mathrm{rank}\, M$. Following [13], we have

$$\mathrm{Prob}\,(X_{n,m} = i) = (\delta_0 P_m^n)(i),$$

where $\delta_0 = (1, 0, 0, \dots)$ and the transition matrix P_m is given by

$$P_m(i,j) = \begin{cases} q^{-1-2i-m}, & \text{if } j = i+1, \\ 1 - (1-q^{-i})(1-q^{-m-i}) - q^{-1-2i-m}, & \text{if } j = i, \\ (1-q^{-i})(1-q^{-m-i}), & \text{if } j = i-1, \\ 0, & \text{otherwise.} \end{cases}$$

The matrix P_m is irreducible, aperiodic, and reversible with respect to the stationary distribution π_m, where

$$\pi_m(i) = \frac{\eta_\infty(q)}{q^{i(i+m)}\eta_i(q)\eta_{i+m}(q)}.$$

We analyze the spectral properties of P_m and prove the following theorem.

Theorem 6. *The operator P_m is compact on $\ell^2(\pi_m)$, with eigenvalues $\{q^{-k} : k \geq 0\}$.*

The eigenspaces for each eigenvalue can be explicitly constructed. Moreover, the theorem remains valid for real $q > 1$ and $m > -1$, with the corresponding transition matrix.

Applying the spectral theorem (Theorem 11), we obtain the following convergence result, which improves the main theorem in [9] (see Remark 16).

Corollary 7. *Let q be a prime power and $m \geq 0$ an integer. Then*

$$\sum_{i=0}^\infty \left|\mathrm{Prob}\left(\mathrm{corank}\ M = i \mid M \in Mat_{n\times(n+m)}(\mathbb{F}_q)\right) - \pi_m(i)\right| = \frac{2\pi_m(0)}{(q-1)q^m} \cdot \frac{1}{q^n} + O\left(\frac{1}{q^{2n}}\right),$$

where the implicit constant is bounded by $(\pi_m(0)^{-2} - 1)^{1/2}$.

The asymptotic expansion can also be extended to higher-order terms of q^{-kn} for $k \geq 2$.

In section 4, we study other matrix spaces, including symmetric, skew-symmetric, and Hermitian matrices following [9]. The spectra of their associated transition operators are as follows.

Theorem 8. *Let P_{sym} (resp. P_{alt}, Q_{alt}, P_{Her}) denote the transition matrix for symmetric (resp. alternating-1, alternating-2, Hermitian) $n \times n$ matrices, analogous to P_m. Then:*

(1) P_{sym} is compact on $\ell^2(\pi_{\mathrm{sym}})$, with eigenvalues $\{\pm q^{-k} : k \geq 0\}\backslash\{-1\}$.

(2) P_{alt} (resp. Q_{alt}) is compact on $\ell^2(\pi_{\mathrm{alt}})$ (resp. $\ell^2(\pi'_{\mathrm{alt}})$), with eigenvalues $\{q^{-2k} : k \geq 0\}$.

(3) P_{Her} is compact on $\ell^2(\pi_{\mathrm{Her}})$, with eigenvalues $\{(-q)^{-k} : k \geq 0\}$.

These spectral results yield analogous asymptotic expansions for the corank distributions, providing sharp convergence rates.

2. Spectral theorem on reversible Markov chain

Let P be a transition matrix defined on a countable set I. Assume that P is irreducible and aperiodic, and that P has a unique stationary distribution denoted by π. According to the basic limit theorem, we have

$$\|\mu P^n - \pi\|_{tv} \to 0, \quad \text{as } n \to \infty,$$

for any nonzero initial distribution μ. Here, the modified total variation distance (without the factor $1/2$) between two distributions μ_1 and μ_2 is defined as follows:

$$\|\mu_1 - \mu_2\|_{tv} := \sum_{i \in I} |\mu_1(i) - \mu_2(i)|.$$

A natural question is how fast μP^n converges to π. Under certain drift conditions (see [10] for details), the convergence rate is generally exponential. Can we derive an explicit asymptotic estimate of the convergence rate for specific P?

2.1. **Reversible Markov chain.** Further assume that P is reversible with respect to π, i.e., $\pi(j)P(j,i) = \pi(i)P(i,j)$ for any $i, j \in I$. Since P is irreducible, $\pi(i) > 0$ for all i. We define the Hilbert space $\ell^2(\pi)$ of complex-valued sequences as follows:

$$\ell^2(\pi) = \left\{ \mu = (\cdots, \mu(i), \cdots) \in \mathbb{C}^I \,\Big|\, \sum_{i \in I} \frac{|\mu(i)|^2}{\pi(i)} < \infty \right\}.$$

The inner product and norm on $\ell^2(\pi)$ are defined as follows:

$$\langle \mu, \nu \rangle_\pi := \sum_{i \in I} \frac{\mu(i)\overline{\nu(i)}}{\pi(i)}, \quad \|\mu\|_\pi := \langle \mu, \mu \rangle^{1/2}.$$

The Cauchy-Schwarz inequality implies $\|\mu\|_{tv} \leq \|\mu\|_\pi$. Indeed,

$$\|\mu\|_{tv}^2 = \left(\sum_{i \in I} \frac{|\mu(i)|}{\sqrt{\pi(i)}} \cdot \sqrt{\pi(i)} \right)^2 \leq \|\mu\|_\pi^2.$$

The operator P naturally acts on $\ell^2(\pi)$ via $P \cdot \mu := \mu P$.

The following spaces are more commonly used in the literature (see [10] for details). For $1 \leq p \leq \infty$, define

$$\ell^p_{\mathrm{old}}(\pi) := \left\{ f = (\cdots, f(i), \cdots)^{\mathrm{T}} \in \mathbb{C}^I \,\Big|\, \|f\|_{\ell^p} < \infty \right\},$$

and

$$P \cdot f := Pf, \quad \|P\|_{\ell^p} := \sup_{\|f\|_{\ell^p}=1} \|Pf\|_{\ell^p}.$$

Here, $\|f\|_{\ell^p} := \left(\sum_{i \in I} |f(i)|^p \pi(i) \right)^{1/p}$ for $p \neq \infty$, and $\|f\|_{\ell^\infty} := \sup_i |f(i)|$. Note that both $\|P\|_{\ell^1}$ and $\|P\|_{\ell^\infty}$ are no greater than one. By Hölder's inequality, we have $\|P\|_{\ell^2} \leq 1$. In particular, $\ell^2_{\mathrm{old}}(\pi)$ is a Hilbert space with the inner product

$$\langle f, g \rangle_{\ell^2} := \sum_{i \in I} f(i)\overline{g(i)}\pi(i).$$

Since P is reversible, there exists an isomorphism between the two Hilbert spaces that is compatible with the action of P:

$$\phi : \ell^2(\pi) \to \ell^2_{\mathrm{old}}(\pi), \quad \mu \mapsto (\cdots, \mu(i)/\pi(i), \cdots)^{\mathrm{T}}.$$

Then

$$\|P\|_\pi := \sup_{\|\mu\|_\pi=1} \|\mu P\|_\pi = \|P\|_{\ell^2} \leq 1.$$

In other words, P is a linear contraction on $\ell^2(\pi)$.

On the other hand, the adjoint operator P^* on $\ell^2(\pi)$ is defined by

$$P^*(i,j) := \frac{P(j,i)\pi(j)}{\pi(i)}.$$

Thus, P is self-adjoint as an operator. By fixing a one-to-one bijection between I and $\mathbb{Z}_{\geq 0}$, we obtain an isomorphism between $\ell^2(\pi)$ and

$$\ell^2 := \left\{ \nu = (\nu(0), \nu(1), \cdots)^{\mathrm{T}} \in \mathbb{C}^{\mathbb{Z}_{\geq 0}} \,\Big|\, \|\nu\| := \left(\sum_{i=0}^{\infty} |\nu(i)|^2 \right)^{1/2} < \infty \right\}$$

by sending μ to $\mu/\sqrt{\pi}$. Hence, $\ell^2(\pi)$ is separable. In summary, P is a bounded, self-adjoint, linear operator on the separable Hilbert space $\ell^2(\pi)$.

2.2. Spectral theory.

Definition 9 (Spectrum of linear operators)**.** *Let T be a linear operator defined on a complex Hilbert space X. The spectrum of T, denoted by $\sigma(T)$, is defined as follows:*

$$\sigma(T) := \{\lambda \in \mathbb{C} : (\lambda I - T) \text{ is not bijective}\}.$$

The spectrum of T is divided into three disjoint subsets:

(a) The point spectrum, or the set of all eigenvalues of T, is defined by

$$\sigma_p(T) = \{\lambda \in \sigma(T) : \mathrm{Ker}(\lambda I - T) \neq 0\}.$$

(b) The continuous spectrum of T is the set defined by

$$\sigma_c(T) = \{\lambda \in \sigma(T) : \mathrm{Ker}(\lambda I - T) = 0, \text{ and } \overline{\mathrm{Im}(\lambda I - T)} = X\}.$$

(c) The residual spectrum of T is the set defined by

$$\sigma_r(T) = \{\lambda \in \sigma(T) : \mathrm{Ker}(\lambda I - T) = 0 \text{ and } \overline{\mathrm{Im}(\lambda I - T)} \subsetneq X\}.$$

Theorem 10 (Spectral theorem)**.** *Let T be a bounded self-adjoint linear operator on an infinite-dimensional separable complex Hilbert space X. Then*

(1) *$\sigma(T)$ is a closed subset in $B(0, \|T\|)$;*
(2) *$\sigma_r(T) = \emptyset$;*
(3) *all eigenvalues of T are real;*
(4) *eigenvectors associated with distinct eigenvalues are orthogonal.*

If T is further assumed to be compact, then

(5) *all eigenspaces of T are finite-dimensional;*
(6) *for any $r > 0$, there are only finitely many eigenvalues of T with absolute value greater than r;*
(7) *$\sigma(T) = \{0\} \cup \sigma_p(T)$, and at least one of $-\|T\|$ or $\|T\|$ is an eigenvalue of T;*
(8) *Arrange all eigenvalues by their absolute value: $\|T\| = |\lambda_0| \geq |\lambda_1| \geq |\lambda_2| \geq \cdots$. Then*

$$\ell^2(\pi) = \overline{\bigoplus_{i \geq 0} V_{\lambda_i}},$$

where V_{λ_i} are the eigenspaces associated with λ_i.

Theorem 11. *Assume further that P is a compact operator on $\ell^2(\pi)$. Let $\lambda_0, \lambda_1, \lambda_2, \cdots$ be all eigenvalues of P with non-increasing absolute value. Then for any $\mu \in \ell^2(\pi)$,*

$$\left\| \mu P^n - \sum_{i=0}^{k} \lambda_i^n \mu_i \right\|_{tv} = O(|\lambda_{k+1}|^n).$$

Here, μ_i is the λ_i-component in the spectral decomposition of μ, and the implicit constant is less than $\|\mu\|_\pi$. In particular,

$$\|\mu P^n - \mu_0\|_{tv} = \begin{cases} \|\mu_1\|_{tv} \cdot |\lambda_1|^n + O(|\lambda_2|^n), & \text{if } |\lambda_1| > |\lambda_2|, \\ (\|\mu_1 + (-1)^n \mu_2\|_{tv}) \cdot |\lambda_1|^n + O(|\lambda_3|^n), & \text{if } |\lambda_1| = |\lambda_2|. \end{cases}$$

Here, $\mu_0 = (\mu \cdot \mathbf{1})\pi$, $\mathbf{1} = (1, 1, \cdots, 1, \cdots)^{\mathrm{T}}$, and the implicit constant does not exceed $\left(\|\mu\|_\pi^2 - (\mu \cdot \mathbf{1})^2\right)^{1/2}$.

This theorem generalizes fact 3 in [23].

Proof. Since P is irreducible, it has a unique stationary distribution π, which is an eigenvector corresponding to the eigenvalue 1. For any $\mu \in V_\lambda$ with $\lambda \neq 1$, we have $\mu \cdot \mathbf{1} = 0$, since $\mu \cdot \mathbf{1} = \mu \cdot P\mathbf{1} = \mu P \cdot \mathbf{1} = \lambda(\mu \cdot \mathbf{1})$. We now prove that $\lambda_1 \neq -1$. If not, let μ be a nonzero eigenvector in V_{-1}, and decompose μ as $\mu = \mu_+ - \mu_-$, where $\mu_\pm \geq 0$. Without loss of generality, we assume $\sum_{i \in I} \mu_+(i) = 1 = \sum_{i \in I} \mu_-(i)$, since $\sum_{i \in I} \mu(i) = \mu \cdot \mathbf{1} = 0$. By the basic limit theorem, $\mu = (-1)^{2n}\mu = \mu P^{2n} = \mu_+ P^{2n} - \mu_- P^{2n}$ converges to $\pi - \pi = 0$ as $n \to \infty$. By similar argument, we have $V_1 = \langle \pi \rangle$.

By the spectral theorem, for any $\mu \in \ell^2(\pi)$, we can write $\mu = \sum_{i \geq 0} \mu_i$, where $\mu_i \in V_{\lambda_i}$. Note that μ_i are orthogonal and $\mu_0 = (\mu \cdot \mathbf{1})\pi$. Then $\mu P^n = \sum_{i \geq 0} \lambda_i^n \mu_i$, and

$$\|\mu P^n - \sum_{i=0}^k \lambda_i^n \mu_i\|_\pi^2 = \sum_{i=k+1}^\infty \lambda_i^{2n} \|\mu_i\|_\pi^2$$

$$= \left(\sum_{i=k+1}^\infty \|\mu_i\|_\pi^2 \cdot \left(\frac{\lambda_i}{\lambda_{k+1}} \right)^{2n} \right) \cdot \lambda_{k+1}^{2n}$$

$$\leq \left(\sum_{i=k+1}^\infty \|\mu_i\|_\pi^2 \right) \cdot \lambda_{k+1}^{2n} = (\|\mu\|_\pi^2 - \sum_{i=0}^k \|\mu_i\|_\pi^2) \cdot \lambda_{k+1}^{2n}.$$

Thus,

$$\|\mu P^n - \sum_{i=0}^k \lambda_i^n \mu_i\|_{tv} = O(|\lambda_{k+1}|^n).$$

In particular, if $|\lambda_1| > |\lambda_2|$, we have

$$\|\mu P^n - \mu_0 - \lambda_1^n \mu_1\|_{tv} = O(|\lambda_2|^n),$$

and hence

$$\|\mu P^n - \mu_0\|_{tv} = \|\mu_1\|_{tv} \cdot |\lambda_1|^n + O(|\lambda_2|^n).$$

If $|\lambda_1| = |\lambda_2|$, we have

$$\|\mu P^n - \mu_0\|_{tv} = (\|\mu_1 + (-1)^n \mu_2\|_{tv}) \cdot |\lambda_1|^n + O(|\lambda_3|^n), \quad \text{as } n \to \infty.$$

All implicit constants are bounded above by $(\|\mu\|_\pi^2 - \|\mu_0\|_\pi^2)^{1/2} = (\|\mu\|_\pi^2 - (\mu \cdot \mathbf{1})^2)^{1/2}$. $\square$

3. Hilbert-Schmidt Markov chains

Recall that a bounded linear operator T on a separable Hilbert space X is called Hilbert-Schmidt if there exists an orthonormal basis $\{e_n : n \geq 0\}$ such that

$$\|T\|_{\mathcal{HS}} := \left(\sum_{i=0}^\infty \|Te_n\|^2 \right)^{\frac{1}{2}} < \infty.$$

A Hilbert-Schmidt operator is always compact. To see this, let P_N be the projection onto the finite-dimensional space spanned by $\{e_1, \cdots, e_N\}$. Then $P_N T$, being a finite-rank operator, is compact and converges to T uniformly.

Note that the Hilbert-Schmidt norm is independent of the choice of orthonormal basis. In our situation, the main idea to prove that an operator is Hilbert-Schmidt is to find an orthonormal basis consisting of eigenvectors and then show that

$$\|T\|_{\mathcal{HS}}^2 = \sum_{i=0}^{\infty} d_i \lambda_i^2 < \infty,$$

where d_i is the dimension of V_{λ_i}.

Definition 12. *Let P be an irreducible, aperiodic, and reversible transition matrix with respect to π. Then P is said to be Hilbert-Schmidt if it is Hilbert-Schmidt as an operator on $\ell^2(\pi)$.*

For any two real numbers $q > 1$ and $m > -1$, consider the transition matrix P_m on $\mathbb{Z}_{\geq 0}$ defined by

$$P_m(i,j) = \begin{cases} q^{-1-2i-m}, & \text{if } j = i+1, \\ 1 - (1-q^{-i})(1-q^{-m-i}) - q^{-1-2i-m}, & \text{if } j = i, \\ (1-q^{-i})(1-q^{-m-i}), & \text{if } j = i-1, \\ 0, & \text{otherwise.} \end{cases}$$

This matrix is irreducible and aperiodic because $P_m(i,i) > 0$ for all $i \geq 0$.

Define the distribution π_m by

$$\pi_m(i) = \frac{\theta_m(q)}{q^{i(i+m)} \eta_i(q) \prod_{j=1}^{i}(1-q^{-m-j})},$$

where

$$\eta_k(q) = \prod_{i=1}^{k}(1-q^{-i}),$$

and

$$\theta_m(q)^{-1} := \sum_{i=0}^{\infty} \frac{1}{q^{i(i+m)} \eta_i(q) \prod_{j=1}^{i}(1-q^{-m-j})} < \infty.$$

Then P_m is reversible with respect to π_m.

We now prove that P_m is Hilbert-Schmidt.

Proof of Theorem 6. First, we prove that for any $k \geq 1$, the real number q^{-k} is an eigenvalue. Define $\pi_m \circ q^i \in \mathbb{C}^{\mathbb{Z}_{\geq 0}}$ by

$$(\pi_m \circ q^i)(k) := \pi_m(k) \cdot q^{ik}.$$

We can check that $\pi_m \circ q^i \in \ell^2(\pi_m)$. We claim that there exist coefficients $a_0, \ldots, a_k$ (depend on k) such that $\sum_{i=0}^{k} a_i \cdot (\pi_m \circ q^i)$ is an eigenvector associated with q^{-k}.

If $\sum_{i=0}^{k} a_i \cdot (\pi_m \circ q^i)$ is an eigenvector associated with q^{-k}, that means for each l,

$$\left(\sum_{i=0}^{k} a_i \cdot (\pi_m \circ q^i) P_m \right)(l) = q^{-k} \left(\sum_{i=0}^{k} a_i \cdot \pi_m(l) \cdot q^{il} \right).$$

By reversibility and $\pi_m(l) \neq 0$, this is equivalent to

$$\sum_{j=l-1}^{l+1} \left(\sum_{i=0}^{k} a_i q^{ij} \right) P_m(l,j) = q^{-k} \left(\sum_{i=0}^{k} a_i q^{il} \right).$$

Since $\sum_{j=l-1}^{l+1} P_m(l,j) = 1$, we have

$$P_m(l,l-1) \sum_{i=0}^{k} a_i (q^{-i} - 1) q^{il} + P_m(l,l+1) \sum_{i=0}^{k} a_i (q^i - 1) q^{il} = (q^{-k} - 1) \left(\sum_{i=0}^{k} a_i q^{il} \right).$$

Substituting the values of $P_m(l,l-1)$ and $P_m(l,l+1)$, we obtain

$$\left(1 - (1+q^{-m})q^{-l} + q^{-m-2l}\right) \sum_{i=0}^{k} a_i (q^{-i} - 1) q^{il} + q^{-1-m-2l} \sum_{i=0}^{k} a_i (q^i - 1) q^{il} = (q^{-k} - 1) \left(\sum_{i=0}^{k} a_i q^{il} \right).$$

Comparing the coefficients of q^{il} on both sides and formally setting $a_{k+1} = a_{k+2} = 0$, we obtain the recurrence relation

$$(q^{-i} - q^{-k})a_i - (1+q^{-m})(q^{-1-i} - 1)a_{i+1} + (q^{-2-i-m} - q^{-m} + q^{i+1-m} - q^{-1-m})a_{i+2} = 0, \quad 0 \leq i \leq k.$$

This recurrence has a unique solution $\{a_0, \ldots, a_k\}$ up to a scalar factor.

Next, we show that these eigenvectors generate the entire space $\ell^2(\pi_m)$.

One can verify that the equation $vP_m = \lambda v$ has only one solution (up to scale) for each eigenvalue λ, meaning all eigenspaces V_λ are one-dimensional. We claim that the P_m-invariant subspace $V := \langle \pi \circ q^i, i \geq 0 \rangle$ is dense in $\ell^2(\pi_m)$, and thus

$$\ell^2(\pi_m) = \overline{V} = \overline{\bigoplus_{i \geq 0} V_{q^{-i}}}.$$

To prove this, it suffices to show that $\delta_0 \in \overline{V}$, where $\delta_i \in \ell^2(\pi_m)$ is defined by $\delta_i(k) = 1$ if $k = i$ and 0 otherwise. Indeed, if $\delta_0 \in \overline{V}$, then δ_1 also lies in $\overline{V}$ because it is a linear combination of δ_0 and $\delta_0 P_m$. By induction, all δ_i (which generate $\ell^2(\pi_m)$) belong to $\overline{V}$.

From the q-series identity (due to Euler [11, eq(19)]), we have

$$\prod_{i=1}^{\infty} \left(1 - q^{-i} t\right) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\prod_{j=1}^{k}(q^j - 1)} t^k.$$

Let $b_k = \frac{(-1)^k}{\prod_{j=1}^{k}(q^j - 1)}$ and $c_k = \frac{b_k}{\eta_\infty(q)\pi_m(0)}$. We claim that

$$\lim_{N \to \infty} \sum_{k=0}^{N} c_k(\pi_0 \circ q^k) = \delta_0 \in \ell^2(\pi_m),$$

which is equivalent to

$$\lim_{N \to \infty} \pi_m(0) \cdot \left| \sum_{k=0}^{N} c_k - \pi_m(0)^{-1} \right|^2 + \frac{1}{\eta_\infty(q)^2 \pi_m(0)^2} \cdot \sum_{i=1}^{\infty} \pi_m(i) \cdot \left| \sum_{k=0}^{N} b_k q^{ki} \right|^2 = 0.$$

By the definition of c_k, the first term converges to 0. Since $\pi_m(i) \ll q^{-i^2 - mi}$ uniformly for all i, it remains to show

$$\lim_{N \to \infty} \sum_{i=1}^{\infty} \frac{1}{q^{i^2 + mi}} \cdot \left| \sum_{k=0}^{N} b_k q^{ki} \right|^2 = 0.$$

Since $\sum_{k=0}^{\infty} b_k q^{ki} = 0$, we have

$$\left| \sum_{k=0}^{N} b_k q^{ki} \right| = \left| \sum_{k=N+1}^{\infty} b_k q^{ki} \right|.$$

Note that

$$\left| \frac{b_k q^{ki}}{b_{k-1} q^{(k-1)i}} \right| = \frac{q^i}{q^k - 1}.$$

To use the property of alternating series, we divide the estimation into two parts:

$$\sum_{i=1}^{N} \frac{1}{q^{i^2 + mi}} \cdot \left| \sum_{k=0}^{N} b_k q^{ki} \right|^2 \quad \text{and} \quad \sum_{i=N+1}^{\infty} \frac{1}{q^{i^2 + mi}} \cdot \left| \sum_{k=0}^{N} b_k q^{ki} \right|^2.$$

On the one hand, we have

$$\sum_{i=1}^{N} \frac{1}{q^{i^2 + mi}} \cdot \left| \sum_{k=N+1}^{\infty} b_k q^{ki} \right|^2 \leq \sum_{i=1}^{N} \frac{1}{q^{i^2 + mi}} b_{N+1}^2 q^{2(N+1)i}$$

$$\leq \frac{1}{\eta_\infty(q)^2} \frac{1}{q^{(N+1)(N+2)}} \sum_{i=1}^{N} \frac{1}{q^{i^2 + mi}} q^{2(N+1)i}$$

$$\ll \frac{1}{q^{(N+1)(N+2)}} \sum_{i=1}^{N+1} q^{2(N+1)i - mi - i^2}.$$

Define $F(N) := \sum_{i=0}^{\infty} q^{2Ni - mi - i^2}$ and $f(N) = \frac{F(N)}{q^{N(N+1)}}$. Then

$$F(N+1) = 1 + q^{2N+1-m} \sum_{i=1}^{\infty} q^{2N(i-1) - m(i-1) - (i-1)^2} = 1 + q^{2N+1-m} F(N).$$

Hence,

$$f(N+1) = \frac{1}{q^{(N+1)(N+2)}} + \frac{f(N)}{q^{m+1}},$$

and $f(N) \to 0$ as $N \to \infty$.

On the other hand,

$$\sum_{i=N+1}^{\infty} \frac{1}{q^{i^2 + mi}} \cdot \left| \sum_{k=0}^{N} b_k q^{ki} \right|^2 \leq \sum_{i=N+1}^{\infty} \frac{1}{q^{i^2 + mi}} b_N^2 q^{2Ni}$$

$$\ll \sum_{i=N+1}^{\infty} q^{-i^2 - mi + 2Ni - N(N+1)}$$

$$\ll \frac{1}{q^{(m+1)N}} \sum_{i=1}^{\infty} q^{-(i+m/2)^2} \to 0 \text{ as } N \to \infty.$$

Hence, all normalized eigenvectors form an orthogonal basis of $\ell^2(\pi_m)$, and

$$\|P_m\|_{\mathcal{HS}}^2 = \sum_{i=0}^{\infty} q^{-2i} = (1 - q^{-2})^{-1}.$$

Therefore, P_m is Hilbert-Schmidt. $\qquad\square$

Remark 13. The eigenvalues of all transition matrices (both in the theorem above and in the next section) can be estimated numerically using matrix truncation methods (see [19]).

From the above proof, we can deduce the following lemma, which is crucial when dealing with different types of matrices in next section.

Lemma 14. *Let $m > -1$ be a real number. If $f(z) = \sum_{i=0}^{\infty} \mu_i z^i \in \mathbb{C}[[z]]$ satisfies $\sum_{i=0}^{\infty} |\mu_i|^2 q^{i^2+mi} < \infty$ and $f(q^k) = 0$ for all $k \geq 0$, then $f = 0$.*

Proof. Let $\mu = (\cdots, \mu_i, \cdots)$. Note that

$$\pi_m(k) \asymp \frac{1}{q^{k(k+m)}}.$$

Thus, $\sum_{i=0}^{\infty} |\mu_i|^2 q^{i^2+mi} < \infty$ if and only if $\mu \in \ell^2(\pi_m)$. On the other hand, $f(q^k) = 0$ for all $k \geq 0$ is equivalent to

$$\langle \mu, \pi_m \circ q^i \rangle_{\pi_m} = 0 \quad \text{for all } i \geq 0.$$

Hence,

$$\mu \in \langle \pi_m \circ q^i \mid i \geq 0 \rangle^{\perp} = \ell^2(\pi_m)^{\perp} = \{0\}.$$

$\qquad\square$

Remark 15.   (1) Unlike the conclusion of Carlson's theorem [8] in complex analysis, the main differences are that we cannot control the growth of $f(z)$, and the points q^k are too sparse.
   (2) If $m < -1$, the lemma does not hold. Take $\mu_i = b_i q^i$, where b_i is defined in the proof of Theorem 6. Then $f(z) = \prod_{k=0}^{\infty}\left(1 - q^{-k}z\right) \neq 0$, but we always have

$$\sum_{i=0}^{\infty} |\mu_i|^2 q^{i^2+mi} < \infty \quad \text{for any } m < -1.$$

4. Corank distribution of random matrices

In this section, we investigate specific Markov chains arising from the corank distributions of different types of matrices over finite fields. These problems have been extensively studied in the literature (see [9, 13]). After proving that these Markov chains are Hilbert-Schmidt, we deduce asymptotic expressions for the corank distributions using the results from earlier sections.

4.1. **Uniform case.** Let q be a prime power and m a non-negative integer. The first example we consider is the uniform distribution on the set of all $n \times (n + m)$ matrices over the finite field $\mathbb{F}_q$. Define the corank of a matrix M as corank $M = n - \text{rank } M$. From [13, Section 1], we know that

$$\text{Prob}\left(\text{corank } M = k \mid M \in \text{Mat}_{n \times (n+m)}(\mathbb{F}_q)\right) = (\delta_0 P_m^n)(k),$$

and the stationary distribution π_m is given by

$$\pi_m(i) = \frac{\eta_{\infty}(q)}{q^{i(i+m)}\eta_i(q)\eta_{i+m}(q)}.$$

Note that this π_m is the same as the one defined in Section 3.

By Theorems 6 and 11, we obtain Corollary 7.

Proof of Corollary 7. By Theorem 6, the maximal non-one eigenvalue of P_m is q^{-1}, and

$$\nu := \pi_m - \frac{\pi_m \circ q}{1 + q^{-m}}$$

is an associated eigenvector (unique up to a scalar since $V_{q^{-1}}$ is one-dimensional). The q^{-1}-component of δ_0 is given by

$$(\delta_0)_{q^{-1}} = \frac{\langle \delta_0, \nu \rangle_{\pi_m}}{\langle \nu, \nu \rangle_{\pi_m}} \nu = \frac{(q^m + 1)^{-1}}{\langle \nu, \nu \rangle_{\pi_m}} \nu.$$

By direct calculation, we have

$$\langle \nu, \nu \rangle_{\pi_m} = \sum_{i=0}^{\infty} \pi_m(i) - \frac{2}{1 + q^{-m}} \sum_{i=0}^{\infty} \pi_m(i) q^i + \frac{1}{(1 + q^{-m})^2} \sum_{i=0}^{\infty} \pi_m(i) q^{2i}$$

$$= M(\pi_m, 0) - \frac{2}{1 + q^{-m}} M(\pi_m, 1) + \frac{1}{(1 + q^{-m})^2} M(\pi_m, 2),$$

where $M(\pi_m, k)$ is the k-th moment of π_m, defined by

$$M(\pi_m, k) = \sum_{i=0}^{\infty} \pi_m(i) q^{ki}.$$

From [3, Example 6.6], it is known that

$$M(\pi_m, 0) = 1, \quad M(\pi_m, 1) = 1 + q^{-m}, \quad M(\pi_m, 2) = 1 + (q+1)q^{-m} + q^{-2m}.$$

Substituting these values, we obtain

$$\langle \nu, \nu \rangle_{\pi_m} = \frac{(q-1)q^m}{(q^m + 1)^2}.$$

In general, note that $M(\pi_m, k) = (\pi_m \circ q^k) \cdot \mathbf{1}$ and $\sum_{i=0}^{k} a_i(\pi_m \circ q^i) \in V_{q^{-k}}$ for some coefficients $a_i \in \mathbb{R}$. Since $V_\lambda \perp \mathbf{1}$ for $\lambda \neq 1$, we can compute $M(\pi_m, k)$ by induction.

Now, we have

$$(\delta_0)_{q^{-1}} = \frac{q^m + 1}{(q-1)q^m} \nu.$$

Note that $\nu(0) > 0$ and $\nu(i) \leq 0$ for all $i > 0$. Since $\nu \cdot \mathbf{1} = \sum_{i=0}^{\infty} \nu(i) = 0$, the total variation norm of ν is

$$\|\nu\|_{tv} = 2\nu(0) = \frac{2\pi_m(0)}{q^m + 1}.$$

Therefore, the total variation norm of $(\delta_0)_{q^{-1}}$ is

$$\|(\delta_0)_{q^{-1}}\|_{tv} = \frac{2\pi_m(0)}{(q-1)q^m}.$$

By Theorem 11, we have

$$\|\delta_0 P_m^n - \pi_m\|_{tv} = \frac{2\eta_\infty(q)/\eta_m(q)}{(q-1)q^m} q^{-n} + O(q^{-2n}),$$

where the implicit constant is bounded above by

$$\left(\|\delta_0\|_\pi^2 - (\delta_0 \cdot \mathbf{1})^2 \right)^{1/2} = \left(\pi_m(0)^{-2} - 1 \right)^{1/2}.$$

$\square$

Remark 16. In [9], Fulman and Goldstein proved that (note that $\|\cdot\|_{tv} = 2\|\cdot\|_{TV}$)

$$\frac{1}{4q^{m+1}}q^{-n} \leq \|\delta_0 P_m^n - \pi_m\|_{tv} \leq \frac{6}{q^{m+1}}q^{-n}.$$

Our estimate improves upon their result, as can be seen from the comparison:

$$\frac{2\eta_\infty(q)/\eta_m(q)}{(q-1)q^m} < \frac{2}{(q-1)q^m} < \frac{6}{q^{m+1}},$$

and

$$\frac{2\eta_\infty(q)/\eta_m(q)}{(q-1)q^m} \geq \frac{2\eta_\infty(q)}{(q-1)q^m} \geq \frac{2\eta_\infty(2)}{(q-1)q^m} > \frac{1}{4q^{m+1}}.$$

Here, $\eta_\infty(2) \approx 0.29$.

Now, we transition from the corank distribution to the cokernel distribution.

Proof of Theorem 4. For a matrix $M \in \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)$, recall that the cokernel of M is defined as the quotient $\mathbb{Z}_p^n/\mathrm{col}(M)$, where $\mathrm{col}(M) := M\mathbb{Z}_p^{n+m}$ denotes the submodule of $\mathbb{Z}_p^n$ generated by the columns of M. For any finite abelian p-group G, the probability measure can be expressed as

$$\mu\left(\{M \in \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p) : \mathrm{coker}\, M \simeq G\}\right) = \sum_{\substack{L \leq \mathbb{Z}_p^n, \\ \mathbb{Z}_p^n/L \simeq G}} \mu(\mathrm{col}^{-1}(L)),$$

where L runs over submodules of $\mathbb{Z}_p^n$.

Fix an $M_0 \in \mathrm{col}^{-1}(L) \subset \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)$. Then the preimage of L can be expressed as

$$\mathrm{col}^{-1}(L) = \{M_0 Q : Q \in \mathrm{GL}_{n+m}(\mathbb{Z}_p)\}.$$

Consider the decomposition $M_0 = P_0 \mathrm{diag}(a_1, a_2, \ldots, a_n) Q_0$, where $P_0 \in \mathrm{GL}_n(\mathbb{Z}_p)$, $Q_0 \in \mathrm{GL}_{m+n}(\mathbb{Z}_p)$, and $\mathrm{diag}(a_1, a_2, \ldots, a_n) \in \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)$ is the diagonal matrix with diagonal elements $a_1, a_2, \ldots, a_n$. Since $\mathrm{col}(M_0)$ has finite index in $\mathbb{Z}_p^n$, all a_i are nonzero and satisfy $|a_1 \cdots a_n|_p = |G|^{-1}$.

Note that $\mu(P_0 \cdot)$ also defines a Haar measure on $\mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)$ with $\mu(P_0 \mathrm{Mat}_{n \times (n+m)}(\mathbb{Z}_p)) = 1$. The uniqueness of the Haar measure implies $\mu(P_0 \cdot) = \mu$. Thus, we obtain

$$\mu(\mathrm{col}^{-1}(L)) = \mu\left(\{\mathrm{diag}(a_1, \ldots, a_n)Q : Q \in \mathrm{GL}_{n+m}(\mathbb{Z}_p)\}\right).$$

Hence,

$$\mu(\mathrm{col}^{-1}(L)) = \mu\left(\{(a_1\alpha_1, \ldots, a_n\alpha_n)^\top : (\alpha_1, \ldots, \alpha_{n+m})^\top \in \mathrm{GL}_{n+m}(\mathbb{Z}_p)\}\right)$$

$$= |a_1|_p^{n+m} \cdots |a_n|_p^{n+m} \mu\left(\{(\alpha_1, \ldots, \alpha_n)^\top : \alpha_i \in \mathbb{Z}_p^{n+m} \setminus \langle p\mathbb{Z}_p^{n+m}, \alpha_1, \ldots, \alpha_{i-1}\rangle\}\right)$$

$$= |G|^{-(n+m)} \prod_{i=m+1}^{n+m} (1 - p^{-i}).$$

Combining this with the submodule counting formula from [3, Proposition 3.1]:

$$\sum_{\substack{L \leq \mathbb{Z}_p^n, \\ \mathbb{Z}_p^n/L \simeq G}} 1 = |G|^n |\mathrm{Aut}(G)|^{-1} \frac{\eta_n(p)}{\eta_{n-r}(p)},$$

where $r = \mathrm{rk}_p(G) := \dim_{\mathbb{F}_p} G/pG$ denotes the p-rank of G. We conclude that

$$\mu\left(\{M : \mathrm{coker}\, M \simeq G\}\right) = |G|^{-m}|\mathrm{Aut}(G)|^{-1}\frac{\eta_{n+m}(p)\eta_n(p)}{\eta_m(p)\eta_{n-r}(p)}.$$

Recalling the classical results (for example, see [9]),

$$\mathrm{Prob}(\mathrm{corank}\, \overline{M} = r \mid \overline{M} \in \mathrm{Mat}_{n\times(n+m)}(\mathbb{F}_p)) = p^{-r(r+m)}\frac{\eta_{n+m}(p)\eta_n(p)}{\eta_{n-r}(p)\eta_r(p)\eta_{r+m}(p)},$$

we establish the following relation:

$$\mu\left(\{M : \mathrm{coker}\, M \simeq G\}\right) = \frac{p^{r(r+m)}\eta_r(p)\eta_{r+m}(p)}{|G|^m|\mathrm{Aut}(G)|\eta_m(p)}\mathrm{Prob}\left(\mathrm{corank}\, \overline{M} = r \mid \overline{M} \in \mathrm{Mat}_{n\times(n+m)}(\mathbb{F}_p)\right).$$

The above process originates from [7, Proposition 1] (for $m = 0$) or [16, Proposition 14.1] (for $m = 1$).

Finally, reformulate the cokernel distribution using the Markov chain:

$$\mu\left(\{M : \mathrm{coker}\, M \simeq G\}\right) = \frac{w_m(G)}{\pi_m(r)}(\delta_0 P_m^n)(r).$$

Thus, to study the asymptotic behavior of the cokernel distribution as $n \to \infty$, we only need to calculate $(\delta_0 P_m^n)(r)$.

In the proof of Corollary 7, we have the decomposition:

$$\delta_0 = \pi_m + \frac{p^m + 1}{(p-1)p^m}\nu + \delta',$$

where

$$\nu = \pi_m - \frac{\pi_m \circ p}{1 + p^{-m}} \in V_{p^{-1}} \quad \text{and} \quad \delta' \in \overline{\bigoplus_{i \geq 2} V_{p^{-i}}}.$$

Hence,

$$\|\delta' P_m^n\|_{tv} \leq \|\delta' P_m^n\|_{\pi_m} \leq \|\delta'\|_{\pi_m}p^{-2n} \leq \left(\pi_m(0)^{-2} - 1\right)^{1/2}p^{-2n},$$

and so $(\delta' P_m^n)(r) = O(p^{-2n})$.

From this,

$$\mu\left(\{M : \mathrm{coker}\, M \simeq G\}\right) = \frac{w_m(G)}{\pi_m(r)}\left(\pi_m(r) + \frac{p^m + 1}{(p-1)p^m}\nu(r)p^{-n} + O(p^{-2n})\right)$$

$$= w_m(G) + \frac{w_m(G)(1 + p^{-m} - p^r)}{p - 1}p^{-n} + O(p^{-2n}).$$

Here, since $w_m(G) \leq \pi_m(r)$, the implicit constant is bounded above by

$$\left(\pi_m(0)^{-2} - 1\right)^{1/2} = \left(\eta_m(p)^2/\eta_\infty(p)^2 - 1\right)^{1/2}.$$

$\square$

4.2. **Skew centrosymmetric case.** Assume q is odd. Consider the space of skew centrosymmetric matrices:

$$\mathrm{Mat}_n^{\mathrm{scs}}(\mathbb{F}_q) := \{M \in \mathrm{Mat}_n(\mathbb{F}_q) : M_{ij} = -M_{ji} = M_{n+1-j,n+1-i}\}.$$

Note that the rank of such matrices is always even (see [9]).

More precisely, we have the following corank distributions:

$$\mathrm{Prob}(\mathrm{corank}\ M = 2k \mid M \in \mathrm{Mat}_{2n}^{\mathrm{scs}}(\mathbb{F}_q)) = \mathrm{Prob}(\mathrm{corank}\ M = k \mid M \in \mathrm{Mat}_n(\mathbb{F}_q))$$

and

$$\mathrm{Prob}(\mathrm{corank}\ M = 2k + 1 \mid M \in \mathrm{Mat}_{2n+1}^{\mathrm{scs}}(\mathbb{F}_q)) = \mathrm{Prob}(\mathrm{corank}\ M = k \mid M \in \mathrm{Mat}_{n\times(n+1)}(\mathbb{F}_q)).$$

Thus, these corank distributions can be directly derived from the uniform case results.

4.3. **Symmetric case.** Let q be a prime power. Consider the space of symmetric matrices:

$$\mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{F}_q) := \{M \in \mathrm{Mat}_n(\mathbb{F}_q) : M^{\mathrm{T}} = M\}.$$

From [13], we have the corank distribution:

$$\mathrm{Prob}(\mathrm{corank}\ M = k \mid M \in \mathrm{Mat}_n^{\mathrm{sym}}(\mathbb{F}_q)) = (\delta_0 P_{\mathrm{sym}}^n)(k),$$

where the transition matrix P_{sym} is defined by:

$$P_{\mathrm{sym}}(i,j) = \begin{cases} q^{-i-1}, & \text{if } j = i+1, \\ q^{-i} - q^{-i-1}, & \text{if } j = i, \\ 1 - q^{-i}, & \text{if } j = i-1, \\ 0, & \text{otherwise.} \end{cases}$$

The Markov chain P_{sym} is irreducible, aperiodic, and reversible with stationary distribution:

$$\pi_{\mathrm{sym}}(k) = \frac{\alpha(q)}{\prod_{i=1}^{k}(q^i - 1)}, \quad \alpha(q) = \prod_{\substack{i=1 \\ i\ \mathrm{odd}}}^{\infty} (1 - q^{-i}).$$

Theorem 17. P_{sym} *is Hilbert-Schmidt on* $\ell^2(\pi_{\mathrm{sym}})$ *with point spectrum:*

$$\sigma_p(P_{\mathrm{sym}}) = \{\pm q^{-k} : k \geq 0\} \setminus \{-1\}.$$

Proof. Similar to the approach in Theorem 6, the eigenvectors associated with eigenvalues $\pm q^{-k}$ can be expressed as linear combinations of π_m and $\{\pi_m \circ (\pm q^i) : i = 1, \ldots, k\}$, with each eigenspace $V_{\pm q^{-k}}$ being 1-dimensional.

The key step is to prove the spectral decomposition:

$$\ell^2(\pi_{\mathrm{sym}}) = \overline{\langle \pi_{\mathrm{sym}} \rangle \oplus \bigoplus_{k \geq 1} V_{\pm q^{-k}}}.$$

This reduces to showing that if $\mu \in \ell^2(\pi_{\mathrm{sym}})$ satisfies both $\mu \perp \pi_{\mathrm{sym}}$ and $\mu \perp (\pi_{\mathrm{sym}} \circ (\pm q^k))$ for all $k \geq 1$, then $\mu = 0$.

The orthogonality condition $\mu \perp (\pi_{\mathrm{sym}} \circ (\pm q^k))$ implies:

$$\sum_{i=0}^{\infty} \mu(i)(\pm 1)^i q^{ki} = 0 \quad \text{for all } k \geq 1.$$

This decouples into two independent conditions:

$$\sum_{i=0}^{\infty} \mu(2i)(q^2)^{ki} = 0 \quad \text{and} \quad \sum_{i=0}^{\infty} \mu(2i+1)(q^2)^{ki} = 0.$$

Since $\mu \in \ell^2(\pi_{\text{sym}})$ is equivalent to $\sum_{i=0}^{\infty} |\mu(i)|^2 q^{i(i+1)/2} < \infty$, we have:

$$\sum_{i=0}^{\infty} |\mu(2i)|^2 (q^2)^{i^2} < \infty \quad \text{and} \quad \sum_{i=0}^{\infty} |\mu(2i+1)|^2 (q^2)^{i^2} < \infty.$$

Define the even and odd parts:

$$\mu_{\text{even}} = (\mu(0), \mu(2), \dots), \quad \mu_{\text{odd}} = (\mu(1), \mu(3), \dots).$$

These satisfy $\mu_{\text{even}}, \mu_{\text{odd}} \in \ell^2(\pi)$ and are orthogonal to $\overline{\oplus_{k \geq 1} V_{(q^2)^{-k}}}$ in $\ell^2(\pi)$, where π is the stationary distribution for the uniform case over $\mathbb{F}_{q^2}$ with $m = 0$.

Similar to the proof of the Lemma 14, we conclude $\mu_{\text{even}}, \mu_{\text{odd}} \in \langle \pi \rangle$. The condition $\mu \perp \pi_{\text{sym}}$ implies:

$$\sum_{i=0}^{\infty} \mu(i) = 0,$$

which forces $\mu_{\text{even}} = a\pi = -\mu_{\text{odd}}$ for some $a \in \mathbb{C}$. Thus $\mu \in \langle \hat{\pi} \rangle$, where

$$\hat{\pi} = (\pi(0), -\pi(0), \pi(1), -\pi(1), \dots).$$

By self-duality of P_{sym}, the orthogonality $\mu \perp \langle \pi_{\text{sym}}, \pi_{\text{sym}} \circ (\pm q^k) : k \geq 1 \rangle$ implies $P_{\text{sym}} \cdot \mu = \mu P_{\text{sym}}$ maintains the same orthogonality. Thus $\mu P_{\text{sym}} \in \langle \hat{\pi} \rangle$. Since $\langle \hat{\pi} \rangle$ is not P_{sym}-invariant, we must have $\mu = 0$. $\qquad \square$

Corollary 18. *The convergence rate is given by:*

$$\|\delta_0 P_{\text{sym}}^n - \pi_{\text{sym}}\|_{tv} = \begin{cases} \frac{2q\alpha(q)}{q^2-1} q^{-n} + O(q^{-2n}), & n \text{ even}, \\ \frac{2q\alpha(q)}{(q^2-1)(q-1)} q^{-n} + O(q^{-2n}), & n \text{ odd}, \end{cases}$$

with implicit constants are less than $(\alpha(q)^{-2} - 1)^{1/2}$.

Proof. The two dominant eigenvalues (excluding 1) are $\lambda_+ = q^{-1}$ and $\lambda_- = -q^{-1}$. Following the approach in Corollary 7, we construct the corresponding eigenvectors:

$$\nu_+ := \pi_{\text{sym}} - \frac{1}{2}(\pi_{\text{sym}} \circ q) \in V_{q^{-1}}, \quad \nu_- := \pi_{\text{sym}} \circ (-q) \in V_{-q^{-1}}.$$

Define the k-th moment of π_{sym} as $M(\pi_{\text{sym}}, k) := \sum_{i=0}^{\infty} \pi_{\text{sym}}(i) q^{ki}$. By induction, we obtain:

$$M(\pi_{\text{sym}}, 0) = 1, \ M(\pi_{\text{sym}}, 1) = 2, \ M(\pi_{\text{sym}}, 2) = 2 + 2q.$$

These yield the following inner products:

$$\langle \nu_+, \nu_+ \rangle_{\pi_{\text{sym}}} = \frac{q-1}{2}, \ \langle \nu_-, \nu_- \rangle_{\pi_{\text{sym}}} = 2 + 2q,$$

Then the spectral projections of δ_0 are:

$$(\delta_0)_+ = \frac{1}{q-1}\nu_+, \ (\delta_0)_- = \frac{1}{2(q+1)}\nu_-.$$

Observe that in $(\delta_0)_+ + (\delta_0)_-$, only the first coordinates is positive, while in $(\delta_0)_+ - (\delta_0)_-$, only the first two coordinates are positive. Hence, the total variation norm of $(\delta_0)_+ + (-1)^n (\delta_0)_-$ is:

$$\|(\delta_0)_+ + (-1)^n (\delta_0)_-\|_{tv} = \begin{cases} \frac{2q\alpha(q)}{q^2-1}, & n \text{ even}, \\ \frac{2q\alpha(q)}{(q^2-1)(q-1)}, & n \text{ odd}. \end{cases}$$

Applying Theorem 11, we obtain the final convergence rate:

$$\|\delta_0 P_{\mathrm{sym}}^n - \pi_{\mathrm{sym}}\|_{tv} = \begin{cases} \frac{2q\alpha(q)}{q^2-1} q^{-n} + O(q^{-2n}), & n \text{ even}, \\ \frac{2q\alpha(q)}{(q^2-1)(q-1)} q^{-n} + O(q^{-2n}), & n \text{ odd}, \end{cases}$$

where the implicit constants are less than $(\alpha(q)^{-2} - 1)^{1/2}$. $\square$

Remark 19. This improves Theorem 4.1 in [9], clarifying that the parity distinction arises from P_{sym} having eigenvalue pairs $\pm q^{-1}$.

4.4. Alternating case. Consider alternating (skew-symmetric) matrices:

$$\mathrm{Mat}_n^{\mathrm{alt}}(\mathbb{F}_q) := \left\{ M \in \mathrm{Mat}_n(\mathbb{F}_q) : M^{\mathrm{T}} = -M \text{ and } M_{ii} = 0 \text{ for all } i \right\}.$$

As established in [9, 13], such matrices always have even rank. The corank distributions are given by:

$$\mathrm{Prob}(\mathrm{corank}\, M = 2j + 1 \mid M \in \mathrm{Mat}_{2n+1}^{\mathrm{alt}}) = (\delta_0 P_{\mathrm{alt}}^n)(j),$$

$$\mathrm{Prob}(\mathrm{corank}\, M = 2j \mid M \in \mathrm{Mat}_{2n}^{\mathrm{alt}}) = (\delta_0 Q_{\mathrm{alt}}^n)(j),$$

with transition matrices:

$$P_{\mathrm{alt}}(i,j) = \begin{cases} q^{-4i-3}, & j = i+1, \\ 1 - q^{-4i-3} - (1 - q^{-2i})(1 - q^{-2i-1}), & j = i, \\ (1 - q^{-2i})(1 - q^{-2i-1}), & j = i-1, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$Q_{\mathrm{alt}}(i,j) = \begin{cases} q^{-4i-1}, & j = i+1, \\ 1 - q^{-4i-1} - (1 - q^{-2i})(1 - q^{-2i+1}), & j = i, \\ (1 - q^{-2i})(1 - q^{-2i+1}), & j = i-1, \\ 0, & \text{otherwise}. \end{cases}$$

The stationary distributions of P_{alt} (resp. Q_{alt}) is:

$$\pi_{\mathrm{alt}}(j) = \frac{\alpha(q)}{q^{2j^2+j} \eta_{2j+1}(q)} \quad \left(\text{resp.} \quad \pi'_{\mathrm{alt}}(j) = \frac{\alpha(q)}{q^{2j^2-j} \eta_{2j}(q)} \right).$$

Theorem 20. *P_{alt} (resp. Q_{alt}) is Hilbert-Schmidt on $\ell^2(\pi_{\mathrm{alt}})$ (resp. $\ell^2(\pi'_{\mathrm{alt}})$) with point spectrum:*

$$\sigma_p(P_{\mathrm{alt}}) = \sigma_p(Q_{\mathrm{alt}}) = \{q^{-2k} : k \geq 0\}.$$

Proof. Similar to the approach in Theorem 6, for P_{alt}, the eigenvectors corresponding to q^{-2k} are linear combinations of $\{\pi_{\mathrm{alt}} \circ (q^{2i}) : 0 \leq i \leq k\}$. To complete the proof, it suffices to show that if $\mu \in \ell^2(\pi_{\mathrm{alt}})$ satisfies:

$$\sum_{i=0}^{\infty} |\mu(i)|^2 (q^2)^{i^2+i/2} < \infty \quad \text{and} \quad \sum_{i=0}^{\infty} \mu(i)(q^2)^{ki} = 0 \quad \text{for all } k \geq 1,$$

then $\mu = 0$. This follows directly from Lemma 14. The proof for Q_{alt} is analogous. □

Corollary 21. *The convergence rates are:*

$$\|\delta_0 P_{\text{alt}}^n - \pi_{\text{alt}}\|_{tv} = \frac{2\alpha(q)}{(q-1)^2(q+1)}q^{-2n} + O(q^{-4n}),$$

$$\|\delta_0 Q_{\text{alt}}^n - \pi'_{\text{alt}}\|_{tv} = \frac{\alpha(q)q}{(q-1)(q+1)}q^{-2n} + O(q^{-4n}),$$

with implicit constants less than $(\eta_1(q)^2\alpha(q)^{-2} - 1)^{1/2}$ and $(\alpha(q)^{-2} - 1)^{1/2}$ respectively.

Proof. Let us first analyze the case for P_{alt}. We begin by constructing the eigenvector associated with the eigenvalue q^{-2}:

$$\nu := \pi_{\text{alt}} - \frac{q}{q+1}(\pi_{\text{alt}} \circ q^2) \in V_{q^{-2}}.$$

Furthermore, we observe that the following combination belongs to the eigenspace $V_{q^{-4}}$:

$$(1+q^2)(1+q^{-1})(\pi_{\text{alt}} - \pi_{\text{alt}} \circ q^2) + \pi_{\text{alt}} \circ q^4 \in V_{q^{-4}}.$$

Proceeding by induction, we establish the moments of the stationary distribution:

$$M(\pi_{\text{alt}}, 0) = 1, \quad M(\pi_{\text{alt}}, 2) = 1 + q^{-1}, \quad M(\pi_{\text{alt}}, 4) = (1+q^2)(1+q^{-1})q^{-1}.$$

These moment calculations lead to two important results. First, the inner product of ν with itself:

$$\langle \nu, \nu \rangle_{\pi_{\text{alt}}} = \frac{q(q-1)}{q+1}.$$

Second, the total variation norm of ν:

$$\|\nu\|_{tv} = \frac{2\alpha(q)}{(q+1)\eta_1(q)}.$$

With these preparations, we can now determine the spectral projection of δ_0 onto $V_{q^{-2}}$:

$$(\delta_0)_{q^{-2}} = q^{-1}(q-1)^{-1}\nu,$$

which consequently gives:

$$\|(\delta_0)_{q^{-2}}\|_{tv} = \frac{2\alpha(q)}{(q-1)^2(q+1)}.$$

Turning now to Q_{alt}, we follow a parallel approach. The corresponding eigenvector is:

$$\nu' := \pi'_{\text{alt}} - \frac{1}{q+1}(\pi'_{\text{alt}} \circ q^2) \in V_{q^{-2}}.$$

Similarly, we identify an element in $V_{q^{-4}}$:

$$\pi'_{\text{alt}} - \frac{1}{q}(\pi'_{\text{alt}} \circ q^2) + \frac{1}{q(q+1)(q^2+1)}(\pi'_{\text{alt}} \circ q^4) \in V_{q^{-4}}.$$

The moment calculations for Q_{alt} yield:

$$M(\pi'_{\text{alt}}, 0) = 1, \quad M(\pi'_{\text{alt}}, 2) = q + 1, \quad M(\pi'_{\text{alt}}, 4) = (q+1)(q^2+1).$$

From these, we derive the key quantities:

$$\langle \nu', \nu' \rangle_{\pi'_{\text{alt}}} = \frac{q(q-1)}{q+1}, \quad \|\nu'\|_{tv} = \frac{\alpha(q)q}{q+1}.$$

Finally, the spectral projection for Q_{alt} satisfies:

$$(\delta_0)_{q^{-2}} = (q-1)^{-1}\nu', \quad \|(\delta_0)_{q^{-2}}\|_{tv} = \frac{\alpha(q)q}{(q-1)(q+1)}.$$

Then the desired results follows from Theorem 11. □

4.5. **Hermitian case.** Let q be a power of an odd prime, and fix $\theta \in \mathbb{F}_{q^2}$ such that $\theta^2 \in \mathbb{F}_q$ but $\theta \notin \mathbb{F}_q$ (see [9]). Every element $\alpha \in \mathbb{F}_{q^2}$ can be expressed as $\alpha = a + b\theta$ with $a, b \in \mathbb{F}_q$, and we define its conjugate as $\overline{\alpha} = a - b\theta$.

For a matrix $M = (\alpha_{ij}) \in \text{Mat}_n(\mathbb{F}_{q^2})$, let $M^* = (\overline{\alpha_{ji}})$ denote its conjugate transpose. The space of Hermitian matrices is:

$$\text{Mat}_n^{\text{Her}}(\mathbb{F}_{q^2}) := \{M \in \text{Mat}_n(\mathbb{F}_{q^2}) : M^* = M\}.$$

The corank distribution is given by:

$$\text{Prob}(\text{corank } M = k \mid M \in \text{Mat}_n^{\text{Her}}(\mathbb{F}_{q^2})) = (\delta_0 P_{\text{Her}}^n)(k),$$

where the transition matrix P_{Her} has entries:

$$P_{\text{Her}}(i,j) = \begin{cases} q^{-2i-1}, & j = i+1, \\ q^{-2i}(1 - q^{-1}), & j = i, \\ 1 - q^{-2i}, & j = i-1, \\ 0, & \text{otherwise,} \end{cases}$$

with stationary distribution:

$$\pi_{\text{Her}}(j) = \frac{\beta(q)}{q^{j^2}\eta_j(q^2)}, \quad \beta(q) = \prod_{i=1,\text{ odd}}^{\infty}(1 + q^{-i})^{-1}.$$

Theorem 22. *The operator P_{Her} is Hilbert-Schmidt on $\ell^2(\pi_{\text{Her}})$ with point spectrum:*

$$\sigma_p(P_{\text{Her}}) = \{(-q)^{-k} : k \geq 0\}.$$

Proof. Following the approach in Theorem 6, we need to verify:

$$\lim_{N \to \infty} \sum_{i=1}^{\infty} \frac{1}{q^{i^2}} \left| \sum_{k=0}^{N} b_k'(-q)^{ki} \right|^2 = 0,$$

where $b_k' = \frac{(-1)^k}{\prod_{j=1}^{k}((-q)^j - 1)}$. The proof decomposes into two cases:

$$\lim_{N \to \infty} \sum_{i=1,odd}^{\infty} \frac{1}{q^{i^2}} \cdot \left| \sum_{k=0}^{N} b_k'(-q)^{ki} \right|^2 = 0 = \lim_{N \to \infty} \sum_{i=1,even}^{\infty} \frac{1}{q^{i^2}} \cdot \left| \sum_{k=0}^{N} b_k'(-q)^{ki} \right|^2.$$

For odd i,

$$\text{sign}(b_k'(-q)^{ki}) \text{ is } \begin{cases} > 0, & \text{if } k \equiv 0, 3 \mod 4 \\ < 0, & \text{if } k \equiv 1, 2 \mod 4. \end{cases}$$

Using the ratio test:

$$\left| \frac{b_k'(-q)^{ki}}{b_{k-1}'(-q)^{(k-1)i}} \right| = \frac{q^i}{q^k - (-1)^k},$$

we establish convergence via alternating series estimates. Indeed, we have

$$\sum_{i=1,odd}^{N} \frac{1}{q^{i^2}} \cdot \left| \sum_{k=N+1}^{\infty} b_k'(-q)^{ki} \right|^2 \ll \frac{1}{q^{(N+1)(N+2)}} \sum_{i=1,odd}^{N} \frac{1}{q^{i^2}} q^{2(N+1)i} \to 0 \text{ as } N \to \infty,$$

and

$$\sum_{i=N+1,odd}^{\infty} \frac{1}{q^{i^2}} \cdot \left| \sum_{k=0}^{N} b_k'(-q)^{ki} \right|^2 \ll \frac{1}{q^N} \to 0 \text{ as } N \to \infty.$$

The even i case follows similarly.

$\square$

Corollary 23. *The convergence rate is:*

$$\|\delta_0 P_{\mathrm{Her}}^n - \pi_{\mathrm{Her}}\|_{tv} = \frac{2\beta(q)}{(q+1)\alpha(q^2)} q^{-n} + O(q^{-2n}),$$

with implicit constant less than $(\beta(q)^{-2} - 1)^{1/2}$.

Proof. The dominant eigenvector is:

$$\nu := \pi_{\mathrm{Her}} \circ (-q) \in V_{-q^{-1}}.$$

Combing with $\pi_{\mathrm{Her}} - \frac{1}{q+1}(\pi_{\mathrm{Her}} \circ q^2) \in V_{q^{-2}}$, we get the inner product $\langle \nu, \nu \rangle_{\pi_{\mathrm{Her}}} = q+1$. Since $\nu \perp \pi_{\mathrm{Her}}$, the total variation norm is:

$$\|\nu\|_{tv} = 2\sum_{j=0}^{\infty} |\nu(2j+1)| = 2\sum_{j=0}^{\infty} \frac{\beta(q)}{q^{(2j+1)^2}\eta_{2j+1}(q^2)} q^{2j+1} = \frac{2\beta(q)}{\alpha(q^2)} \sum_{j=0}^{\infty} \frac{\alpha(q^2)}{(q^2)^{2j^2+j}\eta_{2j+1}(q^2)}$$

Observe that $\left(\cdots, \frac{\alpha(q^2)}{(q^2)^{2j^2+j}\eta_{2j+1}(q^2)}, \cdots \right)$ is the stationary distribution of P_{alt} over $\mathbb{F}_{q^2}$. Hence,

$$\|\nu\|_{tv} = \frac{2\beta(q)}{\alpha(q^2)}.$$

The spectral projection yields:

$$(\delta_0)_{-q^{-1}} = \frac{1}{q+1}\nu, \quad \|(\delta_0)_{-q^{-1}}\|_{tv} = \frac{2\beta(q)}{(q+1)\alpha(q^2)}.$$

Then the desired results follows from Theorem 11.

$\square$

References

[1] Belabas K, Diaz y Diaz F, Friedman E. Small generators of the ideal class group. Math Comp, 2008, 77: 1185-1197
[2] Bhargava M, Taniguchi T, Thorne F. Improved error estimates for the Davenport-Heilbronn theorems. Math Ann, 2024, 389: 3471-3512.
[3] Cohen H, and Lenstra H W. Heuristics on class groups of number fields. In: Proceedings of the Journees Arithmetiques held at Noordwijkerhout, 1983. Lecture Notes in Mathematics, vol. 1068. Berlin: Springer, 1984, 33-62
[4] Cohen H, Martinet J. Étude heuristique des groupes de classes des corps de nombres. J Reine Angew Math, 1990, 404: 39-76
[5] Davenport H, Heilbronn H A. On the density of discriminants of cubic fields. II. Proc R Soc Lond A, 1971, 322: 405-420
[6] Fouvry É, and Klüners J. On the 4-rank of class groups of quadratic number fields. Invent Math, 2007, 167: 455-513
[7] Friedman E, and Washington L C. On the distribution of divisor class groups of curves over a finite field. In: Proceedings of the International Number Theory Conference held at Université Laval, 1987. Théorie des nombres / Number Theory. Berlin: De Gruyter, 1989, 227-239
[8] Fuchs W H J. A generalization of Carlson's theorem. J Lond Math Soc, 1946, 1: 106-110
[9] Fulman J, Goldstein L. Stein's method and the rank distribution of random matrices over finite fields. Ann Probab, 2015, 43: 1274-1314
[10] Gallegos-Herrada M A, Ledvinka D, and Rosenthal J S. Equivalences of Geometric Ergodicity of Markov Chains. J Theor Probab, 2023, 37: 1230-1256
[11] Gasper G, Rahman M. Basic hypergeometric series. Cambridge: Cambridge University Press, 1990

[12] Gerth III F. The 4-class ranks of quadratic fields. Invent Math, 1984, 77: 489-515

[13] Gerth III F. Limit probabilities for coranks of matrices over $GF(q)$. Linear Multilinear A, 1986, 19: 79-93

[14] Hardy G H, Ramanujan S. The normal number of prime factors of a number n. Quarterly J Math, 1917, 48: 76-92

[15] Kim H H. Minimal generators of the ideal class group. J Number Theory, 2021, 222: 157-167

[16] Koymans P, Pagano C. On the distribution of $Cl(K)[\ell^\infty]$ for degree ℓ cyclic fields. J Eur Math Soc, 2022, 24: 1189-1283

[17] Koymans P, Pagano C. Effective convergence of coranks of random Rédei matrices. Acta Arith, 2024, 212: 337-358

[18] Koymans P, Thorner J. Bounds for moments of ℓ-torsion in class groups. Math Ann, 2024, 390: 3221-3237

[19] Kumar K. Truncation method for random bounded self-adjoint operators. Banach J Math Anal, 2015, 9: 98-113

[20] Lewis C, Williams C. Numerical secondary terms in a Cohen-Lenstra conjecture on real quadratic fields. Involve, 2019, 12: 221-233

[21] Mao Y H, Song Y H. Spectral gap and convergence rate for discrete-time Markov chains. Acta Math Sin-English Ser, 2013, 29: 1949-1962

[22] Nunes R M. Squarefree numbers in arithmetic progressions. J Number Theory, 2015, 153: 1-36

[23] Rosenthal J S. Convergence Rates for Markov Chains. Siam Rev, 1995, 37: 387-405

[24] Smith A. 2^∞-Selmer groups, 2^∞-class groups, and Goldfeld's conjecture. arXiv:1702.02325, 2017

[25] Smith A. The distribution of ℓ^∞-selmer groups in degree ℓ twist families I: fixed point selmer groups. arXiv:2207.05674, 2022

[26] Smith A. The distribution of ℓ^∞-selmer groups in degree ℓ twist families II: fixed point selmer groups. arXiv:2207.05143, 2022

[27] Venkatesh A, Ellenberg J S. Statistics of Number Fields and Function Fields. In: Proceedings of the International Congress of Mathematicians 2010. New Delhi: Hindustan Book Agency, 2011, 383-402

[28] Wang W, Wood M M. Moments and interpretations of the Cohen-Lenstra-Martinet heuristics. Comment Math Helv, 2021, 96: 339-387

[29] Wood M M. Random integral matrices and the Cohen-Lenstra heuristics. Am J Math, 2019, 141: 383-398

School of Mathematics and Statistics, Xidian University, 266 Xinglong Section of Xifeng Road, Xi'an, Shaanxi 710126, China

Email address: xuyue@xidian.edu.cn

Beijing Institute of Mathematical Sciences and Applications, Beijing 101408, China.

Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China;

Email address: xwzhu@bimsa.cn