ASYMPTOTIC INDEPENDENCE OF CLASS-GROUP 4-RANKS IN CORRELATED PAIRS OF IMAGINARY QUADRATIC FIELDS

YUE XU AND XIUWU ZHU

ABSTRACT. Fix a squarefree integer $d_0 > 1$, and let d range over the positive squarefree integers coprime to $2d_0$. Although $\mathbb{Q}(\sqrt{-d})$ and $\mathbb{Q}(\sqrt{-d_0 d})$ share all variable ramified primes, we prove that their class-group 4-ranks are asymptotically independent. Over the subfamily $d \le X$, their joint distribution converges in total variation to the product of two copies of the Cohen–Lenstra–Gerth distribution, with error bounded by a negative power of $\log\log X$. We further conjecture that the corrected 2-primary groups $2\,\mathrm{Cl}_{\mathbb{Q}(\sqrt{-d})}[2^\infty]$ and $2\,\mathrm{Cl}_{\mathbb{Q}(\sqrt{-d_0 d})}[2^\infty]$ are asymptotically independent, each with the Cohen–Lenstra distribution.

Suppose in addition that the class number of $\mathbb{Q}(\sqrt{d_0})$ is odd. For a density-one subset of this family, we prove that extension of ideals to $K(d) = \mathbb{Q}(\sqrt{d_0}, \sqrt{-d})$ induces

$$4\,\mathrm{Cl}_{K(d)}[2^\infty] \cong 2\,\mathrm{Cl}_{\mathbb{Q}(\sqrt{-d})}[2^\infty] \oplus 2\,\mathrm{Cl}_{\mathbb{Q}(\sqrt{-d_0 d})}[2^\infty].$$

Together with this decomposition, the group-valued conjecture predicts that $4\,\mathrm{Cl}_{K(d)}[2^\infty]$ is distributed as the direct sum of two independent Cohen–Lenstra 2-groups, giving a corrected Cohen–Lenstra–Martinet distribution for the biquadratic family. Unconditionally, the 8-rank of $\mathrm{Cl}_{K(d)}$ has limiting distribution given by the convolution of two copies of the Cohen–Lenstra–Gerth distribution. The proof combines Smith's box method with quantitative truncated Gaussian-binomial moment inversion for diagonally coupled, fixed-width bordered Rédei matrices.

1. INTRODUCTION

1.1. **Motivation.** The class number of $\mathbb{Q}(\sqrt{d})$ is not a multiplicative function of the squarefree radicand d. One may instead fix a positive squarefree multiplier d_0 and ask whether the class groups of $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{d_0 d})$ remain statistically dependent as d varies. Already for $d_0 = 2$, although $\mathbb{Q}(\sqrt{2})$ has class number one, no pointwise factorization relates the two varying class numbers. The Cohen–Lenstra–Martinet heuristics for biquadratic fields provide the appropriate framework for this joint question. Indeed, for coprime squarefree integers m and n, the fields $\mathbb{Q}(\sqrt{m})$, $\mathbb{Q}(\sqrt{n})$ and $\mathbb{Q}(\sqrt{mn})$ are the quadratic subfields k_0, k_1, k_2 of $K = \mathbb{Q}(\sqrt{m}, \sqrt{n})$. For every odd prime ℓ, extension of ideals and the norm maps give the decomposition

$$\mathrm{Cl}_K[\ell^\infty] \cong \mathrm{Cl}_{k_0}[\ell^\infty] \oplus \mathrm{Cl}_{k_1}[\ell^\infty] \oplus \mathrm{Cl}_{k_2}[\ell^\infty].$$

The conductor–discriminant formula gives $|\Delta_K| = |\Delta_{k_0}\Delta_{k_1}\Delta_{k_2}| \asymp m^2 n^2$, so ordering K by discriminant gives the natural ordering for the simultaneous variation of m and n. Under this ordering, the heuristics predict that, for every odd prime ℓ, the three ℓ-primary class groups are distributed independently [4, 34]. After one quadratic subfield is fixed, the same heuristics predict independence of the two varying subfield class groups.

The prime 2 presents two distinct obstacles. First, there is no comparably general corrected Cohen–Lenstra–Martinet prediction for biquadratic fields. Known results, including the density-one 4-rank formula of Koymans–Morgan–Smit [16], concern lower ranks and do not determine the higher 2-primary structure in the family considered here. Second, the direct-sum decomposition above need not hold integrally at 2. At the level of class numbers, Kuroda's formula records the

2020 *Mathematics Subject Classification.* 11R29, 11R11, 15B52, 60B20.

Key words and phrases. class group, joint distribution, biquadratic field, Cohen–Lenstra–Martinet heuristics, 4-rank.

discrepancy through the Hasse unit index [25]; at the level of group structure, non-split extension data provide a further obstruction.

We specialize to the following fixed-subfield family. Fix a squarefree integer $d_0 > 1$, let d range over positive squarefree integers coprime to $2d_0$, and write

$$k_0 = \mathbb{Q}(\sqrt{d_0}), \qquad k_1(d) = \mathbb{Q}(\sqrt{-d}), \qquad k_2(d) = \mathbb{Q}(\sqrt{-d_0 d}),$$

the three quadratic subfields of $K(d) = \mathbb{Q}(\sqrt{d_0}, \sqrt{-d})$. Every prime dividing d ramifies in both varying imaginary quadratic fields, whereas the primes dividing d_0 supply fixed additional ramification to $k_2(d)$. Genus theory consequently imposes a deterministic relation between the two 2-ranks, so the raw 2-primary class groups cannot be asymptotically independent. Passing to their images under multiplication by 2 removes the genus-theoretic 2-rank contribution and leads instead to the pair

$$2\,\mathrm{Cl}_{k_1(d)}[2^\infty] \qquad \text{and} \qquad 2\,\mathrm{Cl}_{k_2(d)}[2^\infty].$$

The one-field marginals are known: Fouvry–Klüners determined the 4-rank distribution [7], while Smith proved the full one-field Cohen–Lenstra–Gerth distribution in his final two-part treatment [32, Theorem 1.10]; see also the general assembly theorem [33, Theorem 2.14]. The question is whether the two corrected groups are asymptotically independent, equivalently whether their joint distribution converges to the product of the two marginals.

The joint question is better understood when the two fields are tied by an additional arithmetic relation. For $\mathbb{Q}(\sqrt{d})$ and $\mathbb{Q}(\sqrt{-d})$, the reflection principle both forces a correlation and makes the joint distribution accessible. Fouvry and Klüners computed this distribution at the level of 4-ranks and showed that it does not factor as the product of its marginals [8]. Similar special relations underlie work on the 16-ranks and joint-spin statistics of $\mathbb{Q}(\sqrt{-p})$ and $\mathbb{Q}(\sqrt{-2p})$ [20, 21, 22].

The fixed-twist pair studied here lies outside this pattern. Although $k_1(d)$ and $k_2(d)$ share all variable ramified primes, no analogous reflection relation is available. Nor is the family a direct instance of Smith's degree-ℓ twist families [32, 33], since the twisting classes are rational and the compositum remains V_4-Galois over $\mathbb{Q}$. The results of Koymans–Liu on bad parts in odd-degree settings likewise do not apply [19]. We nevertheless prove that the two 4-ranks are asymptotically independent, each with the Cohen–Lenstra–Gerth distribution.

We order the family by the positive squarefree parameter

$$\mathcal{F}_{d_0}(X) := \{0 < d \le X : \mu^2(d) = 1,\ (d, 2d_0) = 1\}. \tag{1}$$

The coprimality condition separates the variable ramification from the fixed ramification contributed by d_0. For a finite abelian group A and $j \ge 1$, write

$$\mathrm{rk}_{2^j} A := \dim_{\mathbb{F}_2}(2^{j-1}A/2^j A), \qquad r_4(n) := \mathrm{rk}_4\,\mathrm{Cl}\big(\mathbb{Q}(\sqrt{n})\big).$$

For $a, b \ge 0$, define the finite-level joint distribution by

$$\mu_X^{(d_0)}(a,b) := \frac{\#\{d \in \mathcal{F}_{d_0}(X) : r_4(-d) = a,\ r_4(-d_0 d) = b\}}{\#\mathcal{F}_{d_0}(X)}. \tag{2}$$

Thus $\mu_X^{(d_0)}$ is a probability measure on $\mathbb{Z}_{\ge 0}^2$. For $q > 1$ and $m \ge 0$, put

$$\eta_m(q) := \prod_{j=1}^{m}(1 - q^{-j}), \qquad \eta_\infty(q) := \prod_{j \ge 1}(1 - q^{-j}),$$

with $\eta_0(q) = 1$. We denote by π_{CL} the Cohen–Lenstra–Gerth distribution on $\mathbb{Z}_{\ge 0}$ [3, 11],

$$\pi_{\mathrm{CL}}(m) = 2^{-m^2}\,\frac{\eta_\infty(2)}{\eta_m(2)^2}. \tag{3}$$

We also write π_{sym} for the limiting corank distribution of uniform symmetric matrices over $\mathbb{F}_2$,

$$\pi_{\mathrm{sym}}(m) = \frac{\eta_\infty(2)}{\eta_\infty(4)} \frac{2^{-m(m+1)/2}}{\eta_m(2)}. \tag{4}$$

1.2. **Main results.** The principal theorem gives effective asymptotic independence in the family (1). For probability measures μ and ν on a countable set S, write

$$d_{\mathrm{TV}}(\mu,\nu) = \frac{1}{2}\sum_{x\in S}\bigl|\mu(\{x\}) - \nu(\{x\})\bigr|.$$

Theorem 1.1. *Let $d_0 > 1$ be squarefree. There exists $c = c(d_0) > 0$ such that*

$$d_{\mathrm{TV}}\bigl(\mu_X^{(d_0)}, \pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}\bigr) \ll_{d_0} (\log\log X)^{-c}. \tag{5}$$

Projecting (5) onto either coordinate recovers the marginal convergence to π_{CL} recalled above. The new assertion is their joint asymptotic independence, expressed by convergence to $\pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}$. The case $d_0 = 2$. Here d is odd. As d ranges over the odd squarefree integers, $\mathbb{Q}(\sqrt{-d})$ and $\mathbb{Q}(\sqrt{-2d})$ run respectively through the families of imaginary quadratic fields with odd and even squarefree radicand; together these two families exhaust all imaginary quadratic fields. Theorem 1.1 shows that, under the natural pairing $d \mapsto 2d$, their class-group 4-ranks are asymptotically independent.

Theorem 1.1 is the 4-rank projection of the following conjectural joint distribution. Let $\mathcal{G}_2$ denote the set of isomorphism classes of finite abelian 2-groups, identifying a group with its isomorphism class, and equip $\mathcal{G}_2$ with the Cohen–Lenstra measure

$$\mu_{\mathrm{CL},2}(G) = \frac{\eta_\infty(2)}{|\operatorname{Aut}(G)|}, \qquad G \in \mathcal{G}_2. \tag{6}$$

The classical identity $\sum_{G\in\mathcal{G}_2} |\operatorname{Aut}(G)|^{-1} = \eta_\infty(2)^{-1}$ shows that this is a probability measure. Gerth's correction predicts that $2\,\mathrm{Cl}_k[2^\infty]$, rather than $\mathrm{Cl}_k[2^\infty]$ itself, is governed by this measure [12]; Smith proved this prediction for imaginary quadratic fields [32, Theorem 1.10]. We propose the following asymptotic independence conjecture for the two varying fields.

Conjecture 1.2. *Let $d_0 > 1$ be squarefree. For every $G_1, G_2 \in \mathcal{G}_2$,*

$$\lim_{X\to\infty} \mathbb{P}_{d\in\mathcal{F}_{d_0}(X)}\bigl(2\,\mathrm{Cl}_{k_i(d)}[2^\infty] \simeq G_i \text{ for } i = 1,2\bigr) = \prod_{i=1}^{2} \mu_{\mathrm{CL},2}(G_i).$$

This conjecture is specific to positive twists. When $d_0 = -1$, reflection produces a genuine correlation between a real and an imaginary quadratic field. For every fixed $d_0 > 1$, Theorem 1.1 proves the 4-rank projection of Conjecture 1.2; the higher 2^j-ranks are not addressed here.

Under an odd-class-number hypothesis on the fixed real quadratic subfield, the following direct-sum decomposition transfers both the theorem and the conjecture to $K(d)$.

Theorem 1.3. *Assume that the class number of $k_0 = \mathbb{Q}(\sqrt{d_0})$ is odd. For all but $o(\#\mathcal{F}_{d_0}(X))$ values $d \in \mathcal{F}_{d_0}(X)$, the Hasse unit index*

$$q_{K(d)} = [\mathcal{O}_{K(d)}^\times : \mathcal{O}_{k_0}^\times \mathcal{O}_{k_1(d)}^\times \mathcal{O}_{k_2(d)}^\times]$$

equals 1, and extension of ideals induces an isomorphism

$$4\,\mathrm{Cl}_{K(d)}[2^\infty] \cong \bigoplus_{i=0}^{2} 2\,\mathrm{Cl}_{k_i(d)}[2^\infty].$$

Theorem 1.3 turns Conjecture 1.2 into the following corrected Cohen–Lenstra–Martinet prediction for the biquadratic field.

Conjecture 1.4. *Assume that the class number of k_0 is odd. For every $H \in \mathcal{G}_2$,*

$$\lim_{X\to\infty} \mathbb{P}_{d\in\mathcal{F}_{d_0}(X)}\big(4\,\mathrm{Cl}_{K(d)}[2^\infty] \simeq H\big) = \sum_{\substack{G_1,G_2\in\mathcal{G}_2\\ G_1\oplus G_2\simeq H}} \mu_{\mathrm{CL},2}(G_1)\mu_{\mathrm{CL},2}(G_2).$$

Taking 4-ranks of the two varying quadratic class groups, and hence the 8-rank in the decomposition above, gives the following unconditional consequence.

Corollary 1.5. *Assume that the class number of k_0 is odd. For all but $o(\#\mathcal{F}_{d_0}(X))$ values $d \in \mathcal{F}_{d_0}(X)$,*

$$\mathrm{rk}_8\,\mathrm{Cl}_{K(d)} = r_4(-d) + r_4(-d_0 d).$$

Moreover, for every $m \geq 0$,

$$\lim_{X\to\infty} \mathbb{P}_{d\in\mathcal{F}_{d_0}(X)}\big(\mathrm{rk}_8\,\mathrm{Cl}_{K(d)} = m\big) = \sum_{a+b=m} \pi_{\mathrm{CL}}(a)\pi_{\mathrm{CL}}(b),$$

and this convergence of distributions holds in total variation.

1.3. **Proof strategy and random-matrix results.** The 4-rank of a quadratic class group is determined by the corank of its Rédei matrix. For the two fields in our family, deleting one variable prime produces a common core: the fixed twist by d_0 appears as a correlated diagonal perturbation, while the prime divisors of $2d_0$ contribute a fixed number of border rows and columns. Smith's box method [31] shows that the symmetrised distribution of the relevant quadratic-residue symbols is close to uniform. Since the corank pair is invariant under relabelling the variable primes, this reduces the arithmetic problem to the joint corank distribution of a pair of coupled, fixed-width bordered Rédei matrices. We use Smith's permutation-averaged equidistribution theorem for a *complete* residue assignment, comprising the pair symbols of the variable primes together with the columns $(-1/p_i)$, $(2/p_i)$ and (q_k/p_i) attached to the fixed integer d_0; the sign vector is thereby part of the equidistributed data rather than a condition imposed on the family. For a published implementation of the box method in the all-1 mod 4 family, see Chan–Koymans–Milovic–Pagano [1, §§4–5]; the mixed-sign specialization needed here is carried out in Section 3.2.

For a single unbordered Rédei matrix, Koymans–Pagano obtained effective convergence by exposing one row and column at a time and comparing the resulting corank jumps with an ergodic Markov chain [17]. A different single-matrix approach, based on the spectrum of the corank transition operator, was developed in our earlier work [36]. These one-parameter methods do not directly close for the coupled bordered matrix pair above: one must simultaneously track two matrices sharing a core and diagonal data, and the borders impose additional kernel conditions that are not preserved by a single corank transition.

Two related moment normalizations have been used to study class-group and random-cokernel distributions. Ordinary moments of group size occur in the work of Fouvry–Klüners and Smith [7, 33]; systematic treatments of surjection moments in the random-group and random-cokernel setting may be found in [35, 34]. Mixed moments also appear in the work of Pagano–Sofos on ray class sequences [28]. In these applications, moments are principally used to identify a limiting distribution through a uniqueness theorem. Here fixed-order mixed moments are not enough: estimates through a growing order must be transferred to the joint distribution while retaining a uniform error. We do this by quantitative truncated inversion of the Gaussian-binomial moment transform.

For a corank variable X, consider the Gaussian binomial moments

$$m_u(X) = \mathbb{E}\begin{bmatrix} X \\ u \end{bmatrix}_2,$$

which count, on average, the u-dimensional subspaces of the kernel. The q-binomial transform taking a distribution to these moments is triangular and has an explicit inverse. On a suitable weighted supremum space, this inverse is bounded into ℓ^1. Concretely, if ρ and ν are two probability distributions on $\mathbb{Z}_{\geq 0}$, $\sigma = \rho - \nu$, and $\theta_u = 2^{-u(u+1)/2}$, then the one-dimensional truncated inversion gives

$$\|\sigma\|_1 \leq C_{\text{inv}} \max_{0 \leq u \leq D} \frac{|m_u(\sigma)|}{\theta_u} + C(D+1)2^{D(D+1)/2} m_{D+1}(|\sigma|).$$

The first term uses only the moments through order D, while the second controls the discarded tail by the $(D+1)$-st absolute moment. For a signed measure on $\mathbb{Z}^2_{\geq 0}$, applying the tensor product of the two one-dimensional inverse operators and estimating the bivariate residual gives an analogous truncated inversion inequality whose inputs are the genuinely mixed moments

$$m_{u,v}(X, Y) = \mathbb{E}\left[\begin{bmatrix} X \\ u \end{bmatrix}_2 \begin{bmatrix} Y \\ v \end{bmatrix}_2\right].$$

These mixed-moment estimates are proved directly for the coupled matrices; they do not follow by multiplying one-dimensional error estimates. Once they are available through a growing order, balancing the truncation depth against the matrix size gives a total-variation bound of the form $C\eta^n$ with $\eta \in (0, 1)$.

The unbordered case already gives an asymptotic independence theorem for diagonal perturbations of a common core. For $\mathbf{u}_{-1} \in \mathbb{F}_2^n$, put

$$\Omega_{\mathbf{u}_{-1}} = \mathbf{u}_{-1}\mathbf{u}_{-1}^\top + \operatorname{diag}(\mathbf{u}_{-1}), \qquad \operatorname{Sym}_n(\Omega_{\mathbf{u}_{-1}}) = \{A \in \operatorname{Mat}_n(\mathbb{F}_2) : A^\top - A = \Omega_{\mathbf{u}_{-1}}\}.$$

For a random element W, write $\operatorname{Law}(W)$ for its probability distribution.

Theorem 1.6. *Let A be uniform in $\operatorname{Sym}_n(\Omega_{\mathbf{u}_{-1}})$, and let $\mathbf{v} \in \mathbb{F}_2^n$ be uniform and independent of A. Set*

$$X = \operatorname{corank} A, \qquad Y = \operatorname{corank}\big(A + \operatorname{diag}(\mathbf{v})\big).$$

There are constants $C < \infty$ and $\eta \in (0, 1)$ such that, if $\mathbf{u}_{-1} = \mathbf{0}$, then

$$\big\|\operatorname{Law}(X, Y) - \pi_{\text{sym}} \otimes \pi_{\text{sym}}\big\|_1 \leq C\eta^n.$$

For every $\alpha > 0$, there are $C_\alpha < \infty$ and $\eta_\alpha \in (0, 1)$ such that, uniformly whenever $\operatorname{rank} \Omega_{\mathbf{u}_{-1}} \geq \alpha n$,

$$\big\|\operatorname{Law}(X, Y) - \pi_{\text{CL}} \otimes \pi_{\text{CL}}\big\|_1 \leq C_\alpha \eta_\alpha^n.$$

When $\mathbf{u}_{-1} = \mathbf{0}$, the matrix A is uniform symmetric and $\operatorname{diag}(\mathbf{v})$ is an independent uniform random diagonal matrix. Thus the first estimate states directly that a random diagonal perturbation makes the two coranks asymptotically independent, with common limiting distribution π_{sym}. This finite-field statement concerns only coranks; the results of Lee and Jung–Lee discussed below instead determine joint cokernel modules for p-adic matrix ensembles.

The diagonal need not be random. If A is a uniform symmetric matrix over $\mathbb{F}_2$, the same mixed-moment estimate gives constants $C < \infty$ and $\eta \in (0, 1)$ such that

$$\big\|\operatorname{Law}\big(\operatorname{corank} A, \operatorname{corank}(A + I_n)\big) - \pi_{\text{sym}} \otimes \pi_{\text{sym}}\big\|_1 \leq C\eta^n.$$

Thus a fixed diagonal shift of full support can make the two coranks asymptotically independent.

Theorem 1.6 is proved in the more precise form of Theorem 4.20. Lee and Jung–Lee determine joint cokernel modules for polynomial or shifted transforms of Haar random $\mathbb{Z}_p$-matrices [23, 15, 24]. Those results concern different p-adic ensembles and determine full joint cokernel modules. Our theorem treats only the coranks of symmetry-defect fibres over $\mathbb{F}_2$ and gives an error at most $C_\alpha \eta_\alpha^n$ in total variation, uniformly when the defect has rank at least αn. In the arithmetic application, $\mathbf{v} = \mathbf{u}_{d_0}$ records the fixed twist by d_0, and $A, A + \operatorname{diag}(\mathbf{u}_{d_0})$ are the common cores of the two Rédei matrices.

In the actual Rédei matrices, every prime-discriminant factor of $2d_0$ contributes border and corner data of bounded width. Theorem 4.13 treats the resulting bordered pairs uniformly in this data and in every resulting 2-adic stratum. Writing $r = \omega(d)$, it gives an ℓ^1-error $O_{d_0}(\vartheta^r)$, for some $\vartheta \in (0,1)$, from $\pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}$ whenever the symmetry defect has rank at least a fixed positive proportion of r. The sign vectors not covered by this rank condition have exponentially small proportion in the uniform residue-symbol distribution. Applying the resulting estimate on all admissible boxes completes the proof of Theorem 1.1.

1.4. **Organization of the paper.** Section 2 proves the Rédei reduction, defines the residue and matrix probability spaces, identifies their uniform measures by an affine bijection, writes the corank-pair pushforward as an explicit sign mixture, and collects the finite-field tools. Section 3 decomposes the family (1) into admissible boxes, proves equidistribution of the symmetrised full residue assignment, and transfers this estimate through the finite pushforward to the original family. Section 4 develops one- and two-dimensional quantitative truncated inversion for Gaussian-binomial moments, proves the bordered joint-corank theorem from conditional mixed moments on a good event, and records the unbordered diagonal-perturbation consequences. Section 5 applies the family-transfer theorem and the bordered joint-corank estimate to prove Theorem 1.1, and then determines the Hasse unit index and the direct-sum defect needed for Theorem 1.3 and Corollary 1.5.

Acknowledgements. The authors thank Yi Ouyang and Ye Tian for their support and helpful discussions. The first author was partially supported by the Tianyuan Fund for Mathematics of the National Natural Science Foundation of China (Grant No. 12526538) and by the Postdoctoral Fellowship Program of CPSF (Grant No. GZC20252038). The authors used generative-AI tools for bibliographic searches, language editing, and technical cross-checking, and take full responsibility for the content.

2. Rédei reduction and finite-field counting

This section collects the algebraic and finite-field preliminaries needed in the proof. We first recall the Rédei–Reichardt description of the 4-rank and the coordinate deletion that converts it into a matrix corank. Applying this construction simultaneously to $-d$ and $-d_0 d$ produces two matrices over $\mathbb{F}_2$ with a common core. We then regard the residue symbols occurring in these matrices as formal variables and introduce finite probability spaces for the resulting pair of coranks. The second part records the finite-field counting results needed to analyze these spaces.

Throughout the remainder of the paper, vectors over finite fields are regarded as column vectors, and we reserve rank and corank for matrices and linear maps. The bracket $\left[\frac{a}{n}\right] \in \mathbb{F}_2$ denotes the additive Kronecker symbol, and hence the additive Jacobi or Legendre symbol in the usual cases.

2.1. **From residue symbols to joint 4-ranks.**

2.1.1. *Rédei matrices and coordinate deletion.* Let Δ be a fundamental discriminant, put $K_\Delta = \mathbb{Q}(\sqrt{\Delta})$, and choose an ordering

$$\Delta = \Delta_1 \cdots \Delta_t$$

of its prime discriminant factors, where $\Delta_i \in \{-4, 8, -8\} \cup \{p^* : p \text{ odd prime}\}$, with $p^* = (-1)^{(p-1)/2} p$. Let p_i be the prime below Δ_i, with $p_i = 2$ at the even prime discriminant. Define the Rédei matrix $R_\Delta \in \mathrm{Mat}_t(\mathbb{F}_2)$ by

$$(R_\Delta)_{ij} = \left[\frac{\Delta_i}{p_j}\right] \quad (i \neq j), \qquad (R_\Delta)_{jj} = \sum_{i \neq j} (R_\Delta)_{ij}. \tag{7}$$

Thus every column of R_Δ sums to zero.

The following criterion is due to Rédei and Reichardt [29].

Theorem 2.1 (Rédei–Reichardt criterion)**.** *The* 4*-rank of the narrow class group of* K_Δ *is given by*

$$\operatorname{rk}_4 \mathrm{Cl}^+(K_\Delta) = \operatorname{corank} R_\Delta - 1. \tag{8}$$

Suppose henceforth that the fundamental discriminant Δ is negative, and retain the factorisation and notation above. The ordinary and narrow class groups of K_Δ then coincide. Let I be the index set of the prime discriminants and write $\Delta_{\mathrm{even}} \in \{1, -4, 8, -8\}$ for the even prime discriminant, with value 1 when none occurs; in that case no even index is introduced. For a condition E, let $\mathbb{1}_{\{E\}} \in \{0, 1\}$ denote its indicator, viewed in $\mathbb{F}_2$ when it occurs in a vector or matrix. Without a subscript, $\mathbf{1}$ denotes the all-ones vector whose length is determined by context. Define

$$\mathbf{u}_{-1}(\Delta) = \big(\mathbb{1}_{\{\Delta_i<0\}}\big)_{i\in I}, \qquad \big(\mathbf{u}_2(\Delta)\big)_i = \begin{cases} \left[\frac{2}{p_i}\right], & p_i \neq 2, \\ 0, & p_i = 2. \end{cases}$$

At an odd index, $\big(\mathbf{u}_{-1}(\Delta)\big)_i = \left[\frac{-1}{p_i}\right]$. At the even index the same formula cannot be used, since $\left[\frac{-1}{2}\right] = 0$ does not distinguish a positive even prime discriminant from a negative one. For any vector $\mathbf{x}$ indexed by a finite set, let $\operatorname{diag}(\mathbf{x})$ denote the diagonal matrix with diagonal entries x_i, and write

$$\Omega_{\mathbf{x}} = \mathbf{x}\mathbf{x}^\top + \operatorname{diag}(\mathbf{x}).$$

If an even index occurs, we label it by 0 and denote its standard basis vector by $\mathbf{e}_0$.

Proposition 2.2. *Let* $\Delta < 0$*. Then*

$$R_\Delta^\top - R_\Delta = \Omega_{\mathbf{u}_{-1}(\Delta)} + \mathbb{1}_{\{\Delta_{\mathrm{even}}=-4\}}\big(\mathbf{e}_0\mathbf{u}_2(\Delta)^\top + \mathbf{u}_2(\Delta)\mathbf{e}_0^\top\big). \tag{9}$$

In particular, the symmetry defect $R_\Delta^\top - R_\Delta$ *is determined by the signs of the prime discriminants and the symbols at* 2*. It contains no pair symbol* $\left[\frac{p_i^*}{p_j}\right]$*.*

Proof. For distinct odd primes p, q, quadratic reciprocity and the definition of the Kronecker symbol give

$$\left[\frac{p^*}{q}\right] + \left[\frac{q^*}{p}\right] = \left[\frac{-1}{p}\right]\left[\frac{-1}{q}\right], \qquad \left[\frac{p^*}{2}\right] = \left[\frac{2}{p}\right],$$
$$\left[\frac{\Delta_{\mathrm{even}}}{p}\right] + \left[\frac{p^*}{2}\right] = \begin{cases} \left[\frac{-1}{p}\right] + \left[\frac{2}{p}\right], & \Delta_{\mathrm{even}} = -4, \\ 0, & \Delta_{\mathrm{even}} = 8, \\ \left[\frac{-1}{p}\right], & \Delta_{\mathrm{even}} = -8. \end{cases} \tag{10}$$

The first identity gives the assertion at two odd indices, and the second line handles an even and an odd index. Only $\Delta_{\mathrm{even}} = -4$ contributes the additional $\mathbf{u}_2(\Delta)$ term. Both sides have zero diagonal. ☐

The defect formula turns the column-sum convention in (7) into the two relations

$$\mathbf{1}^\top R_\Delta = 0, \qquad R_\Delta\mathbf{1} = (R_\Delta^\top - R_\Delta)\mathbf{1}. \tag{11}$$

The first is the defining column-sum relation. Its transpose gives $R_\Delta^\top\mathbf{1} = 0$, and over $\mathbb{F}_2$ this implies the second identity. Together they allow one odd prime discriminant to be removed without changing the relevant corank.

Proposition 2.3. *Let* ρ *be an odd index, so that* $\Delta_\rho = p_\rho^*$*. If* $R_\Delta^{(\rho)}$ *denotes the principal submatrix obtained by deleting row and column* ρ*, then*

$$\operatorname{rk}_4 \mathrm{Cl}(K_\Delta) = \operatorname{corank} R_\Delta^{(\rho)}. \tag{12}$$

Thus the right-hand side is independent of the odd index chosen for deletion.

Proof. Add all columns of R_Δ other than column ρ to column ρ. By (11), the new column is $(R_\Delta^\top - R_\Delta)\mathbf{1}$. The same equation shows that row ρ is the sum of the remaining rows, so deleting it does not change the rank. The resulting $(t-1)\times t$ matrix is

$$\begin{pmatrix} R_\Delta^{(\rho)} & \left((R_\Delta^\top - R_\Delta)\mathbf{1}\right)_{I\setminus\{\rho\}} \end{pmatrix}.$$

Since $\Delta < 0$, the number of negative prime discriminants is odd, and hence $\mathbf{1}^\top \mathbf{u}_{-1}(\Delta) = 1$ in $\mathbb{F}_2$. Formula (9) therefore gives

$$(R_\Delta^\top - R_\Delta)\mathbf{1} = \mathbb{1}_{\{\Delta_{\mathrm{even}}=-4\}}\big(\mathbf{e}_0(\mathbf{1}^\top \mathbf{u}_2(\Delta)) + \mathbf{u}_2(\Delta)\big).$$

If $\Delta_{\mathrm{even}} \neq -4$, this vector is zero. If $\Delta_{\mathrm{even}} = -4$, its restriction to $I \setminus \{\rho\}$ is column 0 of $R_\Delta^{(\rho)}$. Indeed, this follows from $\left[\frac{p_i^*}{2}\right] = \left[\frac{2}{p_i}\right]$ and the column-sum definition of the diagonal. In either case the last column lies in the image of $R_\Delta^{(\rho)}$. Hence

$$\operatorname{rank} R_\Delta = \operatorname{rank} R_\Delta^{(\rho)}, \qquad \operatorname{corank} R_\Delta = \operatorname{corank} R_\Delta^{(\rho)} + 1,$$

and (8) proves the result. $\square$

The oddness of the deleted index is essential. At an even index the last column in the proof need not lie in the image of the principal submatrix.

2.1.2. *The coupled matrix pair.* We apply Proposition 2.3 simultaneously to the Rédei matrices for $-d$ and $-d_0d$. After one common odd-prime coordinate is deleted, the two reduced matrices share a variable block; their remaining coupling lies in diagonal corrections and fixed-width borders.

Fix squarefree integers $d_0 > 1$ and d with $(d, 2d_0) = 1$, and write

$$\begin{aligned} d_0 &= 2^e q_1 \cdots q_s, \quad e \in \{0,1\}, \quad q_1 < \cdots < q_s \text{ odd},\\ d &= p_1 \cdots p_r, \qquad\qquad\qquad\quad p_1 < \cdots < p_r, \qquad r = \omega(d) \geq 1. \end{aligned}$$

For a squarefree integer m, write $\Delta(m)$ for the fundamental discriminant of $\mathbb{Q}(\sqrt{m})$, with $\Delta(1) = 1$. Put

$$\kappa = \#\{i : p_i \equiv 3 \bmod 4\}, \qquad P = \prod_{i\leq r} p_i^* = (-1)^\kappa d.$$

The prime-discriminant factorizations split as

$$\Delta(-d) = P\,\Delta((-1)^{\kappa+1}), \qquad \Delta(-d_0d) = P\,\Delta((-1)^{\kappa+1}d_0). \tag{13}$$

Thus P is the common variable factor, while the remaining factors determine the fixed widths. The first is 1 when κ is odd and -4 when κ is even, so put

$$b_1 = \mathbb{1}_{\{2|\kappa\}}.$$

Factor the second fixed discriminant as

$$\Delta((-1)^{\kappa+1}d_0) = \prod_{k\in\mathcal{D}_2} \Delta_k, \qquad b_2 = |\mathcal{D}_2|,$$

where $\Delta_k = q_k^*$ for $1 \leq k \leq s$. If an even prime discriminant occurs, denote it by Δ_0, index it by 0, and put $q_0 = 2$. Thus the two full Rédei matrices have orders $r + b_1$ and $r + b_2$, and their reductions after deleting one variable coordinate have orders $r - 1 + b_1$ and $r - 1 + b_2$.

The corresponding fixed and variable Rédei blocks are

$$R_{d_0,\kappa} := R_{\Delta((-1)^{\kappa+1}d_0)} \in \operatorname{Mat}_{b_2}(\mathbb{F}_2), \qquad R_d := R_P \in \operatorname{Mat}_r(\mathbb{F}_2),$$

with the rows and columns of $R_{d_0,\kappa}$ indexed by $\mathcal{D}_2$. For every integer λ coprime to d, write

$$u_{\lambda,i} = \left[\frac{\lambda}{p_i}\right], \qquad \mathbf{u}_\lambda = (u_{\lambda,i})_{i\leq r} \in \mathbb{F}_2^r.$$

In particular, $\kappa = |\mathbf{u}_{-1}|$, and

$$R_d^\top - R_d = \Omega_{\mathbf{u}_{-1}}. \tag{14}$$

For a vector indexed by $p_1, \ldots, p_r$, a superscript $^\circ$ means that the last coordinate is deleted. For a matrix whose rows have these indices, it means that the last row is deleted; if its columns have the same indices, the last column is deleted as well.

Theorem 2.4. *For $k \in \mathcal{D}_2$, put*

$$c_k = \left[\frac{P}{q_k}\right] = \begin{cases} \left[\frac{2}{d}\right], & k = 0, \\ \left[\frac{(-1)^\kappa d}{q_k}\right], & 1 \le k \le s. \end{cases} \tag{15}$$

The two 4-ranks are the coranks of the following reductions of the Rédei matrices:

$$r_4(-d) = \operatorname{corank} N_1, \qquad N_1 = \begin{cases} R_d^\circ, & b_1 = 0, \\ \begin{pmatrix} R_d^\circ + \operatorname{diag}(\mathbf{u}_{-1}^\circ) & \mathbf{u}_2^\circ \\ (\mathbf{u}_{-1}^\circ)^\top & \left[\frac{2}{d}\right] \end{pmatrix}, & b_1 = 1, \end{cases}$$

and

$$r_4(-d_0 d) = \operatorname{corank} N_2, \qquad N_2 = \begin{pmatrix} R_d^\circ + b_1 \operatorname{diag}(\mathbf{u}_{-1}^\circ) + \operatorname{diag}(\mathbf{u}_{d_0}^\circ) & (\mathbf{u}_{q_k}^\circ)_{k \in \mathcal{D}_2} \\ \big((\mathbf{u}_{\Delta_k}^\circ)_{k \in \mathcal{D}_2}\big)^\top & R_{d_0,\kappa} + \operatorname{diag}\big((c_k)_{k \in \mathcal{D}_2}\big) \end{pmatrix}.$$

Proof. Between two variable indices, the off-diagonal entries of both full Rédei matrices are those of R_d. The quotient $\Delta(-d)/P$ is 1 or -4, according as $b_1 = 0$ or 1. Its contribution to the diagonal at p_i is therefore $b_1 u_{-1,i}$, which gives the displayed common block.

The two fixed diagonal contributions differ by

$$\left[\frac{\Delta(-d)/P}{p_i}\right] + \left[\frac{\Delta(-d_0 d)/P}{p_i}\right] = \left[\frac{\Delta(-d)\Delta(-d_0 d)/d^2}{p_i}\right] = \left[\frac{d_0}{p_i}\right].$$

Indeed, $\Delta(-d)\Delta(-d_0 d)/(d_0 d^2)$ is one of $1, 4, 16$. Consequently the second variable block differs from the first by $\operatorname{diag}(\mathbf{u}_{d_0}^\circ)$.

It remains to identify the borders and corners. The upper and lower border entries are the vectors $\mathbf{u}_{q_k}$ and $\mathbf{u}_{\Delta_k}$, respectively, by quadratic reciprocity, including at the even index. When $b_1 = 1$, the corresponding vectors for the first matrix are $\mathbf{u}_2$ and $\mathbf{u}_{-1}$, and its corner entry is $\left[\frac{P}{2}\right] = \left[\frac{2}{d}\right]$.

For the second corner block, the column-sum convention gives

$$(R_{\Delta(-d_0 d)})_{kk} = \sum_{i \le r} \left[\frac{p_i^*}{q_k}\right] + \sum_{\substack{\ell \in \mathcal{D}_2 \\ \ell \ne k}} \left[\frac{\Delta_\ell}{q_k}\right] = \left[\frac{P}{q_k}\right] + \sum_{\substack{\ell \in \mathcal{D}_2 \\ \ell \ne k}} \left[\frac{\Delta_\ell}{q_k}\right],$$

which is the displayed lower-right diagonal entry; the off-diagonal entries are $\left[\frac{\Delta_k}{q_\ell}\right]$.

Finally, apply Proposition 2.3 to $\Delta(-d)$ and $\Delta(-d_0 d)$, deleting in each case the coordinate corresponding to p_r. This gives the displayed matrices. $\square$

2.1.3. *Residue assignments and the joint 4-rank map.* Retain the factorisation $d_0 = 2^e q_1 \cdots q_s$ of Section 2.1.2. Since d_0 remains fixed, we suppress it from the notation whenever no ambiguity can arise. We now replace the primes by formal residue assignments and then convert these assignments into the block data of Theorem 2.4. For $r \ge 1$, define

$$\mathsf{Res}_r := \mathbb{F}_2^{\{(i,j):1 \le i < j \le r\}} \times (\mathbb{F}_2^r)^{s+2}. \tag{16}$$

We write its elements as

$$\boldsymbol{\chi} = \big((\chi_{ij})_{i<j}; \mathbf{u}_{-1}, \mathbf{u}_2, \mathbf{u}_{\Delta_1}, \ldots, \mathbf{u}_{\Delta_s}\big).$$

For an ordered tuple of distinct odd primes $(p_1, \dots, p_r)$, all coprime to $2d_0$, its residue assignment is

$$\boldsymbol{\chi}(p_1, \dots, p_r) = \left(\left[\frac{p_i^*}{p_j} \right]_{i<j} ;\ \mathbf{u}_{-1}, \mathbf{u}_2, \mathbf{u}_{\Delta_1}, \dots, \mathbf{u}_{\Delta_s} \right) \in \mathsf{Res}_r. \tag{17}$$

When $d = p_1 \cdots p_r$ and $p_1 < \cdots < p_r$, we write this assignment as $\boldsymbol{\chi}(d)$. Thus

$$\#\mathsf{Res}_r = 2^{\binom{r}{2}+(s+2)r}, \tag{18}$$

and for now Res_r is regarded only as a finite set.

We first pass from residue symbols to matrix coordinates. Fix $r \geq 2$ and a sign vector $\mathbf{u}_{-1} = (u_{-1,1}, \dots, u_{-1,r}) \in \mathbb{F}_2^r$. Let

$$\mathsf{Res}_r(\mathbf{u}_{-1}) := \{\boldsymbol{\chi} \in \mathsf{Res}_r : \text{its sign vector is } \mathbf{u}_{-1}\},$$

and put

$$n = r - 1, \qquad \kappa = |\mathbf{u}_{-1}|, \qquad b_1 = \mathbb{1}_{\{2|\kappa\}}, \qquad \Omega = \Omega_{\mathbf{u}_{-1}^\circ}.$$

Definition 2.5. For $\Omega' \in \mathrm{Mat}_n(\mathbb{F}_2)$, set

$$\mathrm{Sym}_n(\Omega') := \{A \in \mathrm{Mat}_n(\mathbb{F}_2) : A^\top - A = \Omega'\}. \tag{19}$$

Definition 2.6. The coupled-matrix data space over the sign fibre $\mathbf{u}_{-1}$ is

$$\mathsf{MatData}_r(\mathbf{u}_{-1}) := \mathrm{Sym}_n(\Omega) \times (\mathbb{F}_2^n)^{1+s} \times \mathbb{F}_2^{1+s}. \tag{20}$$

We write its elements as

$$x = (A, \mathbf{u}_2, \mathbf{u}_{\Delta_1}, \dots, \mathbf{u}_{\Delta_s}, \mathbf{c}), \qquad \mathbf{c} = (c_0, \dots, c_s).$$

For $\boldsymbol{\chi} \in \mathsf{Res}_r(\mathbf{u}_{-1})$, define $R(\boldsymbol{\chi}) \in \mathrm{Mat}_r(\mathbb{F}_2)$ by

$$R(\boldsymbol{\chi})_{ij} = \begin{cases} \chi_{ij}, & i < j, \\ \chi_{ji} + u_{-1,i}u_{-1,j}, & j < i, \end{cases} \qquad R(\boldsymbol{\chi})_{jj} = \sum_{i \neq j} R(\boldsymbol{\chi})_{ij}.$$

Then $R(\boldsymbol{\chi})^\top - R(\boldsymbol{\chi}) = \Omega_{\mathbf{u}_{-1}}$. Put

$$A(\boldsymbol{\chi}) := R(\boldsymbol{\chi})^\circ + b_1 \,\mathrm{diag}(\mathbf{u}_{-1}^\circ),$$

so that $A(\boldsymbol{\chi}) \in \mathrm{Sym}_n(\Omega)$, and set

$$c_0(\boldsymbol{\chi}) = \sum_{i \leq r} u_{2,i}, \qquad c_k(\boldsymbol{\chi}) = \sum_{i \leq r} \left(u_{\Delta_k,i} + \left[\frac{-1}{q_k} \right] u_{-1,i} \right) \quad (1 \leq k \leq s). \tag{21}$$

The coordinate map is

$$\mathrm{coord}_{\mathbf{u}_{-1}}(\boldsymbol{\chi}) := \big(A(\boldsymbol{\chi}), \mathbf{u}_2^\circ, \mathbf{u}_{\Delta_1}^\circ, \dots, \mathbf{u}_{\Delta_s}^\circ, (c_0(\boldsymbol{\chi}), \dots, c_s(\boldsymbol{\chi}))\big). \tag{22}$$

Lemma 2.7. *The map*

$$\mathrm{coord}_{\mathbf{u}_{-1}} : \mathsf{Res}_r(\mathbf{u}_{-1}) \xrightarrow{\sim} \mathsf{MatData}_r(\mathbf{u}_{-1})$$

is an affine bijection.

Proof. The pair-symbol coordinates not involving r give the off-diagonal entries of A, while those involving r determine its diagonal through the column-sum rule. Hence they parametrize $\mathrm{Sym}_n(\Omega)$ bijectively. For each of the $s+1$ character vectors, its first n entries together with the corresponding coordinate c_k recover its last entry by (21). These disjoint affine changes of coordinates prove the claim. $\square$

We next assemble a point of $\mathsf{MatData}_r(\mathbf{u}_{-1})$ into the two matrices of Theorem 2.4. Put $\kappa_0 = \#\{k : q_k \equiv 3 \bmod 4\}$. We call $(e, \kappa \bmod 2, \kappa_0 \bmod 2)$ the 2-adic stratum; it determines $\mathcal{D}_2$, the even prime discriminant Δ_0 when present, and the border widths. Explicitly,

$$\Delta_0 = \begin{cases} -4, & e=0,\ \kappa+\kappa_0 \text{ even}, \\ 8, & e=1,\ \kappa+\kappa_0 \text{ odd}, \\ -8, & e=1,\ \kappa+\kappa_0 \text{ even}, \end{cases}$$

with no even index when $e=0$ and $\kappa+\kappa_0$ is odd, and

$$b_1 = \mathbb{1}_{\{2|\kappa\}}, \qquad b_2 = s+e+(1-e)\mathbb{1}_{\{2|(\kappa+\kappa_0)\}}. \tag{23}$$

For $x = (A, \mathbf{u}_2, \mathbf{u}_{\Delta_1}, \ldots, \mathbf{u}_{\Delta_s}, \mathbf{c}) \in \mathsf{MatData}_r(\mathbf{u}_{-1})$, define the blocks below. If $0 \in \mathcal{D}_2$, define

$$\mathbf{u}_{\Delta_0} = \begin{cases} \mathbf{u}_{-1}^\circ, & \Delta_0 = -4, \\ \mathbf{u}_2, & \Delta_0 = 8, \\ \mathbf{u}_{-1}^\circ + \mathbf{u}_2, & \Delta_0 = -8. \end{cases}$$

The remaining blocks are

$$\mathbf{u}_{d_0} = b_1 \mathbf{u}_{-1}^\circ + \sum_{k\in\mathcal{D}_2} \mathbf{u}_{\Delta_k}, \tag{24}$$

$$\mathbf{u}_{q_k} = \begin{cases} \mathbf{u}_2, & k=0, \\ \mathbf{u}_{\Delta_k} + \left[\frac{-1}{q_k}\right] \mathbf{u}_{-1}^\circ, & 1 \le k \le s, \end{cases} \qquad (k \in \mathcal{D}_2), \tag{25}$$

and

$$\mathsf{L} = (\mathbf{u}_{\Delta_k})_{k\in\mathcal{D}_2}, \qquad \mathsf{U} = (\mathbf{u}_{q_k})_{k\in\mathcal{D}_2}, \qquad C = R_{d_0,\kappa} + \operatorname{diag}\big((c_k)_{k\in\mathcal{D}_2}\big).$$

If $b_1 = 0$, let $\mathsf{U}_1, \mathsf{L}_1, \mathsf{C}_1$ be empty blocks of compatible sizes; if $b_1 = 1$, put

$$\mathsf{U}_1 = \mathbf{u}_2, \qquad \mathsf{L}_1 = \mathbf{u}_{-1}^\circ, \qquad \mathsf{C}_1 = (c_0).$$

Define

$$N_1(x) = \begin{pmatrix} A & \mathsf{U}_1 \\ \mathsf{L}_1^\top & \mathsf{C}_1 \end{pmatrix}, \qquad N_2(x) = \begin{pmatrix} A + \operatorname{diag}(\mathbf{u}_{d_0}) & \mathsf{U} \\ \mathsf{L}^\top & C \end{pmatrix}. \tag{26}$$

This is the assembly map

$$\mathrm{Asm}_{d_0,r,\mathbf{u}_{-1}} : \mathsf{MatData}_r(\mathbf{u}_{-1}) \longrightarrow \mathrm{Mat}_{n+b_1}(\mathbb{F}_2) \times \mathrm{Mat}_{n+b_2}(\mathbb{F}_2), \qquad x \longmapsto \big(N_1(x), N_2(x)\big). \tag{27}$$

For a matrix pair, write $\operatorname{corank}^{\times 2}(N_1, N_2) = (\operatorname{corank} N_1, \operatorname{corank} N_2)$. The coordinate and assembly maps define

$$\mathrm{Joint4rk}_r|_{\mathsf{Res}_r(\mathbf{u}_{-1})} := \operatorname{corank}^{\times 2} \circ \mathrm{Asm}_{d_0,r,\mathbf{u}_{-1}} \circ \mathrm{coord}_{\mathbf{u}_{-1}}. \tag{28}$$

Since the sign fibres partition Res_r, these composites give a single map

$$\mathrm{Joint4rk}_r : \mathsf{Res}_r \longrightarrow \mathbb{Z}_{\ge 0}^2. \tag{29}$$

For every d in the family with $\omega(d) = r \ge 2$, Theorem 2.4 gives

$$\big(r_4(-d), r_4(-d_0 d)\big) = \mathrm{Joint4rk}_r\big(\boldsymbol{\chi}(d)\big). \tag{30}$$

The cases $r \le 1$ have density zero and are omitted from this formalism.

We finally record how the construction behaves under relabelling and under uniform measure. The symmetric group $\mathfrak{S}_r$ acts on Res_r by relabelling the variable primes; on single-prime coordinates, $(\pi \cdot \mathbf{u}_\lambda)_i = u_{\lambda,\pi^{-1}(i)}$. For $i<j$, quadratic reciprocity gives the pair-coordinate action

$$(\pi\cdot\boldsymbol{\chi})_{ij} = \begin{cases} \chi_{\pi^{-1}(i),\pi^{-1}(j)}, & \pi^{-1}(i) < \pi^{-1}(j), \\ \chi_{\pi^{-1}(j),\pi^{-1}(i)} + u_{-1,\pi^{-1}(i)} u_{-1,\pi^{-1}(j)}, & \pi^{-1}(i) > \pi^{-1}(j). \end{cases} \tag{31}$$

Thus $\pi \cdot \boldsymbol{\chi}(p_1, \dots, p_r) = \boldsymbol{\chi}(p_{\pi^{-1}(1)}, \dots, p_{\pi^{-1}(r)})$, and every π acts by a bijection. The coordinate map itself uses the distinguished index r and is not permutation invariant.

Lemma 2.8. *The map* $\mathrm{Joint4rk}_r$ *is* $\mathfrak{S}_r$*-invariant.*

Proof. Relabelling the variable indices conjugates the two unreduced matrices determined by $\boldsymbol{\chi}$ by the corresponding permutation matrix, with the fixed indices unchanged. Proposition 2.3 makes the coranks of their reductions independent of the deleted variable coordinate, proving the claim. $\square$

For the probabilistic formulation, equip Res_r and each $\mathsf{MatData}_r(\mathbf{u}_{-1})$ with uniform probability measure. Under $\mathrm{Unif}(\mathsf{Res}_r)$, the sign vector is uniform; conditional on its value $\mathbf{u}_{-1}$, Lemma 2.7 identifies the uniform measures on $\mathsf{Res}_r(\mathbf{u}_{-1})$ and $\mathsf{MatData}_r(\mathbf{u}_{-1})$. The latter has independent coordinates

$$A \sim \mathrm{Unif}\big(\mathrm{Sym}_n(\Omega)\big), \qquad \mathbf{u}_2, \mathbf{u}_{\Delta_1}, \dots, \mathbf{u}_{\Delta_s} \sim \mathrm{Unif}(\mathbb{F}_2^n), \qquad \mathbf{c} \sim \mathrm{Unif}(\mathbb{F}_2^{1+s}).$$

Let N_1, N_2 be the resulting random matrices and put

$$\mathbf{Z} = (Z_1, Z_2) := \mathrm{corank}^{\times 2}(N_1, N_2).$$

Then (28) gives

$$(\mathrm{Joint4rk}_r)_* \,\mathrm{Unif}\big(\mathsf{Res}_r(\mathbf{u}_{-1})\big) = \mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}). \tag{32}$$

Averaging (32) over the sign vector gives, for every $r \geq 2$,

$$(\mathrm{Joint4rk}_r)_* \,\mathrm{Unif}(\mathsf{Res}_r) = 2^{-r} \sum_{\mathbf{u}_{-1} \in \mathbb{F}_2^r} \mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}). \tag{33}$$

Permutation invariance also allows us to symmetrise a measure on Res_r without changing its joint 4-rank pushforward. For a probability measure μ on Res_r, put

$$\mu^{\mathrm{sym}} := \frac{1}{r!} \sum_{\pi \in \mathfrak{S}_r} \pi_* \mu.$$

Lemma 2.8 and contraction under pushforward give

$$(\mathrm{Joint4rk}_r)_* \mu^{\mathrm{sym}} = (\mathrm{Joint4rk}_r)_* \mu, \qquad \|(\mathrm{Joint4rk}_r)_* \mu - (\mathrm{Joint4rk}_r)_* \nu\|_1 \leq \|\mu - \nu\|_1 \tag{34}$$

for all probability measures μ, ν on Res_r.

For later use, if $0 < \alpha < 1/2$, Hoeffding's inequality gives

$$2^{-r} \#\{\mathbf{u}_{-1} \in \mathbb{F}_2^r : |\mathbf{u}_{-1}| < \alpha r\} \leq e^{-2(1/2-\alpha)^2 r}. \tag{35}$$

2.2. **Counting for moments and truncated inversion.** The moment calculations in Section 4 use three finite-field ingredients: simultaneous constraints $A\mathbf{x}_i = \mathbf{y}_i$ on a symmetry-defect fibre, Gaussian-binomial inversion, and incidence counts for the tuples occurring in the moment expansion. We record them in that order.

2.2.1. *Joint surjectivity on symmetry-defect fibres.* For $\boldsymbol{\xi} \in \mathbb{F}_2^n$, set $h = |\boldsymbol{\xi}|$. The matrix $\Omega_{\boldsymbol{\xi}}$ defined in Section 2.1.1 has zero diagonal, so it lies in Alt_n. We call $h = 0$ the symmetric case and $h = n$ the all-ones case.

The linear map $A \mapsto A^\top - A$ has kernel Sym_n and image Alt_n. Hence, for every $\Omega \in \mathrm{Alt}_n$, the fibre $\mathrm{Sym}_n(\Omega)$ is an affine translate of Sym_n and has cardinality $2^{\binom{n+1}{2}}$.

To describe these fibres explicitly, let I_m and J_m denote the $m \times m$ identity and all-ones matrices, respectively. Up to coordinate permutation, the fibre attached to $\boldsymbol{\xi}$ depends only on

h: conjugation by a permutation matrix identifies the corresponding fibres and preserves corank. We may therefore order the coordinates so that $\xi_i = 1$ exactly for $i \le h$; then

$$\mathrm{Sym}_n(\Omega_{\boldsymbol{\xi}}) = \left\{ \begin{pmatrix} A_1 & V \\ V^\top & A_2 \end{pmatrix} : A_1 \in \mathrm{Sym}_h(J_h + I_h), \quad A_2 \in \mathrm{Sym}_{n-h}, \quad V \in \mathrm{Mat}_{h\times(n-h)}(\mathbb{F}_2) \right\}.$$

In these coordinates,

$$\operatorname{rank} \Omega_{\boldsymbol{\xi}} \in \{h-1, h\}. \tag{36}$$

Thus the symmetry defect has rank 0 in the symmetric case and rank at least $\alpha n - 1$ whenever $|\boldsymbol{\xi}| \ge \alpha n$. For $n \le 1$, $J_n + I_n = 0$ and the fibre is simply Sym_n.

For the moment calculations, the required input from these fibres is the joint distribution of $A\mathbf{x}_1, \dots, A\mathbf{x}_k$ when $\mathbf{x}_1, \dots, \mathbf{x}_k$ are linearly independent.

Lemma 2.9. *Let* $\mathrm{Sym}_n(\Omega)$ *be as in Definition 2.5, with* $\Omega \in \mathrm{Alt}_n$*, and let* $1 \le k \le n$*. Suppose that* $\mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{F}_2^n$ *are linearly independent, and let* $\mathbf{y}_1, \dots, \mathbf{y}_k \in \mathbb{F}_2^n$ *be arbitrary. Then*

$$\mathbb{P}_{A\sim\mathrm{Unif}(\mathrm{Sym}_n(\Omega))}\big(A\mathbf{x}_i = \mathbf{y}_i \quad \forall i \le k\big) = \mathbb{1}_{\{\mathbf{y}_\bullet \in \Lambda\}} \cdot 2^{-\left(kn - \binom{k}{2}\right)},$$

where the compatibility set is

$$\Lambda := \big\{ \mathbf{y}_\bullet : \ \mathbf{x}_i^\top \mathbf{y}_j + \mathbf{x}_j^\top \mathbf{y}_i = \mathbf{x}_i^\top \Omega \mathbf{x}_j \quad (i<j) \big\}.$$

In particular, if $k = 1$*, then* $\mathbb{P}(A\mathbf{x} = \mathbf{y}) = 2^{-n}$ *for every* $\mathbf{x} \neq \mathbf{0}$ *and every* $\mathbf{y}$*. For general* k*,* $\mathbb{P}(A\mathbf{x}_i = \mathbf{y}_i \, \forall i) \le 2^{-(kn - \binom{k}{2})}$ *for every* Ω*.*

Proof. Fix $A_0 \in \mathrm{Sym}_n(\Omega)$. Since $\mathrm{Sym}_n(\Omega) = A_0 + \mathrm{Sym}_n$, $S := A - A_0$ is uniform on Sym_n, and $A\mathbf{x}_i = \mathbf{y}_i$ reads $S\mathbf{x}_i = \mathbf{z}_i$ with $\mathbf{z}_i := \mathbf{y}_i - A_0\mathbf{x}_i$. Extend $\mathbf{x}_1, \dots, \mathbf{x}_k$ to a basis of $\mathbb{F}_2^n$ and consider the $\mathbb{F}_2$-linear map

$$\Phi : \mathrm{Sym}_n \to (\mathbb{F}_2^n)^k, \qquad S \mapsto (S\mathbf{x}_1, \dots, S\mathbf{x}_k).$$

In the chosen basis, $S \in \ker \Phi$ if and only if the first k rows, and hence by symmetry the first k columns, of the matrix of the bilinear form S vanish; thus $\dim \ker \Phi = \binom{n-k+1}{2}$ and

$$\dim \operatorname{im} \Phi = \binom{n+1}{2} - \binom{n-k+1}{2} = kn - \binom{k}{2}.$$

Since S is symmetric, $\mathbf{x}_i^\top S \mathbf{x}_j = \mathbf{x}_j^\top S \mathbf{x}_i$, so $\operatorname{im} \Phi$ is contained in

$$\Lambda_0 := \{ \mathbf{z}_\bullet : \mathbf{x}_i^\top \mathbf{z}_j = \mathbf{x}_j^\top \mathbf{z}_i \ (i<j) \}.$$

The $\binom{k}{2}$ defining functionals are linearly independent because $\mathbf{x}_1, \dots, \mathbf{x}_k$ are, so $\dim \Lambda_0 = kn - \binom{k}{2} = \dim \operatorname{im} \Phi$. Hence $\operatorname{im} \Phi = \Lambda_0$.

The system is therefore solvable exactly when $\mathbf{z}_\bullet \in \Lambda_0$. After substituting $\mathbf{z}_i = \mathbf{y}_i - A_0\mathbf{x}_i$ and using $A_0^\top - A_0 = \Omega$, this condition becomes $\mathbf{y}_\bullet \in \Lambda$. When it holds, the solution set is a coset of $\ker \Phi$, and hence has probability

$$\frac{|\ker \Phi|}{|\mathrm{Sym}_n|} = 2^{-\left(kn - \binom{k}{2}\right)}.$$

$\square$

For a dependent tuple, apply Lemma 2.9 to a basis of its span. The remaining equations are then either forced or inconsistent. Consequently, the probability exponent depends only on $\dim \operatorname{span}(\mathbf{x}_\bullet)$, while consistency is expressed by pairwise bilinear conditions. This is the form needed in the moment calculation.

2.2.2. *Gaussian-binomial moments and inversion.* Gaussian-binomial moments convert the preceding probabilities into information about corank distributions. We state the required identities for an arbitrary prime power q, although only $q = 2$ is used later. With the notation $\eta_k(q)$ fixed in the introduction, write

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \prod_{i=0}^{k-1} \frac{q^{n-i}-1}{q^{k-i}-1} = q^{k(n-k)} \frac{\eta_n(q)}{\eta_k(q)\eta_{n-k}(q)} \qquad (0 \le k \le n),$$

with $\begin{bmatrix} n \\ 0 \end{bmatrix}_q := 1$ and $\begin{bmatrix} n \\ k \end{bmatrix}_q := 0$ for $k < 0$ or $k > n$. Thus $\begin{bmatrix} n \\ k \end{bmatrix}_q$ is the number of k-dimensional subspaces of an n-dimensional vector space over $\mathbb{F}_q$. We also write

$$|\operatorname{GL}_u(q)| = q^{\binom{u}{2}} \prod_{i=1}^{u} (q^i - 1) = q^{u^2} \eta_u(q), \qquad |\operatorname{GL}_u| := |\operatorname{GL}_u(2)|.$$

For a nonnegative integer-valued random variable X, its u-th Gaussian-binomial moment is

$$\mathbb{E} \begin{bmatrix} X \\ u \end{bmatrix}_2.$$

When X is the corank of a matrix, this is the expected number of u-dimensional subspaces of its kernel. The surjection moments used in random-group and random-cokernel problems satisfy

$$\mathbb{E}\# \operatorname{Sur}(\mathbb{F}_2^X, \mathbb{F}_2^u) = |\operatorname{GL}_u| \, \mathbb{E} \begin{bmatrix} X \\ u \end{bmatrix}_2.$$

Thus the two kinds of moments differ by the factor $|\operatorname{GL}_u|$; see [35, 34] for systematic treatments. We use Gaussian-binomial moments because the q-binomial transform below acts directly on $\mathbb{E}\begin{bmatrix} X \\ u \end{bmatrix}_2$.

For completeness, we record these standard identities and their proofs, so that the truncated-inversion argument used later is self-contained.

Proposition 2.10. *Let q be a prime power.*

(i) (Subspace-counting bounds) *For $0 \le k \le n$, $q^{k(n-k)} \le \begin{bmatrix} n \\ k \end{bmatrix}_q \le \eta_\infty(q)^{-1} q^{k(n-k)}$.*

(ii) (q-Pascal) *For $0 \le k \le n$, $\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k} \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q$.*

(iii) (q-chain) *For $0 \le a \le u \le x$, $\begin{bmatrix} x \\ u \end{bmatrix}_q \begin{bmatrix} u \\ a \end{bmatrix}_q = \begin{bmatrix} x \\ a \end{bmatrix}_q \begin{bmatrix} x-a \\ u-a \end{bmatrix}_q$.*

(iv) (q-binomial inversion) *For $m \ge 0$, $\sum_{j=0}^{m} (-1)^j q^{\binom{j}{2}} \begin{bmatrix} m \\ j \end{bmatrix}_q = \mathbb{1}_{\{m=0\}}$. Equivalently, the infinite lower-triangular matrices*

$$(G_q)_{x,u} := \begin{bmatrix} x \\ u \end{bmatrix}_q, \qquad (G_q^{-1})_{u,a} := (-1)^{u-a} q^{\binom{u-a}{2}} \begin{bmatrix} u \\ a \end{bmatrix}_q \quad (0 \le a \le u \le x) \tag{37}$$

satisfy, for all $x, a \ge 0$,

$$\sum_{u \ge 0} (G_q)_{x,u} (G_q^{-1})_{u,a} = \sum_{u \ge 0} (G_q^{-1})_{x,u} (G_q)_{u,a} = \mathbb{1}_{\{x=a\}}.$$

(v) (q-binomial remainder) *For $N \ge 0$ and $0 \le m \le N$,*

$$\sum_{j=0}^{m} (-1)^j q^{\binom{j}{2}} \begin{bmatrix} N \\ j \end{bmatrix}_q = (-1)^m q^{\binom{m+1}{2}} \begin{bmatrix} N-1 \\ m \end{bmatrix}_q. \tag{38}$$

Proof. For (i), write $\begin{bmatrix} n \\ k \end{bmatrix}_q = q^{k(n-k)} \prod_{i=1}^{k} (1 - q^{-(n-k+i)})/(1 - q^{-i})$. Every factor lies between 1 and $(1 - q^{-i})^{-1}$.

For (ii), classify the k-dimensional subspaces W of an n-dimensional space by whether W is contained in a fixed hyperplane. Formula (iii) follows by counting flags $W \subseteq U$ with $\dim W = a$ and $\dim U = u$ in two orders.

For (iv), substitute $t = -1$ in the q-binomial theorem $\prod_{i=0}^{m-1}(1+q^i t) = \sum_{j=0}^{m} q^{\binom{j}{2}} \begin{bmatrix} m \\ j \end{bmatrix}_q t^j$. Using (iii),

$$(G_q G_q^{-1})_{x,a} = \sum_{u=a}^{x} \begin{bmatrix} x \\ u \end{bmatrix}_q (-1)^{u-a} q^{\binom{u-a}{2}} \begin{bmatrix} u \\ a \end{bmatrix}_q = \begin{bmatrix} x \\ a \end{bmatrix}_q \sum_{j=0}^{x-a} (-1)^j q^{\binom{j}{2}} \begin{bmatrix} x-a \\ j \end{bmatrix}_q = \mathbb{1}_{\{x=a\}}.$$

The same calculation gives $G_q^{-1} G_q = I$.

Finally, (v) follows by induction on m. Expanding the last term with (ii) and using $\binom{m+1}{2} = \binom{m}{2} + m$ cancels the two contributions containing $\begin{bmatrix} N-1 \\ m-1 \end{bmatrix}_q$ and leaves the displayed expression. □

2.2.3. *Subspace-incidence estimates.* The moment calculations reduce to incidence counts for ordered independent tuples relative to fixed subspaces and alternating forms. We record the three estimates used later.

Lemma 2.11. *Let $K \subseteq \mathbb{F}_q^r$ have codimension γ. For $0 \le t \le u \le \gamma$, the number of ordered linearly independent u-tuples $(\mathbf{x}_1, \ldots, \mathbf{x}_u)$ satisfying $\dim(\mathrm{span}(\mathbf{x}_1, \ldots, \mathbf{x}_u) \cap K) = t$ equals*

$$\begin{bmatrix} u \\ t \end{bmatrix}_q q^{ru-t\gamma} \prod_{i=0}^{u-t-1} (1-q^{i-\gamma}) \prod_{i=0}^{t-1} (1-q^{i-(r-\gamma)}) \le \begin{bmatrix} u \\ t \end{bmatrix}_q q^{ru-t\gamma}.$$

In particular, when $t = 0$, this number is

$$q^{ru} \prod_{i=0}^{u-1} (1-q^{i-\gamma}) = q^{ru}(1 + O(q^{u-\gamma})),$$

and the tuples with $t \ge 1$ form an $O(q^{u-\gamma})$ proportion of all ordered independent u-tuples.

Proof. Put $E = \mathbb{F}_q^u$, $V = \mathbb{F}_q^r$, and $m = \dim K = r - \gamma$. Regard the tuple as the linear map $f : E \to V$ defined by $f(e_i) = \mathbf{x}_i$. The tuple is independent exactly when f is injective, and its span meets K in dimension t exactly when $T := f^{-1}(K)$ has dimension t.

Choose T in $\begin{bmatrix} u \\ t \end{bmatrix}_q$ ways. The restriction $f|_T : T \to K$ and the induced map $\overline{f} : E/T \to V/K$ must be injective, giving $\prod_{i=0}^{t-1}(q^m - q^i)$ and $\prod_{i=0}^{u-t-1}(q^\gamma - q^i)$ choices, respectively. Once they are fixed, lift $\overline{f}$ on a complement of T. Each of its $u-t$ basis vectors may be changed by an arbitrary element of K, giving $q^{m(u-t)}$ maps f, all injective with $f^{-1}(K) = T$. The required number is therefore

$$\begin{bmatrix} u \\ t \end{bmatrix}_q q^{m(u-t)} \prod_{i=0}^{t-1}(q^m - q^i) \prod_{i=0}^{u-t-1} (q^\gamma - q^i).$$

Factoring out the powers of q and using $m = r - \gamma$ gives the stated formula and its upper bound.

For $t = 0$, the same formula gives the displayed asymptotic. Finally, the total number of ordered independent u-tuples is $q^{ru} \prod_{i=0}^{u-1}(1 - q^{i-r})$, with the product bounded below by $\eta_\infty(q)$. Using $\begin{bmatrix} u \\ t \end{bmatrix}_q \le \eta_\infty(q)^{-1} q^{t(u-t)}$, the proportion with $t \ge 1$ is at most

$$\eta_\infty(q)^{-2} \sum_{t=1}^{u} q^{t(u-t-\gamma)} = O(q^{u-\gamma}),$$

as required. □

Lemma 2.12. *Let $W \subseteq \mathbb{F}_q^r$ have dimension m, and let $G \subseteq W$ have dimension g. Let $0 \le u \le m$. For an ordered linearly independent u-tuple $(\mathbf{y}_1, \dots, \mathbf{y}_u) \in W^u$, say that $\mathbf{y}_j$* falls in *if $\mathbf{y}_j \in G + \operatorname{span}(\mathbf{y}_1, \dots, \mathbf{y}_{j-1})$. The number of such tuples with exactly f fall-ins, where $0 \le f \le u$, is*

$$|\operatorname{GL}_u(q)|\, q^{(u-f)(g-f)} \begin{bmatrix} g \\ f \end{bmatrix}_q \begin{bmatrix} m-g \\ u-f \end{bmatrix}_q \le \binom{u}{f} q^{m(u-f)} q^{(g+u)f}.$$

Proof. A fall-in increases $\dim(G \cap \operatorname{span}(\mathbf{y}_1, \dots, \mathbf{y}_j))$ by one, whereas a non-fall-in leaves it unchanged. Hence the tuples counted are exactly the ordered bases of the u-dimensional subspaces $U \subseteq W$ with $\dim(U \cap G) = f$. The standard count of such subspaces gives the first formula. For the upper bound, choose the f fall-in positions and bound the number of choices at a fall-in by q^{g+u} and at any other position by q^m. □

Lemma 2.13. *Let B be an alternating bilinear form on $\mathbb{F}_q^r$, let $r_B := \operatorname{rank} B$, and let $\operatorname{rad}(B) := \{x : B(x, y) = 0$ for all $y\}$. A subspace U is* totally isotropic *if $B(x, y) = 0$ for all $x, y \in U$. The number $N_{\text{iso}}(u)$ of u-dimensional totally isotropic subspaces, for $0 \le u \le r$, satisfies*

$$N_{\text{iso}}(u) q^{-(ru - \binom{u}{2})} = \frac{1}{|\operatorname{GL}_u(q)|} \left(1 + O(q^{2u - r_B}) + O(q^{\binom{u+1}{2} - r_B})\right).$$

If B is nondegenerate, that is, $\operatorname{rad}(B) = 0$, the exact formula is

$$N_{\text{iso}}(u) q^{-(ru - \binom{u}{2})} = \frac{1}{|\operatorname{GL}_u(q)|} \prod_{i=0}^{u-1} (1 - q^{2i-r}).$$

Proof. Suppose first that B is nondegenerate. After choosing $\mathbf{x}_1, \dots, \mathbf{x}_i$, the next vector lies in their orthogonal complement but not in their span, giving $q^{r-i} - q^i$ choices. Dividing the product by the number $|\operatorname{GL}_u(q)|$ of ordered bases gives the exact formula.

For general B, one has $\dim \operatorname{rad}(B) = r - r_B$. Separate the isotropic subspaces U according as $U \cap \operatorname{rad}(B)$ is zero or nonzero. In the first case, projection to $\mathbb{F}_q^r / \operatorname{rad}(B)$ is injective, and each isotropic image has $q^{u(r-r_B)}$ lifts. The nondegenerate formula in dimension r_B contributes

$$\frac{1}{|\operatorname{GL}_u(q)|} \prod_{i=0}^{u-1} (1 - q^{2i - r_B}) = \frac{1}{|\operatorname{GL}_u(q)|} (1 + O(q^{2u - r_B})).$$

If $U \cap \operatorname{rad}(B) \ne 0$, then U contains a line in $\operatorname{rad}(B)$; the number of such subspaces is at most $(q^{r-r_B} - 1) \begin{bmatrix} r-1 \\ u-1 \end{bmatrix}_q$. After multiplication by $q^{-(ru - \binom{u}{2})}$, this contributes $|\operatorname{GL}_u(q)|^{-1} O(q^{\binom{u+1}{2} - r_B})$. □

3. From arithmetic boxes to random matrix pairs

For d with $r = \omega(d) \ge 2$, the joint 4-rank is $\operatorname{Joint4rk}_r(\boldsymbol{\chi}(d))$, where $\boldsymbol{\chi}(d) \in \mathsf{Res}_r$ records the relevant residue symbols; see (30). The cases $r \le 1$ have density zero. If instead $\boldsymbol{\chi}$ is uniform on Res_r, then (33) identifies the distribution of $\operatorname{Joint4rk}_r(\boldsymbol{\chi})$ with the corank-pair distribution of the random matrix model. Thus the arithmetic problem is reduced to comparing the assignments $\boldsymbol{\chi}(d)$ arising from primes with a uniform assignment $\boldsymbol{\chi}$. Since $\operatorname{Joint4rk}_r$ is invariant under relabelling and total variation decreases under pushforward, (34) shows that a bound for the permutation-averaged assignment distribution gives the corresponding bound for the joint 4-ranks.

Smith's box method provides this assignment comparison [31, Theorem 5.4, Propositions 6.9–6.10 and Theorem 6.4]. His positive covering reduces the count over $\mathcal{F}_{d_0}(X)$ to prime-coordinate boxes, in which the varying primes range independently over disjoint intervals. On each box satisfying the required size and spacing conditions, the residue-symbol theorem compares the permutation-averaged distribution of $\boldsymbol{\chi}(d)$ with the uniform measure on Res_r. Pushing this estimate forward by Joint4rk$_r$ gives the boxwise comparison with the random matrix model, and summing the box estimates returns the desired comparison over $\mathcal{F}_{d_0}(X)$.

3.1. **Admissible boxes.** Smith's box construction is motivated by the standard model in which, conditional on their number, the $\log\log p$-coordinates of the prime factors of a typical integer resemble the order statistics of uniform points on $[0, \log\log X]$; see [31, §5] and [13]. The precise estimates needed here are summarized in Proposition 3.1.

We now make the construction precise for the integer family. Fix the squarefree integer $d_0 > 1$ and recall $\mathcal{F}_{d_0}(X)$ from (1). All varying prime factors belong to

$$\mathcal{P}_{d_0} = \{p \text{ prime} : p \nmid 2d_0\}.$$

Smith's covering propositions are stated for integers having no prime factor below a fixed cutoff greater than 3. To pass to the family above without changing it, fix once and for all

$$D_* > \max\big(3, \max_{p \mid 2d_0} p\big).$$

Write $d = am$, where a is supported on the finitely many primes $p < D_*$ in $\mathcal{P}_{d_0}$. On each fixed-a slice, m belongs, in Smith's notation, to

$$S_{r-\omega(a)}(X/a, D_*).$$

We apply the covering there and then reinsert the prime factors of a as frozen singleton coordinates. There are $O_{d_0}(1)$ such slices, and the changes $X \mapsto X/a$ and $r \mapsto r - \omega(a)$ are bounded in terms of d_0. Thus all estimates below remain uniform after the slices are summed. References to Smith's positive covering are understood in this slice-wise sense.

For $r \geq 0$, let

$$S_r(X; d_0) = \{d \in \mathcal{F}_{d_0}(X) : \omega(d) = r\}.$$

The Sathe–Selberg and Erdős–Kac theorems place almost all the mass in the range $r = \log\log X + O((\log\log X)^{2/3})$ [30, 5]. For such an r, the box construction uses the following estimates on the sizes and the spacing of the prime factors $p_1 < \cdots < p_r$ of d.

Following Smith [31], d is called *comfortably spaced above* D_1 if

$$2D_1 < p_i < \frac{p_{i+1}}{2} \qquad \text{whenever } i < r \text{ and } p_i > D_1,$$

and *η-regular* if

$$|\log\log p_i - i| < \eta^{1/5} \max\{i, \eta\}^{4/5} \qquad (1 \leq i < r/3).$$

Comfortable spacing supplies disjoint intervals for the active coordinates; regularity supplies the size estimates required by the residue-symbol argument. The choices of D_1 and η used in the covering are fixed as functions of X in Definition 3.4.

Proposition 3.1. *Suppose that*

$$|r - \log\log X| < (\log\log X)^{2/3}.$$

The following estimates hold uniformly in this range.

(i)
$$\#S_r(X; d_0) \asymp_{d_0} \frac{X}{\log X} \frac{(\log\log X)^{r-1}}{(r-1)!},$$

and

$$\sum_{\substack{r \geq 0 \\ |r - \log\log X| \geq (\log\log X)^{2/3}}} \#S_r(X; d_0) \ll \exp\big(-c(\log\log X)^{1/3}\big) \#\mathcal{F}_{d_0}(X).$$

(ii) *For* $D_1 > 3$ *and* $\varepsilon > 0$*, the proportion of elements of* $S_r(X; d_0)$ *not comfortably spaced above* D_1 *is*

$$O_{d_0,\varepsilon}\big((\log D_1)^{-1} + (\log X)^{-1/2+\varepsilon}\big).$$

(iii) *For* $\eta > 1$*, the proportion of elements of* $S_r(X; d_0)$ *that are not* η*-regular is*

$$O_{d_0}\big(e^{-c\eta} + e^{-c(\log\log X)^{1/3}}\big).$$

Proof. Part (i) is the Sathe–Selberg estimate together with the usual Erdős–Kac tail. Parts (ii) and (iii) are [31, Theorem 5.4(1)–(2)], applied on each fixed-a slice described above. The bounded shifts of X and r do not change the displayed ranges or errors, and summing the finitely many slices preserves uniformity. The restriction $p \nmid 2d_0$ changes the relevant prime harmonic sum only by a bounded term:

$$\sum_{\substack{p \le z \\ p \in \mathcal{P}_{d_0}}} \frac{1}{p} = \log\log z + O_{d_0}(1).$$

□

Definition 3.2. Let $D_1 > 3$, $r \ge 1$ and $0 \le m_0 < r$. A *prime-coordinate box above* D_1 is

$$\mathbf{X} = X_1 \times \cdots \times X_r,$$

where $X_i = \{p_i\}$ with $p_i < D_1$ for $i \le m_0$, while for $i > m_0$ the set X_i consists of primes in $\mathcal{P}_{d_0}$ lying in an interval (t_i, t_i'), with

$$t_i' = \left(1 + \frac{1}{e^{\,i-m_0}\log D_1}\right) t_i.$$

For $i \le m_0$, choose auxiliary endpoints $t_i < p_i < t_i' < D_1$ so that p_i is the only prime in (t_i, t_i') and the full endpoint sequence remains strictly increasing. The active intervals lie above D_1, are pairwise disjoint and are arranged in increasing order. The first m_0 coordinates are *frozen* and the remaining coordinates are *active.* The integer image of $\mathbf{X}$ is

$$B_{\mathbf{X}} = \{x_1 \cdots x_r : x_i \in X_i,\ x_1 \cdots x_r \le X\} \subseteq S_r(X; d_0).$$

These are the coordinate widths and frozen coordinates in Smith's construction preceding [31, Proposition 6.9]. The coordinate primes here range over $\mathcal{P}_{d_0}$, and the multiplication map is injective. We restrict to the boxes in the positive covering for which every tuple has product at most X; for these boxes

$$\#B_{\mathbf{X}} = \prod_{i=1}^{r} \#X_i. \tag{39}$$

Thus $\mathbf{X}$ is a space of ordered prime coordinates, whereas $B_{\mathbf{X}}$ is a set of integers; equation (39) identifies their counting measures. Definition 3.2 is parametric in the cutoff D_1; the arithmetic covering uses the X-dependent choice in Definition 3.4.

For $d = p_1 \cdots p_r$, put

$$\omega_{D_1}(d) = \#\{p \mid d : p < D_1\}.$$

The construction freezes these $\omega_{D_1}(d)$ small-prime coordinates.

We must also exclude boxes containing exceptional real characters. Put

$$E = \mathbb{Q}\big(\zeta_8, \sqrt{q_1^*}, \ldots, \sqrt{q_s^*}\big), \qquad d_0 = 2^e q_1 \cdots q_s.$$

Let χ_n be the quadratic character associated with $\mathbb{Q}(\sqrt{n})$. Fix the absolute constant c_{zf}, and let $\mathcal{D}_{\mathrm{exc}}$ be the set of squarefree integers n for which $L(s, \chi_n)$ has a real zero in

$$1 - \frac{c_{\mathrm{zf}}}{\log(|n| + 4)} \le s \le 1.$$

The box $\mathbf{X}$ is called *E-Siegel-free above D_1* if

$$u\prod_{i\in I} x_i \notin \mathcal{D}_{\mathrm{exc}} \tag{40}$$

for every $(x_1,\dots,x_r)\in\mathbf{X}$, every $I\subseteq\{1,\dots,r\}$, every squarefree u whose squareclass is generated by $-1,2,q_1,\dots,q_s$, whenever the absolute value in (40) exceeds D_1. This is [31, Definition 6.2] for $P=\{-1,2,q_1,\dots,q_s\}$.

Proposition 3.3. *Let $D_1>3$. The following estimates hold, the second uniformly for r in the range of Proposition 3.1.*

(i)

$$\frac{\#\{d\in\mathcal{F}_{d_0}(X):\omega_{D_1}(d)>2\log\log D_1\}}{\#\mathcal{F}_{d_0}(X)} \ll_{d_0} (\log D_1)^{-(2\log 2-1)}.$$

(ii)

$$\#\{d\in S_r(X;d_0): d \textit{ occurs in a non-E-Siegel-free box}\} \ll_{d_0} \frac{\#S_r(X;d_0)}{\log D_1}.$$

Proof. For the first estimate, observe that

$$2^{\omega_{D_1}(d)} = \sum_{\substack{a\mid d\\ p\mid a\implies p<D_1}} \mu^2(a).$$

Since $\#\mathcal{F}_{d_0}(X)\asymp_{d_0} X$ and, uniformly for squarefree a supported on primes below D_1 with $(a,2d_0)=1$,

$$\#\{d\in\mathcal{F}_{d_0}(X): a\mid d\}\ll_{d_0}\frac{X}{a},$$

interchanging the two sums and applying Mertens' theorem gives

$$\frac{1}{\#\mathcal{F}_{d_0}(X)}\sum_{d\in\mathcal{F}_{d_0}(X)} 2^{\omega_{D_1}(d)} \ll_{d_0} \prod_{\substack{p<D_1\\ p\nmid 2d_0}}\left(1+\frac{1}{p}\right) \ll_{d_0} \log D_1.$$

Markov's inequality gives part (i). For the second, apply Landau's separation theorem and [31, Proposition 6.10] to the finitely many squareclasses generated by P. $\square$

Smith's residue-symbol theorem uses auxiliary constants $c_1,\dots,c_{12}$. Fix, once and for all,

$$c_{\mathrm{cut}}=\frac{1}{200},\quad c_1=1000,\quad c_2=\frac{1}{1000},\quad c_3=2,\quad c_4=9,\quad c_5=4,\quad c_6=\frac{1}{2000},$$
$$c_7=\frac{1}{100},\quad c_8=\frac{1}{1000},\quad c_9=c_{10}=\frac{1}{50},\quad c_{11}=\frac{1}{100},\quad c_{12}=\frac{1}{200}.$$

These constants satisfy the hypotheses of [31, Theorem 6.4] and the parameter inequalities used in [31, Corollary 6.11]. Choose $0<c_{\mathrm{reg}}<c_{\mathrm{cut}}$ sufficiently small for the regularity estimates in that corollary.

Definition 3.4. Fix the squarefree integer $d_0>1$ and the constants chosen above. For all sufficiently large X, put

$$D_1=D_1(X)=\exp\left(\left(\frac{1}{2}\log\log X\right)^{c_{\mathrm{cut}}}\right),\qquad \eta=\eta(X)=c_{\mathrm{reg}}\log\log\log X. \tag{41}$$

Let $\mathbf{X}=X_1\times\cdots\times X_r$ be a prime-coordinate box above $D_1(X)$, and write m_0 for its number of frozen coordinates. The box $\mathbf{X}$ is *admissible at height X* if the following conditions hold.

(i) It belongs to Smith's positive covering and contains a comfortably spaced, $\eta(X)$-regular integer, with
$$|r - \log\log X| < (\log\log X)^{2/3}.$$
(ii) Its number of frozen coordinates satisfies
$$m_0 \le 2\log\log D_1.$$
(iii) It is E-Siegel-free above D_1.

The scales D_1 and η depend on X, while r and m_0 depend on the box. In particular,
$$m_0 \ll \log\log\log X, \qquad r \asymp \log\log X.$$
Propositions 3.1 and 3.3 show that the losses from atypical prime-factor geometry, excessive freezing and exceptional real characters have total proportion
$$O_{d_0}\big((\log\log X)^{-c}\big)$$
for some $c > 0$.

Smith's positive-covering argument transfers a uniform estimate on admissible boxes to the original family. Smith performs the weighted integration in [31, Proposition 6.9]; the same calculation is written out in [1, Theorem 5.13].

Proposition 3.5. *Fix a squarefree integer $d_0 > 1$ and the constants $c_{\mathrm{cut}}, c_{\mathrm{reg}}$ chosen above. For each sufficiently large X, let $D_1(X)$ and $\eta(X)$ be the scales in (41). There is*
$$c_{\mathrm{box}} = c_{\mathrm{box}}(d_0, c_{\mathrm{cut}}, c_{\mathrm{reg}}) > 0$$
with the following property. Let $\mathcal{A} \subseteq \mathcal{F}_{d_0}(X)$ and $0 \le \theta \le 1$. If, for some $\varepsilon_{\mathrm{box}} \ge 0$,
$$\left| \frac{\#(\mathcal{A} \cap B_{\mathbf{X}})}{\#B_{\mathbf{X}}} - \theta \right| \le \varepsilon_{\mathrm{box}} \tag{42}$$
for every admissible box $\mathbf{X}$ at height X, then
$$\left| \frac{\#\mathcal{A}}{\#\mathcal{F}_{d_0}(X)} - \theta \right| \ll_{d_0, c_{\mathrm{cut}}, c_{\mathrm{reg}}} (\log\log X)^{-c_{\mathrm{box}}} + \varepsilon_{\mathrm{box}}. \tag{43}$$

Consequently, if $Y : \mathcal{F}_{d_0}(X) \to \mathcal{Z}$ is any statistic with countable state space and ρ is a probability measure on $\mathcal{Z}$, then
$$\begin{aligned} &\sum_{\omega\in\mathcal{Z}} \left| \frac{\#\{d \in \mathcal{F}_{d_0}(X) : Y(d) = \omega\}}{\#\mathcal{F}_{d_0}(X)} - \rho(\omega) \right| \\ &\qquad \ll_{d_0, c_{\mathrm{cut}}, c_{\mathrm{reg}}} (\log\log X)^{-c_{\mathrm{box}}} + \sup_{\mathbf{X}} \sum_{\omega\in\mathcal{Z}} \left| \frac{\#\{d \in B_{\mathbf{X}} : Y(d) = \omega\}}{\#B_{\mathbf{X}}} - \rho(\omega) \right|, \end{aligned} \tag{44}$$
where the supremum is over the admissible boxes.

Proof. Let $\mathcal{G}_X$ consist of the integers $d \in \mathcal{F}_{d_0}(X)$ for which $r = \omega(d)$ lies in the range of Proposition 3.1, d is comfortably spaced and $\eta(X)$-regular, $\omega_{D_1}(d) \le 2\log\log D_1$, and d occurs only in E-Siegel-free boxes. Proposition 3.1 and Proposition 3.3 give
$$\#\big(\mathcal{F}_{d_0}(X) \setminus \mathcal{G}_X\big) \ll_{d_0, c_{\mathrm{cut}}, c_{\mathrm{reg}}} (\log\log X)^{-c} \#\mathcal{F}_{d_0}(X)$$
for some $c > 0$. Every box in Smith's positive covering that meets $\mathcal{G}_X$ is admissible.

On each fixed-r, fixed-a, fixed-m_0 slice, restrict Smith's box-parameter measure to boxes meeting $\mathcal{G}_X$ and rescale it by the reciprocal membership factor used in the proof of [31, Proposition 6.9]. Sum these measures over the slices. For the resulting measure ν, put
$$w(d) := \int \mathbb{1}_{\{d \in B_{\mathbf{X}}\}}\, d\nu(\mathbf{X}).$$

The calculation in that proof gives $0 \le w(d) \le 1$, with $w(d) = 1$ for $d \in \mathcal{G}_X$ outside the terminal range
$$X(1 - C/\log D_1) \le d \le X$$
for an absolute $C > 0$. The Sathe–Selberg estimate used there gives
$$\#\{d \in \mathcal{F}_{d_0}(X) : X(1 - C/\log D_1) \le d \le X\} \ll_{d_0} \frac{\#\mathcal{F}_{d_0}(X)}{\log D_1}.$$
Consequently,
$$\sum_{d \in \mathcal{F}_{d_0}(X)} |1 - w(d)| \ll_{d_0, c_{\mathrm{cut}}, c_{\mathrm{reg}}} (\log\log X)^{-c} \#\mathcal{F}_{d_0}(X).$$
Integrating (42) over these admissible boxes gives
$$\left| \sum_{d \in \mathcal{A}} w(d) - \theta \sum_{d \in \mathcal{F}_{d_0}(X)} w(d) \right| \le \varepsilon_{\mathrm{box}} \#\mathcal{F}_{d_0}(X).$$
Replacing the weights by 1 proves (43).

For (44), let the supremum on its right-hand side be δ. For every $\mathcal{E} \subseteq \mathcal{Z}$, the local error for $\mathcal{A} = Y^{-1}(\mathcal{E})$ and $\theta = \rho(\mathcal{E})$ is at most $\delta/2$. Apply the first assertion and then use the identity between total variation and the supremum over events. □

By Proposition 3.5, it remains to prove a uniform boxwise estimate for the residue assignment.

3.2. **Equidistribution of residue assignments.** Fix an admissible box $\mathbf{X} = X_1 \times \cdots \times X_r$, and use the injective coordinate parametrisation in (39) to write $d = p_1 \cdots p_r$ with $p_i \in X_i$. Let
$$\mu_{\mathbf{X}} := \frac{1}{\#B_{\mathbf{X}}} \sum_{d \in B_{\mathbf{X}}} \delta_{\boldsymbol{\chi}(d)}$$
be the empirical distribution on Res_r of the residue assignment (17), and let $\mu_{\mathbf{X}}^{\mathrm{sym}}$ be its permutation average, defined as in Section 2.1.3. Thus $\mu_{\mathbf{X}}$ is the actual symbol distribution on the box, while Smith's theorem controls its symmetrisation.

Proposition 3.6. *Fix $d_0 > 1$. There is $c_{\mathrm{sym}} = c_{\mathrm{sym}}(d_0) > 0$ such that, for every admissible box $\mathbf{X}$ at height X and all sufficiently large X,*
$$\left\| \mu_{\mathbf{X}}^{\mathrm{sym}} - \mathrm{Unif}(\mathsf{Res}_r) \right\|_1 \ll_{d_0} (\log\log X)^{-c_{\mathrm{sym}}}. \tag{45}$$

Proof. Smith's assignment theorem records the symbols between pairs of box coordinates and against a fixed finite set P. In [31, Definition 6.1], take
$$P_0 = \{2, q_1, \ldots, q_s\}, \qquad P = \{-1\} \cup P_0,$$
$$\mathsf{M} = \Big\{ \{i, j\} : 1 \le i < j \le r \Big\}, \qquad \mathsf{M}_P = [r] \times P.$$
For these maximal choices, the assignment space has cardinality $\#\mathsf{Res}_r$, so the uniform main term in Smith's theorem is $1/\#\mathsf{Res}_r$.

Smith records $[\frac{p_i}{p_j}]$ and $[\frac{q_k}{p_i}]$, while (17) uses prime discriminants. The two assignments are exchanged by the involution $\Theta : \mathsf{Res}_r \to \mathsf{Res}_r$ given by
$$\Theta(\boldsymbol{\chi})_{ij} = \chi_{ij} + u_{-1,i} u_{-1,j} \quad (i < j), \qquad \Theta(\mathbf{u}_{\Delta_k})_i = u_{\Delta_k, i} + [\tfrac{-1}{q_k}] u_{-1,i} \quad (1 \le k \le s),$$
with $\mathbf{u}_{-1}$ and $\mathbf{u}_2$ fixed. Quadratic reciprocity shows that Θ converts the raw symbols to those in (17). It is action-equivariant and preserves $\mathrm{Unif}(\mathsf{Res}_r)$, so it suffices to prove (45) for Smith's raw assignment.

The field-theoretic independence hypothesis also holds. Every prime ramified in E divides $2d_0$, whereas every nonempty product of the p_i^* is ramified at some $p_i \nmid 2d_0$. Thus the varying squareclasses are independent over E, and (40) is precisely Smith's Siegel-less condition for the set P above.

We apply [31, Theorem 6.4] with the constants fixed above and the parameter choice of [31, Corollary 6.11]. Namely, take

$$k_0 = m_0, \qquad k_2 = r, \qquad t = D_1, \qquad t' = t'_{m_0+1},$$

so Smith's permutation group is $\mathfrak{S}_r$ and his normalization is $r!$. Let k_1 be the least index at least k_0 for which

$$t'_{k_1+1} > \exp(D_1^{c_6}).$$

The required scale comparison is obtained by putting $L = \log\log X$ and $R = \log L$. Admissibility gives

$$\log\log D_1 = c_{\mathrm{cut}} R + O(1), \qquad r = L + O(L^{2/3}), \qquad m_0 \le 2c_{\mathrm{cut}} R + O(1).$$

Thus c_{cut} plays the role of the cutoff constant in the proof of that corollary. The numerical inequalities among the constants fixed above, together with a sufficiently small choice of c_{reg}, give conditions (1)–(6) of the theorem by the verification in that corollary. The fixed set P is absorbed in the constant terms, [31, Proposition 6.7] handles the frozen singleton coordinates, and the bound $m_0 \le 2\log\log D_1$ is absorbed by the chosen margins. Moreover, regularity at the first coordinate and the choice of c_{reg} give $m_0 \ge 1$ for all sufficiently large X; hence the endpoint convention $t'_{m_0} < D_1 < t'_{m_0+1}$ covers the boundary case $k_1 = k_0$. The remaining Siegel-less hypothesis is condition (iii) of Definition 3.4.

With $k_2 = r$, the conclusion of the theorem reads

$$\sum_{\boldsymbol{\chi}_0 \in \mathsf{Res}_r} \left| \frac{1}{r!} \sum_{\pi \in \mathfrak{S}_r} \frac{\#\{d \in B_{\mathbf{X}} : \pi \cdot \boldsymbol{\chi}(d) = \boldsymbol{\chi}_0\}}{\# B_{\mathbf{X}}} - \frac{1}{\#\mathsf{Res}_r} \right| \le r^{-c_{12}} + (t')^{-c_8},$$

where we used (39) to identify $\#B_{\mathbf{X}}$ with the number of coordinate tuples. Finally $t' = t'_{m_0+1} > D_1(X)$ and $r \asymp \log\log X$, so the right-hand side is $O_{d_0}\big((\log\log X)^{-c_{\mathrm{sym}}}\big)$ for any $c_{\mathrm{sym}} < c_{12}$. □

Corollary 3.7. *For all sufficiently large X, let $\mathbf{X}$ be an admissible box at height X. Then*

$$\sum_{\mathbf{j} \in \mathbb{Z}^2_{\ge 0}} \left| \frac{\#\{d \in B_{\mathbf{X}} : (r_4(-d), r_4(-d_0 d)) = \mathbf{j}\}}{\# B_{\mathbf{X}}} - \big((\mathrm{Joint4rk}_r)_* \operatorname{Unif}(\mathsf{Res}_r)\big)(\mathbf{j}) \right| \ll_{d_0} (\log\log X)^{-c_{\mathrm{sym}}}.$$

Proof. By (30), the empirical distribution in the statement is $(\mathrm{Joint4rk}_r)_*\mu_{\mathbf{X}}$. Therefore

$$\begin{aligned} &\big\|(\mathrm{Joint4rk}_r)_*\mu_{\mathbf{X}} - (\mathrm{Joint4rk}_r)_* \operatorname{Unif}(\mathsf{Res}_r)\big\|_1 \\ &\quad = \big\|(\mathrm{Joint4rk}_r)_*\mu^{\mathrm{sym}}_{\mathbf{X}} - (\mathrm{Joint4rk}_r)_* \operatorname{Unif}(\mathsf{Res}_r)\big\|_1 \\ &\quad \le \big\|\mu^{\mathrm{sym}}_{\mathbf{X}} - \operatorname{Unif}(\mathsf{Res}_r)\big\|_1. \end{aligned}$$

The equality is (34); the inequality is contraction under pushforward. Proposition 3.6 proves the result. □

3.3. **Family transfer.** Combining the boxwise estimate with the positive covering gives the required transfer from the arithmetic family to the matrix model.

Theorem 3.8. *Fix $d_0 > 1$ and $0 < \alpha < 1/2$. There exists $c_{\mathrm{tr}} = c_{\mathrm{tr}}(d_0, \alpha) > 0$ such that, with $\mu_X^{(d_0)}$ as in (2), for every probability measure ρ on $\mathbb{Z}^2_{\ge 0}$ and all sufficiently large X,*

$$\big\|\mu_X^{(d_0)} - \rho\big\|_1 \ll_{d_0,\alpha} (\log\log X)^{-c_{\mathrm{tr}}} + \sup_{\substack{|r - \log\log X| < (\log\log X)^{2/3} \\ |\mathbf{u}_{-1}| \ge \alpha r}} \big\|\mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}) - \rho\big\|_1, \tag{46}$$

where the supremum is over r in the displayed range and $\mathbf{u}_{-1} \in \mathbb{F}_2^r$ of weight at least αr.

Proof. Put $Y(d) = \big(r_4(-d), r_4(-d_0 d)\big)$, and let $\nu_{\mathbf{X}}$ be the distribution of $Y(d)$ for uniform $d \in B_{\mathbf{X}}$. By Corollary 3.7,

$$\big\|\nu_{\mathbf{X}} - (\mathrm{Joint4rk}_r)_* \,\mathrm{Unif}(\mathsf{Res}_r)\big\|_1 \ll_{d_0} (\log\log X)^{-c_{\mathrm{sym}}}$$

uniformly over the admissible boxes at height X; here r is the number of coordinates of $\mathbf{X}$. Apply (44) to Y and to the given ρ, and bound each box term by the triangle inequality through $(\mathrm{Joint4rk}_r)_* \,\mathrm{Unif}(\mathsf{Res}_r)$. Since the admissible boxes at height X have $|r - \log\log X| < (\log\log X)^{2/3}$, this yields the intermediate estimate

$$\big\|\mu_X^{(d_0)} - \rho\big\|_1 \ll_{d_0} (\log\log X)^{-c_{\mathrm{tr}}} + \sup_{|r-\log\log X|<(\log\log X)^{2/3}} \big\|(\mathrm{Joint4rk}_r)_* \,\mathrm{Unif}(\mathsf{Res}_r) - \rho\big\|_1.$$

Finally, (33) writes each finite pushforward as the uniform mixture, over $\mathbf{u}_{-1} \in \mathbb{F}_2^r$, of

$$\mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}).$$

Convexity of the ℓ^1 norm bounds the last norm by

$$\max_{|\mathbf{u}_{-1}|\geq \alpha r} \big\|\mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}) - \rho\big\|_1 + 2 \cdot 2^{-r} \#\{\mathbf{u}_{-1} : |\mathbf{u}_{-1}| < \alpha r\},$$

since the ℓ^1 distance between two probability measures is at most 2. The second term is at most $2e^{-2(1/2-\alpha)^2 r}$ by (35), hence $O_\alpha((\log X)^{-c})$ for some $c = c(\alpha) > 0$ in the displayed range of r. Shrinking c_{tr} proves (46). □

Remark 3.9. The same box construction can be adapted when every prime factor is required to lie in a fixed Frobenius class of a multiquadratic field F. The typical value of r is then $[F : \mathbb{Q}]^{-1} \log\log X$, and $\mathrm{Unif}(\mathsf{Res}_r)$ is replaced by the conditional uniform measure on the affine coset determined by the characters common to E and F. This conditioning can change the induced distribution of the matrix pair: for $F = \mathbb{Q}(i)$ it prescribes the sign vector $\mathbf{u}_{-1}$, and for $F = \mathbb{Q}(\zeta_8)$ it also fixes the symbols at 2. In particular, Theorem 3.8 concerns the full space Res_r. Under a fixed-sign condition its target must be replaced by the corresponding conditional distribution of the residue assignments; its matrix pushforward requires a separate analysis.

4. Joint coranks of bordered matrices

The family transfer (46) reduces the problem to a uniform estimate for the corank pair on the uniform probability space $\mathsf{MatData}_r(\mathbf{u}_{-1})$ introduced in Section 2.1.3. Fix d_0, r and $\mathbf{u}_{-1}$; the 2-adic stratum is determined by d_0 and $|\mathbf{u}_{-1}|$. We prove an ℓ^1 error bound $C\eta^r$, with $\eta \in (0, 1)$, uniformly for $|\mathbf{u}_{-1}| \geq \alpha r$. The complementary sign vectors have exponentially small mass by (35).

This estimate is assembled in three steps. We first prove a quantitative truncated inversion of the one-dimensional Gaussian-binomial moment transform and apply it to a single core block. We then pass to the two-dimensional transform, whose inverse is the tensor power of the one-dimensional inverse but whose truncation residual needs its own bivariate estimate, and extract from it a criterion that converts conditional mixed moments on a good event into a total-variation bound. We finally verify that criterion for the bordered pair by estimating its conditional mixed moments. The last step does not follow from the first: the two fixed border widths need not agree, and the diagonal perturbation $\mathbf{u}_{d_0}$ and the upper border U are determined by the lower-border data and the sign vector rather than sampled independently.

4.1. **One-dimensional truncated inversion for corank distributions.** The normalization here differs from the ordinary size moments occurring, for example, in the work of Fouvry–Klüners and Smith [7, 33]. As recalled in Section 2.2.2, the Gaussian-binomial moment of a corank variable X is a normalized surjection moment:

$$\mathbb{E}\#\operatorname{Sur}(\mathbb{F}_2^X, \mathbb{F}_2^u) = |\operatorname{GL}_u|\,\mathbb{E}\begin{bmatrix} X \\ u \end{bmatrix}_2.$$

Thus it is the surjection-moment normalization used in random finite-group and cokernel problems; see [35, 34]. Equivalently, $\left[\begin{smallmatrix} X \\ u \end{smallmatrix}\right]_2$ is a polynomial of degree u in 2^X, so these moments form a triangular change of basis from ordinary moments of 2^X. This basis counts u-dimensional kernel subspaces, and Lemma 2.9 makes it accessible one order at a time. Fixed-order convergence of these moments, together with a uniqueness argument, is sufficient to identify a limiting distribution.

Fixed-order convergence and uniqueness do not by themselves retain a rate. To obtain a total-variation bound $C\eta^n$ with $\eta \in (0,1)$, we retain moments only through an order D that grows with the matrix size, apply the bounded inverse to this truncated moment sequence, and estimate the residual explicitly. This is the quantitative truncated inversion used below.

We first develop it in one variable. The tensor-product inverse and the multivariable residual are treated in Section 4.2.

4.1.1. *Gaussian-binomial inversion and truncation.* At $q = 2$, set

$$\gamma_{u,a} := (G_2^{-1})_{u,a} = (-1)^{u-a} 2^{\binom{u-a}{2}} \begin{bmatrix} u \\ a \end{bmatrix}_2 \qquad (0 \le a \le u).$$

For a finite signed measure σ on $\mathbb{Z}_{\ge 0}$ and every u for which $m_u(|\sigma|) < \infty$, set

$$m_u(\sigma) := \sum_{x \ge u} \begin{bmatrix} x \\ u \end{bmatrix}_2 \sigma(x). \tag{47}$$

When these absolute moments are finite for every u, write $(\mathcal{G}\sigma)_u = m_u(\sigma)$ and $\mathbf{m}(\sigma) = \mathcal{G}\sigma$ for the full Gaussian-binomial moment transform. For a finitely supported measure the transform is triangular, and finite q-binomial inversion reads

$$\sigma(a) = \sum_{u \ge a} \gamma_{u,a} m_u(\sigma).$$

Set

$$\theta_u := 2^{-u(u+1)/2}, \tag{48}$$

and define

$$\|z\|_{\infty,\theta} := \sup_{u \ge 0} \frac{|z_u|}{\theta_u}, \qquad \ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1) := \{z : \|z\|_{\infty,\theta} < \infty\}.$$

For $z \in \ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1)$, define

$$\big(\mathcal{G}^{-1} z\big)(a) := \sum_{u \ge a} \gamma_{u,a} z_u.$$

Theorem 4.1 shows that this series defines an ℓ^1-sequence. If $\mathbf{m}(|\sigma|) \in \ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1)$, the residual estimate below shows that $\mathcal{G}^{-1}\mathcal{G}\sigma = \sigma$; no inversion is asserted here under the weaker assumption that all moments are merely finite.

Thus $\mathcal{G}$ maps

$$\{\sigma \in \ell^1(\mathbb{Z}_{\ge 0}) : \mathbf{m}(|\sigma|) \in \ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1)\}$$

to $\ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1)$, with formal inverse bounded from $\ell_\theta^\infty(\mathbb{Z}_{\ge 0}^1)$ to $\ell^1(\mathbb{Z}_{\ge 0})$.

Theorem 4.1. *The operator $\mathcal{G}^{-1}$ maps $\ell^\infty_\theta(\mathbb{Z}^1_{\geq 0})$ continuously into $\ell^1(\mathbb{Z}_{\geq 0})$, and*

$$\left\|\mathcal{G}^{-1}z\right\|_1 \leq C_{\mathrm{inv}}\|z\|_{\infty,\theta}, \qquad C_{\mathrm{inv}} := \sum_{u\geq 0}\theta_u\sum_{a=0}^{u}|\gamma_{u,a}| < \infty.$$

Proof. Put $t = u - a$. Proposition 2.10(i) gives

$$\theta_u|\gamma_{u,a}| \leq \eta_\infty(2)^{-1}2^{-u(u+1)/2+\binom{t}{2}+at} = \eta_\infty(2)^{-1}2^{-a(a+1)/2-t} \leq \eta_\infty(2)^{-1}2^{-u}.$$

Thus $C_{\mathrm{inv}} = \sum_u \theta_u \sum_{a\leq u}|\gamma_{u,a}| < \infty$, and

$$\left\|\mathcal{G}^{-1}z\right\|_1 \leq \sum_{a\geq 0}\sum_{u\geq a}|\gamma_{u,a}|\,|z_u| \leq \|z\|_{\infty,\theta}\sum_{u\geq 0}\theta_u\sum_{a\leq u}|\gamma_{u,a}|$$

by Tonelli's theorem. □

Let P_D denote truncation of a sequence: $(P_Dz)_u = z_u$ when $u \leq D$, and zero otherwise. Suppose first that $m_u(|\sigma|) < \infty$ for every u, so that the whole sequence $\mathbf{m}(\sigma)$ is defined, and set

$$Q_D := \mathcal{G}^{-1}P_D\mathcal{G}, \tag{49}$$

and define the residual

$$r_D(\sigma) := (\mathrm{id} - Q_D)\sigma = \sigma - \mathcal{G}^{-1}P_D\mathbf{m}(\sigma). \tag{50}$$

Q_D takes a finite signed measure to its moment sequence, retains the orders $0, \dots, D$, and applies the bounded inverse. Since the composite $\mathcal{G}^{-1}P_D\mathcal{G}$ reads no higher moment, the same formulas define $Q_D\sigma$ and $r_D(\sigma)$ whenever $m_u(|\sigma|) < \infty$ only for $0 \leq u \leq D$, even if the upper-right corner of the full transform is unavailable.

Truncation need not return σ: the decomposition

$$\sigma = Q_D\sigma + r_D(\sigma)$$

isolates in $r_D(\sigma)$ exactly the part of σ that the retained moments do not control. The two terms are estimated separately: Theorem 4.1 bounds $\|Q_D\sigma\|_1$ in terms of the retained moments, and Lemma 4.2 bounds $\|r_D(\sigma)\|_1$ by a single higher moment. These are the two terms of the truncated inversion inequality below.

Lemma 4.2. *There is an absolute constant C such that, for every finite signed measure σ with $m_{D+1}(|\sigma|) < \infty$,*

$$\|r_D(\sigma)\|_1 \leq C(D+1)2^{D(D+1)/2}m_{D+1}(|\sigma|).$$

Proof. Apply the truncated inverse first to a point mass at x. Finite q-binomial inversion gives

$$\mathbb{1}_{\{a=x\}} = \sum_{u=a}^{x}\gamma_{u,a}\begin{bmatrix}x\\u\end{bmatrix}_2.$$

Consequently the residual has the pointwise kernel representation

$$r_D(\sigma)(a) = \sum_x \sigma(x)K_D(x,a), \qquad K_D(x,a) := \sum_{u=\max(a,D+1)}^{x}\gamma_{u,a}\begin{bmatrix}x\\u\end{bmatrix}_2.$$

Every inner sum is finite, so this identity uses only the assumed absolute $(D+1)$-st moment. For $x > D$ and $a \leq D$, the q-chain identity and (38) give

$$|K_D(x,a)| = \begin{bmatrix}x\\a\end{bmatrix}_2 2^{\binom{D-a+1}{2}}\begin{bmatrix}x-a-1\\D-a\end{bmatrix}_2.$$

Summing over $a \le D$ and using Proposition 2.10(i),(iii) gives

$$\sum_{a \le D} |K_D(x,a)| \le C(D+1)2^{D(D+1)/2} \begin{bmatrix} x \\ D+1 \end{bmatrix}_2.$$

For $a > D$, finite inversion gives $K_D(x,a) = \mathbb{1}_{\{a=x\}}$. Consequently

$$\sum_{a \ge 0} |K_D(x,a)| \le 1 + C(D+1)2^{D(D+1)/2} \begin{bmatrix} x \\ D+1 \end{bmatrix}_2,$$

and the first term is absorbed by the second when $x > D$. The result follows after summing over x; for $x \le D$ the residual is identically zero. $\square$

If $\mathbf{m}(|\sigma|) \in \ell^\infty_\theta(\mathbb{Z}^1_{\ge 0})$, then

$$\|r_D(\sigma)\|_1 \ll (D+1)2^{D(D+1)/2}\theta_{D+1} = (D+1)2^{-(D+1)} \longrightarrow 0.$$

At the same time, the tail estimate in the proof of Theorem 4.1 shows that $\mathcal{G}^{-1}P_D\mathcal{G}\sigma$ converges in ℓ^1 to $\mathcal{G}^{-1}\mathcal{G}\sigma$. Hence $\mathcal{G}^{-1}\mathcal{G}\sigma = \sigma$ on the domain displayed above.

Theorem 4.3. *Let σ be a finite signed measure on $\mathbb{Z}_{\ge 0}$ with $m_{D+1}(|\sigma|) < \infty$. Then*

$$\|\sigma\|_1 \le C_{\mathrm{inv}}\|P_D\mathbf{m}(\sigma)\|_{\infty,\theta} + C(D+1)2^{D(D+1)/2}m_{D+1}(|\sigma|), \tag{51}$$

where C is the constant of Lemma 4.2.

Proof. By definition, $\sigma = \mathcal{G}^{-1}P_D\mathbf{m}(\sigma) + r_D(\sigma)$. Theorem 4.1 bounds the first term, and Lemma 4.2 bounds the second. $\square$

The first term on the right of (51) is $C_{\mathrm{inv}}\max_{0\le u\le D}|m_u(\sigma)|/\theta_u$ and uses only the moments through order D; the second controls everything that was discarded. The two terms move in opposite directions as D grows. Balancing them gives the following criterion.

Corollary 4.4. *Let σ_n be finite signed measures on $\mathbb{Z}_{\ge 0}$. Suppose that there are constants $C < \infty$, $C_1 > 1$, $q_{\mathrm{err}} \in (0,1)$ and $\beta_0 > 0$ such that*

(i) *for every n and every $u \le \beta_0 n$,*

$$|m_u(\sigma_n)| \le CC_1^u\theta_u q_{\mathrm{err}}^n;$$

(ii) *the absolute moment sequences are uniformly bounded at the reference scale:*

$$K := \sup_n \|\mathbf{m}(|\sigma_n|)\|_{\infty,\theta} < \infty.$$

Then there are $C' < \infty$ and $\eta \in (0,1)$ such that $\|\sigma_n\|_1 \le C'\eta^n$. More precisely, for every $D \le \beta_0 n$,

$$\|\sigma_n\|_1 \le CC_{\mathrm{inv}}C_1^D q_{\mathrm{err}}^n + C''(D+1)2^{-(D+1)}K.$$

Choosing $D = \lfloor \beta n \rfloor$ with $0 < \beta < \beta_0$ sufficiently small gives $\eta = \max\{2^{-\beta/2}, C_1^\beta q_{\mathrm{err}}\} < 1$. At $u = 0$, part (i) *also controls the difference of the total masses, so no separate normalization is required.*

Proof. Apply Theorem 4.3. Part (i) bounds the truncated term by $CC_{\mathrm{inv}}C_1^D q_{\mathrm{err}}^n$. By part (ii), $m_{D+1}(|\sigma_n|) \le K\theta_{D+1}$, and $2^{D(D+1)/2}\theta_{D+1} = 2^{-(D+1)}$, so the residual is at most $C(D+1)2^{-(D+1)}K$. The stated choice of D balances the two bounds; its polynomial factor is absorbed by decreasing β. $\square$

4.1.2. *Application to a single core block.* For one core block, the zero- and high-rank-defect limits are π_{sym} and π_{CL}, respectively.

Lemma 4.5. *Let $A \sim \mathrm{Unif}(\mathrm{Sym}_n(\Omega))$, put $X := \operatorname{corank} A$ and $\rho_n := \mathrm{Law}(X)$. Then*

$$m_u(\rho_n) = \mathbb{E}\begin{bmatrix} X \\ u \end{bmatrix}_2 \le \eta_\infty(2)^{-1}\theta_u \qquad (n, u \ge 0). \tag{52}$$

Define the target moment

$$m_u(\Omega) := \begin{cases} \dfrac{\theta_u}{\eta_u(2)}, & \Omega = 0, \\ \dfrac{2^{-u^2}}{\eta_u(2)} = \dfrac{1}{|\,\mathrm{GL}_u\,|}, & \Omega \ne 0. \end{cases}$$

For the zero defect this is the limit as $n \to \infty$; for a sequence of nonzero defects it is the limit whenever $\operatorname{rank}\Omega \to \infty$. *More precisely,*

$$\frac{m_u(\rho_n)}{m_u(\Omega)} = \begin{cases} \displaystyle\prod_{i=0}^{u-1}(1 - 2^{i-n}) = 1 + O(2^{u-n}), & \Omega = 0, \\ 1 + O(2^{2u - \operatorname{rank}\Omega}) + O\big(2^{\binom{u+1}{2} - \operatorname{rank}\Omega}\big), & \Omega \ne 0. \end{cases}$$

Thus $m_u(\Omega) \asymp \theta_u$ for $\Omega = 0$, whereas $m_u(\Omega) \asymp 2^{-u^2}$ for $\Omega \ne 0$.

Proof. The first-moment identity is

$$m_u(\rho_n) = \sum_{\dim W = u} \mathbb{P}(W \subseteq \ker A).$$

For a basis $x_1, \dots, x_u$ of W, Lemma 2.9 gives $\mathbb{P}(W \subseteq \ker A) = 2^{-(nu - \binom{u}{2})}$ if W is totally isotropic for the alternating form $B_\Omega(x, y) = x^\top \Omega y$, and zero otherwise. Consequently,

$$m_u(\rho_n) = N_{\mathrm{iso}}(u) 2^{-(nu - \binom{u}{2})}. \tag{53}$$

If $\Omega = 0$, every subspace is isotropic, and

$$m_u(\rho_n) = \begin{bmatrix} n \\ u \end{bmatrix}_2 2^{-(nu - \binom{u}{2})} = \frac{\theta_u}{\eta_u(2)} \frac{\eta_n(2)}{\eta_{n-u}(2)}. \tag{54}$$

For $\Omega \ne 0$, Lemma 2.13 gives

$$m_u(\rho_n) = m_u(\Omega)\left(1 + O(2^{2u - \operatorname{rank}\Omega}) + O\big(2^{\binom{u+1}{2} - \operatorname{rank}\Omega}\big)\right), \qquad m_u(\Omega) = \frac{1}{|\,\mathrm{GL}_u\,|}. \tag{55}$$

The uniform bound follows from $N_{\mathrm{iso}}(u) \le \left[\begin{smallmatrix} n \\ u \end{smallmatrix}\right]_2$ and Proposition 2.10(i). $\square$

The formula (4) is normalized so that its Gaussian-binomial moments are exactly the zero-defect target moments. Indeed, writing $x = u + j$ and using the product formula for $\left[\begin{smallmatrix} x \\ u \end{smallmatrix}\right]_2$ gives

$$\sum_{x \ge u} \begin{bmatrix} x \\ u \end{bmatrix}_2 \pi_{\mathrm{sym}}(x) = \frac{\eta_\infty(2)}{\eta_\infty(4)} \frac{\theta_u}{\eta_u(2)} \sum_{j \ge 0} \frac{\theta_j}{\eta_j(2)} = \frac{\theta_u}{\eta_u(2)}.$$

Here the last equality is Euler's identity

$$\sum_{j \ge 0} \frac{2^{-j(j+1)/2}}{\eta_j(2)} = \frac{\eta_\infty(4)}{\eta_\infty(2)}.$$

Thus bounded Gaussian-binomial inversion identifies the inverse of $(m_u(0))_{u \ge 0}$ with π_{sym}.

Theorem 4.6. *For each n, let $\Omega_n \in \mathrm{Alt}_n$, take $A_n \sim \mathrm{Unif}(\mathrm{Sym}_n(\Omega_n))$, and put $X_n := \operatorname{corank} A_n$.*

(1) *If* $\Omega_n = 0$ *for every* n, *then*

$$\|\operatorname{Law}(X_n) - \pi_{\mathrm{sym}}\|_1 \le C\eta^n$$

for some $\eta \in (0,1)$.

(2) *If* $\Omega_n \neq 0$, *then, for every* $\epsilon \in (0,1)$,

$$\|\operatorname{Law}(X_n) - \pi_{\mathrm{CL}}\|_1 \le C_\epsilon 2^{-(1-\epsilon)\operatorname{rank}\Omega_n}.$$

In particular $\operatorname{Law}(X_n) \to \pi_{\mathrm{CL}}$ *whenever* $\operatorname{rank}\Omega_n \to \infty$; *if* $\operatorname{rank}\Omega_n \ge \alpha n$, *the displayed error is at most* $C_\epsilon(2^{-(1-\epsilon)\alpha})^n$.

Proof. Write $\rho_n := \operatorname{Law}(X_n)$ and set

$$\mu_{\Omega_n} := \mathcal{G}^{-1}(m_u(\Omega_n))_{u\ge 0}, \qquad \sigma_n := \rho_n - \mu_{\Omega_n}.$$

Equation (52) and the corresponding target-moment bound show that all required absolute moments are finite. Converting the ratio estimates of Lemma 4.5 to the θ-weighted norm gives

$$\frac{|m_u(\sigma_n)|}{\theta_u} \ll \begin{cases} 2^u 2^{-n}, & \Omega_n = 0, \\ 2^u 2^{-\operatorname{rank}\Omega_n}, & \Omega_n \neq 0. \end{cases}$$

In the second line the factor $m_u(\Omega_n)/\theta_u = 2^{-\binom{u}{2}}/\eta_u(2)$ cancels the quadratic growth in the relative isotropic-subspace error.

If $\Omega_n = 0$, then $m_u(\Omega_n) = \theta_u/\eta_u(2)$ and $\mu_{\Omega_n} = \pi_{\mathrm{sym}}$. Corollary 4.4, with $D = \lfloor n/2 \rfloor$, gives $\|\sigma_n\|_1 \ll n2^{-n/2}$.

Suppose now that $s_n := \operatorname{rank}\Omega_n > 0$. Since $m_u(\Omega_n) = 1/|\operatorname{GL}_u|$, Gaussian-binomial inversion and Euler's product identity give

$$\begin{aligned}\mu_{\Omega_n}(a) &= \sum_{u\ge a} \gamma_{u,a} m_u(\Omega_n) \\ &= \frac{1}{|\operatorname{GL}_a|}\sum_{j\ge 0} \frac{(1/2)^{\binom{j}{2}}}{\eta_j(2)}\left(-2^{-(a+1)}\right)^j = \frac{2^{-a^2}\eta_\infty(2)}{\eta_a(2)^2}.\end{aligned}$$

Thus $\mu_{\Omega_n} = \pi_{\mathrm{CL}}$. Suppose that

$$\binom{D+2}{2} \le s_n.$$

Then Lemma 4.5 gives

$$m_{D+1}(|\sigma_n|) \ll 2^{-(D+1)^2}.$$

Theorem 4.3 therefore bounds the truncated term by $O(2^D 2^{-s_n})$ and the residual by $O((D+1)2^{-(D+1)(D+2)/2})$. Taking

$$D = \left\lfloor \sqrt{2(1-\epsilon)s_n} \right\rfloor$$

and absorbing the finitely many small values of s_n into the constant gives $\|\sigma_n\|_1 \ll_\epsilon 2^{-(1-\epsilon)s_n}$. $\square$

Remark 4.7 (Comparison with the one-matrix theory). For a nonzero defect, the loss of ϵ in Theorem 4.6 comes from the error term

$$O\left(2^{\binom{u+1}{2} - \operatorname{rank}\Omega}\right)$$

in Lemma 2.13. This term makes the weighted moment error grow with u, so the bounded inverse must be combined with truncation. The spectral method of [36] gives the sharper exponent $2^{-\operatorname{rank}\Omega}$ for the one-matrix ensembles considered there; the estimate above is sufficient for the joint theorem.

If $\operatorname{rank}\Omega$ remains bounded, the second assertion is vacuous and the limiting distribution need not be π_{CL}. For the arithmetic defect $\Omega = \mathbf{u}_{-1}\mathbf{u}_{-1}^\top + \operatorname{diag}(\mathbf{u}_{-1})$, $\operatorname{rank}\Omega = |\mathbf{u}_{-1}| + O(1)$. Thus only the near-symmetric range $|\mathbf{u}_{-1}| = o(r)$ lies outside the range in which the theorem gives uniform convergence to π_{CL}.

4.2. **Multivariable truncated inversion and joint convergence.** The full multivariable moment transform and its inverse are tensor powers of their one-dimensional counterparts. Truncated inversion additionally requires a residual estimate obtained by expanding over nonempty sets of coordinates. The mixed-moment estimates are proved separately for the coupled matrices.

4.2.1. *Tensor-product inversion.* Fix $m \geq 1$. For a finite signed measure σ on $\mathbb{Z}_{\geq 0}^m$ and a multi-index $\mathbf{k} = (k_1, \dots, k_m)$ for which $m_{\mathbf{k}}(|\sigma|) < \infty$, set

$$m_{\mathbf{k}}(\sigma) := \sum_{\mathbf{x} \geq \mathbf{k}} \prod_{i=1}^{m} \begin{bmatrix} x_i \\ k_i \end{bmatrix}_2 \sigma(\mathbf{x}). \tag{56}$$

When these absolute moments are finite for every $\mathbf{k}$, write $(\mathcal{G}\sigma)_{\mathbf{k}} = m_{\mathbf{k}}(\sigma)$ and $\mathbf{m}(\sigma) = \mathcal{G}\sigma$. This is the m-fold tensor power of the one-dimensional transform (47); its inverse is accordingly the tensor power $(\mathcal{G}^{-1})^{\otimes m}$ of the one-dimensional inverse,

$$\big((\mathcal{G}^{-1})^{\otimes m} z\big)(\mathbf{a}) := \sum_{\mathbf{k} \geq \mathbf{a}} \prod_{i=1}^{m} \gamma_{k_i, a_i}\, z_{\mathbf{k}}.$$

Extend the weight (48) multiplicatively,

$$\theta_{\mathbf{k}} := \prod_{i=1}^{m} \theta_{k_i},$$

and define

$$\|z\|_{\infty,\theta} := \sup_{\mathbf{k}} \frac{|z_{\mathbf{k}}|}{\theta_{\mathbf{k}}}, \qquad \ell_\theta^\infty(\mathbb{Z}_{\geq 0}^m) := \{z : \|z\|_{\infty,\theta} < \infty\}.$$

For measures we continue to use the ordinary total-variation norm $\|\sigma\|_1 = \sum_{\mathbf{x}} |\sigma(\mathbf{x})|$. When $m = 1$, this notation reduces to that of Section 4.1.1.

Corollary 4.8. *For every $m \geq 1$, the operator $(\mathcal{G}^{-1})^{\otimes m}$ maps $\ell_\theta^\infty(\mathbb{Z}_{\geq 0}^m)$ continuously into $\ell^1(\mathbb{Z}_{\geq 0}^m)$, and*

$$\big\|(\mathcal{G}^{-1})^{\otimes m} z\big\|_1 \leq C_{\mathrm{inv}}^m \|z\|_{\infty,\theta}, \qquad C_{\mathrm{inv}} = \sum_{u \geq 0} \theta_u \sum_{a=0}^{u} |\gamma_{u,a}|.$$

Proof. For $m = 1$ this is Theorem 4.1, which also shows $C_{\mathrm{inv}} < \infty$. For $m > 1$ the row sums of the tensor product factor coordinatewise, so the one-dimensional bound applies in each coordinate; Tonelli's theorem then gives the stated bound. $\square$

Let P_D now denote coordinatewise truncation: $(P_D z)_{\mathbf{k}} = z_{\mathbf{k}}$ when $\|\mathbf{k}\|_\infty \leq D$, and zero otherwise. The residual of a finite signed measure σ on $\mathbb{Z}_{\geq 0}^m$ for which $m_{\mathbf{k}}(|\sigma|) < \infty$ whenever $\|\mathbf{k}\|_\infty \leq D$ is defined as in (50),

$$r_D(\sigma) := (\mathrm{id}^{\otimes m} - Q_D^{\otimes m})\sigma = \sigma - (\mathcal{G}^{-1})^{\otimes m} P_D \mathbf{m}(\sigma),$$

with Q_D the one-dimensional truncated reconstruction (49); the second equality holds because both the moment transform and the truncation act coordinatewise. Here $P_D\mathbf{m}(\sigma)$ denotes only the displayed finite array of retained moments; the full moment sequence need not exist.

Lemma 4.9. *Let σ be a finite signed measure, and assume that $m_{(D+1)\mathbf{1}_S}(|\sigma|)$ is finite for every nonempty $S \subseteq [m]$. Then*

$$\|r_D(\sigma)\|_1 \leq \sum_{\emptyset \neq S \subseteq [m]} \big[C(D+1)2^{D(D+1)/2}\big]^{|S|} m_{(D+1)\mathbf{1}_S}(|\sigma|). \tag{57}$$

Proof. Recall the one-dimensional operator Q_D of (49) and set $R_D := \mathrm{id} - Q_D$, so that $Q_D + R_D = \mathrm{id}$ in one coordinate. The proof of Lemma 4.2 exhibits K_D as the kernel of R_D and gives the row-sum estimate

$$\sum_a |R_D(x,a)| \le C(D+1)2^{D(D+1)/2}\begin{bmatrix} x \\ D+1 \end{bmatrix}_2.$$

By the definition of the tensor residual above, $r_D(\sigma) = (\mathrm{id}^{\otimes m} - Q_D^{\otimes m})\sigma$. Expanding $Q_D = \mathrm{id} - R_D$ gives

$$\mathrm{id}^{\otimes m} - Q_D^{\otimes m} = \sum_{\emptyset \ne S \subseteq [m]} (-1)^{|S|+1} R_D^{\otimes S} \otimes \mathrm{id}^{\otimes S^c}.$$

Apply the one-dimensional row-sum estimate in the coordinates belonging to S, then sum against $|\sigma|$. Tonelli's theorem gives precisely the term indexed by S in (57). $\square$

If $\mathbf{m}(|\sigma|) \in \ell_\theta^\infty(\mathbb{Z}_{\ge 0}^m)$ and $K := \|\mathbf{m}(|\sigma|)\|_{\infty,\theta}$, the definition of the weighted norm gives

$$m_{(D+1)\mathbf{1}_S}(|\sigma|) \le K\theta_{D+1}^{|S|} \qquad (\emptyset \ne S \subseteq [m]).$$

Consequently, the term in (57) indexed by a set of size k is at most $C^k(D+1)^k 2^{-k(D+1)}K$, since $2^{D(D+1)/2}\theta_{D+1} = 2^{-(D+1)}$.

Theorem 4.10. *Let σ be a finite signed measure, and suppose that all residual absolute moments in (57) are finite. Then*

$$\|\sigma\|_1 \le C_{\mathrm{inv}}^m \|P_D \mathbf{m}(\sigma)\|_{\infty,\theta} + \sum_{\emptyset \ne S \subseteq [m]} \left[C(D+1)2^{D(D+1)/2}\right]^{|S|} m_{(D+1)\mathbf{1}_S}(|\sigma|). \tag{58}$$

For $m = 2$,

$$\begin{aligned} \|\sigma\|_1 \le C_{\mathrm{inv}}^2 \|P_D \mathbf{m}(\sigma)\|_{\infty,\theta} + C(D+1)2^{D(D+1)/2}\big(m_{(D+1,0)}(|\sigma|) + m_{(0,D+1)}(|\sigma|)\big) \\ + \left[C(D+1)2^{D(D+1)/2}\right]^2 m_{(D+1,D+1)}(|\sigma|). \end{aligned} \tag{59}$$

Proof. By definition, $\sigma = (\mathcal{G}^{-1})^{\otimes m} P_D \mathbf{m}(\sigma) + r_D(\sigma)$. Corollary 4.8 bounds the first term, and Lemma 4.9 bounds the second. For $m = 1$ this is Theorem 4.3. $\square$

Combining the tensor-product inverse bound with the multivariable residual estimate gives the following convergence criterion for every fixed m.

Corollary 4.11. *Let σ_n be finite signed measures on $\mathbb{Z}_{\ge 0}^m$. Suppose that there are constants $C < \infty$, $C_1 > 1$, $q_{\mathrm{err}} \in (0,1)$ and $\beta_0 > 0$ such that*

(i) *for every n and every multi-index $\mathbf{k}$ with $\|\mathbf{k}\|_\infty \le \beta_0 n$,*

$$|m_{\mathbf{k}}(\sigma_n)| \le C C_1^{\|\mathbf{k}\|_\infty} \theta_{\mathbf{k}} q_{\mathrm{err}}^n;$$

(ii) *the absolute moment sequences are uniformly bounded in the θ-weighted norm:*

$$K := \sup_n \|\mathbf{m}(|\sigma_n|)\|_{\infty,\theta} < \infty.$$

Then there are $C' < \infty$ and $\eta \in (0,1)$ such that $\|\sigma_n\|_1 \le C'\eta^n$. More precisely,

$$\|\sigma_n\|_1 \le C C_{\mathrm{inv}}^m C_1^D q_{\mathrm{err}}^n + C''(D+1)^m 2^{-(D+1)}K.$$

Choosing $D = \lfloor \beta n \rfloor$ with $0 < \beta < \beta_0$ sufficiently small gives $\eta = \max\{2^{-\beta/2}, C_1^\beta q_{\mathrm{err}}\} < 1$. At $\mathbf{k} = \mathbf{0}$, part (i) *also controls the difference of the total masses, so no separate normalization is required.*

Proof. Apply Theorem 4.10. Part (i) bounds the truncated term, while part (ii) and the weighted-norm estimate preceding Theorem 4.10 bound the residual. The terms with $|S| \geq 2$ are dominated by those with $|S| = 1$. The stated choice of D balances the two bounds; its polynomial factor is absorbed by decreasing β. For $m = 1$ this is Corollary 4.4. $\square$

In particular, convergence of the joint distribution to $\mu_1 \otimes \mu_2$ gives asymptotic independence at the same rate. Indeed, if $\| \operatorname{Law}(X_n, Y_n) - \mu_1 \otimes \mu_2 \|_1 \leq C\eta^n$, then marginalization and the triangle inequality give

$$\| \operatorname{Law}(X_n, Y_n) - \operatorname{Law}(X_n) \otimes \operatorname{Law}(Y_n) \|_1 \leq 3C\eta^n.$$

4.2.2. *Truncated inversion on a good event.* For $m = 2$, the criterion below has two cases. Both restrict the probability measure to a good event, compare its conditional mixed moments with a target array, and apply (59). The cases differ in the control of the three residual moments in that inequality and in the resulting truncation depth. In the applications below, the target array is a product $m_{u,v} = m_u^{(1)} m_v^{(2)}$ of two one-matrix target moment sequences from Lemma 4.5; the corresponding one-dimensional inverses were computed in the proof of Theorem 4.6 and are π_{sym} for a zero symmetry defect and π_{CL} for a defect of high rank.

The required conditional mixed-moment estimate is converted to total variation by the following criterion.

Theorem 4.12. *For each n, let (X_n, Y_n) be a pair of nonnegative integer-valued random variables whose joint distribution is ρ_n, together with an auxiliary random element $\boldsymbol{\beta}_n$ on the same probability space. Let μ be a probability measure on $\mathbb{Z}_{\geq 0}^2$ with Gaussian-binomial moments $m_{u,v}$. Suppose that one of the following two sets of hypotheses holds.*

(I) Exponential loss in D. *The target array $(m_{u,v})$ belongs to $\ell_\theta^\infty(\mathbb{Z}_{\geq 0}^2)$. Set $M_\theta := \|(m_{u,v})\|_{\infty,\theta}$. There are constants $C < \infty$, $C_0 > 1$, $q_{\mathrm{err}} \in (0,1)$, $\delta > 0$ and $\beta_0 > 0$ such that, for every integer $0 \leq D \leq \beta_0 n$, there are an $\boldsymbol{\beta}_n$-measurable event $\mathcal{E}_{n,D}$ and a nonnegative $\boldsymbol{\beta}_n$-measurable error weight $\operatorname{Err}_{n,D}$ with the following properties:*

(a) *for almost every realization $\boldsymbol{\beta} \in \mathcal{E}_{n,D}$ and all $u, v \geq 0$,*

$$\mathbb{E}\left[\begin{bmatrix} X_n \\ u \end{bmatrix}_2 \mid \boldsymbol{\beta} \right] \leq C\theta_u, \qquad \mathbb{E}\left[\begin{bmatrix} Y_n \\ v \end{bmatrix}_2 \mid \boldsymbol{\beta} \right] \leq C\theta_v;$$

(b) *for $0 \leq u, v \leq D+1$ and every $\boldsymbol{\beta} \in \mathcal{E}_{n,D}$,*

$$\left| \mathbb{E}\left[\begin{bmatrix} X_n \\ u \end{bmatrix}_2 \begin{bmatrix} Y_n \\ v \end{bmatrix}_2 \mid \boldsymbol{\beta} \right] - m_{u,v} \right| \leq C_0^D m_{u,v} \operatorname{Err}_{n,D}(\boldsymbol{\beta}),$$

and $\mathbb{E}[\operatorname{Err}_{n,D} \mathbb{1}_{\mathcal{E}_{n,D}}] \leq q_{\mathrm{err}}^n$;

(c) *$\mathbb{P}(\mathcal{E}_{n,D}^c) \leq e^{-\delta n}$.*

(II) Exponential loss in D^2. *The target moments factor as $m_{u,v} = m_u^{(1)} m_v^{(2)}$, where $m_0^{(1)} = m_0^{(2)} = 1$ and, for some $M < \infty$,*

$$m_u^{(i)} \leq M 2^{-u^2} \qquad (i = 1, 2;\ u \geq 0).$$

There are constants $c_{\mathrm{loss}}, \beta_0, C_{\mathrm{err}}, \delta > 0$ and $q_{\mathrm{err}} \in (0,1)$ such that, for every integer $0 \leq D \leq \beta_0 \sqrt{n}$, there are an $\boldsymbol{\beta}_n$-measurable event $\mathcal{E}_{n,D}$ and a nonnegative $\boldsymbol{\beta}_n$-measurable error weight $\operatorname{Err}_{n,D}$ satisfying

(a) *for $0 \leq u, v \leq D+1$ and almost every realization $\boldsymbol{\beta} \in \mathcal{E}_{n,D}$,*

$$\left| \mathbb{E}\left[\begin{bmatrix} X_n \\ u \end{bmatrix}_2 \begin{bmatrix} Y_n \\ v \end{bmatrix}_2 \mid \boldsymbol{\beta} \right] - m_{u,v} \right| \leq 2^{c_{\mathrm{loss}}(D+1)^2} \theta_u \theta_v \operatorname{Err}_{n,D}(\boldsymbol{\beta});$$

(b) *$\mathbb{E}[\operatorname{Err}_{n,D} \mathbb{1}_{\mathcal{E}_{n,D}}] \leq C_{\mathrm{err}} q_{\mathrm{err}}^n$;*

(c) $\mathbb{P}(\mathcal{E}_{n,D}^c) \le e^{-\delta n}$.

Then there are $C' < \infty$ *and* $\eta \in (0,1)$ *such that*

$$\|\rho_n - \mu\|_1 \le C'\eta^n.$$

In case (I)*, if* $m_{u,v} = m_u^{(1)} m_v^{(2)}$ *and the inverse one-dimensional moment sequences are probability measures* μ_1, μ_2*, then* $\mu = \mu_1 \otimes \mu_2$*. The same conclusion, and hence asymptotic independence, holds automatically in case* (II)*.*

Proof. Both cases restrict ρ_n to the good event, compare the conditional mixed moments of the restriction with the target moments, and apply the two-dimensional truncated inversion (59) of Theorem 4.10. We carry out these steps once and then specialize the residual estimates and the truncation depth.

Let $D \ge 0$ be an integer in the range allowed by the case at hand; it is chosen at the end of that case. Write $\mathcal{E} = \mathcal{E}_{n,D}$, $p = \mathbb{P}(\mathcal{E})$, and let $\rho_n^{\mathcal{E}}$ be the restriction of ρ_n to $\mathcal{E}$. For the signed measure

$$\sigma_n := \rho_n^{\mathcal{E}} - p\mu$$

we have

$$\|\rho_n - \mu\|_1 \le \|\sigma_n\|_1 + 2\mathbb{P}(\mathcal{E}^c). \tag{60}$$

Because $\mathcal{E}$ is $\boldsymbol{\beta}_n$-measurable, the moments of σ_n are the integrals over $\mathcal{E}$ of the conditional mixed-moment errors. In case (I), the bound $M_\theta < \infty$ turns the relative bound (I.b) into the absolute bound

$$\left|\mathbb{E}\left[\begin{bmatrix} X_n \\ u \end{bmatrix}_2 \begin{bmatrix} Y_n \\ v \end{bmatrix}_2 \mid \boldsymbol{\beta}\right] - m_{u,v}\right| \le L_D\, \theta_u \theta_v \mathrm{Err}_{n,D}(\boldsymbol{\beta}), \qquad L_D := C_0^D M_\theta,$$

valid for $0 \le u, v \le D+1$ and $\boldsymbol{\beta} \in \mathcal{E}$; in case (II), hypothesis (II.a) is the same inequality with $L_D := 2^{c_{\mathrm{loss}}(D+1)^2}$. Integrating over $\mathcal{E}$ and using $\mathbb{E}[\mathrm{Err}_{n,D}\mathbb{1}_{\mathcal{E}}] \le C_{\mathrm{err}} q_{\mathrm{err}}^n$, with $C_{\mathrm{err}} = 1$ in case (I), gives

$$|m_{u,v}(\sigma_n)| \le C_{\mathrm{err}} L_D\, \theta_u \theta_v q_{\mathrm{err}}^n \qquad (0 \le u, v \le D+1), \tag{61}$$

so that the truncated term of (59) is at most $C_{\mathrm{inv}}^2 C_{\mathrm{err}} L_D q_{\mathrm{err}}^n$. The three residual absolute moments in (59) are estimated from $|\sigma_n| \le \rho_n^{\mathcal{E}} + p\mu$, and their prefactors are handled using $2^{D(D+1)/2}\theta_{D+1} = 2^{-(D+1)}$. Only these residual moments and the choice of D depend on the case.

Case (I). Choose $\beta \in (0, \beta_0)$ so small that $C_0^\beta q_{\mathrm{err}} < 1$, and put $D = \lfloor \beta n \rfloor$. The one-coordinate absolute moments satisfy, by (I.a),

$$m_{D+1,0}(|\sigma_n|) \le (C + M_\theta)\theta_{D+1},$$
$$m_{0,D+1}(|\sigma_n|) \le (C + M_\theta)\theta_{D+1},$$

while condition (I.b) at $(D+1, D+1)$ gives

$$m_{D+1,D+1}(|\sigma_n|) \le \big(2 + C_0^D q_{\mathrm{err}}^n\big) M_\theta \theta_{D+1}^2.$$

Inserting these bounds and (61) into (59) yields

$$\|\sigma_n\|_1 \ll C_0^D q_{\mathrm{err}}^n + (D+1)2^{-(D+1)} + (D+1)^2 2^{-2(D+1)}\big(1 + C_0^D q_{\mathrm{err}}^n\big).$$

For $D = \lfloor \beta n \rfloor$, the right-hand side is at most $C'\eta^n$ for some $\eta \in (0,1)$. Combining this estimate with (60) and (I.c) proves case (I). If the target moments factor as in the final assertion of the theorem, the tensor-product identity for the inverse transform gives $\mu = \mu_1 \otimes \mu_2$.

Case (II). Unlike case (I), this case has no conditional marginal bound and allows a loss of $2^{O(D^2)}$ in the mixed moments. The faster decay 2^{-u^2} of the target moments compensates for this loss at the shorter truncation depth $D \asymp \sqrt{n}$. Choose $\beta \in (0, \beta_0)$ so small that $2^{2c_{\mathrm{loss}}\beta^2} q_{\mathrm{err}} < 1$ and put $D = \lfloor \beta\sqrt{n} \rfloor$. Here (61) reads

$$|m_{u,v}(\sigma_n)| \le C_{\mathrm{err}} 2^{c_{\mathrm{loss}}(D+1)^2} \theta_u \theta_v q_{\mathrm{err}}^n. \tag{62}$$

Applying part (II.a) at $(D+1,0)$, $(0,D+1)$ and $(D+1,D+1)$, and using $m_u^{(i)} \le M2^{-u^2}$ for the target moments, we obtain

$$m_{D+1,0}(|\sigma_n|) + m_{0,D+1}(|\sigma_n|) \ll 2^{-(D+1)^2} + 2^{c_{\text{loss}}(D+1)^2}\theta_{D+1}q_{\text{err}}^n,$$
$$m_{D+1,D+1}(|\sigma_n|) \ll 2^{-2(D+1)^2} + 2^{c_{\text{loss}}(D+1)^2}\theta_{D+1}^2 q_{\text{err}}^n.$$

After multiplication by the residual prefactors in (59), the terms containing q_{err}^n in these two lines are bounded by the truncated signed-moment term and are absorbed into it. Inserting the remaining terms and (62) into (59) gives

$$\|\sigma_n\|_1 \ll 2^{c_{\text{loss}}(D+1)^2}q_{\text{err}}^n + (D+1)2^{-D^2/2+O(D)} + (D+1)^2 2^{-D^2+O(D)}.$$

For $D = \lfloor \beta\sqrt{n} \rfloor$, the sum of the three terms is at most $C'\eta^n$ for some $\eta \in (0,1)$. Part (II.c) and (60) prove convergence in ℓ^1. Let μ_1, μ_2 be the marginals of μ. Their moment sequences are $m_u^{(1)}$ and $m_v^{(2)}$, respectively, whereas $\mu_1 \otimes \mu_2$ has mixed moments $m_u^{(1)} m_v^{(2)} = m_{u,v}$. These arrays belong to $\ell_\theta^\infty(\mathbb{Z}_{\ge 0}^2)$, and bounded tensor-product inversion is injective on this weighted domain. Hence $\mu = \mu_1 \otimes \mu_2$, proving asymptotic independence. $\square$

The good events in both cases may depend on the truncation depth, as the support and cross-rank thresholds grow linearly with D. Only the moments of the good submeasure enter the residual; no high-moment bound is required on the exponentially small bad event.

Case (II) of Theorem 4.12 is applied to the bordered matrix pair on the probability space $\mathsf{MatData}_r(\mathbf{u}_{-1})$ of Section 2.1.3, whose good event is constructed in Section 4.3 and whose conditional mixed moments are computed with a $2^{O(D^2)}$ loss in Section 4.4; case (I) suffices for the symmetric unbordered diagonal perturbations of Section 4.5. In both cases the target array is the product of the two one-matrix target moment sequences, so the limit produced is the corresponding product of one-matrix limiting distributions.

4.3. **The bordered-pair theorem and the good event.** The estimate required in (46) is uniform in the 2-adic stratum and the sign vector.

Theorem 4.13. *Let $d_0 > 1$ be squarefree and $\alpha \in (0, \frac{1}{2})$. There exist $C_{\text{rmt}} = C_{\text{rmt}}(d_0, \alpha) < \infty$ and $\eta = \eta(d_0, \alpha) \in (0,1)$ such that, for every $r \ge 2$ and every sign vector $\mathbf{u}_{-1} \in \mathbb{F}_2^r$ satisfying $|\mathbf{u}_{-1}| \ge \alpha r$, on the probability space $\mathsf{MatData}_r(\mathbf{u}_{-1})$ of Section 2.1.3, the corank pair satisfies*

$$\left\|\mathrm{Law}(Z_1, Z_2) - \pi_{\text{CL}} \otimes \pi_{\text{CL}}\right\|_1 \le C_{\text{rmt}}\,\eta^r.$$

In particular, the limit is independent of the stratum and of the sign vector in the indicated range.

The proof applies case (II) of Theorem 4.12 with the border data as the auxiliary random element. The present subsection fixes that conditioning and constructs the good event $\mathcal{E}_{n,D}$ together with its probability bound (II.c). The remaining input, hypothesis (II.a), is the conditional mixed-moment estimate of Proposition 4.15; the deduction of Theorem 4.13 is therefore completed in Section 4.4, after that proposition has been proved.

Fix d_0, $r \ge 2$ and a sign vector $\mathbf{u}_{-1}$ with $|\mathbf{u}_{-1}| \ge \alpha r$, and take all random variables on $\mathsf{MatData}_r(\mathbf{u}_{-1})$ with the uniform measure of Section 2.1.3. We write

$$n := r-1, \qquad \Omega := \mathbf{u}_{-1}^{\circ}(\mathbf{u}_{-1}^{\circ})^\top + \mathrm{diag}(\mathbf{u}_{-1}^{\circ}), \qquad \Gamma := \mathrm{diag}(\mathbf{u}_{d_0}) + \Omega, \qquad h := |\operatorname{supp} \mathbf{u}_{d_0}|.$$

We call $\boldsymbol{\beta} := (\mathbf{u}_2, \mathbf{u}_{\Delta_1}, \ldots, \mathbf{u}_{\Delta_s}, \mathbf{c})$ the *border data*; it determines $\mathbf{u}_{d_0}$, the two border blocks and the corner block, and the core block $A \sim \mathrm{Unif}(\mathrm{Sym}_n(\Omega))$ is independent of it. The matrices U, L have size $n \times b_2$, and

$$N_2 = \begin{pmatrix} A + \mathrm{diag}(\mathbf{u}_{d_0}) & \mathsf{U} \\ \mathsf{L}^\top & C \end{pmatrix},$$

while N_1 is the matrix of Theorem 2.4, of border width $b_1 \le 1$. Since $d_0 > 1$, one has $b_2 \ge 1$, and consequently $b_1 \le b_2$.

The character vector $\mathbf{u}_{d_0}$ is uniform on $\mathbb{F}_2^n$, independently of the sign vector. Write $d_0 = 2^e q_1 \cdots q_s$, as in Section 2.1.2, and use

$$[\tfrac{q_k}{p}] = [\tfrac{q_k^*}{p}] + [\tfrac{-1}{q_k}][\tfrac{-1}{p}].$$

For each $i \le n$, the definition of $\mathbf{u}_{d_0}$ gives

$$u_{d_0,i} = e\, u_{2,i} + \sum_{k \le s} \big(u_{\Delta_k,i} + [\tfrac{-1}{q_k}] u_{-1,i} \big).$$

This is a nonconstant affine function of $(u_{2,i}, u_{\Delta_1,i}, \dots, u_{\Delta_s,i})$, because $d_0 > 1$ forces $e = 1$ or $s \ge 1$. These coordinate tuples are independent and uniform as i varies, so the coordinates $u_{d_0,i}$ are independent and uniform.

Lemma 4.14. *Let $\alpha \in (0, \frac{1}{2})$ and $|\mathbf{u}_{-1}| \ge \alpha r$. For $H \le n/4$, define*

$$\begin{aligned} \mathcal{E}_{\mathrm{good}}(H) := \{h \ge H\} \cap \{\operatorname{rank} \Gamma \ge H\} \cap \{\operatorname{rank}[\, \mathsf{L}_1^\top \;\; \mathsf{C}_1 \,] = b_1\} \cap \{\operatorname{rank}[\, \mathsf{L}^\top \;\; C \,] = b_2\} \\ \cap \{\ker \mathsf{U} = \mathbf{0}\} \cap \{b_1 = 0 \ \text{or}\ \mathbf{u}_2 \ne \mathbf{0}\}. \end{aligned} \tag{63}$$

There is $\delta = \delta(\alpha, d_0) > 0$ such that, uniformly for $H \le n/4$,

$$\mathbb{P}\big(\mathcal{E}_{\mathrm{good}}(H)^c\big) \le e^{-\delta n}$$

for all sufficiently large n.

Proof. By the preceding calculation, $h = |\operatorname{supp} \mathbf{u}_{d_0}|$ is binomial with parameters $(n, \frac{1}{2})$, so Hoeffding's inequality gives $\mathbb{P}(h < n/4) \le e^{-n/8}$. Also

$$\Gamma = \operatorname{diag}(\mathbf{u}_{d_0} + \mathbf{u}_{-1}^\circ) + \mathbf{u}_{-1}^\circ (\mathbf{u}_{-1}^\circ)^\top,$$

and therefore

$$\operatorname{rank} \Gamma \ge \#\{i \le n : u_{d_0,i} \ne u_{-1,i}\} - 1.$$

The count on the right is again binomial with parameters $(n, \frac{1}{2})$, uniformly in the fixed sign vector. Hence, for every $H \le n/4$,

$$\mathbb{P}(\operatorname{rank} \Gamma < H) \le \mathbb{P}\big(\operatorname{Bin}(n, \tfrac{1}{2}) < H + 1\big) \le e^{-cn}.$$

Together with the estimate for h, this controls the first two conditions, after changing the constants for finitely many n.

For the border rows of N_2, the vectors $\mathbf{u}_{\Delta_1}, \dots, \mathbf{u}_{\Delta_s}$ are independent and uniform in $\mathbb{F}_2^n$. The possible row indexed by 0 has $\mathbb{F}_2^n$-part $\mathbf{u}_{\Delta_0} \in \{\mathbf{u}_{-1}^\circ, \mathbf{u}_2, \mathbf{u}_{-1}^\circ + \mathbf{u}_2\}$. In the last two cases it contains the independent uniform vector $\mathbf{u}_2$; in the first it is the fixed nonzero vector $\mathbf{u}_{-1}^\circ$. Thus every nontrivial linear combination of the b_2 rows has a uniform, or fixed nonzero, $\mathbb{F}_2^n$-part. A union bound gives

$$\mathbb{P}\big(\operatorname{rank}[\, \mathsf{L}^\top \;\; C \,] < b_2\big) \le 2^{b_2 - n}.$$

If $b_1 = 1$, the $\mathbb{F}_2^n$-part of the border row of N_1 is $(\mathbf{u}_{-1}^\circ)^\top$, which is nonzero for all sufficiently large r because $|\mathbf{u}_{-1}| \ge \alpha r$. Thus $\operatorname{rank}[\, \mathsf{L}_1^\top \;\; \mathsf{C}_1 \,] = b_1$; the assertion is vacuous when $b_1 = 0$.

The columns of U are $\mathbf{u}_{q_k}$. For $k \ge 1$, $\mathbf{u}_{q_k} = \mathbf{u}_{\Delta_k} + [\frac{-1}{q_k}] \mathbf{u}_{-1}^\circ$, and, when present, $\mathbf{u}_{q_0} = \mathbf{u}_2$. Every nontrivial linear combination of these columns is uniform in $\mathbb{F}_2^n$, up to a fixed translation. Hence

$$\mathbb{P}(\ker \mathsf{U} \ne \mathbf{0}) \le 2^{b_2 - n}.$$

When $b_1 = 1$, also $\mathbb{P}(\mathbf{u}_2 = \mathbf{0}) = 2^{-n}$. Adding the estimates proves the lemma. $\square$

The bounds for h and $\operatorname{rank}\Gamma$ supply the codimensions of the overlap and cross-radical bad subspaces in Lemma 4.18; the two full-row-rank conditions give $\dim H_1 = \dim H_2 = n$ in Proposition 4.16; and $\ker \mathsf{U} = 0$, together with $\mathbf{u}_2 \neq 0$ when $b_1 = 1$, makes the two coordinate projections injective on every tuple span of nonzero weight.

There are two distinct exceptional sets. The event $\mathcal{E}_{\mathrm{good}}(H)$ constrains the border data alone; its complement has probability at most $e^{-\delta n}$ and is discarded once, in the peeling step (60). After a realization $\boldsymbol{\beta} \in \mathcal{E}_{\mathrm{good}}(H)$ has been fixed, exceptional kernel tuples still occur inside the conditional expectation over the core A. Section 4.4 controls them by the codimensions and ranks guaranteed by (63), rather than by another probability estimate for the border data.

4.4. **Conditional mixed moments of bordered pairs.** Proposition 4.15 is the conditional mixed-moment estimate required by hypothesis (II.a) of Theorem 4.12: for border data in the good event, the conditional Gaussian-binomial mixed moments of the corank pair agree with the product $m_u(\Omega)m_v(\Omega)$ of two one-matrix target moments, up to a factor $2^{O(D^2)}$ times the averaged error variable that supplies (II.b).

Proposition 4.15. *Let $d_0 > 1$ be squarefree and $\alpha \in (0, \frac{1}{2})$. There is a constant $c_{\mathrm{mix}} > 0$, depending only on d_0 and α, with the following property. Fix an integer $r \geq 2$ and a sign vector $\mathbf{u}_{-1}$ with $|\mathbf{u}_{-1}| \geq \alpha r$. Put*

$$b = b_1 + b_2, \qquad H_D = 4D + 6b_2 + b + 12.$$

For every $D \geq 0$ such that

$$\operatorname{rank}\Omega \geq 4D + 8b_2 + 10, \tag{64}$$

all $0 \leq u, v \leq D+1$, and every realization $\boldsymbol{\beta}$ of the border data in $\mathcal{E}_{\mathrm{good}}(H_D)$, the conditional expectation under $\mathsf{MatData}_r(\mathbf{u}_{-1})$ satisfies

$$\begin{aligned} &\Big|\mathbb{E}_A\Big[\begin{bmatrix} Z_1 \\ u \end{bmatrix}_2 \begin{bmatrix} Z_2 \\ v \end{bmatrix}_2 \;\Big|\; \boldsymbol{\beta}\Big] - m_u(\Omega)m_v(\Omega)\Big| \\ &\qquad \leq 2^{c_{\mathrm{mix}}(D+b_2+b+1)^2}\theta_u\theta_v\Big(2^{-h} + 2^{-\operatorname{rank}\Gamma} + 2^{-n} + 2^{-\operatorname{rank}\Omega}\Big). \end{aligned} \tag{65}$$

The conditional moment calculation separates into two steps. First the bordered kernel equations are converted exactly into a weighted sum over two families of tuples. That sum is then estimated by separating pairs with disjoint projected spans from the exceptional pairs. We introduce the common terminology before carrying out the first step.

Let V_1, V_2 be n-dimensional $\mathbb{F}_2$-spaces, let $\psi_i : V_i \to \mathbb{F}_2^n$ be linear maps and Ω_i an alternating form on V_i, for $i = 1, 2$, and let $B_{12} : V_1 \times V_2 \to \mathbb{F}_2$ be bilinear. For ordered linearly independent tuples $\mathbf{w}_\bullet = (\mathbf{w}_1, \ldots, \mathbf{w}_u)$ and $\mathbf{w}'_\bullet = (\mathbf{w}'_1, \ldots, \mathbf{w}'_v)$, put

$$\mathcal{X} := \psi_1(\operatorname{span}\mathbf{w}_\bullet), \qquad \mathcal{Y} := \psi_2(\operatorname{span}\mathbf{w}'_\bullet), \qquad k := \dim(\mathcal{X} + \mathcal{Y}).$$

Given a predicate $\mathcal{R}_{u,v}(\mathbf{w}_\bullet, \mathbf{w}'_\bullet)$, call the tuple pair *compatible* if

(i) ψ_1, ψ_2 are injective on $\operatorname{span}\mathbf{w}_\bullet, \operatorname{span}\mathbf{w}'_\bullet$, respectively;
(ii) the two spans are totally isotropic for Ω_1 and Ω_2, respectively;
(iii) $B_{12}(\mathbf{w}_i, \mathbf{w}'_j) = 0$ for every i, j;
(iv) $\mathcal{R}_{u,v}(\mathbf{w}_\bullet, \mathbf{w}'_\bullet)$ holds.

The associated compatible-pair sum is

$$\mathcal{W}_{u,v} := \frac{1}{|\operatorname{GL}_u||\operatorname{GL}_v|} \sum_{(\mathbf{w}_\bullet, \mathbf{w}'_\bullet)\ \text{compatible}} 2^{-\left(nk - \binom{k}{2}\right)}. \tag{66}$$

Thus compatibility records exactly the algebraic conditions on a tuple pair; once they hold, the weight depends only on the dimension k of the joint projected span.

For $S \subseteq \{1, \ldots, n\}$, let $\pi_S : \mathbb{F}_2^n \to \mathbb{F}_2^S$ be the coordinate projection and put $V_{S^c} := \ker \pi_S$.

Proposition 4.16. *Let $H \ge 0$ and $\boldsymbol{\beta} \in \mathcal{E}_{\mathrm{good}}(H)$, and use $\mathsf{U}_1, \mathsf{L}_1, \mathsf{C}_1$ from the assembly map (27). Thus*

$$N_1 = \begin{pmatrix} A & \mathsf{U}_1 \\ \mathsf{L}_1^\top & \mathsf{C}_1 \end{pmatrix}, \qquad N_2 = \begin{pmatrix} A + \operatorname{diag}(\mathbf{u}_{d_0}) & \mathsf{U} \\ \mathsf{L}^\top & C \end{pmatrix}.$$

Put

$$H_1 := \ker[\,\mathsf{L}_1^\top \ \ \mathsf{C}_1\,] \subseteq \mathbb{F}_2^{n+b_1}, \qquad H_2 := \ker[\,\mathsf{L}^\top \ \ C\,] \subseteq \mathbb{F}_2^{n+b_2},$$

so both spaces have dimension n, and let ψ_1, ψ_2 drop the last b_1, respectively b_2, coordinates. Define the restricted alternating and cross forms directly by

$$\Omega_1 := \begin{pmatrix} \Omega & \mathsf{U}_1 \\ \mathsf{U}_1^\top & 0_{b_1} \end{pmatrix}\Bigg|_{H_1}, \qquad \Omega_2 := \begin{pmatrix} \Omega & \mathsf{U} \\ \mathsf{U}^\top & 0_{b_2} \end{pmatrix}\Bigg|_{H_2}, \qquad B_{12} := \begin{pmatrix} \Gamma & \mathsf{U} \\ \mathsf{U}_1^\top & 0_{b_1 \times b_2} \end{pmatrix}\Bigg|_{H_1 \times H_2}.$$

Finally, with $S = \operatorname{supp} \mathbf{u}_{d_0}$, put

$$\Sigma := \pi_S\{\mathsf{U}_1\mathbf{a} + \mathsf{U}\mathbf{b} : \mathbf{a} \in \mathbb{F}_2^{b_1},\ \mathbf{b} \in \mathbb{F}_2^{b_2}\}, \qquad \dim \Sigma \le b := b_1 + b_2.$$

Set $W_\Sigma := \pi_S^{-1}(\Sigma)$. For tuple pairs $\mathbf{w}_i = (\mathbf{x}_i, \mathbf{a}_i) \in H_1$ and $\mathbf{w}'_j = (\mathbf{y}_j, \mathbf{b}_j) \in H_2$, let $\mathcal{R}_{u,v}$ be the following target-consistency condition: whenever

$$\mathbf{z} = \sum_i \xi_i \mathbf{x}_i = \sum_j \zeta_j \mathbf{y}_j,$$

one has

$$\mathsf{U}_1 \sum_i \xi_i \mathbf{a}_i = \operatorname{diag}(\mathbf{u}_{d_0})\mathbf{z} + \mathsf{U} \sum_j \zeta_j \mathbf{b}_j. \tag{67}$$

In the compatible-pair sum (66), take $V_i = H_i$ and use the maps, alternating forms, and predicate just defined. Then, for all $u, v \ge 0$,

$$\mathbb{E}_A\Big[\begin{bmatrix} Z_1 \\ u \end{bmatrix}_2 \begin{bmatrix} Z_2 \\ v \end{bmatrix}_2 \;\Big|\; \boldsymbol{\beta}\Big] = \mathcal{W}_{u,v},$$

and the predicate has the two properties

$$\mathcal{R}_{u,v} \Longrightarrow \mathcal{X} \cap \mathcal{Y} \subseteq W_\Sigma,$$

$$\mathcal{X} \cap \mathcal{Y} = \{\mathbf{0}\} \text{ and } \psi_1, \psi_2 \text{ injective on the two tuple spans} \Longrightarrow \mathcal{R}_{u,v}.$$

Finally,

$$\dim\ker\psi_1,\ \dim\ker\psi_2 \le b_2, \qquad \operatorname{codim}_{H_1} \psi_1^{-1}(W_\Sigma) \ge h - b_2 - b,$$
$$|\operatorname{rank}\Omega_1 - \operatorname{rank}\Omega|, |\operatorname{rank}\Omega_2 - \operatorname{rank}\Omega| \le 4b_2,$$

and

$$\operatorname{rank} B_{12} \ge \operatorname{rank}\Gamma - 2(b_1 + b_2).$$

Proof. Kernel tuples. The kernel conditions are

$$N_1(\mathbf{x}, \mathbf{a})^\top = 0 \iff A\mathbf{x} = \mathsf{U}_1\mathbf{a}, \quad (\mathbf{x}, \mathbf{a}) \in H_1,$$
$$N_2(\mathbf{y}, \mathbf{b})^\top = 0 \iff A\mathbf{y} = \operatorname{diag}(\mathbf{u}_{d_0})\mathbf{y} + \mathsf{U}\mathbf{b}, \quad (\mathbf{y}, \mathbf{b}) \in H_2.$$

The two full-row-rank conditions in the good event give $\dim H_1 = \dim H_2 = n$.

Expand each Gaussian binomial coefficient by ordered bases:

$$\begin{bmatrix} Z_1 \\ u \end{bmatrix}_2 \begin{bmatrix} Z_2 \\ v \end{bmatrix}_2 = \frac{1}{|\operatorname{GL}_u||\operatorname{GL}_v|} \sum_{\substack{\mathbf{w}_\bullet \text{ independent in } \ker N_1 \\ \mathbf{w}'_\bullet \text{ independent in } \ker N_2}} 1.$$

For a first-family tuple occurring in this sum, a relation $\sum_i \xi_i \mathbf{x}_i = 0$ gives $\mathsf{U}_1 \sum_i \xi_i \mathbf{a}_i = 0$. The map U_1 is injective on the good event. This is vacuous when $b_1 = 0$ and follows from $\mathbf{u}_2 \ne 0$ when $b_1 = 1$. Thus the corresponding relation among the full vectors is zero. Likewise, a relation among the projected second-family vectors gives $\mathsf{U}\sum_j \zeta_j \mathbf{b}_j = 0$, and $\ker \mathsf{U} = 0$ on the good event. Therefore ψ_1 and ψ_2 are injective on every span contributing nonzero weight.

Compatibility with the common core. For fixed kernel tuples, the equations imposed on A have source vectors $\mathbf{x}_i, \mathbf{y}_j$. The two families have target vectors

$$\mathsf{U}_1\mathbf{a}_i, \qquad \operatorname{diag}(\mathbf{u}_{d_0})\mathbf{y}_j + \mathsf{U}\mathbf{b}_j,$$

respectively. If the combined projected family is dependent, these targets define a linear map on $\mathcal{X} + \mathcal{Y}$ precisely when (67) holds. Thus condition (iv) in the definition of compatibility is exactly the consistency of the target assignment, rather than an additional estimate.

Once the target assignment is well defined, the compatibility equations in Lemma 2.9 have three types. Two first-family vectors give $\mathbf{w}_i^\top \Omega_1 \mathbf{w}_j = 0$, and two second-family vectors give $(\mathbf{w}_i')^\top \Omega_2 \mathbf{w}_j' = 0$. A first-family vector against a second-family vector gives

$$\mathbf{x}_i^\top \Gamma \mathbf{y}_j + \mathbf{x}_i^\top \mathsf{U}\mathbf{b}_j + \mathbf{a}_i^\top \mathsf{U}_1^\top \mathbf{y}_j = 0,$$

which is the cross condition for B_{12}. Restriction to H_1 and H_2 gives compatibility conditions (ii) and (iii) for Ω_1, Ω_2, and B_{12}.

Apply Lemma 2.9 to a basis of the joint projected span. If any consistency or compatibility condition fails, the probability is zero; otherwise it is $2^{-(nk-\binom{k}{2})}$, where k is the dimension of the joint projected span; the preceding ordered-basis expansion now gives (66).

Overlap and rank bounds. When the two projections are injective on the corresponding tuple spans, the predicate $\mathcal{R}_{u,v}$ is automatic if the two projected spans have zero intersection: the only common projected relation is then the zero relation on both sides. If $\mathcal{R}_{u,v}$ holds and $z \in \mathcal{X} \cap \mathcal{Y}$, then (67) gives

$$\operatorname{diag}(\mathbf{u}_{d_0})z = \mathsf{U}_1\mathbf{a}_0 + \mathsf{U}\mathbf{b}_0.$$

On $S = \operatorname{supp} \mathbf{u}_{d_0}$ the left side restricts to $\pi_S(z)$, and hence $\pi_S(z) \in \Sigma$. This proves the two asserted properties of $\mathcal{R}_{u,v}$.

The projection kernels have dimensions at most b_1 and b_2. Restricting the coordinate projection to H_1 can lose at most b_1 ranks, and restricting it to H_2 can lose at most b_2 ranks. Passing from V_{S^c} to W_Σ loses at most $\dim \Sigma \le b$ further ranks. Thus

$$\operatorname{codim}_{H_1} \psi_1^{-1}(W_\Sigma) \ge h - b_1 - b \ge h - b_2 - b.$$

Bordering an alternating form by width at most b_2 changes its rank by at most $2b_2$, and restricting it to a subspace of codimension at most b_2 changes the rank by at most another $2b_2$. Hence

$$|\operatorname{rank}\Omega_1 - \operatorname{rank}\Omega| \le 4b_2, \qquad |\operatorname{rank}\Omega_2 - \operatorname{rank}\Omega| \le 4b_2.$$

Finally, adjoining rows and columns cannot lower the rank of the top-left block Γ, while restriction to $H_1 \times H_2$ lowers rank by at most $b_1 + b_2$ on each side. This gives the stated conservative bound $\operatorname{rank} B_{12} \ge \operatorname{rank}\Gamma - 2(b_1 + b_2)$. $\square$

In the bordered setting, the two projected kernel spaces can overlap. The first property of $\mathcal{R}_{u,v}$ in Proposition 4.16 shows that a compatible pair with $\mathcal{X} \cap \mathcal{Y} \neq 0$ must have $\mathcal{X} \cap W_\Sigma \neq 0$, where $W_\Sigma = \pi_S^{-1}(\Sigma)$ has large codimension. The following relative count quantifies the resulting loss. It is the input that controls overlapping compatible pairs in Lemma 4.18.

Lemma 4.17. *Let B be an alternating form on $\mathbb{F}_2^n$, with radical* $\operatorname{rad}(B)$*, and let $K \subseteq \mathbb{F}_2^n$ have codimension γ. Call an ordered linearly independent u-tuple $\mathbf{x}_\bullet$* radical-free isotropic *if its span $\mathcal{X}$ is totally isotropic for B and $\mathcal{X} \cap \operatorname{rad}(B) = \mathbf{0}$. Let $\mathcal{N}$ be the number of such tuples and $\mathcal{N}_{\ge g}$ the number among them with* $\dim(\mathcal{X} \cap K) \ge g$*. If* $\operatorname{rank} B \ge 2u+2$ *and $n \ge 2u+2$, then for $0 \le g \le u$*

$$\mathcal{N}_{\ge g} \;\le\; C^{u}\, 2^{\,g(u-g)+\binom{g}{2}-g\gamma}\, \mathcal{N}, \tag{68}$$

with C absolute.

Proof. Put $R := \mathrm{rad}(B)$. Say a tuple is *K-adapted* if $\mathbf{x}_1, \dots, \mathbf{x}_g \in K$, and let $\mathcal{A}_g$ be the number of radical-free isotropic K-adapted tuples.

Adapted bases. Fix an radical-free isotropic $\mathcal{X}$ with $m := \dim(\mathcal{X} \cap K) \geq g$. Among the $|\,\mathrm{GL}_u\,| = \prod_{i<u}(2^u - 2^i)$ ordered bases of $\mathcal{X}$, those that are K-adapted number $\prod_{i<g}(2^m - 2^i) \cdot \prod_{i=g}^{u-1}(2^u - 2^i)$, so the K-adapted ones make up a fraction $\prod_{i<g} \frac{2^m - 2^i}{2^u - 2^i} \geq \eta_\infty(2)\, 2^{-g(u-m)} \geq \eta_\infty(2)\, 2^{-g(u-g)}$ of them, using $m \geq g$. Summing over the radical-free isotropic $\mathcal{X}$ with $\dim(\mathcal{X} \cap K) \geq g$ gives $\mathcal{A}_g \geq \eta_\infty(2) 2^{-g(u-g)} \mathcal{N}_{\geq g}$.

Counting adapted tuples. Build a tuple one vector at a time. Having chosen $\mathbf{x}_1, \dots, \mathbf{x}_i$ spanning an isotropic U_i with $U_i \cap R = \mathbf{0}$, isotropy confines $\mathbf{x}_{i+1}$ to $U_i^{\perp_B}$, whose dimension is $n - \dim B(U_i) = n - i$ exactly, because $U_i \cap R = \mathbf{0}$ makes B injective on U_i. Hence an unconstrained position offers at most $2^{\,n-i}$ choices, while a position required to lie in K offers at most $|K| = 2^{\,n-\gamma}$. Therefore

$$\mathcal{A}_g \;\leq\; \prod_{i<g} 2^{\,n-\gamma} \cdot \prod_{i=g}^{u-1} 2^{\,n-i} \;=\; 2^{\,nu - \binom{u}{2}} \cdot 2^{\,\binom{g}{2} - g\gamma},$$

since $\sum_{i<g}(n-\gamma) + \sum_{i=g}^{u-1}(n-i) = nu - \binom{u}{2} + \binom{g}{2} - g\gamma$. Conversely, a lower bound for $\mathcal{N}$: at step i one must avoid U_i itself (2^i vectors) and avoid creating an intersection with the radical, i.e. avoid the set $U_i + R$ of at most $2^{\,i+n-\operatorname{rank} B}$ vectors, so the number of choices is at least $2^{\,n-i} - 2^i - 2^{\,i+n-\operatorname{rank} B} \geq \frac{1}{2}\, 2^{\,n-i}$ whenever $i \leq u-1$, $n \geq 2u+2$ and $\operatorname{rank} B \geq 2u+2$. Hence $\mathcal{N} \geq 2^{-u} 2^{\,nu - \binom{u}{2}}$, which gives the required lower bound for $\mathcal{N}$.

Combining the two steps gives (68). $\square$

The comparison is made against the isotropic count $\mathcal{N}$ itself. Both $\mathcal{N}$ and $\mathcal{N}_{\geq g}$ contain only tuples whose spans avoid $\mathrm{rad}(B)$. Tuples meeting the radical are not covered by (68); they are estimated by a separate error term in Lemma 4.18.

Lemma 4.18. *Let $\mathcal{W}_{u,v}$ be the compatible-pair sum (66), formed from n-dimensional spaces V_1, V_2, maps ψ_1, ψ_2, alternating forms Ω_1, Ω_2, a bilinear form B_{12}, and predicates $\mathcal{R}_{u,v}$. Fix a subset $S \subseteq \{1, \dots, n\}$ with $h := |S|$, a subspace $\Sigma \subseteq \mathbb{F}_2^S$ with $\dim \Sigma \leq b$, and a nonzero matrix $\Omega \in \mathrm{Alt}_n$. Put*

$$W_\Sigma := \pi_S^{-1}(\Sigma), \qquad \operatorname{codim} W_\Sigma \geq h - b.$$

Assume, for some fixed $T \geq 0$,

(P1)
$$\dim\ker\psi_1,\ \dim\ker\psi_2 \leq T, \qquad \operatorname{codim}_{V_1} \psi_1^{-1}(W_\Sigma) \geq h - T - b,$$

(P2)
$$|\operatorname{rank}\Omega_1 - \operatorname{rank}\Omega| \leq 4T, \qquad |\operatorname{rank}\Omega_2 - \operatorname{rank}\Omega| \leq 4T.$$

Suppose that the predicates have the two properties

(R1) *$\mathcal{R}_{u,v}$ implies $\mathcal{X} \cap \mathcal{Y} \subseteq W_\Sigma$;*

(R2) *$\mathcal{R}_{u,v}$ holds whenever $\mathcal{X} \cap \mathcal{Y} = \{\mathbf{0}\}$ and ψ_1, ψ_2 are injective on* $\operatorname{span}\mathbf{w}_\bullet, \operatorname{span}\mathbf{w}'_\bullet$*, respectively.*

There is an absolute constant $c_* > 0$ such that, for $0 \leq u, v \leq D+1$, if

$$h \geq 4D + 2T + b + 10, \qquad \operatorname{rank} B_{12} \geq 4D + 2T + 10, \qquad \operatorname{rank}\Omega \geq 4D + 8T + 10,$$

then

$$|\mathcal{W}_{u,v} - m_u(\Omega) m_v(\Omega)| \leq 2^{c_*(D+T+b+1)^2} \theta_u \theta_v \times \left(2^{-h} + 2^{-\operatorname{rank} B_{12}} + 2^{-n} + 2^{-\operatorname{rank}\Omega}\right). \tag{69}$$

Proof. Put the effective depth and combined error

$$D_* := D + T + b + 1, \qquad \varepsilon_* := 2^{-h} + 2^{-\operatorname{rank} B_{12}} + 2^{-n} + 2^{-\operatorname{rank}\Omega}.$$

Throughout the proof, C denotes an absolute constant that may change from line to line. Since $\Omega \neq 0$, the target moments satisfy

$$m_j(\Omega) = \frac{1}{|\operatorname{GL}_j|} \leq C\theta_j.$$

Also put

$$K_\Sigma := \psi_1^{-1}(W_\Sigma), \qquad K_{12} := \{\mathbf{x} \in V_1 : B_{12}(\mathbf{x}, \mathbf{y}) = 0 \text{ for every } \mathbf{y} \in V_2\},$$

so that

$$\operatorname{codim} K_\Sigma \geq h - T - b, \qquad \operatorname{codim} K_{12} = \operatorname{rank} B_{12}.$$

For a compatible pair, property (R1) implies

$$\dim(\mathcal{X} \cap \mathcal{Y}) \leq \dim(\operatorname{span} \mathbf{w}_\bullet \cap K_\Sigma). \tag{70}$$

There are four sources of loss. The subspaces K_Σ and K_{12} control, respectively, overlap of the projected spans and rank loss in the cross equations. The projection kernels have codimension at least $n - T$; they measure tuples omitted when compatibility condition (i) is imposed. Finally, the radicals of the alternating forms are controlled by

$$\operatorname{rank}\Omega_1,\ \operatorname{rank}\Omega_2 \geq \operatorname{rank}\Omega - 4T.$$

The two subspace sources and the radical source contribute compatible pairs outside the main term. The projection kernels instead enter only when the main term is compared with the full one-matrix counts.

Subspace-deficit estimate. Let $K \subseteq V_1$ have codimension γ. Consider compatible pairs for which the first tuple span avoids $\operatorname{rad}(\Omega_1)$ but meets K nontrivially. Their contribution to $\mathcal{W}_{u,v}$ is

$$O\Big(2^{CD_*^2 - \gamma} m_u(\Omega) m_v(\Omega)\Big). \tag{71}$$

To prove (71), put

$$t := \dim(\operatorname{span} \mathbf{w}_\bullet \cap K) \geq 1, \qquad g := \dim(\operatorname{span} \mathbf{w}_\bullet \cap K_{12}), \qquad s := \dim(\mathcal{X} \cap \mathcal{Y}).$$

There are three factors. First, Lemma 4.17 bounds the number of first tuples, relative to the radical-free isotropic count, by

$$2^{-t\gamma + O(D_*^2)}.$$

The hypotheses of that lemma follow from $\operatorname{rank}\Omega_1 \geq \operatorname{rank}\Omega - 4T \geq 2u + 2$ and $n \geq h \geq 2u + 2$; Lemma 2.13 then compares the radical-free count with the baseline isotropic count within a factor $2^{O(D_*^2)}$.

Second, let

$$L := \{\mathbf{y} \in V_2 : B_{12}(\mathbf{w}_i, \mathbf{y}) = 0 \text{ for every } i\}.$$

The map $\operatorname{span} \mathbf{w}_\bullet \to V_2^\vee$ induced by B_{12} has kernel $\operatorname{span} \mathbf{w}_\bullet \cap K_{12}$, of dimension g; hence $\dim L = n - u + g$. Put $G := L \cap \psi_2^{-1}(\mathcal{X})$, so that $\dim G \leq u + T$. Compatibility makes ψ_2 injective on the second tuple span, and therefore

$$\dim(\operatorname{span} \mathbf{w}'_\bullet \cap G) = \dim(\mathcal{X} \cap \mathcal{Y}) = s.$$

Lemma 2.12, applied in L with fall-in parameter s, and with the second-family isotropy condition discarded for an upper bound, gives

$$2^{gv - s(n - 2u + g - T - v) + O(D_*^2)}$$

relative to the isotropic second-tuple count in an $(n-u)$-dimensional space. Here gv records the enlargement of L; the loss from counting unrestricted rather than isotropic tuples is $2^{O(v^2)}$, uniformly in the other parameters, and is absorbed in $2^{O(D_*^2)}$. Indeed,

$$\operatorname{rank}(\Omega_2|_L) \geq \operatorname{rank}\Omega - 4T - 2u \geq 2v + 2,$$

so Lemma 2.13 applies uniformly.

Third, the compatible-pair weight uses $k = u + v - s$, rather than $u + v$. The resulting ratio of weights is exactly

$$2^{\,ns - s(u+v) + \binom{s+1}{2}}.$$

Multiplying the three factors gives, apart from $O(D_*^2)$,

$$-t\gamma + gv - s(n - 2u + g - T - v) + ns - s(u+v) + \binom{s+1}{2} = -t\gamma + gv + s(u - g + T) + \binom{s+1}{2}.$$

This cancellation of every term linear in n is the point of the compatible-pair weight. Since $t \geq 1$ and $g, s, u, v \leq D+1$, the remaining terms contribute $O(D_*^2)$, leaving $2^{-\gamma + O(D_*^2)}$. Restoring the two baseline isotropic counts, the disjoint-span weight, and the $|\operatorname{GL}_u||\operatorname{GL}_v|$ normalization multiplies this relative bound by $m_u(\Omega)m_v(\Omega)$. Summing over t, g, s proves (71).

The disjoint-span main term. Start with the Ω_1-isotropic first tuples. Lemma 2.13, together with (P2), gives their target count and bounds the tuples whose spans meet $\operatorname{rad}(\Omega_1)$ by $2^{CD_*^2 - \operatorname{rank}\Omega}$ on the relative scale. From the remaining radical-free tuples, discard those whose spans meet

$$K_\Sigma, \qquad K_{12}, \qquad \ker\psi_1.$$

Lemma 4.17, summed over positive intersection dimensions, bounds the three discarded proportions by

$$2^{CD_*^2 - h}, \qquad 2^{CD_*^2 - \operatorname{rank} B_{12}}, \qquad 2^{CD_*^2 - n},$$

respectively. Here $\operatorname{codim}\ker\psi_1 \geq n - T$, and the losses $T + b$ are absorbed in CD_*^2. Although $\ker\psi_1 \subseteq K_\Sigma$, we record its stronger codimension separately because it contributes to the 2^{-n} term. Hence the number of remaining first tuples is

$$|\operatorname{GL}_u| m_u(\Omega) 2^{nu - \binom{u}{2}} \left(1 + O(2^{CD_*^2}\varepsilon_*)\right). \tag{72}$$

For these tuples the cross map has rank exactly u, so the second tuples lie in a subspace $L \subseteq V_2$ of dimension $n - u$. The restriction of Ω_2 to L loses rank by at most $2u$. More explicitly,

$$\operatorname{rank}(\Omega_2|_L) \geq \operatorname{rank}\Omega - 4T - 2u \geq 2v + 2,$$

where the last inequality follows from $\operatorname{rank}\Omega \geq 4D + 8T + 10$ and $u, v \leq D+1$. Thus Lemma 2.13 applies to its isotropic v-tuples.

Discard the second tuples whose spans meet $\operatorname{rad}(\Omega_2|_L)$, $\ker\psi_2$, or $L \cap \psi_2^{-1}(\mathcal{X})$. The first discarded proportion is $2^{CD_*^2 - \operatorname{rank}\Omega}$ by Lemma 2.13. The latter two subspaces have codimension in L at least $n - u - T$ and $n - 2u - T$, respectively, so Lemma 4.17 bounds both discarded proportions by $2^{CD_*^2 - n}$. Every remaining second tuple has injective projection and $\mathcal{X} \cap \mathcal{Y} = 0$; property (R2) now makes $\mathcal{R}_{u,v}$ automatic. The number of such tuples is

$$|\operatorname{GL}_v| m_v(\Omega) 2^{(n-u)v - \binom{v}{2}} \left(1 + O(2^{CD_*^2}\varepsilon_*)\right). \tag{73}$$

The bounded rank losses guarantee that all restricted alternating forms are nonzero, so their target moments equal $1/|\operatorname{GL}_j|$.

Every tuple pair retained above has disjoint projected spans, and therefore $k = u + v$. The exponent identity

$$\left(nu - \binom{u}{2}\right) + \left((n-u)v - \binom{v}{2}\right) - \left(n(u+v) - \binom{u+v}{2}\right) = 0$$

shows that (72) and (73), after division by $|\operatorname{GL}_u||\operatorname{GL}_v|$, make the total contribution of these disjoint-span pairs

$$m_u(\Omega)m_v(\Omega) + O\Big(2^{CD_*^2}\theta_u\theta_v\varepsilon_*\Big). \tag{74}$$

We used here $m_j(\Omega)/\theta_j = 2^{-\binom{j}{2}}/\eta_j(2)$, so the relative radical error in Lemma 2.13 is still bounded by the displayed quadratic factor on the $\theta_u\theta_v$ scale.

Compatible pairs outside the main term. By (71), compatible pairs whose first span meets K_Σ or K_{12} contribute at most

$$O\Big(2^{CD_*^2}\theta_u\theta_v(2^{-h} + 2^{-\operatorname{rank} B_{12}})\Big).$$

Pairs whose first span meets $\operatorname{rad}(\Omega_1)$, or whose second span meets the radical of the relevant restriction of Ω_2, contribute

$$O\Big(2^{CD_*^2-\operatorname{rank}\Omega}\theta_u\theta_v\Big)$$

by Lemma 2.13; the possible rank loss $2u+4T$ is absorbed in CD_*^2.

These loci exhaust the compatible pairs outside the disjoint-span contribution in (74). Indeed, compatibility already imposes injectivity of both projections. If a compatible pair avoids the radical loci and K_{12}, then the cross-solution space used above is defined; if it also avoids K_Σ, equation (70) gives $\mathcal{X}\cap\mathcal{Y} = 0$, so the pair belongs to the disjoint-span main term. Thus the projection kernels create the 2^{-n} error only through the tuples omitted from the product count; they are not an additional class of compatible pairs.

Adding the compatible exceptional contribution to (74) proves (69) after enlarging the absolute constant c_*. □

Proof of Proposition 4.15. Let $\boldsymbol{\beta}$ lie in the good event with the threshold H_D of the proposition. Proposition 4.16 identifies the conditional mixed moment with the compatible-pair sum (66). We apply Lemma 4.18 with $T = b_2$ and the cross form B_{12} of Proposition 4.16. Indeed,

$$\operatorname{rank} B_{12} \ge \operatorname{rank}\Gamma - 2b \ge 4D + 2b_2 + 10,$$

because $b \le b_2 + 1$ and $b_2 \ge 1$. The support hypothesis follows from $h \ge H_D$, while the assumed lower bound for $\operatorname{rank}\Omega$ supplies the nonzero-defect hypothesis. Hence (69) applies. Since b_2 and b are bounded in terms of d_0, its quadratic prefactor is at most $2^{c_{\mathrm{mix}}(D+b_2+b+1)^2}$ after enlarging c_{mix}. Finally, $2^{-\operatorname{rank} B_{12}} \le 2^{2b}2^{-\operatorname{rank}\Gamma}$, and the fixed factor is absorbed in the same prefactor. This is (65). □

We can now verify the three hypotheses of case (II) of Theorem 4.12 for the bordered pair, in the order in which that case uses them.

Proof of Theorem 4.13. Fix a sign vector with $|\mathbf{u}_{-1}| \ge \alpha r$, and put $n = r-1$. By (36),

$$\operatorname{rank}\Omega \ge |\mathbf{u}_{-1}^\circ| - 1 \ge \alpha r - 2.$$

After enlarging the final constant, we may assume that r exceeds a constant depending only on (α, d_0). Then $\Omega \ne 0$ and Lemma 4.5, together with the inversion computed in the proof of Theorem 4.6, gives

$$m_u(\Omega) = \frac{1}{|\operatorname{GL}_u|} \ll 2^{-u^2}, \qquad \mathcal{G}^{-1}(m_u(\Omega))_{u\ge 0} = \pi_{\mathrm{CL}}.$$

Thus the target array $m_{u,v} = m_u(\Omega)m_v(\Omega)$ factors and decays as required in case (II), and its inverse is $\pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}$.

Let $b_{\max} = \omega(d_0) + 1$. Choose $\beta_0 > 0$ sufficiently small and, for $0 \le D \le \beta_0\sqrt{n}$, put

$$\mathcal{E}_{n,D} := \mathcal{E}_{\mathrm{good}}(H_D), \qquad \mathrm{Err}_{n,D} := 2^{-h} + 2^{-\operatorname{rank}\Gamma} + 2^{-n} + 2^{-\operatorname{rank}\Omega},$$

with H_D as in Proposition 4.15. Both are measurable with respect to the border data $\boldsymbol{\beta}$. For all sufficiently large n, uniformly in $b_2 \le b_{\max}$,

$$H_D \le \alpha n/4, \qquad \operatorname{rank}\Omega \ge 4D + 8b_2 + 10.$$

Proposition 4.15 therefore supplies part (II.a) of Theorem 4.12, with a prefactor of the required form $2^{c_{\text{loss}}(D+1)^2}$, for a constant depending only on d_0 and α.

For part (II.b), since $h \sim \operatorname{Bin}(n, \frac{1}{2})$,

$$\mathbb{E}[2^{-h}] = \left(\frac{3}{4}\right)^n.$$

Moreover,

$$\operatorname{rank}\Gamma \ge \#\{i \le n : u_{d_0,i} \ne u_{-1,i}\} - 1,$$

where the count is binomial with parameters $(n, \frac{1}{2})$; hence

$$\mathbb{E}[2^{-\operatorname{rank}\Gamma}] \le 2\left(\frac{3}{4}\right)^n.$$

Finally $2^{-n} \le (3/4)^n$ and $2^{-\operatorname{rank}\Omega} \le 4\,2^{-\alpha n}$. Thus

$$\mathbb{E}[\mathrm{Err}_{n,D}\mathbb{1}_{\mathcal{E}_{n,D}}] \le C q_{\text{err}}^n, \qquad q_{\text{err}} := \max\{3/4, 2^{-\alpha}\} < 1,$$

uniformly for $D \le \beta_0\sqrt{n}$. Part (II.c) is the bound $\mathbb{P}(\mathcal{E}_{\text{good}}(H_D)^c) \le e^{-\delta n}$ of Lemma 4.14, which applies because $H_D \le n/4$. Case (II) of Theorem 4.12 now yields

$$\left\|\operatorname{Law}(Z_1, Z_2) - \pi_{\text{CL}} \otimes \pi_{\text{CL}}\right\|_1 \le C\eta^n$$

for some $\eta < 1$. The constants are uniform in the finitely many strata, in $b_2 \le b_{\max}$, and in the sign vector over $|\mathbf{u}_{-1}| \ge \alpha r$. Since $n = r - 1$, this proves Theorem 4.13. ☐

4.5. **Unbordered zero- and high-rank-defect cases.** The zero-defect case has the symmetric rather than the Cohen–Lenstra–Gerth limiting distribution and admits a direct mixed-moment calculation. A high-rank symmetry defect instead uses Lemma 4.18.

Lemma 4.19. *Let $A \sim \operatorname{Unif}(\operatorname{Sym}_n(0))$, fix $\mathbf{v} \in \mathbb{F}_2^n$, put $S = \operatorname{supp}\mathbf{v}$ and $h = |S|$, and set*

$$X = \operatorname{corank} A, \qquad Y = \operatorname{corank}(A + \operatorname{diag}\mathbf{v}).$$

There is an absolute $C > 1$ such that, for $0 \le u, v \le D + 1$ and $h \ge 4D + 8$,

$$\left|\mathbb{E}_A\left[\begin{bmatrix} X \\ u \end{bmatrix}_2 \begin{bmatrix} Y \\ v \end{bmatrix}_2 \mid \mathbf{v}\right] - m_u(0)m_v(0)\right| \le C^D\theta_u\theta_v\left(2^{3D-h} + 2^{3D-n}\right). \tag{75}$$

Proof. Expand the mixed moment over ordered independent kernel tuples and apply Lemma 2.9 to a basis of their joint span. If $\mathcal{X}$ and $\mathcal{Y}$ are the two spans, compatibility says

$$x_S^\top y_S = 0 \quad (x \in \mathcal{X},\ y \in \mathcal{Y}), \qquad \mathcal{X} \cap \mathcal{Y} \subseteq V_{S^c}.$$

Put $t = \dim(\mathcal{X} \cap V_{S^c})$ and $s = \dim(\mathcal{X} \cap \mathcal{Y})$, so $0 \le s \le t$.

For fixed (s, t), Lemma 2.11 gives at most

$$\begin{bmatrix} u \\ t \end{bmatrix}_2 2^{nu - th}$$

ordered first-family tuples. The cross equations cut the second ambient space by $u - t$ dimensions, leaving a space of dimension $n - u + t$. The common span is contained in the t-dimensional space $\mathcal{X} \cap V_{S^c}$; choosing its s-dimensional position and applying Lemma 2.12 bounds the second-family tuples by

$$C^D\, 2^{s(t-s)}\, 2^{(n-u)v}\, 2^{tv}\, 2^{-s(n-u-v)}.$$

Finally, the joint span has dimension $u+v-s$, and Lemma 2.9 assigns the probability

$$2^{-n(u+v-s)+\binom{u+v-s}{2}}.$$

For $s=t=0$, the exact products in the first two counts give, before division by $|\operatorname{GL}_u||\operatorname{GL}_v|$,

$$2^{nu}\,2^{(n-u)v}\,2^{-n(u+v)+\binom{u+v}{2}}=2^{\binom{u}{2}+\binom{v}{2}}.$$

This is $m_u(0)m_v(0)$ after normalization. The omitted factors in the exact independent-tuple products contribute $O(C^D 2^{3D-h})+O(C^D 2^{3D-n})$ relative to this main term.

For $t\ge 1$, use $\begin{bmatrix}u\\t\end{bmatrix}_2\le\eta_\infty(2)^{-1}2^{t(u-t)}$ and

$$\binom{u+v-s}{2}-\binom{u+v}{2}=-s(u+v)+\binom{s+1}{2}.$$

After cancellation of every term involving n, the logarithm to base 2 of the contribution relative to the case $s=t=0$ is at most

$$-t(h-u-v)-t^2+s(t-s)+\binom{s+1}{2}\le -t(h-u-v).$$

The last inequality uses $s\le t$, since $s(t-s)+\binom{s+1}{2}\le\binom{t+1}{2}\le t^2$. Summing over $t\ge 1$ and $0\le s\le t$, and using $u,v\le D+1$ and $h\ge 4D+8$, gives the first error in (75); the possible fall-ins caused by the finite joint span give the second. This proves the lemma. $\square$

For a nonzero defect produced by quadratic reciprocity, the required rank tails can be checked directly. Combining those estimates with Lemma 4.18 gives the following unconditional consequence.

Theorem 4.20. *Let $\mathbf{u}_{-1}\in\mathbb{F}_2^n$, put*

$$\Omega:=\mathbf{u}_{-1}\mathbf{u}_{-1}^\top+\operatorname{diag}(\mathbf{u}_{-1}),$$

and take $A\sim\operatorname{Unif}(\operatorname{Sym}_n(\Omega))$ and $\mathbf{v}\sim\operatorname{Unif}(\mathbb{F}_2^n)$ independently. Set

$$X:=\operatorname{corank}A,\qquad Y:=\operatorname{corank}\big(A+\operatorname{diag}(\mathbf{v})\big).$$

Then the following estimates hold.

1. *If $\mathbf{u}_{-1}=\mathbf{0}$, there are $C_0<\infty$ and $\eta_0\in(0,1)$ such that*
$$\big\|\operatorname{Law}(X,Y)-\pi_{\mathrm{sym}}\otimes\pi_{\mathrm{sym}}\big\|_1\le C_0\eta_0^n.$$
2. *For every $\alpha>0$, there are $C_\alpha<\infty$ and $\eta_\alpha\in(0,1)$ such that, uniformly over $\mathbf{u}_{-1}$ satisfying $\operatorname{rank}\Omega\ge\alpha n$,*
$$\big\|\operatorname{Law}(X,Y)-\pi_{\mathrm{CL}}\otimes\pi_{\mathrm{CL}}\big\|_1\le C_\alpha\eta_\alpha^n.$$

More generally, the same proof applies to a sequence $\Omega_n\in\operatorname{Alt}_n$ for which either $\Omega_n=0$ for all n, or $\operatorname{rank}\Omega_n\ge\alpha n$, provided that $\mathbf{v}_n\sim\operatorname{Unif}(\mathbb{F}_2^n)$ remains independent of the core and there are $c,\delta>0$ such that

$$\mathbb{P}\big(\operatorname{rank}(\operatorname{diag}(\mathbf{v}_n)+\Omega_n)<cn\big)\le e^{-\delta n}.$$

Proof. Condition on $\mathbf{v}$, put $S=\operatorname{supp}\mathbf{v}$, $h=|S|$, and $\Gamma=\operatorname{diag}(\mathbf{v})+\Omega$.

Suppose first that $\mathbf{u}_{-1}=0$. For $0\le D\le\beta_0 n$, with $\beta_0>0$ sufficiently small, take

$$\mathcal{E}_{n,D}=\{h\ge 4D+8\}.$$

Lemma 4.19 supplies part (I.b) of Theorem 4.12, with $\operatorname{Err}_{n,D}=2^{-h}+2^{-n}$. The conditional marginal bounds follow from Lemma 4.5, since diagonal translation preserves the symmetric fibre. Now $h\sim\operatorname{Bin}(n,\frac12)$, $\mathbb{E}[2^{-h}]=(3/4)^n$, and Hoeffding's inequality gives the bad-event bound. Case (I) of Theorem 4.12 therefore gives

$$\big\|\operatorname{Law}(X,Y)-\pi_{\mathrm{sym}}\otimes\pi_{\mathrm{sym}}\big\|_1\le C_0\eta_0^n$$

for some $C_0 < \infty$ and $\eta_0 \in (0,1)$.

Suppose next that $\operatorname{rank}\Omega \geq \alpha n$. In Lemma 4.18 take

$$V_1 = V_2 = \mathbb{F}_2^n, \qquad \psi_1 = \psi_2 = \mathrm{id}, \qquad \Sigma = 0, \qquad T = b = 0,$$

with $\Omega_1 = \Omega_2 = \Omega$ and cross form

$$B_{12}(\mathbf{x}, \mathbf{y}) := \mathbf{x}^\top \Gamma \mathbf{y},$$

and consistency predicate $\mathcal{R}_{u,v} : \mathcal{X} \cap \mathcal{Y} \subseteq V_{S^c}$. Indeed, subtracting the two kernel equations on a common vector gives $\operatorname{diag}(\mathbf{v})\mathbf{z} = 0$, which is precisely this condition. Thus (R1) and (R2) of the weighted-sum lemma hold. For $0 \leq D \leq \beta_0\sqrt{n}$, take

$$\mathcal{E}_{n,D} = \{h \geq 4D + 10\} \cap \{\operatorname{rank}\Gamma \geq 4D + 10\},$$

where β_0 is fixed and n is sufficiently large. The weighted-sum lemma supplies part (II.a) of Theorem 4.12, with

$$\mathrm{Err}_{n,D} = 2^{-h} + 2^{-\operatorname{rank}\Gamma} + 2^{-n} + 2^{-\operatorname{rank}\Omega}.$$

The variable h is $\mathrm{Bin}(n, 1/2)$. After a simultaneous coordinate permutation write $\mathbf{u}_{-1} = (\mathbf{1}_\kappa, \mathbf{0})$. If $h_1 = |\operatorname{supp}\mathbf{v}_1|$ and $h_0 = |\operatorname{supp}\mathbf{v}_0|$, then

$$\operatorname{rank}\Gamma \geq (\kappa - h_1) + h_0 - 1,$$

and the right side plus one has the distribution $\mathrm{Bin}(n, 1/2)$. Hoeffding's inequality gives the bad-event bound, and

$$\mathbb{E}[2^{-h}] = \left(\frac{3}{4}\right)^n, \qquad \mathbb{E}[2^{-\operatorname{rank}\Gamma}] \leq 2\left(\frac{3}{4}\right)^n.$$

Together with $2^{-\operatorname{rank}\Omega} \leq 2^{-\alpha n}$, this verifies the remaining hypotheses of the quadratic-loss criterion. The target moments are $1/|\operatorname{GL}_u|$, so the limit is $\pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}$.

For the final generalization, the assumed lower-tail bound gives

$$\mathbb{E}\big[2^{-\operatorname{rank}(\operatorname{diag}(\mathbf{v}_n) + \Omega_n)}\big] \leq e^{-\delta n} + 2^{-cn}.$$

Together with the binomial estimates for $|\operatorname{supp}\mathbf{v}_n|$ and, in the nonzero case, the assumed linear lower bound for $\operatorname{rank}\Omega_n$, this supplies exactly the averaged error and bad-event bounds used above. $\square$

Remark 4.21 (A fixed identity shift). Let A_n be uniform in $\operatorname{Sym}_n(0)$. Taking $\mathbf{v} = \mathbf{1}$ in Lemma 4.19 and applying Case (I) of Theorem 4.12 gives constants $C < \infty$ and $\eta \in (0,1)$ such that

$$\big\|\mathrm{Law}\big(\operatorname{corank} A_n, \operatorname{corank}(A_n + I_n)\big) - \pi_{\mathrm{sym}} \otimes \pi_{\mathrm{sym}}\big\|_1 \leq C\eta^n.$$

Thus the fixed shift I_n already makes the two coranks asymptotically independent.

Remark 4.22 (Extremal sign vectors, prescribed characters and borders). A condition $p \equiv 1 \bmod 4$ forces $\mathbf{u}_{-1} = 0$, so Theorem 4.20 determines the unbordered core pair. More general character restrictions can place the vectors $(\left[\frac{2}{p}\right], \left[\frac{\Delta_1}{p}\right], \ldots, \left[\frac{\Delta_s}{p}\right])$ in proper affine subspaces, in which case $\mathbf{u}_{d_0}$ need not be uniform and the diagonal perturbation may have bounded support.

Theorem 4.13 includes the all-ones sign vector $\mathbf{u}_{-1} = \mathbf{1}$: its defect is $J_n + I_n$, of rank at least $n-1$, and the two bordered coranks are asymptotically independent with common marginal distribution π_{CL}. The symmetric case is different: its unbordered limit is $\pi_{\mathrm{sym}} \otimes \pi_{\mathrm{sym}}$, but the arithmetic matrices also contain the fixed borders of Theorem 2.4. When the symmetry defect of the core is zero, the augmented alternating forms created by those borders have bounded rank, so tuples meeting their radicals need not be negligible. Even in width one, the corner can change one marginal away from π_{sym}. For arbitrary fixed width the limiting mixed moments have not been computed here. Thus the full squarefree family uses the range $|\mathbf{u}_{-1}| \geq \alpha r$ of Theorem 3.8; bordered sign vectors with $|\mathbf{u}_{-1}|/r \to 0$, including $\mathbf{u}_{-1} = 0$, require separate analysis.

5. Proofs of the main results

5.1. Asymptotic independence of the two 4-ranks.

Proof of Theorem 1.1. Fix once and for all $\alpha = 1/4$. By Theorem 3.8, with $\rho = \pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}}$, it suffices to estimate the distribution of the bordered corank pair uniformly for

$$|r - \log\log X| < (\log\log X)^{2/3}, \qquad |\mathbf{u}_{-1}| \geq \alpha r.$$

Theorem 4.13, uniformly in $\mathbf{u}_{-1}$ (and hence in the stratum it determines), gives constants $C_{d_0} < \infty$ and $\vartheta \in (0,1)$ such that

$$\left\| \mathrm{Law}_{\mathsf{MatData}_r(\mathbf{u}_{-1})}(\mathbf{Z}) - \pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}} \right\|_1 \leq C_{d_0} \vartheta^r \ll_{d_0} \vartheta^{(\log\log X)/2} \tag{76}$$

for all sufficiently large X.

Substituting (76) into (46) gives

$$\left\| \mu_X^{(d_0)} - \pi_{\mathrm{CL}} \otimes \pi_{\mathrm{CL}} \right\|_1 \ll_{d_0} (\log\log X)^{-c_{\mathrm{tr}}} + \vartheta^{(\log\log X)/2}.$$

Total variation is one half of the ℓ^1 distance, so shrinking the exponent gives (5) for some $c = c(d_0) > 0$. $\square$

5.2. The biquadratic application.

Let $K := \mathbb{Q}(\sqrt{d_0}, \sqrt{-d})$, a biquadratic field containing the three quadratic subfields

$$k_0 = \mathbb{Q}(\sqrt{d_0}) \text{ (real)}, \qquad k_1 = \mathbb{Q}(\sqrt{-d}), \qquad k_2 = \mathbb{Q}(\sqrt{-d_0 d}).$$

The dependence of k_1, k_2 and K on d is suppressed throughout this subsection. For k_0, and only there, Cl_{k_0} denotes the ordinary class group; write h_{k_0} for its order.

All density statements below are relative to the family (1). For quantities A_d and B_d on this family, write

$$A_d \overset{\mathbb{P}}{=} B_d \quad \Longleftrightarrow \quad \mathbb{P}_{d \in \mathcal{F}_{d_0}(X)}(A_d = B_d) \longrightarrow 1 \quad (X \to \infty).$$

The superscript $\mathbb{P}$ has the same meaning over any other relation; in particular, $G_d \overset{\mathbb{P}}{\cong} H_d$ means that the displayed isomorphism exists with probability tending to one. Pointwise identities retain the ordinary equality sign. Since $\#\mathcal{F}_{d_0}(X) \asymp_{d_0} X$, any exceptional set of squarefree parameters of size $o(X)$ remains negligible after restriction to this family. We use this observation for the external density-one inputs below.

5.2.1. *Class-group maps and class-number formulas.*

Let

$$j : \prod_{i=0}^{2} \mathrm{Cl}_{k_i} \longrightarrow \mathrm{Cl}_K, \qquad N : \mathrm{Cl}_K \longrightarrow \prod_{i=0}^{2} \mathrm{Cl}_{k_i}$$

be extension of ideals and the product of the three norm maps, respectively. Extension and norm of ideals give $N \circ j = 2\,\mathrm{id}$ and $j \circ N = 2\,\mathrm{id}$. Kuroda's class number formula gives

$$h_K = \frac{1}{2} q_K h_{k_0} h_{k_1} h_{k_2}, \tag{77}$$

where

$$q_K = [\mathcal{O}_K^\times : \mathcal{O}_{k_0}^\times \mathcal{O}_{k_1}^\times \mathcal{O}_{k_2}^\times]$$

is the Hasse unit index. This is the $k = \mathbb{Q}$, imaginary case printed immediately after [25, equation (1.5)]. We also use the ambiguous class number formula for a cyclic extension L/k with Galois group G:

$$|\,\mathrm{Cl}_L^G\,| = \frac{|\,\mathrm{Cl}_k\,| \prod_{v \leq \infty} e_v}{[L:k]\,[\mathcal{O}_k^\times : \mathcal{O}_k^\times \cap N_{L/k}(L^\times)]}. \tag{78}$$

Lemma 5.1. *Let M/k be a fixed abelian extension and let $m \geq 1$. For a density-one set of squarefree integers d, every element of $\mathrm{Gal}(M/k)$ occurs as $\mathrm{Art}_{M/k}(\mathfrak{p})$ for at least m distinct primes $\mathfrak{p} \mid d$ of k.*

This is [16, Corollary 2.3].

5.2.2. *The Hasse unit index.* The field K is totally imaginary with unique real quadratic subfield k_0. Write $E_F := \mathcal{O}_F^\times$, let ε_0 be the fundamental unit of k_0, and let μ_K be the group of roots of unity of K.

The totally imaginary signature makes the unit index particularly simple. First discard the finitely many d for which $\mu_K \neq \{\pm 1\}$. Indeed, extra roots of unity force K to contain $\mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$; since the only imaginary quadratic subfields of K are k_1 and k_2, this determines d up to finitely many possibilities. This omission does not affect any density-one statement. For the remaining d, the unit group E_K has $\mathbb{Z}$-rank one and

$$E_{k_1} = E_{k_2} = \mu_K = \{\pm 1\}, \qquad E_{k_0}E_{k_1}E_{k_2} = \langle \mu_K, \varepsilon_0 \rangle.$$

For every $\varepsilon \in E_K$,

$$\prod_{i=0}^{2} N_{K/k_i}(\varepsilon) = N_{K/\mathbb{Q}}(\varepsilon)\varepsilon^2,$$

and therefore $E_K^2 \subseteq E_{k_0}E_{k_1}E_{k_2}$. Thus

$$E_K = \langle \mu_K, \varepsilon_0 \rangle \quad \text{or} \quad E_K = \langle \mu_K, \sqrt{u\varepsilon_0} \rangle \quad \text{for some } u \in \mu_K,$$

so $q_K \in \{1, 2\}$, with $q_K = 2$ precisely in the second case.

Theorem 5.2. *Let $d_0 > 1$ be squarefree. Then*

$$q_K \stackrel{\mathbb{P}}{=} 1.$$

Proof. Suppose $q_K = 2$. Then, for some $u \in \mu_K$, $\eta := \sqrt{u\varepsilon_0}$ belongs to $E_K \setminus k_0$. Hence $K = k_0\big(\sqrt{u\varepsilon_0}\big)$, and since $K = k_0(\sqrt{-d})$ we get $-d \equiv u\varepsilon_0$ modulo $(k_0^\times)^2$. Complex conjugation sends u to u^{-1}, and hence

$$N_{K/k_0}(\eta)^2 = N_{K/k_0}(u\varepsilon_0) = \varepsilon_0^2.$$

Therefore $N_{K/k_0}(\eta) = \pm\varepsilon_0$, so either ε_0 or $-\varepsilon_0$ is a norm from K.

Now K/k_0 is the quadratic extension cut out by $-d$, so a prime $\mathfrak{p}$ of k_0 with $\mathfrak{p} \mid d$ is ramified in K/k_0, and for an element of $k_0^\times$ that is a local norm everywhere one has, at every $\mathfrak{p} \nmid 2\infty$ unramified in $k_0(\sqrt{\pm\varepsilon_0})/k_0$,

$$\left(\frac{\pm\varepsilon_0}{\mathfrak{p}}\right) = 1, \qquad \text{i.e.} \qquad \mathrm{Art}_{k_0(\sqrt{\pm\varepsilon_0})/k_0}(\mathfrak{p}) = 1.$$

Apply Lemma 5.1 with $k = k_0$, with the fixed abelian extension $M = k_0\big(\sqrt{\varepsilon_0}\big) \cdot k_0\big(\sqrt{-\varepsilon_0}\big)$, which is independent of d, and with $m = 1$. For a density-one set of d, every element of $\mathrm{Gal}(M/k_0)$ occurs at a prime $\mathfrak{p} \mid d$. Choosing one that is nontrivial on both quadratic subextensions shows that neither ε_0 nor $-\varepsilon_0$ is a norm from K, contradicting $q_K = 2$. $\square$

5.2.3. *The direct-sum defect.* Since $j \circ N = 2 \cdot \mathrm{id}$ on Cl_K, one has $2\,\mathrm{Cl}_K \subseteq \mathrm{im}\, j$. Multiplying once more by 2 and passing to primary parts gives

$$4\,\mathrm{Cl}_K[2^\infty] \subseteq j\Big(\prod_i 2\,\mathrm{Cl}_{k_i}[2^\infty]\Big), \qquad \big|4\,\mathrm{Cl}_K[2^\infty]\big| \leq \prod_i \big|2\,\mathrm{Cl}_{k_i}[2^\infty]\big|.$$

Indeed, let $x \in 4\,\mathrm{Cl}_K[2^\infty]$ and write $x = j(y)$ with $y \in \prod_i 2\,\mathrm{Cl}_{k_i}$. Split $y = y_2 + y'$ into its 2-primary and odd parts. Then $j(y_2)$ is 2-primary and $j(y')$ has odd order, and $x = j(y_2) + j(y')$ is 2-primary, so $j(y') = 0$ and $x = j(y_2)$ with $y_2 \in \prod_i 2\,\mathrm{Cl}_{k_i}[2^\infty]$, the splitting into primary parts being compatible with multiplication by 2.

Definition 5.3. The preceding inclusion allows us to define the *direct-sum defect* by

$$\lambda_K := \log_2 \frac{\prod_{i=0}^{2} \left|2\operatorname{Cl}_{k_i}[2^\infty]\right|}{\left|4\operatorname{Cl}_K[2^\infty]\right|} \;=\; \sum_{i=0}^{2}\sum_{n\geq 2} \operatorname{rk}_{2^n}\operatorname{Cl}_{k_i} - \sum_{n\geq 3}\operatorname{rk}_{2^n}\operatorname{Cl}_K \;\geq\; 0.$$

Abbreviate $S := \prod_{i=0}^{2} 2\operatorname{Cl}_{k_i}[2^\infty]$ and $Q := 4\operatorname{Cl}_K[2^\infty]$, both finite. We have $Q \subseteq j(S)$ and trivially $|j(S)| \leq |S|$, so

$$|Q| \;\leq\; |j(S)| \;\leq\; |S|.$$

Thus $\lambda_K = \log_2(|S|/|Q|) \geq 0$, and $\lambda_K = 0$ holds if and only if both inequalities are equalities. Equality on the right is equivalent to the injectivity of $j|_S$, while equality on the left, given $Q \subseteq j(S)$, is equivalent to $Q = j(S)$. Hence $\lambda_K = 0$ if and only if j restricts to an isomorphism $j|_S : S \to Q$.

It remains to compute λ_K. The ambiguous class number formula gives the 2-rank of Cl_K; combining it with the 2-part of Kuroda's formula relates λ_K to the 4-rank of Cl_K. The density-one formula of Koymans–Morgan–Smit evaluates that 4-rank independently and leaves only dyadic data.

Let $\Delta_{d_0} := \Delta(d_0)$ be the discriminant of k_0. Thus, for $p \nmid 2d_0$, the additive symbol $[\frac{\Delta_{d_0}}{p}]$ vanishes exactly when p splits in $k_0/\mathbb{Q}$.

Proposition 5.4. *Let $d_0 > 1$ be squarefree with h_{k_0} odd, and let*

$$\omega_+(d) := \#\{p \mid d :\ [\tfrac{\Delta_{d_0}}{p}] = 0\}, \qquad \omega_-(d) := \#\{p \mid d :\ [\tfrac{\Delta_{d_0}}{p}] = 1\}$$

be the numbers of prime divisors of d that split, respectively remain inert, in $k_0/\mathbb{Q}$, and let $\delta = \delta(d) \in \{0,1,2\}$ be the number of primes of k_0 above 2 that ramify in K/k_0. Then

$$\operatorname{rk}_2 \operatorname{Cl}_K \;\overset{\mathbb{P}}{=}\; 2\omega_+(d) + \omega_-(d) + \delta - 1. \tag{79}$$

Proof. Write $G = \operatorname{Gal}(K/k_0) = \langle\sigma\rangle$ and $A = \operatorname{Cl}_K[2^\infty]$. For the quadratic extension K/k_0 one has $\iota_{k_0\to K} \circ N_{K/k_0} = 1 + \sigma$ on Cl_K, where ι denotes extension of ideals. Hence $(1+\sigma)\operatorname{Cl}_K \subseteq \iota(\operatorname{Cl}_{k_0})$. The latter group has odd order, whereas A is a 2-group, and therefore

$$(1+\sigma)A = 0, \qquad \left(\operatorname{Cl}_K^G\right)[2^\infty] = A^G = A[2] = \operatorname{Cl}_K[2].$$

In particular, $|(\operatorname{Cl}_K^G)[2^\infty]| = 2^{\operatorname{rk}_2 \operatorname{Cl}_K}$.

For the unit norm index, the four classes in $E_{k_0}/E_{k_0}^2$ are represented by $1, -1, \varepsilon_0, -\varepsilon_0$. Thus

$$[E_{k_0} : E_{k_0} \cap N_{K/k_0}(K^\times)] \leq 4,$$

with strict inequality only if one of the three nontrivial representatives u_0 is a norm. In that case u_0 is a local norm at every ramified prime $\mathfrak{p} \mid d$, and hence

$$\operatorname{Art}_{k_0(\sqrt{u_0})/k_0}(\mathfrak{p}) = 1 \qquad (\mathfrak{p} \mid d).$$

Apply Lemma 5.1 to the fixed abelian extension

$$M = k_0(\sqrt{-1})\,k_0(\sqrt{\varepsilon_0})\,k_0(\sqrt{-\varepsilon_0}).$$

For each $u_0 \in \{-1, \varepsilon_0, -\varepsilon_0\}$, choose $\sigma_{u_0} \in \operatorname{Gal}(M/k_0)$ whose restriction to $k_0(\sqrt{u_0})$ is nontrivial. For a density-one set of d, the lemma supplies a prime $\mathfrak{p}_{u_0} \mid d$, possibly depending on u_0, with Artin symbol σ_{u_0}. Thus no nontrivial unit class is a norm, and consequently

$$[E_{k_0} : E_{k_0} \cap N_{K/k_0}(K^\times)] \overset{\mathbb{P}}{=} 4. \tag{80}$$

Apply (78) to K/k_0 and count the ramified places.

Infinite places. k_0 is real quadratic, so it has two real places; K is totally imaginary, so both become complex, and each contributes $e_v = 2$: a factor 2^2.

Primes dividing d. Let $p \mid d$. Since $\gcd(d, 2d_0) = 1$, p is unramified in $k_0/\mathbb{Q}$, and it splits or is inert according as $[\frac{\Delta_{d_0}}{p}]$ is 0 or 1; the number of primes of k_0 above p is therefore 2 or 1. Each such prime $\mathfrak{p}$ satisfies $v_{\mathfrak{p}}(-d) = 1$ because d is squarefree and p is unramified in k_0, so $\mathfrak{p}$ ramifies in $K = k_0(\sqrt{-d})$. Together these contribute $2^{\,2\omega_+(d)+\omega_-(d)}$.

Primes above 2. By definition these contribute 2^{δ}; that δ depends only on the 2-adic stratum follows from the local square-class computation in Lemma 5.7 below.

No other place ramifies. Hence $\prod_v e_v = 2^{\,2\omega_+ + \omega_- + \delta + 2}$, and (78), (80), and $[K : k_0] = 2$ give, on a density-one subset,

$$\left|\mathrm{Cl}_K^G\right| = \frac{h_{k_0} \cdot 2^{\,2\omega_+ + \omega_- + \delta + 2}}{2 \cdot 4} = h_{k_0} \cdot 2^{\,2\omega_+ + \omega_- + \delta - 1}.$$

Since h_{k_0} is odd, the 2-primary part of Cl_K^G has order $2^{\,2\omega_+ + \omega_- + \delta - 1}$. The first paragraph of the proof identifies this order with $2^{\mathrm{rk}_2 \mathrm{Cl}_K}$. □

Taking 2-parts in (77) gives the identity on which the comparison of ranks turns:

$$\sum_{n \geq 1} \mathrm{rk}_{2^n} \mathrm{Cl}_K \;=\; \log_2 q_K - 1 + \sum_{i=0}^{2} \sum_{n \geq 1} \mathrm{rk}_{2^n} \mathrm{Cl}_{k_i}\,. \tag{81}$$

We need a negative-twist correction to the real-base formula of Koymans–Morgan–Smit. Recall that $\Delta(n)$ is the discriminant of $\mathbb{Q}(\sqrt{n})$, and let Δ_F be the discriminant of F. For an integer m, write $\omega_{\mathrm{in}}(m)$ for the number of its distinct prime divisors that are inert in $F/\mathbb{Q}$.

Lemma 5.5. *Let $F/\mathbb{Q}$ be real quadratic. For a density-one set of negative squarefree integers n,*

$$\mathrm{rk}_4 \mathrm{Cl}_{F(\sqrt{n})} = \omega_{\mathrm{in}}(\Delta(n)) + \omega(\Delta_F) + \mathrm{rk}_2 \mathrm{Cl}_F - 2. \tag{82}$$

Proof. Let χ_n be the quadratic character of $F(\sqrt{n})/F$, and use the Selmer groups of [16, §3.1]. The kernel of restriction from $H^1(G_F, \mathbb{Z}/2\mathbb{Z})$ to $H^1(G_{F(\sqrt{n})}, \mathbb{Z}/2\mathbb{Z})$ is $\{0, \chi_n\}$. At every real place the local condition defining $\mathrm{Sel}_{\chi_n}(G_F, \mathbb{Z}/2\mathbb{Z})$ is zero. Since $n < 0$, the character χ_n is nontrivial at both real places, and hence

$$\mathrm{Sel}_{\chi_n}(G_F, \mathbb{Z}/2\mathbb{Z}) \cap \{0, \chi_n\} = \{0\}.$$

Thus the last sentence in the proof of [16, Theorem 3.4], which inserts the one-dimensional kernel $\{0, \chi_n\}$, must be omitted in this sign range. On the same generic set used in their proof of Theorem 6.1, the dimension identity is consequently

$$\begin{aligned} \dim_{\mathbb{F}_2} 2\,\mathrm{Sel}(G_{F(\sqrt{n})}, \mathbb{Z}/4\mathbb{Z}) = \;& \dim_{\mathbb{F}_2} \mathrm{Sel}_{\chi_n}(G_F, \mathbb{Z}/2\mathbb{Z}) \\ &+ \dim_{\mathbb{F}_2} \mathrm{Sel}(G_F, \mathbb{Z}/2\mathbb{Z}), \end{aligned}$$

without the final -1 printed there. Proposition 3.10 of [16] is unchanged and, for real F, gives

$$\dim_{\mathbb{F}_2} \mathrm{Sel}_{\chi_n}(G_F, \mathbb{Z}/2\mathbb{Z}) = \omega_{\mathrm{in}}(\Delta(n)) + \omega(\Delta_F) - 2.$$

Their Lemma 3.1 identifies the other two dimensions with $\mathrm{rk}_4 \mathrm{Cl}_{F(\sqrt{n})}$ and $\mathrm{rk}_2 \mathrm{Cl}_F$, proving (82).

Finally, the exceptional sets in the inputs to their proof have size $o(X)$ among squarefree $|n| \leq X$. Their intersection with $n < 0$, and hence with the fixed coprimality subfamily used here, is still $o(X)$. □

Write $\iota_2 = \mathbb{1}_{\{2 \text{ is inert in } k_0/\mathbb{Q}\}}$ and $w = \omega(\Delta_{d_0}) - s \in \{0, 1\}$. Since the odd prime divisors of $\Delta(-d)$ are those of d,

$$\omega_{\mathrm{in}}(\Delta(-d)) = \omega_-(d) + b_1 \iota_2.$$

Proposition 5.6. *Let $d_0 > 1$ be squarefree with h_{k_0} odd. Then*

$$\lambda_K \overset{\mathbb{P}}{=} b_1(\iota_2 - 1) + w - (b_2 - s) + \delta, \tag{83}$$

where b_1, b_2 are the border widths from Section 2.1.2 and δ is as in Proposition 5.4.

Proof. The hypothesis that h_{k_0} is odd gives $\operatorname{rk}_2 \operatorname{Cl}_{k_0} = 0$ identically. On the density-one set supplied by Theorem 5.2 and Proposition 5.4, one also has $q_K = 1$. Split the two sides of (81) into their 2-, 4- and higher-rank contributions, and cancel the higher tails by the definition of λ_K. This gives

$$\operatorname{rk}_4 \operatorname{Cl}_K = \sum_{i=0}^{2} \operatorname{rk}_2 \operatorname{Cl}_{k_i} - \operatorname{rk}_2 \operatorname{Cl}_K - 1 + \lambda_K.$$

Genus theory gives $\operatorname{rk}_2 \operatorname{Cl}_{k_1} = r + b_1 - 1$ and $\operatorname{rk}_2 \operatorname{Cl}_{k_2} = r + b_2 - 1$. Substituting (79) and $r = \omega_+(d) + \omega_-(d)$ yields

$$\operatorname{rk}_4 \operatorname{Cl}_K \overset{\mathbb{P}}{=} \omega_-(d) + b_1 + b_2 - 2 - \delta + \lambda_K. \tag{84}$$

On the common density-one set, apply (82) with $F = k_0$ and $n = -d$, use $\operatorname{rk}_2 \operatorname{Cl}_{k_0} = 0$ and $\omega(\Delta_{d_0}) = s + w$, and compare the result with (84). This gives (83). $\square$

The right side of (83) is determined by the following local calculation at 2.

Lemma 5.7. *Let $M/\mathbb{Q}_2$ be a ramified quadratic extension and let $u \in \mathbb{Z}_2^\times$.*

- (i) *If $M = \mathbb{Q}_2(\sqrt{v})$ with $v \in \mathbb{Z}_2^\times$ (equivalently, $v \equiv 3 \bmod 4$), then $M(\sqrt{u})/M$ is unramified for every $u \in \mathbb{Z}_2^\times$.*
- (ii) *If $M = \mathbb{Q}_2(\sqrt{2v})$ with $v \in \mathbb{Z}_2^\times$, then $M(\sqrt{u})/M$ is ramified precisely when $u \equiv 3 \bmod 4$.*

Proof. The unramified quadratic extension of $\mathbb{Q}_2$ is $\mathbb{Q}_2(\sqrt{5})$, so the unramified quadratic extension of M is $M(\sqrt{5})$; hence for a unit u the extension $M(\sqrt{u})/M$ is unramified (possibly trivial) if and only if u lies in $(M^\times)^2 \cup 5\,(M^\times)^2$. Now a unit u is a square in M if and only if $\mathbb{Q}_2(\sqrt{u}) \subseteq M$, i.e. if and only if $\mathbb{Q}_2(\sqrt{u})$ is $\mathbb{Q}_2$ or M.

In case (i), $M = \mathbb{Q}_2(\sqrt{v})$ with v a unit, so the units that are squares in M are those with $u \equiv 1$ or $u \equiv v$ modulo $(\mathbb{Z}_2^\times)^2$, that is, the classes $\{1, v\}$ in $\mathbb{Z}_2^\times/(\mathbb{Z}_2^\times)^2 = \{1,3,5,7\} \bmod 8$. Adjoining the classes $5 \cdot \{1, v\} = \{5, 5v\}$ and using $v \in \{3, 7\} \bmod 8$, one gets $\{1,3,5,7\}$ in either case: $\{1,3\} \cup \{5,7\}$ for $v \equiv 3$, and $\{1,7\} \cup \{5,3\}$ for $v \equiv 7$. So every unit is covered and no unit gives a ramified extension.

In case (ii), M is generated by the non-unit $\sqrt{2v}$, so $M \neq \mathbb{Q}_2(\sqrt{u})$ for any unit u and the only unit square class in M is $\{1\}$. The unramified class is $\{5\}$, and the remaining classes $\{3, 7\}$, equivalently $u \equiv 3 \bmod 4$, give ramified extensions. $\square$

5.2.4. *The direct-sum decomposition and the 8-rank distribution.*

Proof of Theorem 1.3. The Hasse unit-index assertion is Theorem 5.2. By the characterization following Definition 5.3, it remains to prove $\lambda_K \overset{\mathbb{P}}{=} 0$ for the group assertion. Proposition 5.6 reduces this to showing that

$$E = b_1(\iota_2 - 1) + w - (b_2 - s) + \delta$$

vanishes in every admissible 2-adic stratum.

Throughout, $b_1 = 1$ exactly when κ is even, i.e. when $-d \equiv 3 \bmod 4$. There are three cases according to the behaviour of 2 in k_0.

2 *splits* ($d_0 \equiv 1 \bmod 8$*; then* $e = 0$*,* κ_0 *even,* $w = 0$*,* $\iota_2 = 0$). Formula (23) gives $b_2 - s = b_1$. Both completions are $\mathbb{Q}_2$, and $\mathbb{Q}_2(\sqrt{u})/\mathbb{Q}_2$ is unramified for $u \equiv 1 \bmod 4$ and ramified for $u \equiv 3 \bmod 4$; so $\delta = 2b_1$. Hence $E = -b_1 - b_1 + 2b_1 = 0$.

2 *is inert* ($d_0 \equiv 5 \bmod 8$*; then* $e=0$*,* κ_0 *even,* $w=0$*,* $\iota_2=1$). Again $b_2 - s = b_1$. The completion is the unramified quadratic extension $L/\mathbb{Q}_2$, in which 5 becomes a square, so $L(\sqrt{u})/L$ is unramified for $u \equiv 1, 5 \bmod 8$ and ramified for $u \equiv 3, 7 \bmod 8$; thus $\delta = b_1$. Hence $E = -b_1 + b_1 = 0$.

2 *ramifies.* Then $\iota_2 = 0$ and $w = 1$, and there is one dyadic prime $\mathfrak{p}$ of k_0, so $\delta \in \{0,1\}$ and $E = 1 - b_1 - (b_2 - s) + \delta$.

- $e = 0$ and $d_0 \equiv 3 \bmod 4$ (κ_0 odd): here $k_{0,\mathfrak{p}} = \mathbb{Q}_2(\sqrt{d_0})$ with d_0 a unit, so Lemma 5.7(i) gives $\delta = 0$ for every d. Since κ_0 is odd, $b_2 - s = 1 - b_1$, and hence $E = 1 - b_1 - (1 - b_1) = 0$.
- $e = 1$: here $k_{0,\mathfrak{p}} = \mathbb{Q}_2(\sqrt{d_0})$ with $d_0 = 2q_1 \cdots q_s$, so Lemma 5.7(ii) gives $\delta = 1$ exactly when $-d \equiv 3 \bmod 4$, i.e. $\delta = b_1$. Also $b_2 - s = 1$, so $E = 1 - b_1 - 1 + b_1 = 0$.

This exhausts the three possible behaviours of 2 in k_0. □

Proof of Corollary 1.5. Taking 2-ranks in the isomorphism of Theorem 1.3 gives

$$\operatorname{rk}_8 \operatorname{Cl}_K \overset{\mathbb{P}}{=} \sum_{i=0}^{2} \operatorname{rk}_4 \operatorname{Cl}_{k_i}.$$

Since h_{k_0} is odd, $\operatorname{rk}_4 \operatorname{Cl}_{k_0} = 0$, which proves the first assertion.

For uniform $d \in \mathcal{F}_{d_0}(X)$, put

$$X_d := \operatorname{rk}_8 \operatorname{Cl}_K, \qquad Y_d := r_4(-d) + r_4(-d_0 d).$$

The first assertion couples these variables so that

$$d_{\mathrm{TV}}\big(\mathrm{Law}(X_d), \mathrm{Law}(Y_d)\big) \le \mathbb{P}(X_d \ne Y_d) = o(1).$$

By Theorem 1.1 and contraction of total variation under the sum map, the distribution of Y_d converges in total variation to the convolution of π_{CL} with itself, whose mass at m is

$$\sum_{a+b=m} \pi_{\mathrm{CL}}(a)\pi_{\mathrm{CL}}(b).$$

The triangle inequality gives the same convergence for X_d. □

Remark 5.8 (Related biquadratic results). Prior results for biquadratic fields include the Dirichlet setting of Fouvry–Koymans and Fouvry–Koymans–Pagano [6, 9], and the density-one formula of Koymans–Morgan–Smit for $F(\sqrt{n})$ with F quadratic [16]. In the real-base, negative-twist correction used above, the archimedean local condition changes the displayed constant by one. Arithmetic statistics of the Hasse unit index in other biquadratic families were studied by Chan–Milovic, Milovic, and Koymans–Pagano [2, 27, 18]. Those works concern different signatures, a different fixed quadratic subfield, or prime-parametrized families.

The 2-torsion of biquadratic and, more generally, multiquadratic fields is treated by Fröhlich and by Koymans–Pagano [10, 14]. The 2-rank of imaginary bicyclic biquadratic fields was studied by McCall–Parry–Ranalli [26]. Proposition 5.4 is the direct specialization of the ambiguous class number formula to the present family. The resulting 4-rank identity is likewise a specialization of Kuroda's formula. For $K_n = \mathbb{Q}(\sqrt{n}, \sqrt{-n})$, Fouvry–Koymans [6, equations (3.1) and (3.2)] obtain

$$\operatorname{rk}_4 \operatorname{Cl}_{K_n} = \operatorname{rk}_2 \operatorname{Cl}\big(\mathbb{Q}(\sqrt{n})\big) + \operatorname{rk}_2 \operatorname{Cl}\big(\mathbb{Q}(\sqrt{-n})\big) - \operatorname{rk}_2 \operatorname{Cl}_{K_n} + \log_2 Q(n) - 1$$

under hypotheses that make the 8-ranks and all higher ranks vanish. In the present family with fixed real subfield k_0, the higher-rank tails cannot be discarded a priori. Their difference is recorded by the nonnegative defect λ_K, which is shown to vanish only after the dyadic calculation.

References

[1] S. Chan, P. Koymans, D. Milovic, C. Pagano, *The 8-rank of the narrow class group and the negative Pell equation*, Forum Math. Sigma **10** (2022), e46, 1–46; doi:10.1017/fms.2022.40.

[2] S. Chan, D. Milovic, *Kuroda's formula and arithmetic statistics*, Math. Z. **300** (2022), 1509–1527; doi:10.1007/s00209-021-02823-6.

[3] H. Cohen, H. W. Lenstra, Jr., *Heuristics on class groups of number fields*, in: Number Theory (Noordwijkerhout 1983), Lecture Notes in Math. **1068**, Springer, 1984, 33–62.

[4] H. Cohen, J. Martinet, *Étude heuristique des groupes de classes des corps de nombres*, J. Reine Angew. Math. **404** (1990), 39–76.

[5] P. Erdős, M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. **62** (1940), 738–742.

[6] É. Fouvry, P. Koymans, *On Dirichlet biquadratic fields*, J. Théor. Nombres Bordeaux **34** (2022), no. 3, 637–646; doi:10.5802/jtnb.1220.

[7] É. Fouvry, J. Klüners, *On the 4-rank of class groups of quadratic number fields*, Invent. Math. **167** (2007), 455–513.

[8] É. Fouvry, J. Klüners, *On the Spiegelungssatz for the 4-rank*, Algebra Number Theory **4** (2010), 493–508.

[9] É. Fouvry, P. Koymans, C. Pagano, *On the 4-rank of class groups of Dirichlet biquadratic fields*, J. Inst. Math. Jussieu **21** (2022), 1543–1570; doi:10.1017/S1474748020000651.

[10] A. Fröhlich, *Central Extensions, Galois Groups, and Ideal Class Groups of Number Fields*, Contemp. Math. **24**, Amer. Math. Soc., 1983.

[11] F. Gerth III, *The 4-class ranks of quadratic fields*, Invent. Math. **77** (1984), 489–515.

[12] F. Gerth III, *Extension of conjectures of Cohen and Lenstra*, Expo. Math. **5** (1987), 181–184.

[13] A. Granville, *Prime divisors are Poisson distributed*, Int. J. Number Theory **3** (2007), no. 1, 1–18; erratum, ibid. **3** (2007), no. 4, 649–651.

[14] P. Koymans, C. Pagano, *Higher genus theory*, Int. Math. Res. Not. IMRN 2022, no. 4, 2772–2823; doi:10.1093/imrn/rnaa196.

[15] J. Jung, J. Lee, *Joint distribution of the cokernels of random p-adic matrices II*, Forum Math. **36** (2024), 1119–1145; doi:10.1515/forum-2023-0131.

[16] P. Koymans, A. Morgan, H. Smit, *The 4-rank of class groups of* $K(\sqrt{n})$, arXiv:2101.03407, 2021.

[17] P. Koymans, C. Pagano, *Effective convergence of coranks of random Rédei matrices*, Acta Arith. **212** (2024), 337–358; doi:10.4064/aa230318-4-1.

[18] P. Koymans, C. Pagano, *On Stevenhagen's conjecture*, arXiv:2201.13424, 2022; to appear in Acta Math.

[19] P. Koymans, Y. Liu, *Statistics of bad parts of class groups*, arXiv:2512.22849, 2025.

[20] P. Koymans, *The 16-rank of* $\mathbb{Q}(\sqrt{-p})$, Algebra Number Theory **14** (2020), no. 1, 37–65; doi:10.2140/ant.2020.14.37.

[21] P. Koymans, D. Milovic, *On the 16-rank of class groups of* $\mathbb{Q}(\sqrt{-2p})$ *for primes* $p \equiv 1 \pmod 4$, Int. Math. Res. Not. IMRN 2019, no. 23, 7406–7427; doi:10.1093/imrn/rny010.

[22] P. Koymans, D. Milovic, *Joint distribution of spins*, Duke Math. J. **170** (2021), no. 8, 1723–1755; doi:10.1215/00127094-2020-0068.

[23] J. Lee, *Joint distribution of the cokernels of random p-adic matrices*, Forum Math. **35** (2023), 1005–1020; doi:10.1515/forum-2022-0209.

[24] J. Lee, *Mixed moments and the joint distribution of random groups*, J. Algebra **641** (2024), 49–84; doi:10.1016/j.jalgebra.2023.10.038.

[25] F. Lemmermeyer, *Kuroda's class number formula*, Acta Arith. **66** (1994), 245–260; doi:10.4064/aa-66-3-245-260.

[26] T. M. McCall, C. J. Parry, R. R. Ranalli, *The 2-rank of the class group of imaginary bicyclic biquadratic fields*, Canad. J. Math. **49** (1997), no. 2, 283–300; doi:10.4153/CJM-1997-014-2.

[27] D. Z. Milovic, *On Hasse's unit index*, arXiv:2001.07213, 2020.

[28] C. Pagano, E. Sofos, *4-ranks and the general model for statistics of ray class groups of imaginary quadratic number fields*, arXiv:1710.07587, 2017.

[29] L. Rédei, H. Reichardt, *Die Anzahl der durch vier teilbaren Invarianten der Klassengruppe eines beliebigen quadratischen Zahlkörpers*, J. Reine Angew. Math. **170** (1934), 69–74.

[30] A. Selberg, *Note on a paper by L. G. Sathe*, J. Indian Math. Soc. (N.S.) **18** (1954), 83–87.

[31] A. Smith, 2^∞*-Selmer groups,* 2^∞*-class groups, and Goldfeld's conjecture*, arXiv:1702.02325v2, 7 June 2017.

[32] A. Smith, *The distribution of* ℓ^∞*-Selmer groups in degree* ℓ *twist families I*, J. Amer. Math. Soc. **39** (2026), 1–72; doi:10.1090/jams/1062.

[33] A. Smith, *The distribution of ℓ^∞-Selmer groups in degree ℓ twist families II*, J. Amer. Math. Soc. **39** (2026), 453–514; doi:10.1090/jams/1063.

[34] W. Wang, M. M. Wood, *Moments and interpretations of the Cohen–Lenstra–Martinet heuristics*, Comment. Math. Helv. **96** (2021), 339–387; doi:10.4171/CMH/514.

[35] M. M. Wood, *Random integral matrices and the Cohen–Lenstra heuristics*, Amer. J. Math. **141** (2019), no. 2, 383–398; doi:10.1353/ajm.2019.0008.

[36] Y. Xu, X. Zhu, *The error term in the Cohen–Lenstra heuristic via the random matrix approach*, Sci. China Math. **69** (2026), no. 4, 1011–1032; doi:10.1007/s11425-024-2455-y.

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

Email address: `xwzhu@bimsa.cn`