Genus Distribution of Random q-ary Lattices

Peter Bruin [*2], Léo Ducas [†1,2], and Shane Gibbons [†1,2]

[1]Cryptology Group, CWI, Amsterdam, The Netherlands,
firstname.lastname@cwi.nl
[2]Mathematical Institute, Leiden University, Leiden, The Netherlands,
P.J.Bruin@math.leidenuniv.nl

March 2022

Abstract

The genus is an efficiently computable arithmetic invariant for lattices up to isomorphism. Given the recent proposals of basing cryptography on the lattice isomorphism problem, it is of cryptographic interest to classify relevant families of lattices according to their genus.

We propose such a classification for q-ary lattices, and also study their distribution. In particular, for an odd prime q, we show that random q-ary lattices are mostly concentrated on two genera.

Because the genus is *local*, this also provides information on the distribution for general odd q. The case of q a power of 2 is also studied, although we only achieve a partial classification.

1 Introduction

Motivation The lattice isomorphism problem (LIP) is the problem of finding, assuming it exists, an isometry sending one lattice to another. It plays a role in the construction of special lattices [PP85], as well as the enumeration of perfect quadratic forms and determination of the densest lattice packing of a given dimension [Sch09] via Voronoï's algorithm. It is a problem of interest in complexity theory, and has also been proven to admit statistical zero-knowledge proofs [HR14]. The best known algorithm in the worst-case [HR14] has super-exponential complexity $n^{\Theta(n)}$, though other algorithms are also used in practice [PS97, SHVvW20].

The problem also comes in a decisional variant, namely determining whether such an isometry between two given lattices exists. For this decisional version,

[*]Supported by the Dutch Research Council (NWO/OCW) as part of the Quantum Software Consortium programme (project number 024.003.037)
[†]Supported by ERC Starting Grant 947821 (ARTICULATE)

one can look at invariants, and answer in the negative if these invariant mismatch. While geometric invariants appear to be hard to compute (requiring one to enumerate short vectors), there is also an efficiently computable arithmetic invariant known as the genus of a lattice (or a quadratic form) [CS13]. Indeed, while the lattice isomorphism can be rephrased as the question of equivalence of quadratic form over the integers, the genus provides a coarser but efficiently computable classification, namely, equivalence of quadratic forms over p-adic integers.

In other words, for two lattices to be plausibly hard to distinguish up to isomorphism, the pair of input lattices must be in the same genus; otherwise, the lattices are obviously not isomorphic. The construction of pairs of lattices that are hard to distinguish up to isomorphism has recently become of cryptographic interest [DvW21, BGPSD21]. While [DvW21] proposes a suitable construction of pairs of lattices in the same genus, this construction is somewhat inefficient (involving the doubling of dimension and increasing geometric gaps). A simpler approach is suggested, where one lattice would be a remarkably decodable lattice, and the second lattice would be chosen at random conditioned on it being in the same genus. Hence the question studied in this work: how is the genus of a random lattice distributed?

Contributions To answer the question, one must first settle on a model of random lattices. Following the cryptographic literature [Ajt96], we choose to focus on the case of q-ary lattices, i.e. integer lattices $\Lambda \subset \mathbb{Z}^n$ defined by m independent equations modulo q (thus fixing the determinant to $\det(\Lambda) = q^m$). Section 2.4 remarks that the *locality* of the genus carries over to the modulus q; this means that we can reduce our study to the case $q = p^k$ for an individual prime p. That is, if $q = q_1 q_2$ where q_1 and q_2 are coprime, the genus distribution of q-ary lattice is simply given by the Cartesian product of the distributions of genera of q_1-ary and q_2-ary lattices.

1. Let us start with the simplest case, namely subset sum lattices modulo an odd prime $q = p$, that is when $m = 1$ and $k = 1$. Lemma 3 shows that there are exactly three possible genera among the p^{n-1} choices to define that lattice. Lemma 5 relates the exact count of each genus to an exact formula [Car53]. Asymptotically (growing p and n), two genera each account for almost a half of the lattices, while the third genus only accounts for about $1/p$ of them.

2. Generalizing to $q = p^k$ for $k \geq 1$ but still for subset sum lattices ($m = 1$), the same method also fully determines the genera, and their distributions can be partially described also, as stated in Theorem 1: there are $2k + 1$ genera, two of them accounting for almost half of the lattices each, two more accounting for about a $1/2p$ fraction of the lattices, two more accounting for about a $1/2p^2$ fraction, and so on until the last one, accounting for about a $1/q$ fraction of the lattices.

3. When q is a power of 2, Lemma 6 shows that either two or four genera each account for about $\frac{1}{4}$ or $\frac{1}{8}$ of these lattices, depending on whether $q = 2$ or $q > 2$, respectively. The rest of the genera are usually smaller, but are more difficult to fully classify.

4. Finally, we treat the general case of lattices given by a random full-rank parity check matrix $A \in (\mathbb{Z}/q\mathbb{Z})^{n \times m}$ for odd q and arbitrary $m \leq n$, and obtain a similar answer. That is, we show again in Theorem 2 that two genera account for about half of the lattices each, and that the remaining genera account for about $1/p$ of them. The analysis is however quite different, in particular resorting to approximations. After relating the genus to the quantity $\det(A^T A) \bmod p$, we rewrite $A^T A \bmod p$ as a *random walk* over the additive group of $m \times m$ symmetric matrices modulo p. We then show that it rapidly converges to the uniform distribution using *harmonic analysis* and *Gauss sums*. Then, we invoke a result of Carlitz [Car54, Theorem 3] (see also MacWilliams [Mac69, Theorem 2]) to conclude on the distribution of the determinant.

Conclusions The bottom line for cryptographic applications is that the genus of a q-ary lattice does not carry so much information. For odd q, it provides about 1 bit of Shannon entropy for each odd prime p dividing q.

A more formal conclusion for hardness is that one can efficiently sample random lattices conditioned on being in one of the large genera, simply by rejection sampling. Because this conditioning is mild, it preserves the presumed average-case of decoding random q-ary lattices. This enables the suggestion of [DvW21] for instantiating distinguish LIP with a pair of lattices consisting on one hand of a remarkable q-ary lattice falling in one of the largest genera, and on the other hand of a random q-ary lattice conditioned on having the same genus.

2 Preliminaries

2.1 Notation and Terminology

Let $\mathbb{F}_p$ be the field $\mathbb{Z}/p\mathbb{Z}$. We use $\mathbb{Z}_p$ for the p-adic integers, **not** the ring $\mathbb{Z}/p\mathbb{Z}$. The p-adic integers are the completion of the integers with respect to the p-adic metric $|\cdot|_p : \mathbb{Z} \to \mathbb{R}_{\geq 0}$, given by $|x|_p = p^{-c}$, where p^c is the largest power of p that divides x. Recall that the p-adic integers can be shown to be isomorphic to the set

$$\mathbb{Z}_p \cong \left\{ \sum_{r=0}^{\infty} a_r p^r : 0 \leq a_r < p \right\},$$

with addition and multiplication defined with powers of p 'carrying' as they do in regular addition, and with convergence defined with respect to the p-adic metric. So, intuitively one may think of the p-adic integers this way. It is also worth noting that there is a canonical inclusion $\mathbb{Z} \hookrightarrow \mathbb{Z}_p$, with each integer

being represented in $\mathbb{Z}_p$ by its base-p expansion. Therefore the integers are a subring of the p-adic integers, and in the context of LIP, it is natural to study equivalence over rings that are larger than $\mathbb{Z}$.

Definition 1 (p-part, p'-part, p-adic order)**.** *Let p be prime, and let $\alpha \in \mathbb{Z}_p \supset \mathbb{Z}$. Then α can be written in the form*

$$\alpha = p^s \beta$$

for some $s \in \{0, 1, \ldots\}$ and β coprime to p. The p-part *of α is p^s, while the* p'-part *of α is β. The* p-adic order *of α is s.*

An important characteristic of numbers modulo a prime p that we will use later on is *quadratic residuosity*, defined by the Legendre symbol at p.

Definition 2. *The* Legendre symbol *at an odd prime p, $\left(\frac{\cdot}{p}\right) : \mathbb{Z} \to \{0, \pm 1\}$ is given by*

$$\left(\frac{n}{p}\right) = \begin{cases} 0 & \text{if } p \mid n \\ 1 & \text{if } \exists\ a \in \mathbb{F}_p^* \text{ such that } a^2 = n \mod p \\ -1 & \text{if } \nexists\ a \in \mathbb{F}_p^* \text{ such that } a^2 = n \mod p. \end{cases}$$

The Legendre symbol is not defined at $p = 2$; instead there is an analogous symbol that we require.

Definition 3. *The* Kronecker symbol $\left(\frac{\cdot}{2}\right) : \mathbb{Z} \to \{0, \pm 1\}$ *is given by*

$$\left(\frac{n}{2}\right) := \begin{cases} 0 & \text{if } n \text{ is even,} \\ 1 & \text{if } n \equiv \pm 1 \mod 8, \\ -1 & \text{if } n \equiv \pm 3 \mod 8. \end{cases}$$

2.2 Quadratic Forms, Equivalences and Lattices

Quadratic forms have many equivalent definitions in many contexts. For our purposes, we use the following.

Definition 4. *Let Q be an $n \times n$ matrix over a ring R. The* quadratic form *defined by the matrix Q is the map $q_Q : R^n \to R$ given by*

$$x \mapsto x^T Q x.$$

If $R = \mathbb{R}$, then such a form is called positive definite *if for all non-zero $x \in \mathbb{R}^n$ we have $q_Q(x) > 0$. An* integral *quadratic form is a quadratic form over $\mathbb{Z}$.*

Definition 5. *Let q_1, q_2 be quadratic forms over a ring R. Then q_1 is* equivalent to q_2 over R *if there exists a matrix $H \in GL_n(R)$ such that for all $x \in R^n$,*

$$q_1(x) = q_2(Hx).$$

Given two symmetric matrices Q_1 and Q_2, the corresponding quadratic forms are equivalent over R if and only if there exists $H \in GL_n(R)$ such that

$$Q_1 = H^T Q_2 H.$$

A quadratic form Q can also be interpreted as a degree-2 polynomial in n variables, via

$$\sum_{1 \leq i,j, \leq n} Q_{ij} X_i X_j.$$

With this in mind, the idea of a direct sum of quadratic forms can be established.

Definition 6. *Let q, q' be two quadratic forms with corresponding matrices Q,Q' of dimension m and n respectively. The direct sum $q \oplus q'$ can be interpreted as the $m+n$ dimensional quadratic form formed by the sum*

$$\sum_{1 \leq i,j, \leq m} Q_{ij} X_i X_j + \sum_{1 \leq i,j, \leq n} Q'_{ij} X_{i+m} X_{j+m}.$$

In the matrix setting, we have

$$q \oplus q' = \begin{pmatrix} Q & 0 \\ 0 & Q' \end{pmatrix}$$

A lattice is a discrete additive subgroup of $\mathbb{R}^n$, and q-ary lattices are a family of integer lattice that can be randomly sampled from a finite set.

Definition 7. *A q-ary lattice Λ of dimension n is a lattice such that*

$$q\mathbb{Z}^n \subseteq \Lambda \subseteq \mathbb{Z}^n.$$

A q-ary lattice Λ can be expressed in a number of ways. The most convenient for the aim of this paper is to consider Λ as the kernel of a linear map $A : \mathbb{Z}^n \to (\mathbb{Z}/q\mathbb{Z})^m$. This map can be characterised by a matrix A (or vector if $m = 1$) called a *parity check matrix* (resp. *subset sum*); a name inherited from linear codes over $\mathbb{F}_2$. First we consider q-ary parity check lattices, with a distinction made between a parity check vector and the more familiar parity check matrix.

Definition 8. *Let $q \in \mathbb{N}$, and $v \in (\mathbb{Z}/q\mathbb{Z})^n$. Then the* subset sum *lattice is given by:*

$$\Lambda_q^{\perp}(v) = \{x \in \mathbb{Z}^n : x \cdot v = 0 \mod q\}.$$

If $A \in (\mathbb{Z}/q\mathbb{Z})^{n \times m}$, then the parity check *lattice is given by:*

$$\Lambda_q^{\perp}(A) = \{x \in \mathbb{Z}^n : A^T x = 0 \mod q\}.$$

When classifying the genus of a lattice, we first check its determinant. Therefore, to compare the distribution of genera, we first ensure that all lattices we discuss have the same determinant. We want $\det(\Lambda) = q^m$, which happens if and only if the parity check matrix is full rank modulo each prime divisor of

q (or for every prime divisor p of q, the vector v has at least one element coprime to p). Therefore, we will only consider vectors v and matrices A with this condition. This lets us easily put a basis of the lattice into Hermite Normal Form (HNF), which, as we see later, makes it possible to count the 'size' of each genus.

Definition 9. *An integer lattice with basis B has a* corresponding quadratic form*, whose defining matrix is given by BB^T.*

This matrix BB^T is the *Gram matrix* of the basis B, i.e. the matrix G whose ij-entry is the inner product of the i^{th} and i^{th} basis vectors. Two lattices with bases B, B' that differ by an orthonormal transform $O \in \mathcal{O}_n(\mathbb{R})$ via $B = B'O$ give the same quadratic form

$$BB^T = B'O(B'O)^T = B'OO^TB'^T = B'B'^T.$$

So if we work in the domain of quadratic forms, we may focus only on basis transformations.

In this text we often use 'lattice' and 'quadratic form' interchangeably when talking about the genus. But it is important to remember that while any *integer* lattice has a corresponding *integral* quadratic form, every *integral* quadratic form only has a corresponding *integral* (not necessarily integer) lattice, i.e. all norms and inner products are integers. Consider the integral form X^2+XY+Y^2, which has corresponding Gram matrix

$$G = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix},$$

which does not come from an integer lattice. This example shows that the set of integer lattices is not bijective with the set of quadratic forms.

Quadratic Forms of q-ary lattices Now suppose q is a power of a prime p. Our parity check vectors $v \in (\mathbb{Z}/q\mathbb{Z})^n$ have at least one entry coprime to p. After permuting the coordinates, we may assume the first entry, v_1, is in $(\mathbb{Z}/q\mathbb{Z})^*$. Scaling by $-v_1^{-1}$, we see that without loss of generality we can take v of the form $\begin{pmatrix} -1 \\ v_0 \bmod q \end{pmatrix}$ for some $v_0 \in \mathbb{Z}^{n-1}$ with entries in $\{0, 1, \ldots, q-1\}$. The lattice $\Lambda_q^\perp(v)$ then has the following (row) basis in HNF:

$$B = \begin{pmatrix} q & 0 \\ v_0 & \mathbb{I}_{n-1} \end{pmatrix}.$$

The corresponding Gram matrix is

$$BB^T = \begin{pmatrix} q^2 & qv_0^T \\ qv_0 & v_0v_0^T + \mathbb{I}_{n-1} \end{pmatrix}.$$

We can do the same for a parity check matrix $A \in (\mathbb{Z}/q\mathbb{Z})^{n\times m}$ as follows. If A has full rank modulo p, then after permuting the rows and applying column

operations, we may assume $A = \begin{pmatrix} -\mathbb{I}_m \\ A_0 \bmod q \end{pmatrix}$ for some $A_0 \in \mathbb{Z}^{(n-m)\times m}$ with entries in $\{0, 1, \ldots, q-1\}$. The lattice $\Lambda_q^{\perp}(A)$ has HNF basis

$$B = \begin{pmatrix} q\mathbb{I}_m & 0 \\ A_0 & \mathbb{I}_{n-m} \end{pmatrix}$$

and the corresponding Gram matrix is

$$BB^T = \begin{pmatrix} q^2\mathbb{I}_m & qA_0^T \\ qA_0 & A_0A_0^T + \mathbb{I}_{n-m} \end{pmatrix}.$$

2.3 Genus Symbol

Locally, we may consider the equivalence class of a quadratic form at a single prime p. The *Jordan decomposition* of a quadratic form f at a prime p is described in [CS13, Chapter 15]. For any odd finite prime p, a quadratic form can always be diagonalised over the p-adic integers $\mathbb{Z}_p$ as the direct sum:

$$f = f_1 \oplus pf_p \oplus p^2 f_{p^2} \oplus \ldots, \tag{1}$$

where each f_{p^i} is a quadratic form over the p-adic integers and whose determinant is coprime to p. The Jordan decomposition at -1 (where -1 is the intuitive notation for the infinte prime or ∞) is the decomposition

$$f = f_1 \oplus (-1)f_{-1},$$

where both f_1 and f_{-1} are positive definite. Throughout our discussion, we only consider positive definite quadratic forms which are always equivalent over $\mathbb{R}$. This is because for any full rank basis B, and any $0 \neq x \in \mathbb{R}^n$, the quadratic form BB^T has the condition that $x^TBB^Tx = \|B^Tx\|^2 > 0$. The Jordan decomposition at $p = 2$ is a block diagonalisation. The blocks are either of the form

$$(qx) \text{ or } \begin{pmatrix} qa & qb \\ qb & qc \end{pmatrix}$$

with $x, b, ac - b^2$ coprime to 2, and a, c divisible by 2. This block diagonalisation is a direct sum of quadratic forms

$$f = f_1 \oplus 2f_2 \oplus 4f_4 \oplus \ldots \oplus 2^r f_{2^r} \tag{2}$$

with each f_q coprime to 2.

Definition 10 (Genus)**.** *Two quadratic forms Q_1 and Q_2 lie in the same* genus *if they are equivalent over $\mathbb{R}$ and the p-adic integers $\mathbb{Z}_p$ for all finite primes p.*

A Jordan decomposition has an associated p-adic *symbol*, and any two forms with the same p-adic symbol are equivalent over $\mathbb{Z}_p$ [Cas78, O'M71].

Definition 11. *For $p \neq 2$, the* symbol at p *of a quadratic form with Jordan decomposition*

$$f = f_1 \oplus pf_p \oplus p^2 f_{p^2} \oplus \ldots \oplus p^r f_{p^r}$$

is the sequence

$$1^{\varepsilon_1 n_1} \quad p^{\varepsilon_p n_p} \quad \ldots \quad (p^r)^{\varepsilon_{p^r} n_{p^r}}$$

where $\varepsilon_q = \left(\frac{\det f_q}{p}\right)$ *and* $n_q = \dim f_q$.

Two quadratic forms are equivalent over $\mathbb{Z}_p$ for an odd p if and only if they have the same genus symbol at p [Cas78, O'M71]. The symbol at $p = 2$ is more complicated to define, but we know sufficient and necessary conditions for two forms to be equivalent over $\mathbb{Z}_2$ [CS13, O'M71]. To any component f_s in a quadratic form f over $\mathbb{Z}_2$, one can associate an *oddity* t_s, which is the trace modulo 8 of the diagonalised f_s. We refer to [CS13, Chapter 15, 5.1] for the definition and the properties that we need. Suppose f has Jordan decomposition

$$f = f_1 \oplus 2f_2 \oplus 4f_4 \oplus \ldots \oplus 2^r f_{2^r}.$$

The *sign* ε_s of f_s is the Kronecker symbol $\left(\frac{\det(f_s)}{2}\right) \in \{\pm 1\}$, see Definition 3. The *type* of f_s is I if the matrix representation of f_s has an odd number in its main diagonal, and II otherwise. To such a Jordan decomposition one associates a *genus symbol* depending on the dimension, sign, type and oddity of the f_s.

$$1_{t_s}^{\varepsilon_1 n_1} \quad 2_{t_2}^{\varepsilon_2 n_2} \quad \cdots \quad (2^r)_{t_{2^r}}^{\varepsilon_{2^r} n_{2^r}}$$

A form over $\mathbb{Z}_2$ may have multiple Jordan decompositions with different signs and oddities of the f_s. Still, one may attach a *canonical symbol* to each form in such a way that two forms are equivalent over $\mathbb{Z}_2$ if and only if their canonical symbols agree. This involves grouping parts of the Jordan decomposition together, conditioned on their type. See [CS13, Chapter 15, 7.3–7.6] for a complete description.

2.4 Reduction to the Prime Power Case

The following lemma shows that the genus of a q-ary lattice only depends on local properties at the prime divisors of q.

Lemma 1. *Let q be a positive integer, and let $A, A' \in (\mathbb{Z}/q\mathbb{Z})^{n \times m}$ be parity check matrices. Then for any prime $l \nmid q$, the quadratic forms associated to $\Lambda_q^\perp(A)$ and $\Lambda_q^\perp(A')$ are equivalent over $\mathbb{Z}_l$.*

Proof. Let B be any basis matrix for $\Lambda_q^\perp(A)$. Then B has determinant q^m and is thus invertible over $\mathbb{Z}_l$. Therefore BB^T as a quadratic form can be transformed by $B^{-1} \in GL_n(\mathbb{Z}_l)$ to the identity. The same holds for A', and the claim follows. □

Therefore the genus is independent of any primes not dividing q. Given any q-ary lattice, we next show how to reduce to the case where q is a prime power.

Lemma 2. *Let q be a positive integer, and let p^e be the highest power of a prime p dividing q. Let $A \in (\mathbb{Z}/q\mathbb{Z})^{n\times m}$ be a parity check matrix. Then $\Lambda_q^\perp(A)$ is equivalent to $\Lambda_{p^e}^\perp(A \bmod p^e)$ over $\mathbb{Z}_p$.*

Proof. The lattice $\Lambda_q^\perp(A)$ is contained in $\Lambda_{p^e}^\perp(A \bmod p^e)$, with index prime to p. Given bases of these two lattices, the base change matrix relating the two is therefore in $GL_n(\mathbb{Z}_p)$. This implies the claim. □

Thus, via the Chinese Remainder Theorem, a parity check matrix equation modulo $q = p_1^{e_1} \dots p_r^{e_r}$ can be expressed as multiple separate parity check matrix equations modulo each prime power $p_i^{e_i}$, and vice versa. Combining the above two results, we see that the genus of a q-ary parity check matrix depends only on the symbols at p_i of the respective $p_i^{e_i}$-ary parity check lattices.

3 Subset Sum Lattices

In Section 2.2 we stated that we are assuming our lattice is full rank modulo every prime divisor of q. We may now proceed to prove our results about lattices generated as subset sums of vectors of the form $(1, v_0)$, knowing they apply to all subset sum lattices $\Lambda_q^\perp(v)$ of determinant q.

Section 3.1 considers the simplest case: when q is an odd prime. There is a discussion about the number of genera and the closed form formula counting the exact number of $\Lambda_q^\perp(v)$ whose quadratic forms fall into each genus. The odd prime power case in Section 3.2 can be proven using the same strategy as that in Section 3.1, and the conditions describing the lattices that fall into each genus are similar. Finally, Section 3.3 describes the distribution of the four largest genera when q is a power of 2. This is not a complete classification.

3.1 Odd Prime Modulus

Lemma 3. *Let p be an odd prime, and let $v, u \in \mathbb{F}_p^n$ be non-zero. Then $\Lambda_p^\perp(v)$ and $\Lambda_p^\perp(u)$ are in the same genus if and only if the following Legendre symbols match:*

$$\left(\frac{v \cdot v}{p}\right) = \left(\frac{u \cdot u}{p}\right).$$

However, Lemma 1 tells us that for a q-ary lattice, the only difference in the genus symbol will be at those primes p dividing q. Lemma 3 is about the genus, while Lemma 4 below is specifically about $\mathbb{Z}_p$-equivalence.

Lemma 4. *Let p be an odd prime, and let $v, u \in \mathbb{F}_p^n$ be non-zero. Then $\Lambda_p^\perp(v)$ and $\Lambda_p^\perp(u)$ are equivalent over $\mathbb{Z}_p$ if and only if the following Legendre symbols match:*

$$\left(\frac{v \cdot v}{p}\right) = \left(\frac{u \cdot u}{p}\right).$$

Proof. As noted in Section 2.2, we may assume that $\Lambda_p^\perp(v)$ has a Gram matrix of the form

$$G = \begin{pmatrix} p^2 & pv_0^T \\ pv_0 & v_0v_0^T + \mathbb{I}_{n-1} \end{pmatrix}$$

for some $v_0 \in \mathbb{Z}^{n-1}$. Since we are looking for the Jordan decomposition of this Gram matrix, we need to consider how many terms in its diagonalisation are divisible by p. First recall the elementary result, the "matrix determinant lemma", which gives us

$$(\det(v_0v_0^T + \mathbb{I}_{n-1}) \bmod p) = (\|v_0\|^2 + 1 \bmod p) = v \cdot v \mod p.$$

Second, note that the diagonalisation of $v_0v_0^T + \mathbb{I}_{n-1}$ is

$$\begin{pmatrix} \|v_0\|^2 + 1 & 0 \\ 0 & \mathbb{I}_{n-2} \end{pmatrix}.$$

If $\det(\mathbf{v_0v_0^T} + \mathbb{I}) \neq \mathbf{0}$ mod p: If $M := v_0v_0^T + \mathbb{I}$ has non-zero determinant mod p, then it is invertible over $\mathbb{Z}_p$. Also, v_0 is an eigenvector of M, with eigenvalue $\|v_0\|^2 + 1$. So v_0^T is an eigenvector of M^{-1}, with eigenvalue $\frac{1}{\|v_0\|^2+1}$. This aids with the calculations below.

Now, consider the Gram matrix

$$G = \begin{pmatrix} p^2 & pv_0^T \\ pv_0 & v_0v_0^T + \mathbb{I}_{n-1} \end{pmatrix},$$

and the transformation

$$H = \begin{pmatrix} 1 & 0 \\ -\frac{p}{\|v_0\|^2+1}v_0 & \mathbb{I}_{n-1} \end{pmatrix}. \tag{3}$$

Since $p \nmid \|v_0\|^2 + 1$ by assumption, the matrix H has entries in $\mathbb{Z}_p$ and is invertible. Therefore, G is in the same $\mathbb{Z}_p$-equivalence class as

$$\begin{aligned} G' =& H^T G H \\ =& \begin{pmatrix} \frac{p^2}{\|v_0\|^2+1} & 0 \\ 0 & v_0v_0^T + \mathbb{I}_{n-1} \end{pmatrix}. \end{aligned}$$

The submatrix $v_0v_0^t + \mathbb{I}$ diagonalises to $Diag\{\|v_0\|^2 + 1, 1, \ldots, 1\}$, giving us a Jordan decomposition

$$f_1 \oplus p^2 f_{p^2} = X_1^2 + \ldots X_{n-2}^2 + \left(\|v_0\|^2 + 1\right) X_{n-1}^2 + p^2 \left(\frac{1}{\|v_0\|^2 + 1}\right) X_n^2.$$

The signs of f_1 and f_{p^2} are both equal to the Legendre symbol

$$\left(\frac{\det(f_1)}{p}\right) = \left(\frac{\|v_0\|^2 + 1}{p}\right) = \left(\frac{\det(f_{p^2})}{p}\right)$$

And indeed $\dim f_1 = n - 1, \dim f_{p^2} = 1$, regardless of the residue class. So if $v \cdot v$ and $u \cdot u$ are in the same residue class modulo p, then the corresponding $\Lambda_p^\perp(v)$ and $\Lambda_p^\perp(u)$ are equivalent over $\mathbb{Z}_p$.

If $\det(\mathbf{v_0 v_0^T} + \mathbb{I}) = \mathbf{0} \bmod \mathbf{p}$: Firstly, if $p^2 \nmid \|v_0\|^2 + 1$, the transformation H from Equation 3 remains in $\mathbb{Z}_p$, so the proof goes through as in the $\det(v_0 v_0^T + \mathbb{I}) \neq 0$ case; however, the diagonalised form is

$$G'' = \begin{pmatrix} p\frac{p}{\|v_0\|^2+1} & 0 & 0 \\ 0 & p\frac{\|v_0\|^2+1}{p} & 0 \\ 0 & 0 & \mathbb{I}_{n-2} \end{pmatrix}.$$

The decomposition of this form is then

$$f_1 \oplus p f_p,$$

where f_1 has dimension $n-2$ and sign 1, and f_p has dimension 2 and sign

$$\left(\frac{\det f_p}{p}\right) = 1.$$

Now, in the case when $p^2 \mid \|v_0\|^2 + 1$, we must first apply the transform that maps $(v_1, \ldots, v_i, \ldots, v_{n-1}) \mapsto (v_1, \ldots, v_i + p, \ldots, v_{n-1})$, where $p \nmid v_i$. Such a v_i exists since v_0 is not the zero vector (recall $p \mid \|v_0\|^2 + 1$). This is achieved by first applying the transform

$$\begin{pmatrix} 1 & 0 \\ b & \mathbb{I}_{n-1} \end{pmatrix},$$

where b is the zero column-vector with a 1 in the ith position. We thus have reduced to the previous case, since if $p^2 \mid \|v_0\|^2 + 1$, then p^2 cannot possibly divide $\|v_0\|^2 + 1 + p^2 + 2pv_i$ (recall, $v_i < p$). □

By Lemma 1, the genera of all possible $\Lambda_p^\perp(v)$ with v of fixed length depend only on the primes dividing the modulus q. Lemma 4 therefore tells us that the genus of $\Lambda_p^\perp(v)$ is determined by the residue class of $v \cdot v$ modulo p. This concludes the proof of Lemma 3.

Distribution We immediately see that any $\Lambda_p^\perp(v)$ can fall into one of only three genera, since $\left(\frac{v \cdot v}{p}\right) = -1, 0$ or 1. In this simplest case (random p-ary subset sum lattices for odd p), we have an exact classification of the size of the genera. The following result [Car53] gives an exact closed-form formula for the number of distinct $v \in \mathbb{F}_p^n$ with $v \cdot v = r \mod p$ for any given r.

Lemma 5. *Let p be an odd prime. For any $r \in \mathbb{F}_p$, let S_r be the set of points $v \in \mathbb{F}_p^n$ with $v \cdot v = r$. If n is odd,*

$$|S_r| = \begin{cases} p^{n-1} + \left(\frac{(-1)^{\frac{n-1}{2}} r}{p}\right)\sqrt{p^{n-1}} & \text{if } r \neq 0 \\ p^{n-1} & \text{if } r = 0. \end{cases}$$

Whereas when n is even,

$$|S_r| = \begin{cases} p^{n-1} - \left(\frac{(-1)^{\frac{n}{2}}}{p}\right)\sqrt{p^{n-2}} & \text{if } r \neq 0 \\ p^{n-1} + \left(\frac{(-1)^{\frac{n}{2}}}{p}\right)(p-1)\sqrt{p^{n-2}} & \text{if } r = 0. \end{cases}$$

Since the lattices are partitioned into genera via the value of $\left(\frac{v \cdot v}{p}\right)$, we can count the amount of v's that have $v \cdot v = r \mod p$ via the above lemma, and we are left with the relative 'size' of each genus in this context. If g_v is the genus within which $\Lambda_q^\perp(v)$ lies, then the relative size of g_v is given by

$$|g_v|_{\text{rel}} = \sum_{\substack{r \in \mathbb{F}_p \\ (\frac{r}{p}) = (\frac{v \cdot v}{p})}} |S_r|$$

Note that this is *not* a statement about the number of quadratic forms or equivalence subclasses in g_v, but a count of how many randomly chosen v's have corresponding quadratic forms in g_v. Heuristically, in the asymptotic setting as $n \to \infty$ the proportion of such v's with $v \cdot v = r$ is approximately $1/p$. Euler's Criterion tells us that there are two cosets of equal size in $\mathbb{F}_p^* / \left(\mathbb{F}_p^*\right)^2$. Therefore, with Lemma 3, we know the genera are distributed as follows:

1. The largest two genera (corresponding to those v with $\left(\frac{v \cdot v}{p}\right) = \pm 1$) each account for slightly less than half of this family of lattices.
2. The smallest genus (where p divides $v \cdot v$) accounts for approximately $\frac{1}{p}$.

Note that the orbit of each basis B under the action of $GL_m(\mathbb{F}_p)$ is of equal size. So our counting argument for lattices in Hermite Normal Form extends to all full-rank bases of q-ary lattices, and therefore the distribution of genera of all lattices $\Lambda_p^\perp(A)$ is equal to the distribution of the genera of the BB^T where B is in HNF.

3.2 Modulo an Odd Prime Power

A more refined statement to Lemma 3 exists for prime powers.

Theorem 1. *Let $q = p^e$ be an odd prime power, and let $v, u \in (\mathbb{Z}/q\mathbb{Z})^n$ have at least one entry coprime to p. Then $\Lambda_q^\perp(v)$ and $\Lambda_q^\perp(u)$ are equivalent over $\mathbb{Z}_p$ if and only if*

1. *the p-parts of $v \cdot v$ and $u \cdot u$ are equal or are both at least q, and*
2. *the p'-parts of $v \cdot v$ and $u \cdot u$ lie in the same residue class mod p.*

The proof works in much the same way as in Section 3.1, with a few important distinctions. We may assume a Gram matrix for $\Lambda_q^\perp(v)$ is

$$G = \begin{pmatrix} q^2 & q v_0^T \\ q v_0 & v_0 v_0^T + \mathbb{I}_{n-1} \end{pmatrix}$$

with $v_0 \in \mathbb{Z}^{n-1}$. The Jordan decomposition will depend entirely on $\|v_0\|^2 + 1 = \det(v_0 v_0^T + \mathbb{I}_{n-1})$. More explicitly, it will depend on both the p-adic order and the residuosity of the p'-part of this determinant. Firstly, however, if the p-part

of $\|v_0\|^2+1$ is greater than q, then we must first apply the transform that maps $(v_1,\ldots,v_i,\ldots,v_{n-1}) \mapsto (v_1,\ldots,v_i+q,\ldots,v_{n-1})$, where $p \nmid v_i$. To do so, we use the same transform as at the end of the proof of Lemma 4:

$$\begin{pmatrix} 1 & 0 \\ b & \mathbb{I}_{n-1} \end{pmatrix},$$

where b is the zero column-vector with a 1 in the ith position. This replaces $\|v_0\|^2+1$ with $\|v_0\|^2+1+q^2+2qv_i$, which clearly has p-part equal to q. Since the p-part of $\|v_0\|^2+1$ is no greater than q, the matrix

$$H = \begin{pmatrix} 1 & 0 \\ \frac{q}{\|v_0\|^2+1} v_0 & \mathbb{I}_{n-1} \end{pmatrix},$$

is in $GL_n(\mathbb{Z}_p)$, so G is $\mathbb{Z}_p$-equivalent to

$$H^T G H = \begin{pmatrix} \frac{q^2}{\|v_0\|^2+1} & 0 \\ 0 & v_0 v_0^T + \mathbb{I}_{n-1} \end{pmatrix},$$

which then diagonalises to

$$G' = \begin{pmatrix} \frac{q^2}{\|v_0\|^2+1} & 0 & 0 \\ 0 & \|v_0\|^2+1 & 0 \\ 0 & 0 & \mathbb{I}_{n-2} \end{pmatrix}$$

So if $\det(v_0 v_0^T + \mathbb{I}_{n-1})$ has p-part p^i, then the Jordan decomposition is

$$f = f_1 \oplus p^i f_{p^i} \oplus p^{e-i} f_{p^{e-i}}$$

where $f_{p^{e-i}}$ has dimension 1, f_{p^i} has dimension 0 if $i=0$, and dimension 1 if $i>0$. In the former case, the sign of f_1 is

$$\left(\frac{\det(f_1)}{p}\right) = \left(\frac{\|v_0\|^2+1}{p}\right)$$

and the sign of f_{p^e} is

$$\left(\frac{\det(f_{p^e})}{p}\right) = \left(\frac{\|v_0\|^2+1}{p}\right).$$

In the latter case, the sign of f_1 is 1, the sign of f_{p^i} is

$$\left(\frac{\det(f_{p^i})}{p}\right) = \left(\frac{\left(\|v_0\|^2+1\right)/p^i}{p}\right),$$

and the sign of $f_{p^{e-i}}$ is

$$\left(\frac{\det(f_{p^{e-i}})}{p}\right) = \left(\frac{\left(\|v_0\|^2+1\right)/p^i}{p}\right).$$

The sign of each form depends on the p'-part of this determinant. Thus any two such forms whose values for $\|v_0\|^2+1$ both have equal p-parts and p'-parts in the same residue class mod p are in the same $\mathbb{Z}_p$ equivalence class.

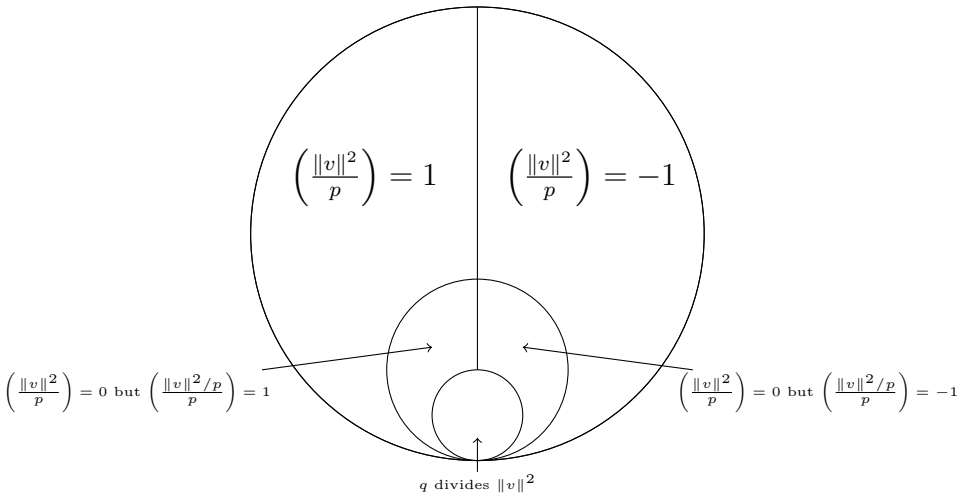

Figure 1: Example distribution when $q = p^2$

Distribution Assuming that $v \cdot v$ is approximately uniform in $\mathbb{Z}/q\mathbb{Z}$, the genera are distributed as follows:

1. The two largest genera each account for just less than $\frac{1}{2}$ of all lattices, and correspond to when $\left(\frac{v \cdot v}{p}\right) = 1$ or -1.
2. The two next-largest genera account for approximately $\frac{1}{2p}$ of the lattices, and correspond to when $\left(\frac{(v \cdot v)/p}{p}\right) = 1$ or -1.
3. The genera reduce in size, with approximately $\frac{1}{2p^i}$ lattices being in each of the two genera corresponding with $\left(\frac{(v \cdot v)/p^i}{p}\right) = 1$ or -1.
4. The smallest genus accounts for about $\frac{1}{q}$ of all lattices: those where $v \cdot v$ is divisible by q.

In total, we therefore expect $2e + 1$ different genera. Figure 1 demonstrates this when $q = p^2$.

3.3 Modulo a Power of 2

Lemma 6. *Let $q = 2^e$, where $e > 0$, and let $v, u \in (\mathbb{Z}/q\mathbb{Z})^n$. If $v \cdot v, u \cdot u$ are odd, and the coordinates of v and u are not all odd, then the quadratic forms associated to $\Lambda_q^\perp(v)$ and $\Lambda_q^\perp(u)$ are equivalent over $\mathbb{Z}_2$ if and only if*

$$e > 1 \text{ and } v \cdot v = u \cdot u \mod 8, \text{ or}$$
$$e = 1 \text{ and } v \cdot v = u \cdot u \mod 4.$$

Note that Lemma 6 only covers odd values of $v \cdot v$. The genus symbol at 2 is in general more complicated, so when q is a power of 2, the genera are not easy to classify fully, short of an exhaustive search. If we assume the values of $v \cdot v \mod 8$ are approximately uniform in the set $\{0, 1, \ldots, 7\}$ when v is chosen uniformly randomly, then we expect the 4 different genera of odd values of $v \cdot v$ to each include about $\frac{1}{8}$ of all such lattices. The conditions for equivalence when $v \cdot v$ is even are efficiently calculable, but the resulting genera usually amount to much less than $\frac{1}{8}$ of all possible lattices. Therefore, to get a broader idea of the genus distribution, we only consider odd values of $v \cdot v \mod 8$. We can also deliberately avoid those v's whose coordinates are all odd, via the condition $n \not\equiv v \cdot v \mod 8$. This is not too restrictive, since the set of such v's with $v \cdot v \equiv n \mod 8$ is of size 2^{-n}.

Proof of Lemma 6. Let f be the quadratic form defined by $\Lambda_q^\perp(v)$. As noted in Section 2.2, we can assume without loss of generality that a Gram matrix for $\Lambda_q^\perp(v)$ is

$$G = \begin{pmatrix} q^2 & q v_0^T \\ q v_0 & \mathbb{I}_{n-1} + v_0 v_0^T \end{pmatrix}.$$

The matrix

$$H = \begin{pmatrix} 1 & 0 \\ -\frac{q}{\|v_0\|^2+1} v_0 & \mathbb{I}_{n-1} \end{pmatrix}$$

is in $GL_2(\mathbb{Z}_2)$, and defines a change of basis transforming G to

$$\begin{aligned} G' &= H^T G H \\ &= \begin{pmatrix} \frac{q^2}{\|v_0\|^2+1} & 0 \\ 0 & \mathbb{I}_{n-1} + v_0 v_0^T \end{pmatrix}. \end{aligned}$$

Over $\mathbb{Z}_2$, the form f therefore has Jordan decomposition

$$f = f_1 + q^2 f_{q^2},$$

where f_1 is defined by the matrix $\mathbb{I}_{n-1} + v_0 v_0^T$ and $f_{q^2} = \frac{1}{\|v_0\|^2+1} X^2$. By Section 2.3, equivalence of such quadratic forms is determined by the canonical symbol at 2.

The form $f_{q^2} = \frac{1}{\|v_0\|^2+1} X^2$ is of type I, with oddity $t(v) = 1/(\|v_0\|^2+1) = v \cdot v \mod 8$ and sign $\varepsilon(v) = \left(\frac{v \cdot v}{2}\right)$. These three invariants depend only on $v \cdot v \mod 8$.

Now consider the form f_1. By assumption, v_0 has at least one even coordinate. Therefore $\mathbb{I}_{n-1} + v_0 v_0^T$ has an odd entry in its main diagonal, so f_1 has type I. The oddity formula [CS13, Chapter 15, 5.1] implies that f has oddity $n \bmod 8$, using the fact that f is equivalent to the trivial form at all primes $p \neq 2$ (and the so-called *p-excess* at odd p equals 0). Moreover, $q^2 f_{q^2}$ has the same oddity as f_{q^2}, namely $t(v) = v \cdot v \bmod 8$. It follows that f_1 has oddity $n - t(v) = (n - v \cdot v) \bmod 8$. By the matrix determinant lemma, the sign is $\left(\frac{\det(\mathbb{I}_{n-1} + v_0 v_0^T)}{2}\right) = \left(\frac{v \cdot v}{2}\right)$. Again, the invariants depend only on $v \cdot v \mod 8$.

In the notation of [CS13, Chapter 15, 7.4]) the 2-adic genus symbol of the Jordan decomposition $f = f_1 + q^2 f_{q^2}$ is therefore $1^{\varepsilon(v)(n-1)}_{n-t(v)} \quad (q^2)^{\varepsilon(v)1}_{t(v)}$. By the results of [CS13, Chapter 15, 7.5–7.6], the canonical 2-adic symbol is

$$\begin{cases} [1^{\varepsilon(v)(n-1)}]_{n-t(v)} \ : \ [(q^2)^{\varepsilon(v)1}]_{t(v)} & \text{if } e > 1 \\ [1^{+(n-1)}]_{n-1} \quad [4^{+1}]_1 & \text{if } e = 1 \text{ and } t(v) \in \{1, 5\}, \\ [1^{+(n-1)}]_{n-7} \quad [4^{+1}]_7 & \text{if } e = 1 \text{ and } t(v) \in \{3, 7\}. \end{cases}$$

This implies the claim. □

Distribution Assume that $v \cdot v \mod 8$ is approximately uniform in $\mathbb{Z}/8\mathbb{Z}$. Then the conditions in Lemma 6 mean that there are four 'large' genera, each containing approximately $\frac{1}{8}$ of all possible $\Lambda_q^\perp(v)$'s for a given dimension n. One genus is relatively small, and only exists if n is even. It contains those $\Lambda_q^\perp(v)$'s wherein v has only odd entries. This occurs with frequency 2^{-n} if n is odd, and never if n is even.

4 Parity Check Matrices

In this section we describe a partial classification of q-ary lattices from parity check matrices A when $q = p^e$ is an odd prime power. Lemma 7 describes a sufficient condition for two parity check lattices to be in the same genus. If we assume that $\det(A^T A) \mod p$ is approximately uniform in $\{0, \dots p-1\}$, then the two largest genera each contain just less than $\frac{1}{2}$ of all such lattices, with the rest all together making up around $\frac{1}{p}$ of the lattices. Classifying the smaller genera is unwieldy, but we are able to describe the two largest genera for any prime power q, and prove that they tend to be the largest. With A chosen uniformly randomly, Lemma 9 gives an upper bound on the total variation distance between $A^T A$ and a uniformly random symmetric matrix. Lemma 8 [Car54, Theorem 3] then gives a closed form probability that $\det(S) = d \mod p$ for any d for a uniform symmetric matrix $S \in \mathbb{F}_p^{m \times m}$. These results lead us to Theorem 2, which describes most of the genus distribution for parity check matrices modulo a prime power.

Classification First, we show that the Legendre symbol of $\det(A^T A)$ at p is sufficient to classify any lattice when this determinant is non-zero modulo p. The proof of this follows the same strategy as Section 3.

Lemma 7. *Let $q = p^e$ be an odd prime power, let $m \leq n$ be non-negative integers, and let $A, A' \in (\mathbb{Z}/q\mathbb{Z})^{n \times m}$ be full rank over $\mathbb{F}_p$. If*

$$\left(\frac{\det(A^T A)}{p}\right) = \left(\frac{\det(A'^T A')}{p}\right) \neq 0,$$

then $\Lambda_q^\perp(A)$ and $\Lambda_q^\perp(A')$ are in the same genus.

Proof. It suffices to prove that the genus symbols of the lattices at p match. As noted in Section 2.2, a Gram matrix for $\Lambda_q^\perp(A)$ is

$$G = \begin{pmatrix} q^2\mathbb{I}_m & qA_0^T \\ qA_0 & \mathbb{I}_{n-m} + A_0A_0^T \end{pmatrix}.$$

for some $A_0 \in \mathbb{Z}^{(n-m)\times m}$ with $A^TA = \mathbb{I}_m + A_0^TA_0 \bmod q$. By hypothesis, we have $p \nmid \det(A^TA) = \det(\mathbb{I}_m + A_0^TA_0) \bmod q$, so the matrix $\mathbb{I}_m + A_0^TA_0$ is invertible over $\mathbb{Z}_p$. By the Weinstein–Aronszajn identity, we have

$$\det(\mathbb{I}_{n-m} + A_0A_0^T) = \det(\mathbb{I}_m + A_0^TA_0),$$

so $\mathbb{I}_{n-m} + A_0A_0^T$ is also invertible over $\mathbb{Z}_p$. A computation shows that a block-diagonalisation of G is thus

$$G' = \begin{pmatrix} q^2\left(\mathbb{I}_m - A_0^T(\mathbb{I}_{n-m} + A_0A_0^T)^{-1}A_0\right) & 0 \\ 0 & \mathbb{I}_{n-m} + A_0A_0^T \end{pmatrix}.$$

The Woodbury matrix identity

$$\mathbb{I}_m - A_0^T(\mathbb{I}_{n-m} + A_0A_0^T)^{-1}A_0 = (\mathbb{I}_m + A_0^TA_0)^{-1}$$

allows us to write the block-diagonalisation above as

$$G' = \begin{pmatrix} q^2\left(\mathbb{I}_m + A_0^TA_0\right)^{-1} & 0 \\ 0 & \mathbb{I}_{n-m} + A_0A_0^T \end{pmatrix}.$$

The matrix G' corresponds to the quadratic form

$$f_1 \oplus q^2 f_{q^2},$$

where f_1 has dimension $n-m$ and f_{q^2} has dimension m. The Weinstein–Aronszajn identity tells us that both have sign $\left(\frac{\det(\mathbb{I}_m + A_0^TA_0)}{p}\right) = \left(\frac{\det(A^TA)}{p}\right)$. Therefore, any two such quadratic forms generated by parity check matrices with the same residuosity of $\det(A^TA) \bmod p$ are in the same genus. □

Distribution For parity check matrices, the argument is similar to that of Section 3.1, but with more technical details. For uniformly random $A \in \mathbb{F}_p^{n\times m}$, we will show that the distribution of $\det(A^TA)$ is approximately uniform in $\mathbb{F}_p$ for large n.

We take p odd for simplicity. By Lemma 9 the total variation distance between the distribution of A^TA and the uniform distribution on the additive group of symmetric matrices decreases exponentially as $n \to \infty$. We know which genera are largest; again it is those $\Lambda_p^\perp(A)$ with either $\left(\frac{\det(A^TA)}{p}\right) = \pm 1$, each of which occur about one half of the time (by Euler's Criterion).

Theorem 2. *Let $q = p^e$ be an odd prime power, and let $n, m \in \mathbb{Z}$ with $n > m$. If A is chosen uniformly at random from $(\mathbb{Z}/q\mathbb{Z})^{n\times m}$, the quadratic form corresponding to $\Lambda_q^\perp(A)$ can fall into one of three cases with the following approximate probabilities.*

1. *Two large genera corresponding to $\left(\frac{\det(A^TA)}{p}\right) = \pm 1$. These cases both occur with probability in the range*

$$\left[\left(\frac{1}{2} - \frac{1}{2p}\right) - p^{-3} - \varepsilon, \left(\frac{1}{2} - \frac{1}{2p}\right) + p^{-3} + \varepsilon\right]$$

   (i.e. almost all such quadratic forms fall into these cases).

2. *A number of smaller genera, corresponding to when $\left(\frac{\det(A^TA)}{p}\right) = 0$. This occurs with probability in the range*

$$\left[\frac{1}{p} - p^{-4} - \varepsilon, \frac{1}{p} + p^{-4} + \varepsilon\right].$$

where $\varepsilon = \frac{p^{m-n}}{4} + \frac{p^{2m-2n}}{2}$.

Proof. Lemma 8 [Car54, Theorem 3] (see [BM10] for the version used here) tells us that the distribution of determinants of symmetric random matrices is near to uniform as $n \to \infty$, but does not necessarily converge to uniform. For each possible determinant d, we have $\left|\mathcal{P}\left[\det(C) = d \mod p\right] - \frac{1}{p}\right| \leq 2p^{-4}$. By Euler, there are $(p-1)/2$ values for d in each residuosity class, so the probability that $\left(\frac{\det(C)}{p}\right) = \pm 1$ is bounded by $2p^{-4}((p-1)/2) \leq p^{-3}$. Lemma 9 states that the total variation distance between A^TA and the uniform distribution on symmetric matrices is bounded by $\frac{p^{m-n}}{4} + \frac{p^{2m-2n}}{2}$. The projection of a uniformly random $A \in (\mathbb{Z}/q\mathbb{Z})^{n\times m}$ onto $(\mathbb{Z}/p\mathbb{Z})^{n\times m}$ is uniformly random in $(\mathbb{Z}/p\mathbb{Z})^{n\times m}$. And if $\det(A^TA)$ is coprime to q, then $\left(\frac{\det(A^TA)}{p}\right) = \pm 1$, so Lemma 9 applies. Lemma 7 and the combination of Lemmas 8 and 9 via the data processing inequality gives us Theorem 2. □

The probabilities in Theorem 2 get close to $\left(\frac{1}{2} - \frac{1}{2p}\right)$ and $\frac{1}{p}$ respectively, as either p increases or the gap between n and m increases. Now we prove the lemmas required in the proof of Theorem 2.

Lemma 8. *Let p be a prime, let $C \in \mathbb{F}_p^{n\times n}$ be a uniformly random symmetric square matrix, and let $d \in \mathbb{F}_p^*$. Define $\lambda := 1/p$, $k := \lceil n/2 \rceil$, and*

$$s := \begin{cases} 0 & \text{if } p = 2 \text{ or } n \text{ is odd,} \\ \left(\frac{d}{p}\right)(-1)^{k(p-1)/2} & \text{otherwise.} \end{cases}$$

Then

$$\mathcal{P}\left[\det(C) = d \mod p\right] = \left(\frac{\lambda}{1-\lambda}\right) \frac{\prod_{2k}(\lambda)}{\prod_k(\lambda^2)} \left(1 + s\lambda^k\right)$$

where $\prod_k(x) := (1-x)(1-x^2)\dots(1-x^k)$.

Furthermore,

$$\left|\mathcal{P}\left[\det(C) = d \mod p\right] - \frac{1}{p}\right| \leq 2p^{-4}.$$

So, in the asymptotic setting, as $n \to \infty$, this probability is not far from $\frac{1}{p}$ for each d, hence our uniformity assumption. The above result applies for random symmetric matrices. Note that for each symmetric matrix $C \in \mathbb{F}^{m\times m}$ there is an explicit formula [Car54, Theorem 1] for the exact number of different $A \in \mathbb{F}^{n\times m}$ such that $A^TA = C$. Instead of using these formulae, the following results demonstrate that the matrix A^TA for a uniformly random A is itself near to a uniformly random symmetric matrix.

The distribution of A^TA as a random walk Let p be a prime number. Let m, n be positive integers. Let $G \subseteq \mathbb{F}_p^{m\times m}$ be the $\mathbb{F}_p$-vector space of symmetric $m \times m$-matrices over $\mathbb{F}_p$.

We take $A \in \mathbb{F}_p^{n\times m}$ uniformly randomly and study the distribution of

$$A^TA \in G.$$

Let $a_1, \dots, a_n$ be the columns of A^T. Then $a_1, \dots, a_n$ are independently uniformly distributed in $\mathbb{F}_p^m$. We have

$$A^TA = \sum_{i=1}^{n} a_i a_i^T.$$

Thus A^TA can be seen as the outcome of a random walk in G starting at 0 and with n steps, each consisting of translation by an element vv^T with $v \in \mathbb{F}_p^m$ uniformly random.

Let f_n denote the probability distribution of this random walk after n steps. Then we have

$$\begin{aligned} f_0 &= \delta_0, \\ f_1 &= p^{-m} \sum_{v\in\mathbb{F}_p^m} \delta_{vv^T}, \\ f_n &= f_1^{*n} \quad \text{for } n \geq 2 \end{aligned}$$

where δ_X is the indicator function of $X \in G$, and f_1^{*n} is the n-th convolution power of f_1.

We recall that the *total variation distance* between two probability distributions f and f' on G is the quantity

$$\|f - f'\|_1 = \frac{1}{2} \sum_{X \in G} |f(X) - f'(X)| \in [0, 1].$$

For $0 \leq r \leq m$, we write $N_{m,r}(\mathbb{F}_p)$ for the number of symmetric $m \times m$-matrices of rank r over $\mathbb{F}_p$. By a result of Carlitz [Car54, Theorem 3] (see also MacWilliams [Mac69, Theorem 2]), we have

$$N_{m,r}(\mathbb{F}_p) = \prod_{i=0}^{r-1} (p^{m-i} - 1) \Big/ \prod_{i=1}^{\lfloor r/2 \rfloor} (1 - p^{-2i}).$$

Lemma 9. *If p is odd, then we have*

$$\begin{aligned} 4\|f_n - u\|_1^2 &\leq \sum_{r=1}^{m} N_{m,r}(\mathbb{F}_p) p^{-nr} \\ &= (p^m - 1)p^{-n} + \frac{(p^m - 1)(p^{m-1} - 1)}{1 - p^{-2}} p^{-2n} + \cdots . \end{aligned}$$

Furthermore, if $n > m$,

$$\|f_n - u\|_1^2 \leq \frac{p^{m-n}}{4} + \frac{p^{2m-2n}}{2}$$

and in particular the total variation distance between f_n and uniform symmetric matrices decreases as $n - m \to \infty$.

The proof will be given after a series of required lemmas. Let us write $\hat{G}$ for the group of homomorphisms $G \to \mathbb{C}^\times$. Given a function $f \colon G \to \mathbb{C}$, we define the Fourier transform of f as

$$\hat{f}(\phi) = \sum_{X \in G} f(X)\phi(X).$$

We define a homomorphism $\psi \colon \mathbb{F}_p \to \mathbb{C}^\times$ by

$$\psi(x) = \exp(2\pi i x/p).$$

We identify $\hat{G}$ with the space of upper triangular $m \times m$-matrices over $\mathbb{F}_p$ by sending such a matrix Y to the homomorphism $\phi_Y \colon G \to \mathbb{C}^\times$ defined by

$$\phi_Y(X) = \psi(\operatorname{tr}(XY)).$$

Lemma 10. *For all $Y \in \hat{G}$, we have*

$$\hat{f}_1(\phi_Y) = p^{-m} \sum_{v \in \mathbb{F}_p^m} \psi(q_Y(v)),$$

where q_Y is the quadratic form defined by Y as in Definition 4.

Proof. For all $v \in \mathbb{F}_p^m$, the cyclic property of the trace gives

$$\begin{aligned}\phi_Y(vv^T) &= \psi(\mathrm{tr}(vv^T Y)) \\ &= \psi(q_Y(v)).\end{aligned}$$

This implies

$$\begin{aligned}\hat{f}_1(\phi_Y) &= \sum_{X \in G} f_1(X)\phi_Y(X) \\ &= p^{-m} \sum_{v \in \mathbb{F}_p^m} \phi_Y(vv^T) \\ &= p^{-m} \sum_{v \in \mathbb{F}_p^m} \psi(q_Y(v)),\end{aligned}$$

as claimed. □

From now on we suppose that p is odd.

If q is a quadratic form over $\mathbb{F}_p$, we denote by $\mathrm{rk}\, q$ the rank of the *symmetric* matrix defining q. Note that if q_Y is the quadratic form defined by some $Y \in \hat{G}$, then $\mathrm{rk}\, q_Y$ does *not* in general equal the rank of the upper triangular matrix Y.

Lemma 11. *For every $Y \in \hat{G}$, we have*

$$|\hat{f}_1(\phi_Y)| = p^{-(\mathrm{rk}\, q_Y)/2}.$$

Proof. After a change of variables, we may assume the quadratic form q_Y is diagonal, with coefficients $c_1, \ldots, c_m \in \mathbb{F}_p$. Applying Lemma 10 gives

$$\begin{aligned}\hat{f}_1(\phi_Y) &= p^{-m} \sum_{v \in \mathbb{F}_p^m} \psi(q_Y(v)) \\ &= p^{-m} \sum_{v_1, \ldots, v_m \in \mathbb{F}_p} \psi(c_1 v_1^2 + \cdots + c_m v_m^2) \\ &= p^{-m} \prod_{i=1}^{m} h(c_i),\end{aligned}$$

where for $c \in \mathbb{F}_p$ we define

$$h(c) = \sum_{x \in \mathbb{F}_p} \psi(cx^2).$$

For $c \in \mathbb{F}_p^\times$, we compute (using $\sum_{t\in\mathbb{F}_p} \psi(t) = 0$)

$$\begin{aligned} h(c) &= \sum_{t\in\mathbb{F}_p} \#\{x \in \mathbb{F}_p \mid cx^2 = t\}\psi(t) \\ &= \sum_{t\in\mathbb{F}_p} \left(1 + \left(\frac{t/c}{p}\right)\right)\psi(t) \\ &= \sum_{t\in\mathbb{F}_p} \left(\frac{t/c}{p}\right)\psi(t) \\ &= \left(\frac{c}{p}\right) \sum_{t\in\mathbb{F}_p} \left(\frac{t}{p}\right)\psi(t). \end{aligned}$$

The last sum is a quadratic Gauss sum, and a well-known result of Gauss (see for example [Apo98, Theorem 8.15]) gives

$$|h(c)| = \sqrt{p} \quad \text{for all } c \in \mathbb{F}_p^\times.$$

Since furthermore $h(0) = p$, it follows that

$$\begin{aligned} |\hat{f}_1(\phi_Y)| &= p^{-m} \cdot p^{m-\operatorname{rk} q_Y} \cdot p^{(\operatorname{rk} q_Y)/2} \\ &= p^{-(\operatorname{rk} q_Y)/2}, \end{aligned}$$

as claimed. □

Proof of Lemma 9. Because the Fourier transform converts convolution into pointwise multiplication, we have

$$\widehat{f_n} = (\hat{f}_1)^n.$$

The upper bound lemma of Diaconis and Shahshahani [Dia88, Chapter 3B, Lemma 1] therefore gives the following bound on the total variation distance between f_n and the uniform distribution u:

$$\|f_n - u\|_1^2 \le \frac{1}{4} \sum_{Y\in\hat{G}\setminus\{0\}} |\hat{f}_1(\phi_Y)^n|^2.$$

The lemma now follows from Lemma 11 and the definition of $N_{m,r}(\mathbb{F}_p)$. □

References

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 99–108, 1996.

[Apo98] T.M. Apostol. *Introduction to Analytic Number Theory.* Undergraduate Texts in Mathematics. Springer New York, 1998.

[BGPSD21] Huck Bennett, Atul Ganju, Pura Peetathawatchai, and Noah Stephens-Davidowitz. Just how hard are rotations of $\mathbb{Z}^n$? algorithms and cryptography with the simplest lattice. Cryptology ePrint Archive, Paper 2021/1548, 2021. `https://eprint.iacr.org/2021/1548`.

[BM10] Richard P. Brent and Brendan D. McKay. On determinants of random symmetric matrices over $\mathbb{Z}_m$. ArXiv, 2010. `https://arxiv.org/abs/1004.5440`.

[Car53] Leonard Carlitz. Weighted quadratic partitions over a finite field. *Can J. Math.*, 5:317–323, 1953.

[Car54] L. Carlitz. Representations by quadratic forms in a finite field. *Duke Math. J.*, 21:123–137, 1954.

[Cas78] John William Scott Cassels. *Rational Quadratic Forms*. Academic Press, New York, 1978.

[CS13] John Horton Conway and Neil James Alexander Sloane. *Sphere packings, lattices and groups*, volume 290. Springer Science & Business Media, 2013.

[Dia88] Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988.

[DvW21] Léo Ducas and Wessel van Woerden. On the lattice isomorphism problem, quadratic forms, remarkable lattices, and cryptography, 2021. `https://eprint.iacr.org/2021/1332`.

[HR14] Ishay Haviv and Oded Regev. On the lattice isomorphism problem. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 391–404. SIAM, 2014.

[Mac69] Jessie MacWilliams. Orthogonal matrices over finite fields. *Amer. Math. Monthly*, 76:152–164, 1969.

[O'M71] Onorato Timothy O'Meara. *Introduction to quadratic forms*. Springer Verlag, 1971.

[PP85] Wilhelm Plesken and Michael Pohst. Constructing integral lattices with prescribed minimum. i. *mathematics of computation*, 45(171):209–221, 1985.

[PS97] Wilhelm Plesken and Bernd Souvignier. Computing isometries of lattices. *Journal of Symbolic Computation*, 24(3-4):327–334, 1997.

[Sch09] Achill Schurmann. *Computational geometry of positive definite quadratic forms: Polyhedral reduction theories, algorithms, and applications*, volume 48. American Mathematical Soc., 2009.

[SHVvW20] Mathieu Dutour Sikiric, Anna Haensch, John Voight, and Wessel PJ van Woerden. A canonical form for positive definite matrices. *ANTS XIV*, page 179, 2020.