Lattice-Based Vulnerabilities in Lee Metric Post-Quantum Cryptosystems

Anna-Lena Horlemann[1[0000−0003−2685−2343]], Karan Khathuria[2[0000−0002−9886−2770]], Marc Newman[1[0009−0000−2818−3492]], Amin Sakzad[3[0000−0003−4569−3384]], and Carlos Vela Cabello[1[0000−0003−3362−8817]]

[1] University of St.Gallen, St. Gallen, Switzerland
{anna-lena.horlemann, marc.newman, carlos.velacabello}@unisg.ch
[2] Quantinuum, Partnership House, Carlisle Place, London SW1P 1BX, United Kingdom
karan.khathuria@quantinuum.com
[3] Monash University, Clayton, Australia
amin.sakzad@monash.edu

Abstract. Post-quantum cryptography has gained attention due to the need for secure cryptographic systems in the face of quantum computing. Code-based and lattice-based cryptography are two prominent approaches, both heavily studied within the NIST standardization project. Code-based cryptography—most prominently exemplified by the McEliece cryptosystem—is based on the hardness of decoding random linear error-correcting codes. Despite the McEliece cryptosystem having been unbroken for several decades, it suffers from large key sizes, which has led to exploring variants using metrics than the Hamming metric, such as the Lee metric. This alternative metric may allow for smaller key sizes, but requires further analysis for potential vulnerabilities to lattice-based attack techniques. In this paper, we consider a generic Lee metric based McEliece type cryptosystem and evaluate its security against lattice-based attacks.

Keywords: code-based cryptography · Lee metric · Hamming metric · lattice-based cryptography · ℓ_1-norm.· ℓ_2-norm.

1 Introduction

In response to the threat posed by quantum computing to traditional cryptographic systems, post-quantum cryptography has gained significant attention over the last few years. Among the various approaches within post-quantum cryptography, code-based and lattice-based cryptography are two of the most widely studied research directions and constitute a majority of the current proposals in the NIST standardization project.

Code-based cryptography is founded on the hardness of decoding (random linear) error-correcting codes, a problem that remains intractable for both classical and quantum computers in its general form. This branch of cryptography

has its roots in the McEliece cryptosystem, proposed in 1978, a system that remains unbroken to date and, therefore, promises high security guarantees. On the other hand, it suffers from the drawback of requiring large public key sizes. Consequently, one of the main research tasks in code-based cryptography is to establish variants of the McEliece cryptosystem with smaller keys. One way of doing so is to use decoding metrics other than the originally proposed Hamming metric, e.g., the rank or the Lee metric, of which the latter is the main topic of this work.

Lattice-based cryptography, on the other hand, relies on the difficulty of solving problems in (high-dimensional) lattices, such as the Shortest Vector Problem (SVP) and the Learning With Errors (LWE) problem. Lattice-based schemes offer several compelling advantages, including strong security proofs and practical efficiency. The seminal works of Ajtai and Dwork in the late 1990s laid the groundwork for this domain, leading to the development of numerous cryptographic protocols that are both theoretically sound and practically viable.

The use of the Lee metric in code-based cryptography was first suggested in [7] and has since been studied from a coding-theoretic perspective in, e.g., [2–4, 16]. These results suggest that the generic Lee syndrome decoding problem is (much) harder than its Hamming metric counter-part, which would imply that smaller codes could be used in a code-based cryptosystem when using the Lee metric instead of the Hamming metric. This would, in turn, lead to a reduced public key size.

To complement the coding-theoretic perspective, it is well-known that the Lee metric over modular integer rings is the analog of the ℓ_1-norm over the integers. It is therefore important to analyze the security of a Lee metric code-based cryptosystem with respect to lattice techniques (in the ℓ_1- or ℓ_2-norm). Exactly this approach has recently been used in [8] to break the signature scheme FuLeeca [11], which was submitted to the NIST standardization project. The attack exploits several properties of those Lee metric codes which arise from the specific parameters that were suggested—in particular, a large modulus and small minimum distance of the error vector.

In this paper, we first consider a McEliece type cryptosystem over the Lee metric and then study the attackability of such Lee metric code-based cryptosystems with lattice techniques more generally. For this, we will focus on public key encryption schemes (and not on digital signatures). In particular, we will derive complexity reductions to and from several known lattice problems including the bounded distance decoding problem (BDD), the Lee-distance decoding problem (LeeDP), and the unique shortest vector problem (uSVP); as shown in Fig. 1. We will then analyze and find the values and parameters for which lattice reduction algorithms could be applied to Lee metric codes embedded in a lattice and compare the marginal error distributions of the Lee metric, the Hamming metric, and the ℓ_1- and the ℓ_2-norms for both the Laplace and Gaussian distributions.

The paper is structured as follows. Section 2 provides all the necessary preliminaries, definitions and results needed for the rest of the paper. It also includes the Lee-McEliece cryptosystem. In Section 3, we first establish the relation between

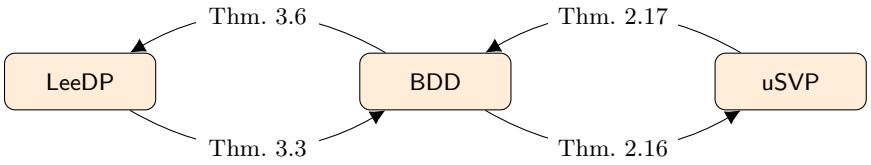

Fig. 1. Scheme of the reductions for full rank integer lattices in the ℓ_1-norm.

the shortest vector (in ℓ_1-norm) of a lattice constructed based on Construction A and the minimum Lee distance of its underlying code. We further establish a two way complexity reduction between LeeDP, BDD, and uSVP, see Theorems 3.3 and 3.7. In Section 4, we study when the techniques in the FuLeakage attack [8] that were applied to FuLeeca [11] can and cannot be applied to the cryptosystem in Section 2. Finally, we establish connections between the Lee metric and the Laplace distribution and use it to compare Laplace and discrete Gaussian distributions in terms of Rényi divergence.

2 Preliminaries

We denote by $\mathbb{Z}_q$ the ring of integers modulo q. We will switch between two different representations of the elements of $\mathbb{Z}_q$, namely the standard representation $\{0, 1, 2, \ldots, q-1\}$, and the representation centered at zero $\{-\lfloor (q-1)/2 \rfloor, \ldots, 0, \ldots, \lfloor q/2 \rfloor\}$. If not specified, we will use $\mathbb{Z}_q = \{0, 1, 2, \ldots, q-1\}$.

For a convex set $S \subseteq \mathbb{R}^n$ that spans a k-dimensional subspace, we will denote the k-dimensional relative volume of S by $\mathrm{Vol}_k(S)$, i.e., the volume of S in the linear space spanned by S. Given a set U of k vectors over $\mathbb{R}$, we will denote the span of U in $\mathbb{R}$ by $\mathrm{Span}_{\mathbb{R}}(U) := \left\{ \sum_{i=1}^{k} x_i \mathbf{u}_i \mid \ x_i \in \mathbb{R}, \mathbf{u}_i \in U \right\}$.

Definition 2.1 *Let* $\mathbf{A}$ *be a* $n \times n$ *invertible matrix and define* $M_{i,j}$ *to be the determinant of the* $(n-1) \times (n-1)$ *matrix obtained by removing the* i*th row and* j*th column from* $\mathbf{A}$*. Then the* adjugate *of* $\mathbf{A}$ *is defined to be*

$$\mathrm{adj}(\mathbf{A}) := \left[(-1)^{i+j} M_{j,i} \right]_{1 \leqslant i,j \leqslant n}.$$

It is a well-known property of the adjugate that

$$\mathrm{adj}(\mathbf{A}) \cdot \mathbf{A} = \det(\mathbf{A}) \cdot \mathbf{I}_n = \mathbf{A} \cdot \mathrm{adj}(\mathbf{A}).$$

2.1 Lee metric codes and the Lee-McEliece system

Definition 2.2 *For* $x \in \mathbb{Z}_q$ *we define the* Lee weight *to be*

$$\mathrm{wt}_L(x) := \min\{ \mid x \mid, \mid q - x \mid \},$$

Then, for $\mathbf{x} \in \mathbb{Z}_q^n$*, we define the* Lee weight *to be the sum of the Lee weights of its coordinates,*

$$\mathrm{wt}_L(\mathbf{x}) := \sum_{i=1}^{n} \mathrm{wt}_L(x_i)\,.$$

We define the Lee distance *of* $\mathbf{x}$, $\mathbf{y} \in \mathbb{Z}_q^n$ *as*

$$d_L(\mathbf{x}, \mathbf{y}) := \mathrm{wt}_L(\mathbf{x} - \mathbf{y})\,.$$

Note that for $q = 2, 3$ the Lee weight is equal to the Hamming weight wt_H in $\mathbb{Z}_q^n$, which is defined as

$$\mathrm{wt}_H(\mathbf{x}) := |\{i \mid x_i \neq 0\}| \quad \text{for all } \mathbf{x} \in \mathbb{Z}_q^n.$$

Both for practical and theoretical reasons the following marginal distributions per coordinate of a vector with constant Lee—respectively, Hamming—weight t will be useful.

Lemma 2.3 *(a) [2, Lemma 1] Let* $\mathbf{x} \in \mathbb{Z}_q^n$ *be a uniformly random vector with* $\mathrm{wt}_L(\mathbf{x}) = t = Tn$ *for some* $T \in [0, \lfloor q/2 \rfloor]$*. Further, let* E *be the random variable representing a coordinate of* $\mathbf{x}$*. Then, as* n *tends to infinity, for any* $j \in \mathbb{Z}_q$*,*

$$F_T(j) := \Pr(E = j) = \frac{\exp(-\beta \mathrm{wt}_L(j))}{\sum_{i=0}^{q-1} \exp(-\beta \mathrm{wt}_L(i))}, \tag{1}$$

where β *is the unique real solution to the constraint*

$$T = \sum_{i=0}^{q-1} \mathrm{wt}_L(i) \cdot \Pr(E = i). \tag{2}$$

(b) Let $\mathbf{x} \in \mathbb{Z}_q^n$ *be a uniformly random vector with* $wt_H(\mathbf{x}) = t = \delta n$ *for some* $\delta \in [0, 1]$*. Further, let* E *be the random variable representing a coordinate of* $\mathbf{x}$*. Then, as* n *tends to infinity, for any* $j \in \mathbb{Z}_q$*,*

$$H_\delta(j) := \Pr(E = j) = \begin{cases} 1 - \delta & \text{if } j = 0 \\ \frac{\delta}{q-1} & \text{otherwise} \end{cases}.$$

Note that, even though the Lee and Hamming marginal distribution is an asymptotic result for growing n, the de facto distribution for small n only differs by something very small. Therefore, we will use the marginals from above in our analysis. Furthermore, we will use T for the relative Lee distance and δ for the relative Hamming distance.

Remark 2.4 *For* $q = 2, 3$ *we get*

$$\beta = \log\left(\frac{1 - \delta}{\delta}(q - 1)\right)$$

above in the Lee distribution, and hence the Lee distribution equals the Hamming distribution.

Definition 2.5 *Let q be a positive integer.*

1. *A* code *over $\mathbb{Z}_q$ of length n is a subset of $\mathbb{Z}_q^n$.*
2. *A* (ring-)linear code *over $\mathbb{Z}_q$ of length n is a $\mathbb{Z}/q\mathbb{Z}$-submodule of $\mathbb{Z}_q^n$.*
3. *The* minimum Lee distance $d_L(\mathcal{C})$ *of a code $\mathcal{C} \subseteq \mathbb{Z}_q^n$ is the minimum of all Lee distances of distinct codewords of $\mathcal{C}$:*

$$d_L(\mathcal{C}) = \min\{d_L(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in \mathcal{C} \text{ with } \mathbf{x} \neq \mathbf{y}\}.$$

Linear codes can be completely represented through a generator or a parity-check matrix.

Definition 2.6 *A matrix $\mathbf{G}$ is called a* generator matrix *for a (ring-)linear code $\mathcal{C}$ if its row space corresponds to $\mathcal{C}$. In addition, we call a matrix $\mathbf{H}$ a* parity-check matrix *for $\mathcal{C}$ if its kernel corresponds to $\mathcal{C}$.*

Note that such generator and parity-check matrices are not unique. If q is not prime, even the number of rows of such matrices is not unique.

The general security assumption of code-based cryptography is based on the hardness of the syndrome decoding problem (SDP).[4] The Lee metric version of this is as follows:

Problem 2.7 (Lee syndrome decoding problem (LeeSDP_t)) *Given a linear code $\mathcal{C}$ over $\mathbb{Z}_q$ of length n with parity check matrix $\mathbf{H} \in \mathbb{Z}_q^{(n-k)\times n}$, a syndrome $\mathbf{s} \in \mathbb{Z}_q^{n-k}$ and a positive integer $t \in \mathbb{N}$, find $\mathbf{e} \in \mathbb{Z}_q^n$ such that $\mathrm{wt}_L(\mathbf{e}) \leqslant t$ and $\mathbf{e}\mathbf{H}^\top = \mathbf{s}$, where $^\top$ denotes the transposition operation.*

Note that this problem is equivalent to the general decoding problem for linear codes:

Problem 2.8 (Lee decoding problem (LeeDP_t)) *Given a linear code $\mathcal{C}$ over $\mathbb{Z}_q$ of length n, a vector $\mathbf{r} \in \mathbb{Z}_q^n$ and a positive integer $t \in \mathbb{N}$, find $\mathbf{c} \in \mathcal{C}$ such that $\mathrm{wt}_L(\mathbf{r} - \mathbf{c}) \leqslant t$.*

It was shown in [16] that the syndrome decoding problem is NP-complete for any additive weight function, which includes the Lee metric. It is therefore a cryptographically interesting computationally hard problem to be used in public key cryptosystems. Algorithm 1 shows a general setup of a McEliece-type public key encryption scheme with Lee metric codes.

Remark 2.9 *The Lee isometries are generally not transitive on the sphere of vectors with a fixed Lee weight. To prevent partial information leakage about the error vector during the encryption, this should be considered when choosing the secret linear code and generator matrix. Furthermore, the isometry φ could be replaced by a near-isometry (i.e., maps that possibly change the weight of the*

[4] De facto this is not true for the McEliece cryptosystem, since the codes used are not random. However, we will not go into detail about this issue in this paper.

Algorithm 1 Lee-McEliece cryptosystem

Secret key: The generator matrix $\mathbf{G}_{sec} \in \mathbb{Z}_q^{k\times n}$ of an efficiently decodable Lee metric code with error-correction capacity $w \in \mathbb{N}$, and a Lee-isometry φ.
Public key: The generator matrix $\mathbf{G}_{pub} = \varphi(\mathbf{G}_{sec}) \in \mathbb{Z}_q^{k\times n}$ and w.
Encryption: To encrypt the message $\mathbf{m} \in \mathbb{Z}_q^k$ choose an error vector $\mathbf{e}$ of Lee weight w uniformly at random and create the cipher

$$\mathbf{c} = \mathbf{m}\mathbf{G}_{pub} + \mathbf{e}.$$

Decryption: Decode

$$\varphi^{-1}(\mathbf{c})$$

in the secret code to retrieve $\varphi^{-1}(\mathbf{m})\mathbf{G}_{sec}$. Recover $\mathbf{m}$ through linear algebra operations and application of φ.

vector by at most some prescribed value t), and the error weight in the encryption should be chosen to be $w-t$, such that the receiver can still uniquely decrypt. It is not the topic of this paper to analyze this issue, however it will be of paramount importance when suggesting a specific instance of such a cryptosystem.

There are two main types of attacks that need to be analyzed in this setting: key recovery attacks, where the attacker can recover the secret linear code and its efficient decoding algorithm; and message recovery attacks, where the intruder recovers the message $\mathbf{m}$ from the ciphertext $\mathbf{c}$ without recovering the secret key. In this paper we will focus on the latter, by using known lattice techniques to recover the message.

2.2 Lattice theory

We assume that the space $\mathbb{R}^n$ is equipped with the ℓ_1-norm $\|\mathbf{v}\|_1 := \sum_{i=1}^n |v_i|$. Note that this differs from the classical approach, where the ℓ_2-norm is used.

Definition 2.10 *The ℓ_1 distance between two vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ is denoted by*

$$d_1(\mathbf{v}, \mathbf{w}) := \|\mathbf{v} - \mathbf{w}\|_1 .$$

Definition 2.11 *Given m linearly independent vectors $\mathbf{b}_1, \ldots, \mathbf{b}_m \in \mathbb{R}^n$, the lattice generated by them is given by*

$$\mathcal{L}(\mathbf{b}_1, \ldots, \mathbf{b}_m) := \left\{ \sum_{i=1}^m x_i \mathbf{b}_i : x_i \in \mathbb{Z} \right\}.$$

For a vector $\mathbf{r} \in \mathbb{R}^n$, the distance between $\mathbf{r}$ and $\mathcal{L}$ is given by $d_1(\mathbf{r}, \mathcal{L}) := \inf\{d_1(\mathbf{r}, \mathbf{v}) : \mathbf{v} \in \mathcal{L}\}$. The shortest vector of a lattice $\mathcal{L}$ is the vector in $\mathcal{L}$ having the smallest ℓ_1 norm. The length of the shortest vector is denoted by $\lambda_1(\mathcal{L})$, the length of the shortest lattice vector that is not a multiple of the shortest vector is denoted by $\lambda_2(\mathcal{L})$.

We can now state the two lattice problems in the ℓ_1-norm that are of interest for us:

Problem 2.12 (α-Bounded distance decoding problem (BDD_α)) *Given an integer lattice $\mathcal{L}$ and a vector $\mathbf{r} \in \mathbb{Z}^n$ such that $d_1(\mathbf{r}, \mathcal{L}) < \alpha\lambda_1(\mathcal{L})$, find $\mathbf{v} \in \mathcal{L}$ such that $d_1(\mathbf{v}, \mathbf{r}) < \alpha\lambda_1(\mathcal{L})$.*

Problem 2.13 (γ-unique shortest vector problem (uSVP_γ)) *Given an integer lattice $\mathcal{L}$ such that $\lambda_2(\mathcal{L}) > \gamma\lambda_1(\mathcal{L})$, find a non-zero vector $\mathbf{v} \in \mathcal{L}$ of length $\lambda_1(\mathcal{L})$.*

The connection between those two problems has already been studied in [9] as follows:

Theorem 2.14 *[9, Theorem 1] For any $\gamma \geqslant 1$, there is a polynomial time reduction from $\mathsf{BDD}_{1/2\gamma}$ to uSVP_γ.*

Theorem 2.15 *[9, Theorem 2] For any polynomially bounded $\gamma(n) = n^{O(1)}$, there is a polynomial time reduction from uSVP_γ to $\mathsf{BDD}_{1/\gamma}$.*

We remark that the results in [9] were proven for the ℓ_2-norm over $\mathbb{R}^n$; however, it was noted that the same proofs will hold for any other ℓ_p-norm as well. Without loss of generality, we also assume that the above results hold for integer lattices and target vectors. We can thus use the ℓ_1-versions as follows:

Theorem 2.16 *For any $\gamma \geqslant 1$, there is a polynomial time reduction from $\mathsf{BDD}_{1/(2\gamma)}$ to uSVP_γ over the ℓ_1-norm.*

Theorem 2.17 *For any polynomially bounded $\gamma(n) = n^{O(1)}$, there is a polynomial time reduction from uSVP_γ to $\mathsf{BDD}_{1/\gamma}$ over the ℓ_1-norm.*

Lastly, for our results in Section 4 we will make use of the following two known results.

Theorem 2.18 *[15] Let $C_n = [-\frac{1}{2}, \frac{1}{2}]^n \subseteq \mathbb{R}^n$, i.e., the n-dimensional unit cube centered at the origin. Let $P_k \subseteq \mathbb{R}^n$ be any k-dimensional linear subspace. Then $\mathrm{Vol}_k(C_n \cap P_k) \geqslant 1$.*

Theorem 2.19 *Let $\mathcal{L}$ be a k-dimensional lattice in $\mathbb{R}^n$ and let $S \subseteq \mathrm{Span}_{\mathbb{R}}(\mathcal{L})$ be a convex set symmetric about the origin (i.e., $\mathbf{x} \in S$ implies $-\mathbf{x} \in S$). Suppose that $\mathrm{Vol}_k(S) > m \cdot 2^k \cdot \det(\mathcal{L})$. Then there are m different pairs of vectors $\pm\mathbf{z}_1, \ldots, \pm\mathbf{z}_m \in S \cap \mathcal{L}\backslash\{0\}$.*

The above theorem is an extension of Minkowski's convex body theorem. Since the standard form of Minkowski's theorem is for full-dimensional lattices and $m = 1$, we provide the proof of this version in Appendix A for completeness. Our proof is based on the proofs from [10, Theorem 20-21] and [14, Theorem 5-6].

2.3 Distributions

Let F be a probability distribution over the sample space X. Then, we denote the support of F by $\mathrm{Supp}(F) := \{x \in X \mid F(x) \neq 0\}$. Throughout the paper, we may interchangeably use the same symbol to denote both the probability distribution and its density function.

We define a continuous Gaussian distribution over $\mathbb{R}$ by its density function $\mathrm{D}_{\mathbb{R},\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-x^2/\sigma^2)$ and over a lattice $\mathcal{L} \subseteq \mathbb{R}^n$ as follows:

Definition 2.20 (Discrete Gaussian) *For a lattice $\mathcal{L}$, the discrete Gaussian distribution $\mathrm{D}_{\mathcal{L},\sigma}$ is defined by the probability density function*

$$\mathrm{D}_{\mathcal{L},\sigma}(\mathbf{x}) := \frac{\exp(-\|\mathbf{x}\|_2^2/2\sigma^2)}{\sum_{\mathbf{y}\in\mathcal{L}} \exp(-\|\mathbf{y}\|_2^2/2\sigma^2)},$$

for every $\mathbf{x} \in \mathcal{L}$.

Similarly, we define a continuous Laplace distribution over $\mathbb{R}$ by its density function $\mathrm{Lap}_{\mathbb{R},b}(x) = \frac{1}{2b} \exp(-|x|/b)$ and over a lattice $\mathcal{L} \subseteq \mathbb{R}^n$ as follows:

Definition 2.21 (Discrete Laplace) *For a lattice $\mathcal{L}$, the discrete Laplace distribution $\mathrm{Lap}_{\mathcal{L},b}$ with $b > 0$ is defined by its probability density function*

$$\mathrm{Lap}_{\mathcal{L},b}(\mathbf{x}) := \frac{\frac{1}{2b} \exp(-\|\mathbf{x}\|_1/b)}{\sum_{\mathbf{y}\in\mathcal{L}} \frac{1}{2b} \exp(-\|\mathbf{y}\|_1/b)},$$

for every $\mathbf{x} \in \mathcal{L}$.

Given the integer lattice $\mathbb{Z}^n$, it is easy to check that $\mathrm{D}_{\mathbb{Z}^n,\sigma} = \prod_{i=1}^n \mathrm{D}_{\mathbb{Z},\sigma}$ and $\mathrm{Lap}_{\mathbb{Z}^n,b} = \prod_{i=1}^n \mathrm{Lap}_{\mathbb{Z},b}$.

We use the Rényi and Kullback-Leibler divergence to measure the closeness of two distributions.

Definition 2.22 *Let F and G be discrete probability distributions with $Supp(F) \subseteq Supp(G)$. Then,*

1. ***(Rényi divergence)** for any $a \in (1, \infty]$, the Rényi divergence of order a between F and G is given by:*

$$R_a(F||G) := \begin{cases} \left(\sum\limits_{x\in Supp(F)} \frac{F(x)^a}{G(x)^{a-1}} \right)^{\frac{1}{a-1}} & \text{for } a \in (1,\infty) \\ \max\limits_{x\in Supp(F)} \frac{F(x)}{G(x)} & \text{for } a = \infty \end{cases}$$

2. ***(Kullback-Leibler divergence)** the Kullback-Leibler (KL) divergence between F and G is given by*

$$KL(F||G) := \sum_{x\in Supp(F)} F(x) \log\left(\frac{F(x)}{G(x)}\right)$$

The definitions are extended in a natural way to continuous distributions using integrals instead of the summations. Note that the Kullback-Leibler divergence is a logarithm of the limit of Rényi divergence of order a as a goes to 1, i.e.,

$$KL(F||G) = \log\left(\lim_{a\to 1} R_a(F||G)\right).$$

See [6] for a proof. We will use the following properties of Rényi and Kullback-Leibler divergence. We again refer to [6] for proofs.

Lemma 2.23 *Let F and G be probability distributions with $Supp(F) \subseteq Supp(G)$. Further, let $F^{(n)} = F \times \cdots \times F$ and $G^{(n)} = G \times \cdots \times G$ be the product of n independent and identical copies of F and, respectively, G. Then,*

1. ***Multiplicativity of Rényi divergence**:*

$$R_a\left(F^{(n)}||G^{(n)}\right) = \prod_{i=1}^{n} R_a(F||G).$$

2. ***Additivity of Kullback-Leibler divergence**:*

$$KL\left(F^{(n)}||G^{(n)}\right) = \sum_{i=1}^{n} KL(F||G).$$

3 Complexity Reductions of Lee Metric Decoding Problems

In this section we show that for bounded error vectors the Lee metric decoding problem (Problem 2.8) over linear codes reduces to the bounded distance decoding problem (Problem 2.12) over lattices in the ℓ_1-norm, and vice versa. All the results from this section are also summarized in Fig. 1.

In general, we can always associate a lattice to a given linear code. One of the most common approaches is known as Construction A, which takes a linear code in $\mathbb{Z}_q^n$ and translates it over $\mathbb{Z}^n$ using the vectors from $q\mathbb{Z}^n$.

Definition 3.1 (Construction A) *Let $\mathcal{C}$ be a linear code in $\mathbb{Z}_q^n$ and let $\mathbf{G}$ be a $k \times n$ generator matrix of $\mathcal{C}$. Then the* Construction A *lattice associated to $\mathcal{C}$ is given by:*

$$\mathcal{L}_{\mathsf{A}}(\mathcal{C}) = \{\mathbf{c} \in \mathbb{Z}^n : \mathbf{c} = \mathbf{G}^\top \mathbf{x} \bmod q \text{ for some } \mathbf{x} \in \mathbb{Z}^k\}.$$

It can be easily seen that $\mathcal{L}_{\mathsf{A}}(\mathcal{C}) = \mathcal{C} + q\mathbb{Z}^n$, and hence $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$ does not depend on the choice of the generator matrix $\mathbf{G}$. If the code $\mathcal{C}$ is clear from the context, we will simply denote $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$ by $\mathcal{L}_{\mathsf{A}}$.

With the representation of $\mathbb{Z}_q^n$ centered around the origin, i.e., $\mathbb{Z}_q^n = \{-\lfloor (q-1)/2 \rfloor, \ldots, 0, \ldots, \lfloor q/2 \rfloor\}^n$, we obtain that the construction of the lattice $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$ preserves the metric structure on $\mathcal{C}$, i.e., the length of the shortest ℓ_1-norm vector in $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$ relates to the minimum Lee distance of $\mathcal{C}$.

Proposition 3.2 *Let $\mathcal{C}$ be a linear code in $\mathbb{Z}_q^n$. Then the ℓ_1-norm of the shortest vector in the* Construction A *lattice $\mathcal{L}_{\mathsf{A}}$ is given by*

$$\lambda_1(\mathcal{L}_{\mathsf{A}}) = \min\{q, d_L(\mathcal{C})\},$$

where $d_L(\mathcal{C})$ is the minimum Lee distance of $\mathcal{C}$.

This proposition has previously appeared in [1] and [12] without a proof. Thus, for completeness we give the proof below.

Proof. For simplicity we assume that q is odd and let $M = \lfloor q/2 \rfloor$. For an even q, the proof would be similar with only minor changes in the representation of $\mathbb{Z}_q$.

As described earlier, we represent elements of $\mathbb{Z}_q$ in $\mathbb{Z}$ by $\{-M, \ldots, M\}$. Using this representation, we get a one-to-one correspondence between the codewords in $\mathcal{C}$ and the lattice points of $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$ inside the n-cube $[-M, M]^n$. Note that each codeword $\mathbf{c} \in \mathcal{C}$ and its representative, say $\tilde{\mathbf{c}}$, in $\mathcal{L}_{\mathsf{A}}$ satisfy $\mathrm{wt}_L(\mathbf{c}) = \|\tilde{\mathbf{c}}\|_1$. This implies that $\lambda_1(\mathcal{L}_{\mathsf{A}}) \leqslant d_L(\mathcal{C})$. Moreover, since $(q, 0, \ldots, 0) \in \mathcal{L}_{\mathsf{A}}$, we get $\lambda_1(\mathcal{L}_{\mathsf{A}}) \leqslant q$, and hence $\lambda_1(\mathcal{L}_{\mathsf{A}}) \leqslant \min\{q, d_L(\mathcal{C})\}$.

Now, to show that $\lambda_1(\mathcal{L}_{\mathsf{A}}) \geqslant \min\{q, d_L(\mathcal{C})\}$, it is enough to show that $\lambda_1(\mathcal{L}_{\mathsf{A}}) \geqslant q$ or $\lambda_1(\mathcal{L}_{\mathsf{A}}) \geqslant d_L(\mathcal{C})$. Let $\mathbf{x} \in \mathcal{L}_{\mathsf{A}}$ be a lattice point such that $\|\mathbf{x}\|_1 = \lambda_1(\mathcal{L}_{\mathsf{A}})$. If $\mathbf{x} \bmod q = \mathbf{0}$, then $\lambda_1(\mathcal{L}_{\mathsf{A}}) \geqslant q$ as q is the smallest ℓ_1-norm for a non-zero point in $\mathbb{Z}^n$. Now, if $\mathbf{x} \bmod q \neq \mathbf{0}$, then $\mathbf{x} \in \mathcal{L}_{\mathsf{A}} \cap [-M, M]^n$ because, if $|x_i| > M$ for any i, then by either subtracting or adding q to x_i one can obtain another lattice point with ℓ_1-norm strictly smaller than $\|\mathbf{x}\|_1 = \lambda_1(\mathcal{L}_{\mathsf{A}})$, which is a contradiction. Since $\mathbf{x} \in \mathcal{L}_{\mathsf{A}} \cap [-M, M]^n$, we get a codeword in $\mathcal{C}$ that corresponds to $\mathbf{x}$ and has Lee weight equal to $\|\mathbf{x}\|_1 = \lambda_1(\mathcal{L}_{\mathsf{A}})$. This implies that $d_L(\mathcal{C}) \leqslant \lambda_1(\mathcal{L}_{\mathsf{A}})$.

We remark that a similar result for the Hamming distance and the ℓ_2-norm for Construction A lattices has been given in [13] (see Corollary 2 therein).

Theorem 3.3 *Let $\mathcal{C}$ be a linear code over $\mathbb{Z}_q$ with minimum Lee distance $d_L(\mathcal{C})$. Then, for any $t = \alpha \min\{q, d_L(\mathcal{C})\} \in \mathbb{Z}$ for some $\alpha \in (0, 1)$, there is a polynomial time reduction from* LeeDP_t *on $\mathcal{C}$ to* BDD_α *in the ℓ_1-metric on $\mathcal{L}_{\mathsf{A}}(\mathcal{C})$.*

Proof. We consider an instance of LeeDP_t on $\mathcal{C}$ with $\mathbf{r}$ being the vector in $\mathbb{Z}_q^n$ to be decoded. Note that we can write $\mathbf{r} = \mathbf{c} + \mathbf{e}$, where $\mathbf{c} \in \mathcal{C}$ is the closest codeword to $\mathbf{r}$ and $\mathbf{e} \in \mathbb{Z}_q^n$ is the corresponding error vector. Let $\tilde{\mathbf{r}}, \tilde{\mathbf{c}}, \tilde{\mathbf{e}} \in [-M, M]^n$ be the corresponding representatives of $\mathbf{r}, \mathbf{c}, \mathbf{e}$, respectively, in $\mathcal{L}_{\mathsf{A}} \cap [-M, M]^n$. Since $\tilde{\mathbf{r}} - \tilde{\mathbf{c}} = \tilde{\mathbf{e}} \mod q$, we get that

$$\tilde{\mathbf{r}} - \tilde{\mathbf{c}} = \tilde{\mathbf{e}} + \mathbf{v}q\mathbf{I}_n$$

for some $\mathbf{v} \in \mathbb{Z}^n$ (in fact, it is $\mathbf{v} \in \{-1, 0, 1\}^n$). Then $\bar{\mathbf{c}} := \tilde{\mathbf{c}} - \mathbf{v}q\mathbf{I}_n$ is an element of $\mathcal{L}_{\mathsf{A}}$ and fulfills $\tilde{\mathbf{r}} - \bar{\mathbf{c}} = \tilde{\mathbf{e}}$. This implies $d_1(\tilde{\mathbf{r}}, \mathcal{L}_{\mathsf{A}}) \leqslant \|\tilde{\mathbf{r}} - \bar{\mathbf{c}}\|_1 = \|\tilde{\mathbf{e}}\|_1 = \mathrm{wt}_L(\mathbf{e}) \leqslant t = \alpha\lambda_1(\mathcal{L}_{\mathsf{A}})$. Hence, we get an instance of BDD_α for a received vector $\tilde{\mathbf{r}}$ with $d_1(\tilde{\mathbf{r}}, \mathcal{L}_{\mathsf{A}}) \leqslant \alpha\lambda_1(\mathcal{L}_{\mathsf{A}})$. The BDD_α oracle now gives a lattice vector $\mathbf{x}$ satisfying $d_1(\tilde{\mathbf{r}}, \mathbf{x}) \leqslant \alpha\lambda_1(\mathcal{L}_{\mathsf{A}}) = t$. Let $\mathbf{c}_x := \mathbf{x} \pmod q$, then we have that $\mathbf{c}_x \in \mathcal{C}$ (according to the definition of Construction A lattices) and $\mathrm{wt}_L(\mathbf{r} - \mathbf{c}_x) \leqslant t$.

Remark 3.4 *In the case when we have a* LeeDP_t *instance with* $t = \alpha d_L(\mathcal{C})$ *and* $d_L(\mathcal{C}) > q$*, the reduction to* BDD_α *does not hold. Note that in this case, we cannot directly apply the* BDD_α *oracle, like we did in the proof of Theorem 3.3, because we may not satisfy* $d_1(\tilde{\mathbf{r}}, \mathcal{L}_\mathsf{A}) \leqslant \alpha\lambda_1(\mathcal{L}_A) = \alpha q$ *for any* $\alpha \in (0, 1)$*.*

Lemma 3.5 *Let* $\mathcal{L} \subseteq \mathbb{Z}^n$ *be a full rank integer lattice with basis vectors* $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$*, let*

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{bmatrix},$$

and let $q = \det(\mathbf{B})$*. Let* $\mathcal{C}_\mathsf{A}(\mathbf{B}) \subseteq \mathbb{Z}_q^n$ *be the code generated by the vectors of* $\mathbf{b}_i \pmod q$*. Then* $\mathcal{L}_\mathsf{A}(\mathcal{C}_\mathsf{A}(\mathbf{B})) = \mathcal{L}$*.*

Proof. Let $\breve{\mathbf{b}}_i \in \mathbb{Z}_q^n$ be the coordinate-wise reduction of $\mathbf{b}_i$ modulo q and let $\widetilde{\mathbf{b}}_i \in \mathbb{Z}^n$ be $\breve{\mathbf{b}}_i$ considered as an integer vector. For $1 \leqslant i \leqslant n$, let $\mathbf{q}_i \in \mathbb{Z}^n$ be the vectors with all zeros except for a q in the ith coordinate. Then, for all i,

$$\mathbf{b}_i = \widetilde{\mathbf{b}}_i + \sum_{j=1}^{n} c_{i,j}\mathbf{q}_j$$

for some integers $c_{i,j}$. By definition, $\mathcal{L}_\mathsf{A}(\mathcal{C}_\mathsf{A}(\mathbf{B}))$ is generated by the vectors $\widetilde{\mathbf{b}_1}, \ldots, \widetilde{\mathbf{b}_n}, \mathbf{q}_1, \ldots, \mathbf{q}_n$ so $\mathbf{b}_i \in \mathcal{L}_\mathsf{A}(\mathcal{C}_\mathsf{A}(\mathbf{B}))$ for all i, and therefore $\mathcal{L} \subseteq \mathcal{L}_\mathsf{A}(\mathcal{C}_\mathsf{A}(\mathbf{B}))$. Conversely, because $\mathrm{adj}(\mathbf{B}) \cdot \mathbf{B} = q\mathbf{I}_n$ (or, alternatively, see [10, Theorem 16]), each $\mathbf{q}_j \in \mathcal{L}$ and then, for all i, $\widetilde{\mathbf{b}_i} = \mathbf{b}_i - \sum_{j=1}^{n} c_{i,j} \cdot \mathbf{q}_j \in \mathcal{L}$ so we have $\mathcal{L}_\mathsf{A}(\mathcal{C}_\mathsf{A}(\mathbf{B})) \subseteq \mathcal{L}$.

Theorem 3.6 *Let* $\mathcal{L} \subseteq \mathbb{Z}^n$ *be a full-rank integer lattice with basis* $\mathbf{B}$*. Then for some* $\alpha \in (0, 1)$ *and* $t = \alpha\lambda_1(\mathcal{L})$ *there exists a polynomial time reduction from* BDD_α *with received vector* $\mathbf{r} \in \mathbb{Z}^n$ *in the* ℓ_1*-norm on* $\mathcal{L}$ *to* LeeDP_t *on some* $\mathcal{C} \subseteq \mathbb{Z}_q^n$*, for some* q*.*

Proof. Given $\mathbf{r} \in \mathbb{Z}^n$ such that $d_1(\mathbf{r}, \mathcal{L}) < \alpha\lambda_1(\mathcal{L})$, we know there exists some $\mathbf{v} \in \mathcal{L}$ such that $d_1(\mathbf{v}, \mathbf{r}) < \alpha\lambda_1(\mathcal{L})$. Let $\mathcal{C} \subseteq \mathbb{Z}_q^n$ be the code obtained (as above) by the reducing the lattice modulo $q = \det(\mathbf{B})$ where we represent coordinates of the vectors in $\mathbb{Z}_q^n$ with integers between $\lceil -q/2 \rceil$ and $\lfloor q/2 \rfloor$ (omitting the value $\lceil -q/2 \rceil$ in the case that q is even). Let $\breve{\mathbf{r}} = \mathbf{r} \pmod q$ and $\breve{\mathbf{v}} = \mathbf{v} \pmod q$. Then $d_L(\breve{\mathbf{v}}, \breve{\mathbf{r}}) \leqslant d_1(\mathbf{v}, \mathbf{r}) < t$.

Given input $\mathcal{C}$ and $\breve{\mathbf{r}}$, LeeDP_t outputs a codeword $\breve{\mathbf{c}} \in \mathcal{C}$ such that $d_L(\breve{\mathbf{c}}, \breve{\mathbf{r}}) < t$. Now, consider $\breve{\mathbf{e}} = \breve{\mathbf{r}} - \breve{\mathbf{c}}$, and let $\widetilde{\mathbf{e}}, \widetilde{\mathbf{r}}, \widetilde{\mathbf{c}} \in \mathbb{Z}^n$ be the vectors $\breve{\mathbf{e}}, \breve{\mathbf{r}}, \breve{\mathbf{c}}$ considered as an integer vectors, and set $\mathbf{s} = \mathbf{r} - \widetilde{\mathbf{e}}$. By Lemma 3.5, we know that $\widetilde{\mathbf{s}} = \widetilde{\mathbf{r}} - \widetilde{\mathbf{e}} \in \mathcal{L}$ and $\mathbf{q}_1, \ldots, \mathbf{q}_n \in \mathcal{L}$. Additionally, we know that there exist $c_1, \ldots, c_n \in \mathbb{Z}$ such that $\mathbf{r} = \widetilde{\mathbf{r}} + \sum_{i=1}^{n} c_i\mathbf{q}_i$. Then

$$\mathbf{r} - \widetilde{\mathbf{e}} = \widetilde{\mathbf{r}} + \sum_{i=1}^{n} c_i\mathbf{q}_i - \widetilde{\mathbf{e}} = \widetilde{\mathbf{s}} + \sum_{i=1}^{n} c_i\mathbf{q}_i \in \mathcal{L}.$$

Lastly, note that

$$\|\widetilde{\mathbf{e}}\|_1 = \sum_{i=1}^{n} |\widetilde{e_i}| = \sum_{i=1}^{n} \min\{|\breve{e_i}|, |q - \breve{e_i}|\} = \mathrm{wt}_L(\breve{\mathbf{e}}) < t = \alpha\lambda_1(\mathcal{L}).$$

Thus $\mathbf{v} = \mathbf{r} - \widetilde{\mathbf{e}}$ is a valid solution to BDD_α.

4 Containment of Finite Codes in Construction A Lattices

In this section we will study when a code over $\mathbb{Z}_q$ is completely contained in the lattice generated by a given generator matrix of the code. We remark that the containment was one of the crucial factors in the FuLeakage attack [8] on the signature scheme FuLeeca [11] since this allowed them to reduce the attack complexity by reducing the lattice dimension. This would also work as a message-recovery attack on a Lee-McEliece cryptosystem if the majority of the codewords are contained in a lower dimensional sublattice of Construction A. We therefore analyze the cardinality of the intersection of the code with the lattice generated by a generator matrix of the code (which is always a sublattice of Construction Aand the Construction A lattice consists of a union of affine shifts of the sublattice).

We first introduce a fixed notation for the lattice generated by the generator matrix of the code:

Definition 4.1 (Construction $\mathsf{A}_\mathbf{G}$) *Let $\mathcal{C}$ be a linear code in $\mathbb{Z}_q^n$, let $\mathbf{G}$ be a $k \times n$ generator matrix of $\mathcal{C}$, and let $\mathbf{g}_i$ with $i \in \{1, \ldots, k\}$ be its rows. Then, the* Construction $\mathsf{A}_\mathbf{G}$ *lattice associated to $\mathbf{G}$ is given by:*

$$\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}) = \left\{ \sum_{i=1}^{k} z_i \mathbf{g}_i \ : \ z_i \in \mathbb{Z} \right\}.$$

In the FuLeeca attack in [8] it was experimentally shown that the secret codewords (very short vectors) and signatures of the scheme in Construction A are both always contained in the Construction $\mathsf{A}_\mathbf{G}$ lattice. However, this is not true in general and running BDD or SVP solvers on the whole Construction A lattice is usually not feasible. Therefore, we would like to know when $\mathcal{C}$ is contained in $\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})$, or—if not—how many elements of the code are contained in $\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})$.

We will use a generalized version of Minkowski's bound (see Theorem 2.19) to derive a lower bound on the cardinality of $\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})$:

Theorem 4.2 *Let $\mathcal{C}$ be a linear code in $\mathbb{Z}_q^n$, let $\mathbf{G}$ be a generator matrix of $\mathcal{C}$ considered in $\mathbb{Z}^{k \times n}$, and let*

$$M := \begin{cases} \frac{q-1}{2} & \text{for } q \text{ odd} \\ \frac{q}{2} - 1 & \text{for } q \text{ even} \end{cases}.$$

Then

$$|\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})| \geqslant 2m + 1,$$

where m is the largest positive integer strictly less than $\frac{(2M)^k}{2^k\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}}$.

Proof. Let $S := [-M, M]^n \cap \mathrm{Span}_{\mathbb{R}}(\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}))$ where $\mathrm{Span}_{\mathbb{R}}(\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})) \subseteq \mathbb{R}^n$ is the k-dimensional $\mathbb{R}$-subspace spanned by $\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})$. Note that in the case where q is even, due to the requirement that S be symmetric, we are omitting some possible lattice points from our lower bound by excluding those whose coordinates take values of $q/2$. By Theorem 2.18, we have that $\mathrm{Vol}_k(S) \geqslant (2M)^k$.

Let m be the largest positive integer strictly less than $\frac{(2M)^k}{2^k\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}}$. Then

$$\frac{\mathrm{Vol}_k(S)}{2^k \det(\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}))} = \frac{\mathrm{Vol}_k(S)}{2^k\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}} \geqslant \frac{(2M)^k}{2^k\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}} > m.$$

By Theorem 2.19, we know that there are then at least m pairs of non-zero vectors in $S \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}) \subseteq \mathcal{C}$. Including the zero vector gives us the bound.

Remark 4.3 *Note that the lower bound introduced in Theorem 4.2 is inversely proportional to $\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}$, i.e., it maximizes when $\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}$ is minimal. It is well-known that $\sqrt{\det(\mathbf{G}\mathbf{G}^\top)}$ is minimized for unimodular even or Type II lattices (which are closely related with self-dual codes), see e.g., [5]. This is an indication that self-dual codes might admit a large number of codewords in $\mathcal{L}_{\mathsf{A}_\mathbf{G}}$ and would hence be cryptographically insecure. Similary, the bound increases for growing q, indicating that a very large q will likely be insecure.*

In the above theorem we establish a lower bound for the cardinality of the intersection of a code $\mathcal{C}$ over $\mathbb{Z}_q^n$. Naturally, we would also like to derive an upper bound on this number. A trivial upper bound is the cardinality of $\mathcal{C}$, that is, $|\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})| \leqslant |\mathcal{C}|$. We remark that there exists a reverse Minkowski bound, which could be used to derive another upper bound—however, it turns out that doing so results in a bound above the trivial bound, which is not useful.

In general, the lower bound derived in Theorem 4.2 and the trivial upper bound are not tight, however in special cases they are. In the following examples we illustrate this fact.

Example 4.4 *Let $\mathcal{C}_1$ and $\mathcal{C}_2$ be linear codes in $\mathbb{Z}_7^2$ with generator matrices $\mathbf{G}_1 = \begin{pmatrix}1 & 1\end{pmatrix}$ and $\mathbf{G}_2 = \begin{pmatrix}1 & 2\end{pmatrix}$ respectively. Fig. 2 depicts these two codes and their corresponding lattices $\mathcal{L}_{\mathsf{A}}$ and $\mathcal{L}_{\mathsf{A}_\mathbf{G}}$. Note that $|\mathcal{C}_1 \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}_1)| = 7$ and $|\mathcal{C}_2 \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}_2)| = 3$. Both codes have 7 elements, which is also the trivial upper bound for $|\mathcal{C}_i \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}_i)|$, for $i = 1, 2$. We see that the trivial upper bound for $|\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})|$ is attained for $\mathcal{C}_1$, but not for $\mathcal{C}_2$. Now, the lower bound from Theorem 4.2 for these two cases is*

$$|\mathcal{C}_1 \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}_1)| \geqslant 2\left\lfloor \frac{3}{\sqrt{2}} \right\rfloor + 1 = 5,$$

$$|\mathcal{C}_2 \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}_2)| \geqslant 2\left\lfloor \frac{3}{\sqrt{5}} \right\rfloor + 1 = 3$$

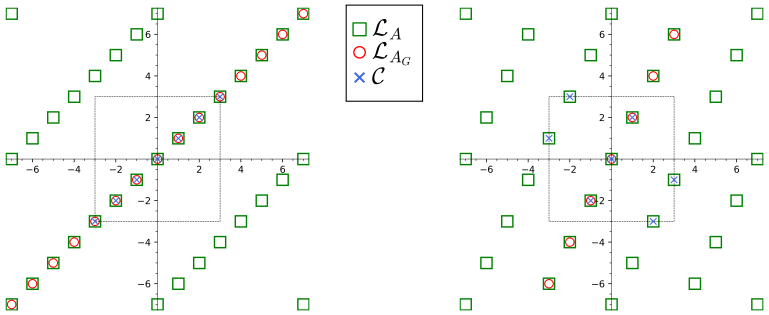

Fig. 2. Lattices $\mathcal{L}_{\mathsf{A}}$ and $\mathcal{L}_{\mathsf{A_G}}$ for $\mathcal{C}_1$ and $\mathcal{C}_2$ in Example 4.4.

respectively. We see that the lower bound for $|\mathcal{C} \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C})|$ *is attained for* $\mathcal{C}_2$*, but not for* $\mathcal{C}_1$*.*

Example 4.5 *Let* $\mathcal{C}_3 \subseteq \mathbb{Z}_7^3$ *and* $\mathcal{C}_4 \subseteq \mathbb{Z}_{13}^5$ *be linear codes with generator matrices*

$$\mathbf{G}_3 = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 2 & 3 \end{pmatrix} \text{ and } \mathbf{G}_4 = \begin{pmatrix} 3 & 1 & 2 & 5 & -4 \\ 3 & 2 & 3 & 6 & -1 \\ -1 & 2 & 5 & -5 & 6 \end{pmatrix},$$

respectively. In these two cases, both the trivial upper bound and the lower bound from Theorem 4.2 are not tight since $|\mathcal{C}_3 \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C}_3)| = 19$ *and* $|\mathcal{C}_4 \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C}_4)| = 17$ *and the bounds give*

$$49 = |\mathcal{C}_3| \geqslant |\mathcal{C}_3 \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C}_3)| \geqslant 2\left\lfloor \frac{9}{\sqrt{19}} \right\rfloor + 1 = 5,$$
$$2197 = |\mathcal{C}_4| \geqslant |\mathcal{C}_4 \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C}_4)| \geqslant 2\left\lfloor \frac{6^3}{\sqrt{23804}} \right\rfloor + 1 = 3.$$

Remark 4.6 *Both the lower bound from Theorem 4.2, and the actual number* $|\mathcal{C} \cap \mathcal{L}_{\mathsf{A_G}}(\mathcal{C})|$ *generally depend on the choice of generator matrix* $\mathbf{G}$ *and are not a code invariant (see Example 4.8). It would be useful to find a characterization or tighter bounds to understand when a generator matrix leads to a big (or small) intersection number. This seems to be a complex task, since the number of zeros (i.e., the Hamming weight), the number of different Lee weights, and the largest Lee weight (i.e., the* ℓ_∞*-norm) of the basis vectors have an impact on the wrap-around behavior (at the boundaries) of the code over* $\mathbb{Z}_q$*, when represented over* $\mathbb{Z}$*.*

To give more insight into the wrap-around behavior we describe it in the one-dimensional case. However, already for codes of dimension 2, the situation is much more complex and is left as an open problem for future work.

Proposition 4.7 *Let us represent $\mathbb{Z}_q$ centered at zero, and let $M = \lfloor q/2 \rfloor$. Further, let $\mathcal{C}$ be a one-dimensional linear code on $\mathbb{Z}_q^n$ and $\mathbf{G}$ a $1 \times n$ generator matrix. Then we have that:*

1. *if $|\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})| = q$, then all non-zero entries of $\mathbf{G}$ are ± 1. This is the only case where $\mathcal{C} \subseteq \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})$.*
2. *if $||\mathbf{G}||_\infty = t \in \{0, 1, \ldots, M\}$ then*

$$|\mathcal{C} \cap \mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C})| = \begin{cases} 2\lfloor M/t \rfloor + 1 & \text{if } q \text{ odd} \\ \lfloor M/t \rfloor + \lfloor (M-1)/t \rfloor + 1 & \text{if } q \text{ even} \end{cases}.$$

Proof. We assume for simplicity that q is odd (the even case is analogous).

1. We can easily count the $\mathbb{Z}$-multiples of an entry g_i in $\mathbf{G}$ within $\mathbb{Z}_q$ (for $\lambda \in \mathbb{Z} \backslash \{0\}$):
$$-\frac{q-1}{2} \leqslant \lambda g_i \leqslant \frac{q-1}{2} \iff |g_i| \leqslant \left| \frac{q-1}{2\lambda} \right|.$$
Thus, for $q-1$ non-zero multiples of $\mathbf{G}$ to be in $\mathbb{Z}_q^n$ we need, in particular for $\lambda = \pm(q-1)/2$, to have $|g_i| \leqslant \left|\frac{q-1}{2\lambda}\right| = 1$, which implies that $g_i \in \{0, \pm 1\}$.
2. Denote by g_i the i-th entry of $\mathbf{G}$. With a similar counting argument as above we get that for $\lambda \in \mathbb{Z}$, we have $|\lambda g_i| \leqslant |\lambda t|$ and
$$|\lambda t| \leqslant \frac{q}{2} \iff |\lambda| \leqslant \frac{q}{2t} = \frac{M}{t},$$
i.e., exactly for $\lambda \in \{-\lfloor M/t \rfloor, \ldots, \lfloor M/t \rfloor\}$ the λ-multiple of $\mathbf{G}$ is contained in $\mathcal{L}_{\mathsf{A}_\mathbf{G}}(\mathcal{C}) \cap \mathbb{Z}_q^n$, which implies the statement.

Example 4.8 *Let $\mathcal{C}_5$ be a linear code in $\mathbb{Z}_{11}^2$ and $\mathbf{G}_5 = \begin{pmatrix} 1 & 2 \end{pmatrix}$ and $\mathbf{G}_5' = \begin{pmatrix} 5 & -1 \end{pmatrix}$ be two generator matrices. It is easy to see that $|\mathcal{C}_5| = 11$ but when looking at the cardinality of the intersection with the lattice we have that*

$$\left|\mathcal{C}_5 \cap \mathcal{L}_{\mathsf{A}_{\mathbf{G}_5}}(\mathcal{C}_5)\right| = 5 \text{ and } \left|\mathcal{C}_5 \cap \mathcal{L}_{\mathsf{A}_{\mathbf{G}_5'}}(\mathcal{C}_5)\right| = 3.$$

Now, let $\mathcal{C}_3 \subseteq \mathbb{Z}_7^3$ be the same code as in Example 4.5, then $\mathcal{C}_3$ can also be generated with

$$\mathbf{G}_3' = \begin{pmatrix} 0 & 1 & 1 \\ 3 & 0 & 1 \end{pmatrix} \text{ and } \mathbf{G}_3'' = \begin{pmatrix} 0 & 2 & 2 \\ 3 & 2 & 3 \end{pmatrix}.$$

Again, when checking the cardinality of the intersection we obtain

$$\left|\mathcal{C}_3 \cap \mathcal{L}_{\mathsf{A}_{\mathbf{G}_3'}}(\mathcal{C}_3)\right| = 20 \text{ and } \left|\mathcal{C}_3 \cap \mathcal{L}_{\mathsf{A}_{\mathbf{G}_3''}}(\mathcal{C}_3)\right| = 9.$$

This illustrates the dependency on the choice of generator matrix.

5 Comparison of Error Distributions

In this section we will compare the different error distributions related to the metrics described before. First, we will compare the Hamming and the Lee metric. Then we will show the connection between the Lee metric and the Laplace distribution, which motivates us to compare the Laplace and the Gaussian distribution (to compare the behavior of the ℓ_1- and ℓ_2-norm). For the discrete distributions we will use Rényi divergence, whereas for the continuous distributions, we will use Kullback-Leibler convergence.

Let us recall from Lemma 2.3 that for a uniformly random vector $\mathbf{x} \in \mathbb{Z}_q^n$ with normalized Lee distance T, the marginal Lee distribution is given by

$$F_T(j) := \Pr(E = j) = \frac{\exp(-\beta \mathrm{wt}_L(j))}{\sum_{i=0}^{q-1} \exp(-\beta \mathrm{wt}_L(i))},$$

where β is the unique real solution to the constraint

$$T = \sum_{i=0}^{q-1} \mathrm{wt}_L(i) \Pr(E = i).$$

We can extend this distribution for length n vectors over $\mathbb{Z}_q$ by assuming that each coordinate is independent and identically distributed and we obtain for any $\mathbf{x} \in \mathbb{Z}_q^n$:

$$F_T^{(n)}(\mathbf{x}) := \prod_{i=1}^{n} F_T(x_i) = \frac{\exp(-\beta \mathrm{wt}_L(\mathbf{x}))}{\sum_{\mathbf{y} \in \mathbb{Z}_q^n} \exp(-\beta \mathrm{wt}_L(\mathbf{y}))}.$$

5.1 Lee vs. Hamming distribution

Remember that for a given normalized Hamming weight δ, we get the following marginal distribution $H_\delta(j)$ for $j \in \mathbb{F}_q$:

$$H_\delta(j) = \begin{cases} 1-\delta & \text{if } j = 0 \\ \frac{\delta}{(q-1)} & \text{otherwise} \end{cases}.$$

Similar to the Lee metric distribution, we can extend the Hamming distribution for length n vectors over $\mathbb{Z}_q$ by assuming that each coordinate is independent and identically distributed, i.e., for each $\mathbf{x} \in \mathbb{Z}_q$

$$H_\delta^{(n)}(\mathbf{x}) = \prod_{i=1}^{n} H_\delta(x_i) = \left(\frac{\delta}{q-1}\right)^{|\mathrm{Supp}(\mathbf{x})|} (1-\delta)^{n-|\mathrm{Supp}(\mathbf{x})|}.$$

Theorem 5.1 *Let F_T denote the asymptotic marginal Lee distribution and let H_δ denote the asymptotic marginal Hamming distribution from Lemma 2.3, both over $\mathbb{Z}_q$. For any $0 < T < \lfloor q/2 \rfloor$, let β be the corresponding value from (2) and $c_1 := (\sum_{i=0}^{q-1} \exp(-\beta \mathrm{wt}_L(i)))^{-1}$. Let $\delta \in (0,1)$, then*

1. *the Rényi divergence of order ∞ between F_T and H_δ is given by:*

$$R_\infty(F_T||H_\delta) = \max\left\{\frac{c_1}{1-\delta}, \frac{c_1 e^{-\beta\nu(\beta)}(q-1)}{\delta}\right\},$$

where $\nu(\beta) = 1$ if $\beta \geqslant 0$ and $\nu(\beta) = \lfloor q/2 \rfloor$ if $\beta < 0$.

2. *for any given T, the Rényi divergence $R_\infty(F_T||H_\delta)$ is minimized at $\delta = \frac{e^{-\beta\nu(\beta)}(q-1)}{1+e^{-\beta\nu(\beta)}(q-1)}$, giving the lower bound:*

$$R_\infty(F_T||H_\delta) \geqslant c_1 + c_1 e^{-\beta\nu(\beta)}(q-1). \tag{3}$$

3. *Assuming the coordinates are independent and identically distributed, the Rényi divergence of F_T and H_δ for length n vectors over $\mathbb{Z}_q$ is as follows:*

$$R_\infty\left(F_T^{(n)}||H_\delta^{(n)}\right) \geqslant \left(c_1 + c_1 e^{-\beta\nu(\beta)}(q-1)\right)^n.$$

Proof. We have

$$\begin{aligned} R_\infty(F_T||H_\delta) &:= \max_{\lceil -q/2 \rceil \leqslant j \leqslant \lfloor q/2 \rfloor} \frac{F_T(j)}{H_\delta(j)} \\ &= \max\left\{\frac{c_1}{1-\delta}, \max_{1 \leqslant j \leqslant \lfloor q/2 \rfloor} \frac{c_1 e^{-\beta j}(q-1)}{\delta}\right\}. \end{aligned}$$

It is easy to see that the second term is maximal at $j = 1$ if $\beta \geqslant 0$, or $j = \lfloor q/2 \rfloor$ if $\beta < 0$. Hence we get

$$R_\infty(F_T||H_\delta) = \max\left\{\frac{c_1}{1-\delta}, \frac{c_1 e^{-\beta\nu(\beta)}(q-1)}{\delta}\right\},$$

where $\nu(\beta) = 1$ if $\beta \geqslant 0$ and $\nu(\beta) = \lfloor q/2 \rfloor$ if $\beta < 0$.

Note that the first term increases as δ increases, whereas the second term decreases as δ increases. The maximum between the two minimizes when the first term is equal to the second term, i.e., $\frac{c_1}{1-\delta} = \frac{c_1 e^{-\beta\nu(\beta)}(q-1)}{\delta}$ or $\delta = \frac{e^{-\beta\nu(\beta)}(q-1)}{1+e^{-\beta\nu(\beta)}(q-1)}$. Thus, we get

$$R_\infty(F_T||H_\delta) \geqslant c_1 + c_1 e^{-\beta\nu(\beta)}(q-1).$$

For vectors of length n, with independent and identically distributed coordinates, the inequality follows using the multiplicative property of Rényi divergence (Lemma 2.23), i.e., $R_\infty\left(F_T^{(n)}||H_\delta^{(n)}\right) = \prod_{i=1}^n R_\infty(F_T||H_\delta)$.

Remark 5.2 *Plugging in the values for β and c_1 for $q \in \{2, 3\}$ in Equation (3), we get*

$$R_\infty(F_T||H_\delta) \geqslant 1$$

showing that the bound is tight (since the two distributions coincide) in these cases.

As q increases, we can observe that the distributions F_T and H_δ become quite different from each other. We confirm this behavior by showing that the Rényi divergence between them goes to infinity as q goes to infinity.

Proposition 5.3 *Let $T \in (0, \lfloor q/2 \rfloor]$ and $\delta \in (0,1)$ be constant (with respect to q) real numbers. Then, the Rényi divergence $R_\infty(F_T||H_\delta)$ goes to infinity as q goes to infinity.*

Proof. Let β be the corresponding value from (2) and $c_1 := (\sum_{i=0}^{q-1} \exp(-\beta \mathrm{wt}_L(i)))^{-1}$. As q goes to infinity, we can rewrite (2) as follows:

$$\begin{aligned} T &= \lim_{q\to\infty} \frac{\sum_{i=0}^{q} \mathrm{wt}_L(i) \exp(-x\ \mathrm{wt}_L(i))}{\sum_{i=0}^{q} \exp(-x\ \mathrm{wt}_L(i))} \\ &= \lim_{M\to\infty} \frac{2\sum_{i=1}^{M} i \exp(-x\ i)}{1 + 2\sum_{i=1}^{M} \exp(-x\ i)} \\ &= \frac{2\exp(-x)}{(1-\exp(-x))^2} \cdot \frac{1-\exp(-x)}{\exp(-x)+1} \\ &= \frac{2\exp(x)}{\exp(2x)-1}. \end{aligned}$$

Therefore, as q goes to infinity, β converges to the positive real solution of $T = 2\exp(x)/(\exp(2x)-1)$, i.e.,

$$\exp(\beta) \to \frac{1+\sqrt{1+T^2}}{T} \text{ as } q \to \infty.$$

Using this, it is easy to check that, as $q \to \infty$,

$$c_1 \to \frac{\exp(\beta)-1}{\exp(\beta)+1} = \frac{(1-T)+\sqrt{1+T^2}}{(1+T)+\sqrt{1+T^2}}.$$

As a conclusion, we note that for a fixed T (constant with respect to q), both $e^{-\beta}$ and c_1 converge to a constant as q tends to infinity. Hence, $R_\infty(F_T||H_\delta) \geqslant c_1 + c_1 e^{-\beta}(q-1)$ diverges as q goes to infinity. This confirms the intuition that the Lee metric diverges away from the Hamming metric with growing q.

5.2 Laplace vs. Gaussian distribution

We motivate this section by first showing the connection between the Lee metric and the discrete Laplace distribution. Recall that the discrete Laplace distribution over $\mathbb{Z}$ with parameter $b > 0$ is defined as

$$\mathrm{Lap}_{\mathbb{Z},b}(x) = \frac{\exp\left(-|x|/b\right)}{\sum_{y\in\mathbb{Z}} \exp(-|y|/b)}.$$

On the other hand, by considering the representation of $\mathbb{Z}_q$ centered at origin, the marginal Lee distribution $F_T(j)$ can be rewritten as

$$F_T(j) = \frac{\exp(-\beta\,|j|)}{\sum_{i\in\mathbb{Z}_q}\exp(-\beta\,|i|)}, \tag{4}$$

for each $j \in \mathbb{Z}_q = \{-\lfloor (q-1)/2 \rfloor, \ldots, \lfloor q/2 \rfloor\}$. We can observe that the above two distributions would coincide when $b = 1/\beta$ and q goes to infinity. We can deduce the same for length n vectors as well, by assuming that for each coordinate the distributions are independent and identical.

Lemma 5.4 *Let F_T be the Lee distribution over $\mathbb{Z}_q^n$ from Lemma 2.3, and let $\mathrm{Lap}_{\mathbb{Z}^n,b}$ be the discrete Laplace distribution over $\mathbb{Z}^n$. Assuming that each coordinate is independent and identically distributed, we get that*

$$\lim_{q\to\infty} F_T(\mathbf{x}) = \mathrm{Lap}_{\mathbb{Z}^n,\frac{1}{\beta}}(\mathbf{x}),$$

for every $\mathbf{x} \in \mathbb{Z}_q^n$.

Proof. From Equation (4), we have

$$\begin{aligned} F_T(\mathbf{x}) &= \frac{1}{\left(\sum_{i\in\mathbb{Z}_q}\exp(-\beta\,|i|)\right)^n}\prod_{j=1}^{n}\exp(-\beta\,|x_j|) \\ &= \frac{1}{\sum_{\mathbf{y}\in\mathbb{Z}_q^n}\exp(-\beta\,\|\mathbf{y}\|_1)}\exp(-\beta\,\|\mathbf{x}\|_1). \end{aligned}$$

Now, if we take the limit of q to infinity, then we see that the above Lee distribution converges to the discrete Laplace distribution $\mathrm{Lap}_{\mathbb{Z},b}$ with parameter $b = 1/\beta$.

The above lemma says that the Lee metric distribution is close to the Laplace distribution for large q. Therefore, in the case of large q, it would make sense to compare Laplace and Gaussian distribution in order to compare Lee and Euclidean error distributions. In other words, such a comparison would give us a good understanding when lattice techniques in the ℓ_2-norm are beneficial to solve ℓ_1-norm lattice problems or Lee metric decoding problems.

We start by comparing the continuous version of Laplace and Gaussian distribution by computing the Kullback-Leibler divergence between them.

Theorem 5.5 *The Kullback-Leibler divergence between the Laplace distribution $\mathrm{Lap}_{\mathbb{R},b}$ and the Gaussian distribution $\mathrm{D}_{\mathbb{R},\sigma}$ is given by*

$$KL(\mathrm{Lap}_{\mathbb{R},b}\,||\,\mathrm{D}_{\mathbb{R},\sigma}) = \log\left(\frac{\sigma\sqrt{\pi/2}}{b}\right) + \frac{b^2}{\sigma^2} - 1.$$

For vectors of length n, we assume the distributions are independent and identical and obtain:

$$KL(\mathrm{Lap}_{\mathbb{R}^n,b}\,||\,\mathrm{D}_{\mathbb{R}^n,\sigma}) = n\left(\log\left(\frac{\sigma\sqrt{\pi/2}}{b}\right) + \frac{b^2}{\sigma^2} - 1\right).$$

Proof. We compute the KL divergence as follows:

$$\begin{aligned} KL(\mathrm{Lap}_{\mathbb{R},b} \,||\, \mathrm{D}_{\mathbb{R},\sigma}) &:= \int_{-\infty}^{\infty} \mathrm{Lap}_{\mathbb{R},b}(x) \log\left(\frac{\mathrm{Lap}_{\mathbb{R},b}(x)}{\mathrm{D}_{\mathbb{R},\sigma}(x)}\right) dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2b} e^{-|x|/b} \left(\log\left(\frac{1}{2b} e^{-|x|/b}\right) - \log\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}\right)\right) dx \\ &= \int_{-\infty}^{0} \frac{1}{2b} e^{x/b} \left(\frac{x^2}{2\sigma^2} + \frac{x}{b} + \log\left(\frac{\sigma\sqrt{\pi/2}}{b}\right)\right) dx \\ &+ \int_{0}^{\infty} \frac{1}{2b} e^{-x/b} \left(\frac{x^2}{2\sigma^2} - \frac{x}{b} + \log\left(\frac{\sigma\sqrt{\pi/2}}{b}\right)\right) dx. \end{aligned}$$

Using integration by parts multiple times, we obtain

$$KL(\mathrm{Lap}_{\mathbb{R},b} \,||\, \mathrm{D}_{\mathbb{R},\sigma}) = \log\left(\frac{\sigma\sqrt{\pi/2}}{b}\right) + \frac{b^2}{\sigma^2} - 1.$$

For vectors of length n, the result follows using the additive property of KL divergence (Lemma 2.23), i.e., $KL(\mathrm{Lap}_{\mathbb{R}^n,b} \,||\, \mathrm{D}_{\mathbb{R}^n,\sigma}) = \sum_{i=1}^{n} KL(\mathrm{Lap}_{\mathbb{R},b} \,||\, \mathrm{D}_{\mathbb{R},\sigma})$.

Using standard analytical tools, we can find the minimum KL divergence between the Laplace and Gaussian distribution.

Corollary 5.6 *For any given $b > 0$, the KL divergence $KL(\mathrm{Lap}_{\mathbb{R},b} \,||\, \mathrm{D}_{\mathbb{R},\sigma})$ has exactly one local minimum at $\sigma = b\sqrt{2}$. In this case we obtain*

$$KL(\mathrm{Lap}_{\mathbb{R},b} \,||\, \mathrm{D}_{\mathbb{R},\sigma}) = \frac{\log(\pi) - 1}{2} \approx 0.072365.$$

The above corollary shows that for any given Laplace distribution with parameter b, the Gaussian distribution with variance $\sigma = b\sqrt{2}$ is the closest. Moreover, the minimal divergence value is independent of the parameter b. Therefore, even though $b = \frac{1}{\beta}$ when using the Laplace distribution as an approximation of the Lee distribution, and β depends on q and δ, the similarity of the ℓ_2-norm of lattice vectors and the Lee weight of codewords is generally independent of q and δ (for very large q).

Remark 5.7 *Theorem 5.5 shows that the Kullback-Leibler divergence between the continuous Gaussian and the Laplacian distribution (which we take as a good approximation of the Lee distribution for large q) grows linearly in n. This indicates that as n increases the Laplace distribution diverges away from the Gaussian distribution.*

Recall that we took a continuous (Laplace) distribution to represent a discrete (Lee) distribution. Naturally, it would be a better approximation to take the discrete Laplace distribution as an approximation of the Lee distribution. To finalize this section, we will compute the divergence between discrete Laplace and the discrete Gaussian distribution to illustrate the behavior of them.

Remark 5.8 *Note that the Rényi divergence between discrete Laplace* $\mathrm{Lap}_{\mathbb{Z},b}$ *and discrete Gaussian* $\mathrm{D}_{\mathbb{Z},\sigma}$ *is* ∞*, for any given parameters* b *and* σ*. This easily follows from the following calculations:*

$$R_2(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) = \sum_{x\in\mathbb{Z}} \frac{\left(\frac{1}{S_1(b)} \exp\left(-|x|/b\right)\right)^2}{\left(\frac{1}{S_2(\sigma)} \exp\left(-x^2/2\sigma^2\right)\right)},$$

where $S_1(b) = \sum_{y\in\mathbb{Z}} e^{-|y|/b}$ *and* $S_2(\sigma) = \sum_{y\in\mathbb{Z}} e^{-y^2/2\sigma^2}$*. This evaluates to:*

$$R_2(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) = \frac{S_2(\sigma)}{(S_1(b))^2}\left(1 + 2\sum_{x\geqslant 1} \exp\left(\frac{x^2}{2\sigma^2} - \frac{2x}{b}\right)\right).$$

It is easy to see that the summation goes to infinity, because $\exp\left(\frac{x^2}{2\sigma^2} - \frac{2x}{b}\right)$ *goes to infinity as* x *goes to infinity. This also implies that the Rényi divergence* R_a *is infinity for all order* $a \in (1,\infty]$*, because* $R_a(F||G)$ *is non-decreasing as a function of* a*.*

Unlike Rényi divergence, the Kullback-Leibler divergence between discrete Laplace and discrete Gaussian is finite.

Theorem 5.9 *Let* $b, \sigma > 0$ *be real numbers. The Kullback-Leibler divergence between the discrete Laplace distribution* $\mathrm{Lap}_{\mathbb{Z},b}$ *and the discrete Gaussian distribution* $\mathrm{D}_{\mathbb{Z},\sigma}$ *is given by*

$$KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) = \log\left(\frac{S_2(\sigma)}{S_1(b)}\right) + \frac{1}{S_1(b)}\left(\frac{e^{1/b}(e^{1/b}+1)}{(e^{1/b}-1)^3\sigma^2} - \frac{2e^{1/b}}{b(e^{1/b}-1)^2}\right),$$

where $S_1(b) = \sum_{y\in\mathbb{Z}} e^{-|y|/b}$ *and* $S_2(\sigma) = \sum_{y\in\mathbb{Z}} e^{-y^2/2\sigma^2}$*.*

Proof. We compute the KL divergence as follows:

$$\begin{aligned} KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) &:= \sum_{x\in\mathbb{Z}} \mathrm{Lap}_{\mathbb{Z},b}(x) \log\left(\frac{\mathrm{Lap}_{\mathbb{Z},b}(x)}{\mathrm{D}_{\mathbb{Z},\sigma}(x)}\right) \\ &= \sum_{x\in\mathbb{Z}} \frac{e^{-|x|/b}}{S_1(b)}\left(\log\left(\frac{e^{-|x|/b}}{S_1(b)}\right) - \log\left(\frac{e^{-x^2/2\sigma^2}}{S_2(\sigma)}\right)\right), \end{aligned}$$

where $S_1(b) = \sum_{y\in\mathbb{Z}} e^{-|y|/b}$ and $S_2(\sigma) = \sum_{y\in\mathbb{Z}} e^{-y^2/2\sigma^2}$. Note that both $S_1(b)$ and $S_2(\sigma)$ are positive finite real numbers for any given $b, \sigma > 0$. In particular, $S_1(b) = \frac{e^{1/b}+1}{e^{1/b}-1}$, and $S_2(\sigma)$ is an evaluation of the theta function θ_3[5].

$$\begin{aligned} KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) &= \frac{1}{S_1(b)} \sum_{x\in\mathbb{Z}} e^{-|x|/b}\left(\log\left(\frac{S_2(\sigma)}{S_1(b)}\right) + \frac{x^2}{2\sigma^2} - \frac{|x|}{b}\right) \\ &= \frac{1}{S_1(b)}\left(\log\left(\frac{S_2(\sigma)}{S_1(b)}\right) + 2\sum_{x\geqslant 1} e^{-x/b}\left(\frac{x^2}{2\sigma^2} - \frac{x}{b} + \log\left(\frac{S_2(\sigma)}{S_1(b)}\right)\right)\right) \end{aligned}$$

[5] We note that $S_2(\sigma) = \theta_3(0, e^{-1/2\sigma^2})$, where $\theta_3(u,q) = 1 + 2\sum_{n=1}^{\infty} q^{n^2}\cos(2nu)$ is a theta function.

To proceed forward, we apply the following identities:

$$\sum_{x\geqslant 1} e^{-x/b} = \frac{1}{e^{1/b}-1}; \sum_{x\geqslant 1} xe^{-x/b} = \frac{e^{1/b}}{(e^{1/b}-1)^2}; \sum_{x\geqslant 1} x^2e^{-x/b} = \frac{e^{1/b}(e^{1/b}+1)}{(e^{1/b}-1)^3},$$

and obtain

$$KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma}) = \log\left(\frac{S_2(\sigma)}{S_1(b)}\right) + \frac{1}{S_1(b)}\left(\frac{e^{1/b}(e^{1/b}+1)}{(e^{1/b}-1)^3\sigma^2} - \frac{2e^{1/b}}{b(e^{1/b}-1)^2}\right).$$

Lower bound on KL divergence between discrete Laplace and discrete Gaussian: Given a fixed $b > 0$, we can numerically find $\sigma > 0$ that minimizes the KL divergence between $\mathrm{Lap}_{\mathbb{Z},b}$ and $\mathrm{D}_{\mathbb{Z},\sigma}$. Using Theorem 5.9, we can observe that for a fixed b the KL divergence $KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma})$ is a continuous and differentiable function of $\sigma \in (0,\infty)$. Moreover, it has exactly one minimum point $\sigma_{\min} \in (0,\infty)$, which is the positive real root of

$$\frac{1}{S_2(\sigma)}\frac{\partial S_2(\sigma)}{\partial \sigma} - \frac{2e^{1/b}}{(e^{1/b}-1)^2\sigma^3} = 0. \tag{5}$$

In Table 1, we provide some numerical estimates of the minimum point $\sigma_{\min}$ and the corresponding KL divergence, for different given values of b. We note that the minimum KL divergence converges to $\frac{\log(\pi)-1}{2} \approx 0.072365$ as the Laplace width b increases (see Figure 3). Recall that the constant $\frac{\log(\pi)-1}{2}$ is the minimum KL divergence between continuous Laplace and Gaussian distributions, as seen in Corollary 5.6. Thus, similar to the continuous case, we can conclude that the Laplace distribution diverges from the Gaussian distribution for growing length n.

Laplace width b	Minimum point $\sigma_{\min}$	Minimum divergence $KL(\mathrm{Lap}_{\mathbb{Z},b} \,\|\, \mathrm{D}_{\mathbb{Z},\sigma})$
0.1	0.223609	7.83 x 10^{-8}
0.5	0.607753	0.0886053
1.0	1.35696	0.101332
2.0	2.79918	0.0819178
4.0	5.64215	0.0749139
8.0	11.3063	0.0730125

Table 1. Numerical estimates of the minimal KL divergence between discrete Laplace $\mathrm{Lap}_{\mathbb{Z},b}$ and discrete Gaussian distribution $\mathrm{D}_{\mathbb{Z},\sigma}$, for various given values of parameter b. Here, the minimum point $\sigma_{\min}$ is a solution to Equation (5) that minimizes $KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma})$ for the given b.

6 Conclusions

Due to the recent development in Lee metric code-based cryptography, in particular the lattice-based attack on the NIST submission FuLeeca, we analyzed

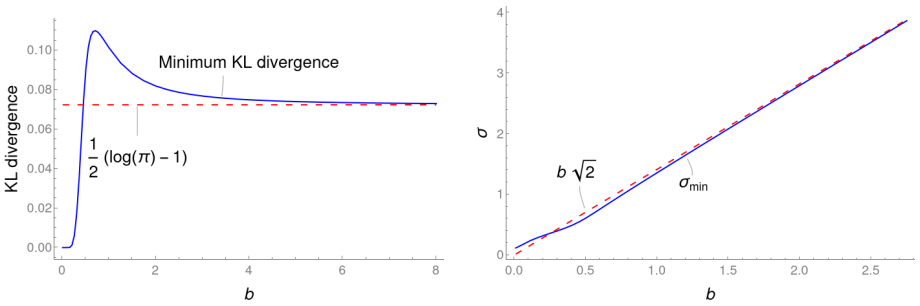

Fig. 3. Numerical estimates of the minimal KL divergence between Laplace and Gaussian distribution. In the left figure, we plot the minimum $KL(\mathrm{Lap}_{\mathbb{Z},b} \,|\, \mathrm{D}_{\mathbb{Z},\sigma})$ (solid blue line) as a function of b and compare it with the constant $\frac{\log(\pi)-1}{2}$ (dashed red line) corresponding to the continuous case (Corollary 5.6). In the right figure, we plot the corresponding $\sigma_{\min}$ where the minimum divergence is achieved. Here again we compare $\sigma_{\min}$ with the minimum sigma $\sigma = b\sqrt{2}$ obtained in the continuous case (Corollary 5.6).

the connection of Lee metric code-based cryptosystems and the corresponding lattice problems in the ℓ_1- and the ℓ_2-norm.

In particular, we showed that there are polynomial time reductions in both directions between the Lee decoding problem over $\mathbb{Z}_q$ and the unique shortest vector problem (via the bounded distance decoding problem) over $\mathbb{Z}$ with respect to the ℓ_1-norm (where for the reduction from LeeDP_t to BDD_α we require that $t < q$). Moreover, we gave a lower bound on the number of points that are contained in the lattice generated by a given code basis, showing that this number depends on q and the actual choice of basis. The bound suggests that the success likelihood of an attack by finding vectors in this lower dimensional lattice increases for growing q. Furthermore, we studied the divergence behavior of various probability distributions connected to the Lee and Hamming weight, as well as the ℓ_1- and ℓ_2-norm. Our results show that the behavior of the Lee metric diverges from the one of the Hamming metric for growing modulus q, and that the Laplace distribution diverges from the Gaussian distribution for growing vector length n (for large q).

These results show that when (Hamming and) lattice techniques can be used to break Lee metric code-based cryptosystems. Hence, this can tell us which parameters should be avoided when designing new public key encryption schemes or digital signatures using Lee metric error correcting codes. In particular, when q is chosen extremely small, then Hamming-based coding techniques can be used to attack Lee metric cryptographic schemes. On the other hand, for large q, ℓ_1-lattice techniques might be applicable, however, using ℓ_2-techniques will most likely not work. Moreover, when q is very large, then the idea of the FuLeakage attack of using the lattice generated by a basis of the code (instead of the whole Construction A lattice) has a higher success probability than for smaller q. This means that in general, q should be chosen large enough such that Hamming

attacks are not applicable but also small enough that a large part of the $\mathcal{L}_{\mathbf{A}_\mathbf{G}}$ lattice is not contained in $\mathcal{L}_\mathbf{A}$. Furthermore, n should be chosen large enough that the ℓ_2-norm is not a good approximation of the ℓ_1-norm. The exact parameters depend on the actual cryptosystem, the relationship of modulus, length, and minimum distance of the code and need to be investigated when designing a new Lee metric code-based cryptosystem.

References

1. Antonio, C.d.A., Jorge, G.C., Costa, S.I.: Decoding q-ary lattices in the lee metric. In: 2011 IEEE Information Theory Workshop. pp. 220–224. IEEE (2011)
2. Bariffi, J., Bartz, H., Liva, G., Rosenthal, J.: On the properties of error patterns in the constant Lee weight channel. In: International Zurich Seminar on Information and Communication (IZS) (2022)
3. Bariffi, J., Khathuria, K., Weger, V.: Information set Decoding for Lee-metric codes using restricted balls. In: Code-Based Cryptography Workshop. pp. 110–136. Springer (2022)
4. Chailloux, A., Debris-Alazard, T., Etinski, S.: Classical and quantum algorithms for generic syndrome decoding problems and applications to the Lee metric. In: Post-Quantum Cryptography: 12th International Workshop, PQCrypto 2021, Daejeon, South Korea, July 20–22, 2021, Proceedings 12. pp. 44–62. Springer (2021)
5. Conway, J.H., Sloane, N.J.A.: Lattices with few distances. Journal of number theory **39**(1), 75–90 (1991). https://doi.org/10.1016/0022-314X(91)90035-A
6. van Erven, T., Harremos, P.: Rényi Divergence and Kullback-Leibler Divergence. IEEE Transactions on Information Theory **60**(7), 3797–3820 (2014). https://doi.org/10.1109/TIT.2014.2320500
7. Horlemann-Trautmann, A.L., Weger, V.: Information set decoding in the Lee metric with applications to cryptography. Advances in Mathematics of Communications **15**(4) (2021)
8. Hörmann, F., van Woerden, W.: Fuleakage: Breaking FuLeeca by Learning Attacks. Cryptology ePrint Archive, Paper 2024/353 (2024), https://eprint.iacr.org/2024/353
9. Lyubashevsky, V., Micciancio, D.: On bounded distance decoding, unique shortest vectors, and the minimum distance problem. In: Annual International Cryptology Conference. pp. 577–594. Springer (2009)
10. Micciancio, D.: Lecture notes on Introduction to Lattices, https://cseweb.ucsd.edu/classes/wi12/cse206A-a/lec1.pdf
11. Ritterhoff, S., Maringer, G., Bitzer, S., Weger, V., Karl, P., Schamberger, T., Schupp, J., Wachter-Zeh, A.: FuLeeca: A Lee-Based Signature Scheme. In: Code-Based Cryptography - 11th International Workshop, CBCrypto 2023. Lecture Notes in Computer Science, vol. 14311, pp. 56–83. Springer (2023). https://doi.org/10.1007/978-3-031-46495-9_4
12. Rush, J.A., Sloane, N.J.A.: An improvement to the Minkowski-Hiawka bound for packing superballs. Mathematika **34**(1), 8–18 (1987). https://doi.org/https://doi.org/10.1112/S0025579300013231
13. Sakzad, A., Sadeghi, M.: On cycle-free lattices with high rate label codes. Adv. Math. Commun. **4**(4), 441–452 (2010). https://doi.org/10.3934/AMC.2010.4.441

14. Shmonin, G.: Lecture notes on Minkowski's theorem and its applications, https://www.epfl.ch/labs/disopt/wp-content/uploads/2018/09/minkowski.pdf
15. Vaaler, J.: A geometric inequality with applications to linear forms. Pacific Journal of Mathematics **83**(2), 543–553 (1979)
16. Weger, V., Khathuria, K., Horlemann, A.L., Battaglioni, M., Santini, P., Persichetti, E.: On the hardness of the Lee syndrome decoding problem. Advances in Mathematics of Communications **18**(1), 233–266 (2024). https://doi.org/10.3934/amc.2022029

A Proof of Minkowski's convex body theorem (Theorem 2.19)

In order to prove Theorem 2.19, we use Blichfeldt's theorem on non-full dimensional lattices. In the following, our proofs are based on the proofs from [10, Theorem 20-21] and [14, Theorem 5-6].

Theorem A.1 (Blichfeldt). *Let $\mathcal{L}$ be a k-dimensional lattice in $\mathbb{R}^n$ and $S \subseteq \mathrm{Span}_{\mathbb{R}}(\mathcal{L})$ be a convex set symmetric about the origin (i.e., $\mathbf{x} \in S$ implies $-\mathbf{x} \in S$). Suppose that $\mathrm{Vol}_k(S) > m \cdot \det(\mathcal{L})$, for some integer m. Then, there are $m+1$ vectors $\mathbf{z}_1, \ldots, \mathbf{z}_{m+1}$ in S such that $\mathbf{z}_i - \mathbf{z}_j \in \mathcal{L}$ for each i, j.*

Proof. Let $\mathbf{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_k\}$ be a basis of $\mathcal{L}$ and $\Pi(\mathbf{B})$ be the fundamental parallelepiped associated to $\mathbf{B}$ defined as $\Pi(\mathbf{B}) := \left\{\sum_{i=1}^{k} x_i \mathbf{b}_i \mid x_i \in [0,1)\right\}$. Consider the sets $S_{\mathbf{x}} := S \cap \{\mathbf{y} + \mathbf{x} \mid \mathbf{y} \in \Pi(\mathbf{B})\}$ for each $\mathbf{x} \in \mathcal{L}$. We note that these sets form a partition of S, i.e., they are pairwise disjoint and $S = \bigcup_{\mathbf{x}\in\mathcal{L}} S_{\mathbf{x}}$. Thus, we have $\mathrm{Vol}_k(S) = \sum_{\mathbf{x}\in\mathcal{L}} \mathrm{Vol}_k(S_{\mathbf{x}})$.

Now consider the shifted sets $S_{\mathbf{x}} - \mathbf{x} := \{\mathbf{y} - \mathbf{x} \mid \mathbf{y} \in S_{\mathbf{x}}\}$. We note that $S_{\mathbf{x}} - \mathbf{x} = (S - \mathbf{x}) \cap \Pi(\mathbf{B})$. Now, since $\mathrm{Vol}_k(S_{\mathbf{x}}) = \mathrm{Vol}_k(S_{\mathbf{x}} - \mathbf{x})$, we have that

$$\sum_{\mathbf{x}\in\mathcal{L}} \mathrm{Vol}_k(S_{\mathbf{x}} - \mathbf{x}) = \sum_{\mathbf{x}\in\mathcal{L}} \mathrm{Vol}_k(S_{\mathbf{x}}) = \mathrm{Vol}_k(S) > m \cdot \det(\mathcal{L}) = m \cdot \mathrm{Vol}(\Pi(\mathbf{B})).$$

From $\sum_{\mathbf{x}\in\mathcal{L}} \mathrm{Vol}_k(S_{\mathbf{x}} - \mathbf{x}) > m \cdot \mathrm{Vol}(\Pi(\mathbf{B}))$ and $S_{\mathbf{x}} - \mathbf{x} \subseteq \Pi(\mathbf{B})$, we deduce that there exist $m+1$ distinct points $\mathbf{x}_1, \ldots, \mathbf{x}_{m+1} \in \mathcal{L}$ such that $\bigcap_{i=1}^{m+1}(S_{\mathbf{x}_i} - \mathbf{x}_i)$ is non-empty. Let $\mathbf{y} \in \bigcap_{i=1}^{m+1}(S_{\mathbf{x}_i} - \mathbf{x}_i)$ and $\mathbf{z}_i = \mathbf{y} + \mathbf{x}_i \in S_{\mathbf{x}_i} \subseteq S$ for each $i \in \{1, \ldots, m+1\}$. Thus, we have $m+1$ vectors $\mathbf{z}_1, \ldots, \mathbf{z}_{m+1} \in S$ such that $\mathbf{z}_i - \mathbf{z}_j = \mathbf{x}_i - \mathbf{x}_j \in \mathcal{L}$ for each i, j.

Now, we prove Theorem 2.19 as a corollary to Blichfeldt's theorem.

Theorem 2.19 *Let $\mathcal{L}$ be a k-dimensional lattice in $\mathbb{R}^n$ and let $S \subseteq \mathrm{Span}_{\mathbb{R}}(\mathcal{L})$ be a convex set symmetric about the origin (i.e., $\mathbf{x} \in S$ implies $-\mathbf{x} \in S$). Suppose that $\mathrm{Vol}_k(S) > m \cdot 2^k \cdot \det(\mathcal{L})$, for some integer m. Then there are m different pairs of vectors $\pm\mathbf{z}_1, \ldots, \pm\mathbf{z}_m \in S \cap \mathcal{L}\backslash\{0\}$.*

Proof. Consider the set $\frac{1}{2}S = \{\mathbf{x} \mid 2\mathbf{x} \in S\}$, then it is easy to note that $\mathrm{Vol}_k\left(\frac{1}{2}S\right) = \frac{1}{2^k}\mathrm{Vol}_k(S)$. This follows since $S \subseteq \mathrm{Span}_{\mathbb{R}}(\mathcal{L})$ is contained in a k-dimensional subspace of $\mathbb{R}^n$ and we can apply an orthogonal transformation to embed S in $\mathbb{R}^k$ without changing its volume.

Now, since $\mathrm{Vol}_k\left(\frac{1}{2}S\right) > m \cdot \det(\mathcal{L})$, we apply Theorem A.1 to obtain $m+1$ vectors $\frac{1}{2}\mathbf{x}_1, \ldots, \frac{1}{2}\mathbf{x}_{m+1} \in \frac{1}{2}S$ such that $\frac{1}{2}\mathbf{x}_i - \frac{1}{2}\mathbf{x}_j \in \mathcal{L}$ for all i, j. We assume that $\mathbf{x}_1$ is the smallest vector with respect to the lexicographic order $\prec$.

Define $\mathbf{z}_i = \frac{1}{2}\mathbf{x}_{i+1} - \frac{1}{2}\mathbf{x}_1 \in \mathcal{L}$ for each $i \in \{1, \ldots, m\}$. Clearly, $\mathbf{z}_i$'s are distinct vectors, and since $0 \prec \mathbf{z}_i$ for all i, we have $\mathbf{z}_i \neq -\mathbf{z}_j$ for all i, j. Finally, since S is convex and symmetric, $\mathbf{z}_i = \frac{1}{2}\mathbf{x}_{i+1} + \frac{1}{2}(-\mathbf{x}_1) \in S$ for all i.