The Rényi Smoothing Parameter and Its Applications in Lattice-Based Cryptography

Cong Ling[1], Laura Luzzi[2], and Hao Yan[1]

[1] Imperial College London
{h.yan22, c.ling}@imperial.ac.uk
[2] CY Cergy Paris Université, ENSEA, CNRS
laura.luzzi@ensea.fr

Abstract. The smoothing parameter is a cornerstone concept in lattice-based cryptography. Traditionally defined using the L^∞ distance, this standard formulation can be overly stringent compared to the L^1 (or statistical) distance more commonly employed in cryptographic contexts. Recent work has proposed relaxed definitions based on Kullback-Leibler (KL) divergence and L^1 distance, thereby loosening the constraints required for the distance to vanish. However, the additive nature of the L^1 distance can be limiting for cryptographic applications where probability preservation is essential. In this paper, we introduce the Rényi smoothing parameter of a lattice, based on Rényi divergence, to address this limitation. The advantages of Rényi divergence in cryptographic settings are well known thanks to its multiplicative nature. The Rényi smooting parameter provides a tunable framework that interpolates between the L^1 and L^∞ distances, offering enhanced flexibility. We present two complementary methods to study the averaging behavior of the Rényi flatness factor: one uses classical tools such as the Minkowski-Hlawka ensemble and Rogers' formula for computing lattice function moments; the other employs Construction A lattices derived from random codes. Finally, we illustrate how this new perspective yields improvements in lattice-based cryptographic constructions.

Keywords: Lattice-based cryptography · Minkowski-Hlawka theorem · Rényi divergence · Rogers' formula · smoothing parameter

1 Introduction

The *smoothing parameter* of a lattice Λ is defined as the minimum amount of Gaussian noise that, when added to the lattice, produces a distribution that is close to uniform over the fundamental domain $\mathbb{R}^n/\Lambda$ [23]. This concept plays a central role in both the theory and practice of lattice-based cryptography. Theoretically, it underpins the security proofs of a wide range of cryptographic constructions based on lattices [23, 28]; lattice Gaussian sampling with width close to the smoothing parameter enables solutions to the approximate closest vector problem (CVP) and shortest vector problem (SVP) [1, 2, 24]. Practically,

the smoothing parameter guides concrete parameter selection in several lattice-based schemes [13], such as the FALCON signature algorithm [27]. It is also important to note that closeness to uniformity can be evaluated using different distance metrics.

In cryptography, the L^1 (statistical) distance is the standard metric. However, its additive nature can be limiting in security proofs where preservation of probabilities is crucial. To address this, Rényi divergence has been proposed as a more suitable alternative [3]. Owing to its multiplicative property, Rényi divergence offers several benefits in cryptographic settings, including tighter security reductions and more efficient parameter selection. As a result, it has seen growing adoption in lattice-based cryptography (see, e.g., [4, 21, 26, 31]).

Interestingly, the situation is reversed in the context of the smoothing parameter. Micciancio and Regev originally defined the smoothing parameter using the L^∞ distance [23], which corresponds to Rényi divergence of order ∞. This is a stricter notion than the L^1 distance, requiring a larger Gaussian parameter for the folded distribution to approximate uniformity. To mitigate this, an L^1-based version of the smoothing parameter was proposed in [7, 20], potentially enabling more efficient implementations through smaller parameter choices. Indeed, many lattice-based protocols require sampling at or above the smoothing parameter. Tighter sampling enhances security by resisting attacks that exploit large approximation factors. However, the L^1-based smoothing parameter lacks the multiplicative property, and thus inherits the same limitations as the L^1 distance.

This highlights a gap between the L^1 and L^∞ definitions of the smoothing parameter. To bridge this gap, we introduce the *Rényi smoothing parameter*[3], defined using the Rényi divergence of order $1 \leq \alpha \leq \infty$. This new definition interpolates between the two extremes and forms the central focus of this paper. In particular, choosing an order α close to 1—such as $\alpha = 2$—allows the smoothing parameter to retain both the multiplicative property and improved parameter efficiency, offering the best of both worlds.

1.1 Our Contributions

Given a lattice Λ and a noise probability density function (pdf) $\rho(\boldsymbol{x})$, define the Λ-periodic function

$$\rho_\Lambda(\boldsymbol{x}) = \sum_{\boldsymbol{\lambda}\in\Lambda} \rho(\boldsymbol{x}+\boldsymbol{\lambda}).$$

In this paper, we use the Rényi divergence to measure the closeness between the folded distribution $\rho_\Lambda(\boldsymbol{x})$ and the uniform distribution $U(\boldsymbol{x}) = 1/V(\Lambda)$, where

[3] Rényi divergence-based smoothing was previously considered in [3], but their analysis was limited to the case of Rényi divergence of order ∞, which coincides with the standard smoothing parameter [23].

$V(\Lambda)$ is the covolume of Λ, over a fundamental domain $\mathcal{R}(\Lambda)$ of Λ:

$$R_\alpha(\rho_\Lambda \| U) = \left(\int_{\mathcal{R}(\Lambda)} \frac{\rho_\Lambda(\boldsymbol{x})^\alpha}{1/V(\Lambda)^{\alpha-1}} \, d\boldsymbol{x} \right)^{\frac{1}{\alpha-1}}, \quad \alpha > 1.$$

We define $R_1(\rho_\Lambda \| U)$ and $R_\infty(\rho_\Lambda \| U)$ as the limits of $R_\alpha(\rho_\Lambda \| U)$ as $\alpha \to 1$ and $\alpha \to \infty$, respectively.

Typically, we consider the n-dimensional Gaussian distribution with parameter $s > 0$:

$$\rho_s(\boldsymbol{x}) = \frac{1}{s^n} \exp\left(-\frac{\pi \|\boldsymbol{x}\|^2}{s^2} \right).$$

In this case, we write $R_\alpha(\rho_\Lambda \| U)$ more explicitly as $R_\alpha(\rho_{s,\Lambda} \| U)$.

Definition 1. *Let $\rho_s(\boldsymbol{x})$ be a Gaussian noise pdf and $\varepsilon \geq 0$. The Rényi smoothing parameter $\eta_\varepsilon^{(\alpha)}(\Lambda)$ of a lattice Λ with order $\alpha \in [1, \infty]$ is defined as the smallest $s > 0$ such that $R_\alpha(\rho_{s,\Lambda} \| U) \leq 1 + \varepsilon$.*

In Section 2, we derive a closed-form expression for $R_\alpha(\rho_{s,\Lambda} \| U)$ when $\alpha \in \mathbb{N}_{\geq 2}$, which allows the Rényi smoothing parameter $\eta_\varepsilon^{(\alpha)}(\Lambda)$ to be explicitly defined as the smallest $s > 0$ satisfying:

$$\sum_{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{\alpha-1} \in \Lambda^*} \exp\left(-\pi s^2 (\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A} \otimes \boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_{\alpha-1}^T)^T \right) \leq (1+\varepsilon)^{\alpha-1}, \tag{1}$$

where Λ^* is the dual lattice and $\boldsymbol{A} = \boldsymbol{I}_{\alpha-1} + \boldsymbol{J}_{\alpha-1}$, with $\boldsymbol{I}_k$ the $k \times k$ identity matrix and $\boldsymbol{J}_k$ the $k \times k$ all-one matrix. Here, $\otimes$ denotes the Kronecker product.

For $\alpha = 2$, this simplifies to the particularly elegant expression:

$$\sum_{\boldsymbol{\lambda} \in \Lambda^*} \exp\left(-2\pi s^2 \|\boldsymbol{\lambda}\|^2 \right) \leq 1 + \varepsilon.$$

Interestingly, this resembles the expression for the standard L^∞ smoothing parameter, defined as [23]:

$$\sum_{\boldsymbol{\lambda} \in \Lambda^*} \exp\left(-\pi s^2 \|\boldsymbol{\lambda}\|^2 \right) \leq 1 + \varepsilon.$$

The only difference is a factor of 2 in the exponent. Hence, we obtain the exact relation:

$$\eta_\varepsilon^{(2)}(\Lambda) = \frac{\sqrt{2}}{2} \eta_\varepsilon^{(\infty)}(\Lambda). \tag{2}$$

This highlights the advantage of using the Rényi smoothing parameter: for order $\alpha = 2$, it yields a $\frac{\sqrt{2}}{2}$ reduction in s. Although seemingly modest, this gain can have a significant impact in lattice-based cryptography. For example, the security of the GPV signature scheme is highly sensitive to the value of s [26].

While this is encouraging, we should also be aware of the inherent limit of Rényi smoothing. As shown in Section 3, on average the smoothing parameter of a lattice with unit volume (as the lattice dimension $n \to \infty$) is given by:

$$\eta_\varepsilon^{(\alpha)}(\Lambda) \to \alpha^{-1/2(\alpha-1)} = \begin{cases} \frac{1}{\sqrt{e}}, & \text{as } \alpha \to 1; \\ \frac{1}{\sqrt{2}}, & \text{for } \alpha = 2; \\ \frac{1}{\sqrt[4]{3}}, & \text{for } \alpha = 3; \\ \vdots & \\ 1, & \text{as } \alpha \to \infty. \end{cases}$$

Therefore, the maximum possible reduction is a factor of $\sqrt{e}$, achieved in the limit as $\alpha \to 1$, in agreement with the results in [7, 20].

Remark 1. By defining Rényi flatness factor $\epsilon_\Lambda^{(\alpha)}(s) \triangleq R_\alpha(\rho_{s,\Lambda} \| U) - 1$, the connection between $\epsilon_\Lambda^{(\alpha)}(s)$ and $\eta_\varepsilon^{(\alpha)}(\Lambda)$ is as follows. For $\alpha \in \mathbb{N}_{\geq 2} \cup \{\infty\}$, we have

$$\epsilon_\Lambda^{(\alpha)}(s) = \varepsilon \Leftrightarrow s = \eta_\varepsilon^{(\alpha)}(\Lambda).$$

Specifically $\eta_\varepsilon^{(\infty)}(\Lambda) = \eta_\varepsilon(\Lambda)$ where $\eta_\varepsilon(\Lambda)$ is the commonly defined smoothing parameter in [23].

Theorem 1 (Informal)**.** *For $\alpha \in \mathbb{N}_{\geq 2}$ and random lattices of unit volume, we have*

$$\mathbb{E}_\Lambda[\epsilon_\Lambda^{(\alpha)}(s)] \to 0$$

exponentially fast as $n \to \infty$, provided:

$$s > \alpha^{-1/(2\alpha-2)},$$

which is optimal.

Here 'optimal' means no lattice of unit volume exists for the smoothing behavior unless the above bounds is satisfied, which is stated in the following.

Theorem 2 (Informal)**.** *For $\alpha \in \mathbb{N}_{\geq 2}$ and any lattice Λ of unit volume, we have*

$$\epsilon_\Lambda^{(\alpha)}(s) \to \infty,$$

exponentially fast as $n \to \infty$, provided:

$$s < \alpha^{-1/(2\alpha-2)}.$$

Thus the bound $s > \alpha^{-1/(2\alpha-2)}$ is sharp: the Rényi flatness factor of order α of a lattice cannot vanish for any $s < \alpha^{-1/(2\alpha-2)}$. In particular, the L^∞ flatness factor explodes exponentially for any $s < 1$. This can also be seen from (5):

$$\epsilon_\Lambda^{(\infty)}(s) > \frac{1}{s^n} - 1$$

since the theta series $\Theta_\Lambda(\tau) > 1$ for any $\tau > 0$. Thus, as s becomes smaller than 1, the L^∞ flatness factor $\approx \frac{1}{s^n}$, but the Rényi flatness factor can still be brought under control. This demonstrates the advantage of the Rényi flatness factor.

We also obtain vanishing rates with explicit exponents of the Rényi flatness factor.

Theorem 3 (Informal)**.** *For $\alpha \geq 2$, $s > \alpha^{-1/(2\alpha-2)}$ and random lattices of unit volume,*

$$\mathbb{E}_\Lambda[\epsilon_\Lambda^{(\alpha)}(s)] = O(e^{-nE_s^{(\alpha)}})$$

where

$$E_s^{(\alpha)} = \min\left\{\frac{1}{2}\log\alpha + (\alpha-1)\log s, \frac{1}{2}\log 2s^2\right\}.$$

Our proofs estimate the average behavior of random lattice ensembles. When the average behavior vanishes, there must exist some lattice in the ensemble with the same property. We use both the Minkowski–Hlawka ensemble and the Construction A ensemble, which are detailed in later sections. The formal versions of these results appear in Theorem 5, Theorem 6 and Theorem 7.

1.2 Technical Overview

Random Lattice Ensembles In our subsequent proofs establishing the existence of smoothing-good lattices, we analyze the average behavior of lattices drawn from certain random ensembles. In particular, we consider two ensembles: the Minkowski–Hlawka ensemble and Loeliger's ensemble via Construction A. These two ensembles form the basis for our analysis of smoothing-good lattices. Their averaging properties allow us to deduce that there exist lattices whose average behavior meets the requirements for smoothing.

Rogers' Formula. To study the asymptotic behavior of lattices drawn from the Minkowski–Hlawka ensemble, one must evaluate expectations of the form

$$\mathbb{E}_\Lambda\big[f(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_k)\big].$$

This requires the Rogers averaging formula, which generalizes Siegel's mean-value theorem to the multivariate setting. In the resulting expansion, the leading terms are straightforward to compute, but the remaining terms are notoriously difficult to handle. In this work, we treat each term in the Gaussian case in turn and derive explicit asymptotic convergence bounds.

Loeliger's Formula for Higher-Order Moments. To derive the Rényi flatness factor of order $\alpha \geq 2$ for the Construction-A lattice ensemble, we extend Loeliger's averaging lemma to α-fold correlations, mirroring the multivariate Rogers formula. Recall that a Construction-A lattice $\Lambda(\mathcal{C}) \subset \mathbb{R}^n$ is obtained by lifting a linear code $\mathcal{C} \subset \mathbb{F}_q^n$ modulo p. By grouping ordered α-tuples of codewords according to the dimension k of their span, one shows that the probability a random

α-tuple spans a k-dimensional subspace of $\mathbb{F}_q^n$ is related to the Gaussian binomial coefficient. Substituting this decomposition into the ensemble average produces a finite sum of Gaussian integrals indexed by $k = 1, \ldots, \min(\alpha, n)$. Each term can then be bounded asymptotically, yielding explicit convergence rates for the α-th moment.

1.3 Related Work and Open Problems

Our research builds upon and complements several recent advances in information theory and lattice-based cryptography. Luzzi, Ling and Bloch [20] analyzed the Kullback–Leibler (KL) and L^1 smoothing behavior of Construction A ensembles under Gaussian noise. Debris-Alazard, Ducas, Resch and Tillich [7] performed an analysis of L^1 smoothing using both Minkowski–Hlawka and Construction A ensembles, considering both Gaussian and ball-shaped noise distributions. They both [7, 20] showed the existence of lattices whose L^1 smoothing parameter $\eta_\varepsilon^{(1)}(\Lambda) \to V(\Lambda)^{1/n}/\sqrt{e}$ for a suitable sequence $\varepsilon_n \to 0$. This improves upon the result on L^∞ smoothing parameter $\eta_{\varepsilon_n}(\Lambda) \to V(\Lambda)^{1/n}$.

Pouly and Shen [24] employed Rogers' formula to study the L^∞ smoothing properties of the Minkowski–Hlawka ensemble under Gaussian noise, though their analysis is limited to the first and second moments. Rogers' formula has also been used in [11, 12] to prove probabilistic bounds for the shortest vectors in module lattices.

Interestingly, the left-hand side of (1) can be identified with the degree-$(\alpha-1)$ Siegel theta series [10] associated to the dual lattice Λ^*:

$$\Theta_{\Lambda^*}^{(\alpha-1)}\left(\boldsymbol{A}s^2\right) = \sum_{(\boldsymbol{\lambda}_1,\cdots,\boldsymbol{\lambda}_{\alpha-1})\in\Lambda^*\times\cdots\times\Lambda^*} e^{-\pi s^2 \operatorname{tr}(\boldsymbol{S}_{\alpha-1}\boldsymbol{A})},$$

where $\operatorname{tr}(\cdot)$ denotes the trace of a matrix, and the Gram matrix $\boldsymbol{S}_{\alpha-1} = (S_{ij})$ where $S_{ij} = \langle \boldsymbol{\lambda}_i, \boldsymbol{\lambda}_j \rangle, i, j = 1, \ldots, \alpha-1$. Moreover, the infinite sum appearing in our Proposition 1—which involves matrix determinants—bears a resemblance to a variant of the Selberg zeta function, as studied by Koecher [33]:

$$\zeta_{\boldsymbol{A}}(k) = \sum_{\boldsymbol{0}\neq\boldsymbol{D}\in\mathbb{Z}^{m\times n}} |\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|^{-k}$$

where $\boldsymbol{A}$ is positive definite symmetric and the sum is over a complete set of representatives for equivalence relation $\boldsymbol{D} \sim \boldsymbol{D}'$ if $\boldsymbol{D}' = \boldsymbol{D}\boldsymbol{U}$ with $\boldsymbol{U} \in \mathrm{GL}_n(\mathbb{Z})$. This function extends the classical Epstein zeta function from quadratic forms over vectors to quadratic forms over matrices. Exploring these connections may offer a fruitful direction for future research.

While we have illustrated some applications of the Rényi smoothing parameter, we believe that many more problems in lattice-based cryptography could benefit from this framework, particularly given the central role played by Rényi divergence in recent cryptographic analyses.

1.4 Roadmap

The remainder of this paper is organized as follows: Section 2 introduces the basic definitions of lattices, discrete Gaussian distributions, information theoretic measures, and the Rényi smoothing parameter. Section 3 develops our first main result by applying Rényi divergence to analyze the average behavior of the Minkowski–Hlawka ensemble. Section 4 extends our analysis to the Construction A ensemble. Finally, we present selected applications in lattice-based cryptography in Section 5.

Throughout the paper, we use standard asymptotic notation: $f(x) = O(g(x))$ if $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$, $f(x) = o(g(x))$ if $\limsup_{x\to\infty} |f(x)/g(x)| = 0$, and $f(x) = \omega(g(x))$ if $\limsup_{x\to\infty} |g(x)/f(x)| = 0$. $O(\cdot)$ always denotes growth in n alone. Whenever we write $o(\cdot)$ with respect to parameters beyond n, those additional variables will be clearly indicated.

2 Definitions

2.1 Information-Theoretic Measures and Inequalities

We define the Rényi divergence and other distance in the same way as [3].

Let $p(\boldsymbol{x})$ and $q(\boldsymbol{x})$ be probability density functions on $\mathbb{R}^n$. Define

$$\Delta(p,q) = \frac{1}{2}\,\|p-q\|_1 = \frac{1}{2}\int \big|\, p(\boldsymbol{x}) - q(\boldsymbol{x})\big|\, d\boldsymbol{x},$$
$$\|p-q\|_\infty = \sup_{\boldsymbol{x}} \big|\, p(\boldsymbol{x}) - q(\boldsymbol{x})\big|.$$

For $\alpha \neq 1$, the Rényi divergence is define by [4]

$$R_\alpha\big(p\|q\big) = \left(\int \frac{p(\boldsymbol{x})^\alpha}{q(\boldsymbol{x})^{\alpha-1}}\, d\boldsymbol{x}\right)^{1/(\alpha-1)}.$$

The limit $\alpha \to 1$ corresponds to the KL divergence. In the limit $\alpha \to \infty$ [9],

$$R_\infty\big(p\|q\big) = \operatorname{ess\,sup}_p \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})}.$$

Here, ess sup denotes the essential supremum, i.e., the supremum taken after ignoring sets of measure zero. These satisfy the following monotonicity for $\alpha_1 \leq \alpha_2$:

$$R_{\alpha_1}(p\|q) \;\leq\; R_{\alpha_2}(p\|q).$$

We will also use the following key inequalities:

[4] In cryptography, the convention for Rényi divergence typically refers to the exponential of the classical Rényi divergence, which is originally defined using the logarithm.

Lemma 1 (Pinsker's Inequality)**.** *For $\alpha \in [1, \infty]$,*

$$R_\alpha(p\|q) \geq e^{2\,\Delta(p,q)^2}.$$

We now present the properties of statistical distance including probability preservation and data processing inequality.

Lemma 2. *Let p and q be two probability densities on $(\mathcal{X}, \mathcal{F})$. Then for any measurable $E \subseteq \mathcal{X}$,*

$$\int_E q(\boldsymbol{x})\, d\boldsymbol{x} \geq \int_E p(\boldsymbol{x})\, d\boldsymbol{x} - \Delta(p, q).$$

Moreover, for any measurable map $f : \mathcal{X} \to \mathcal{Y}$, letting p_f and q_f denote the pushforward densities on $\mathcal{Y}$, one has

$$\Delta(p_f, q_f) \leq \Delta(p, q).$$

Lemma 3 ([3, Lemma 2.9])**.** *Let p, q be two probability density functions over a measurable space $\mathcal{X}$, and let $(p_i)_i, (q_i)_i$ be two families of density functions. The Rényi divergence satisfies the following properties:*

- ***Data processing inequality:*** *For any measurable function f, and let p_f, q_f denote the pushforward densities of p, q through f, we have*

$$R_\alpha(p_f\|q_f) \leq R_\alpha(p\|q).$$

- ***Multiplicativity:*** *For product densities,*

$$R_\alpha\left(\prod_i p_i \middle\| \prod_i q_i\right) = \prod_i R_\alpha(p_i\|q_i).$$

- ***Probability preservation:*** *For any measurable event $E \subseteq \mathrm{Supp}(q)$ and $\alpha \in (1, +\infty)$, we have*

$$\int_E q(\boldsymbol{x})\, d\boldsymbol{x} \geq \frac{\left(\int_E p(\boldsymbol{x})\, d\boldsymbol{x}\right)^{\alpha/(\alpha-1)}}{R_\alpha(p\|q)}.$$

2.2 Lattice Gaussian Distribution

An n-dimensional lattice Λ in the Euclidean space $\mathbb{R}^n$ is a set defined by

$$\Lambda = \mathcal{L}(\boldsymbol{B}) = \{\boldsymbol{B}\boldsymbol{x} : \boldsymbol{x} \in \mathbb{Z}^n\}$$

where the columns of the basis matrix $\boldsymbol{B} = [\boldsymbol{b}_1, \cdots, \boldsymbol{b}_n]$ are assumed to be linearly independent. The dual lattice Λ^* of a lattice Λ is defined as the set of vectors $\boldsymbol{v} \in \mathbb{R}^n$ such that $\langle \boldsymbol{v}, \boldsymbol{\lambda} \rangle \in \mathbb{Z}$, for all $\boldsymbol{\lambda} \in \Lambda$ (see, e.g., [6]).

A measurable set $\mathcal{R}(\Lambda) \subset \mathbb{R}^n$ is a fundamental region of the lattice Λ if $\cup_{\boldsymbol{\lambda} \in \Lambda}(\mathcal{R}(\Lambda) + \boldsymbol{\lambda}) = \mathbb{R}^n$ and if $(\mathcal{R}(\Lambda) + \boldsymbol{\lambda}) \cap (\mathcal{R}(\Lambda) + \boldsymbol{\lambda}')$ has measure 0 for

any $\boldsymbol{\lambda} \neq \boldsymbol{\lambda}'$ in Λ. The volume of a fundamental region is equal to that of the Voronoi cell $V(\Lambda) = \sqrt{|\det(\boldsymbol{B}^T\boldsymbol{B})|}$.

For $s > 0$ and $\boldsymbol{v} \in \mathbb{R}^n$, the usual Gaussian pdf of parameter s centered at $\boldsymbol{v} \in \mathbb{R}^n$ is given by

$$\rho_{s,\boldsymbol{v}}(\boldsymbol{x}) = \frac{1}{s^n} e^{-\frac{\pi\|\boldsymbol{x}-\boldsymbol{v}\|^2}{s^2}},$$

for all $\boldsymbol{x} \in \mathbb{R}^n$. For convenience, we write $\rho_s(\boldsymbol{x}) = \rho_{s,\boldsymbol{0}}(\boldsymbol{x})$.

Consider the Λ-periodic function

$$\rho_{s,\Lambda}(\boldsymbol{x}) = \sum_{\boldsymbol{\lambda}\in\Lambda} \rho_{s,\boldsymbol{\lambda}}(\boldsymbol{x}) = \frac{1}{s^n} \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\frac{\pi\|\boldsymbol{x}-\boldsymbol{\lambda}\|^2}{s^2}}, \tag{3}$$

for all $\boldsymbol{x} \in \mathbb{R}^n$. Observe that $\rho_{s,\Lambda}$ restricted to the fundamental region $\mathcal{R}(\Lambda)$ is a probability density.

We define the *discrete Gaussian distribution* over Λ centered at $\boldsymbol{v} \in \mathbb{R}^n$ as the following discrete distribution taking values in $\boldsymbol{\lambda} \in \Lambda$:

$$D_{\Lambda,s,\boldsymbol{v}}(\boldsymbol{\lambda}) = \frac{\rho_{s,\boldsymbol{v}}(\boldsymbol{\lambda})}{\rho_{s,\boldsymbol{v}}(\Lambda)}, \quad \forall \boldsymbol{\lambda} \in \Lambda,$$

where $\rho_{s,\boldsymbol{v}}(\Lambda) \triangleq \sum_{\boldsymbol{\lambda}\in\Lambda} \rho_{s,\boldsymbol{v}}(\boldsymbol{\lambda}) = \rho_{s,\Lambda}(\boldsymbol{v})$. Again for convenience, we write $D_{\Lambda,s} = D_{\Lambda,s,\boldsymbol{0}}$.

In some sense, the continuous distribution $\rho_{s,\Lambda}$ and the discrete distribution $D_{\Lambda,s}$ are the Fourier dual of each other. To see this, note that since $\rho_{s,\Lambda}(\boldsymbol{x})$ is Λ-periodic, it has the Fourier expansion on the dual lattice Λ^*

$$\rho_{s,\Lambda}(\boldsymbol{x}) = \frac{1}{V(\Lambda)} \sum_{\boldsymbol{\lambda}^*\in\Lambda^*} \hat{\rho}_s(\boldsymbol{\lambda}^*) e^{j2\pi\langle\boldsymbol{\lambda}^*,\boldsymbol{x}\rangle}$$

where

$$\hat{\rho}_s(\boldsymbol{y}) = \int \rho_s(\boldsymbol{x}) e^{-\mathrm{i}2\pi\langle\boldsymbol{x},\boldsymbol{y}\rangle} = e^{-\pi s^2\|\boldsymbol{y}\|^2} \tag{4}$$

is the Fourier transform. Thus, the Fourier coefficients $\hat{\rho}_s(\boldsymbol{\lambda}^*)$ have a discrete Gaussian distribution over the dual lattice Λ^* (upon normalization).

Definition 2 (L^∞ Flatness Factor [18]). *For a lattice Λ and for a parameter s, the L^∞ flatness factor is defined by:*

$$\epsilon_\Lambda(s) \triangleq \max_{\boldsymbol{x}\in\mathcal{R}(\Lambda)} \left|V(\Lambda)\rho_{s,\Lambda}(\boldsymbol{x}) - 1\right|.$$

The expression of $\epsilon_\Lambda(s)$ can be established by theta series (see [6]) defined as

$$\Theta_\Lambda(\tau) = \sum_{\boldsymbol{\lambda}\in\Lambda} e^{-\pi\tau\|\boldsymbol{\lambda}\|^2}.$$

Lemma 4 (Expression of $\epsilon_\Lambda(s)$ [18])**.**

$$\epsilon_\Lambda(s) = \frac{V(\Lambda)}{s^n}\Theta_\Lambda\left(\frac{1}{s^2}\right) - 1 = \Theta_{\Lambda^*}(s^2) - 1. \tag{5}$$

The flatness factor is shown to be equivalent to the notion of smoothing parameter.

Definition 3 (L^∞ Smoothing Parameter [23, 28]). *For a lattice Λ and for $\varepsilon > 0$, the L^∞ smoothing parameter $\eta_\varepsilon(\Lambda)$ is the smallest $s > 0$ such that*

$$\sum_{\boldsymbol{\lambda} \in \Lambda^*} e^{-\pi s^2 \|\boldsymbol{\lambda}\|^2} \leq 1 + \varepsilon.$$

Lemma 5 ([18, Prop. 3]). *If $s = \eta_\varepsilon(\Lambda)$, then $\epsilon_\Lambda(s) = \varepsilon$.*

2.3 Rényi Smoothing Parameter

In this section, we generalize both the smoothing parameter and flatness factor to their Rényi versions.

Definition 4 (Rényi Flatness Factor). *Given a lattice Λ, and Gaussian noise pdf $\rho(\boldsymbol{x})$, the Rényi flatness factor with order $\alpha > 1$ is defined as:*

$$\epsilon_\Lambda^{(\alpha)}(s) \triangleq R_\alpha(\rho_{s,\Lambda} \| U) - 1 = \left(\int_{\mathcal{R}(\Lambda)} \frac{\rho_{s,\Lambda}(\boldsymbol{x})^\alpha}{1/V(\Lambda)^{\alpha-1}} d\boldsymbol{x} \right)^{\frac{1}{\alpha-1}} - 1.$$

For $\alpha \in \mathbb{N}_{\geq 2}$ we obtain the following expression of $\epsilon_\Lambda^{(\alpha)}(s)$.

Theorem 4. *Let Λ be an n-dimensional lattice, with Λ^* its dual lattice, and $\alpha \in \mathbb{N}_{\geq 2}$. Then*

$$\epsilon_\Lambda^{(\alpha)}(s) = \left(\sum_{\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{\alpha-1} \in \Lambda^*} e^{-\pi s^2 (\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A} \otimes \boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_{\alpha-1}^T)^T} \right)^{\frac{1}{\alpha-1}} - 1$$

*where $\boldsymbol{A} = \boldsymbol{I}_{\alpha-1} + \boldsymbol{J}_{\alpha-1}$ with $\boldsymbol{J}_{\alpha-1}$ the all-one matrix (*i.e.*, all entries are 1) of size $(\alpha-1) \times (\alpha-1)$ and $\otimes$ the Kronecker product.*

The infinite sum on the right-hand side is a theta function. Note that the matrix $\boldsymbol{A}$ has one eigenvalue α and other eigenvalues 1, with determinant $|\boldsymbol{A}| = \alpha$. It is worth pointing out that $\boldsymbol{A}$ is the Gram matrix of the $A_{\alpha-1}$ lattice whose squared minimum distance is 2.

Proof. We need a basic property of the Fourier transform:

$$\int_{\mathcal{R}(\Lambda)} e^{j2\pi\langle \boldsymbol{\lambda}, \boldsymbol{x} \rangle} d\boldsymbol{x} = \begin{cases} V(\Lambda), & \boldsymbol{\lambda} = \boldsymbol{0}; \\ 0, & \boldsymbol{\lambda} \in \Lambda^* \setminus \{\boldsymbol{0}\}. \end{cases}$$

Using the Fourier expansion

$$V(\Lambda)\rho_{s,\Lambda}(\boldsymbol{x}) = \sum_{\boldsymbol{\lambda} \in \Lambda^*} \hat{\rho}_s(\boldsymbol{\lambda}) e^{j2\pi\langle \boldsymbol{\lambda}, \boldsymbol{x} \rangle},$$

we have

$$\begin{aligned} V(\Lambda)^{\alpha}\int_{\mathcal{R}(\Lambda)}\rho_{s,\Lambda}(\boldsymbol{x})^{\alpha}d\boldsymbol{x} &= \int_{\mathcal{R}(\Lambda)}\left(\sum_{\boldsymbol{\lambda}\in\Lambda^*}\hat{\rho}_s(\boldsymbol{\lambda})e^{j2\pi\langle\boldsymbol{\lambda},\boldsymbol{x}\rangle}\right)^{\alpha}d\boldsymbol{x} \\ &= \int_{\mathcal{R}(\Lambda)}\sum_{\boldsymbol{\lambda}_i\in\Lambda^*}\hat{\rho}_s(\boldsymbol{\lambda}_1)\cdots\hat{\rho}_s(\boldsymbol{\lambda}_\alpha)e^{j2\pi\langle\boldsymbol{\lambda}_1+\cdots\boldsymbol{\lambda}_\alpha,\boldsymbol{x}\rangle}d\boldsymbol{x} \\ &= \sum_{\boldsymbol{\lambda}_i\in\Lambda^*}\hat{\rho}_s(\boldsymbol{\lambda}_1)\cdots\hat{\rho}_s(\boldsymbol{\lambda}_\alpha)\int_{\mathcal{R}(\Lambda)}e^{j2\pi\langle\boldsymbol{\lambda}_1+\cdots\boldsymbol{\lambda}_\alpha,\boldsymbol{x}\rangle}d\boldsymbol{x} \\ &= V(\Lambda)\sum_{\boldsymbol{\lambda}_i\in\Lambda^*,\sum\boldsymbol{\lambda}_i=\boldsymbol{0}}\hat{\rho}_s(\boldsymbol{\lambda}_1)\cdots\hat{\rho}_s(\boldsymbol{\lambda}_\alpha). \end{aligned}$$

Recall that $\hat{\rho}_s(\boldsymbol{y}) = e^{-\pi s^2\|\boldsymbol{y}\|^2}$, we deduce

$$\begin{aligned} V(\Lambda)^{\alpha-1}\int_{\mathcal{R}(\Lambda)}\rho_{s,\Lambda}(\boldsymbol{x})^{\alpha}d\boldsymbol{x} &= \sum_{\boldsymbol{\lambda}_i\in\Lambda^*,\sum\boldsymbol{\lambda}_i=\boldsymbol{0}}e^{-\pi s^2(\|\boldsymbol{\lambda}_1\|^2+\cdots\|\boldsymbol{\lambda}_\alpha\|^2)} \\ &= \sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*}e^{-\pi s^2(\|\boldsymbol{\lambda}_1\|^2+\cdots\|\boldsymbol{\lambda}_{\alpha-1}\|^2+\|\boldsymbol{\lambda}_1+\cdots\boldsymbol{\lambda}_{\alpha-1}\|^2)} \\ &= \sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*}e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T}. \end{aligned}$$

The proof is completed by substituting the above expression into the definition of $\epsilon_{\Lambda}^{(\alpha)}(s)$. □

From this theorem we are able to define Rényi smoothing parameter $\eta_{\varepsilon}^{(\alpha)}(\Lambda)$ and establish its connection with Rényi flateness factor $\epsilon_{\Lambda}^{(\alpha)}(s)$.

Definition 5 (Rényi Smoothing Parameter for $\alpha \in \mathbb{N}_{\geq 2}$). *Given a lattice Λ, and $\boldsymbol{A}$ as defined previously, the Rényi smoothing parameter $\eta_{\varepsilon}^{(\alpha)}(\Lambda)$ with order $\alpha \in \mathbb{N}_{\geq 2}$ is defined as the smallest $s > 0$ such that*

$$\sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*}e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T} \leq (1+\varepsilon)^{\alpha-1}. \tag{6}$$

Setting $\alpha = 2$, we recognize the left-hand side

$$\sum_{\boldsymbol{\lambda}\in\Lambda^*}e^{-2\pi s^2\|\boldsymbol{\lambda}\|^2} = \Theta_{\Lambda^*}(2s^2) \tag{7}$$

which coincides with the theta series of the dual lattice Λ^*.

When we set $\alpha = 3$, the original series reduces to

$$\sum_{\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\in\Lambda^*}\exp\Big(-2\pi s^2\big(\|\boldsymbol{\lambda}_1\|^2+\|\boldsymbol{\lambda}_2\|^2+\langle\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2\rangle\big)\Big).$$

Moreover, this double sum can be identified with the degree-2 Siegel theta series [10] associated to the dual lattice Λ^*:

$$\Theta^{(2)}_{\Lambda^*}\big(\boldsymbol{A}s^2\big) \;=\; \sum_{(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2)\in\Lambda^*\times\Lambda^*} e^{-\pi s^2\,\mathrm{tr}(\boldsymbol{S}_2\boldsymbol{A})},$$

where the Gram matrix $\boldsymbol{S}_2=(S_{ij})_{2\times 2}$, $S_{ij}=\big\langle\boldsymbol{\lambda}_i,\boldsymbol{\lambda}_j\big\rangle, i,j=1,2$.

In general, the left-hand side of (6) can be identified with the degree-$(\alpha-1)$ Siegel theta series $\Theta^{(\alpha-1)}_{\Lambda^*}\big(\boldsymbol{A}s^2\big)$.

For any integer $\alpha\geq 2$, the Rényi smoothing parameter and the Rényi flatness factor satisfy

$$\epsilon^{(\alpha)}_{\Lambda}(s)=\varepsilon \quad\Longleftrightarrow\quad s=\eta^{(\alpha)}_{\varepsilon}(\Lambda)\,.$$

Moreover, by Lemma 4 and the definition of the Rényi divergence of order ∞, we have

$$\epsilon^{(\infty)}_{\Lambda}(s)=\max_{\boldsymbol{x}\in\mathcal{R}(\Lambda)}\{V(\Lambda)\,\rho_{s,\Lambda}(\boldsymbol{x})\}-1=\epsilon_{\Lambda}(s)\,.$$

Combining this with Lemma 5 yields

$$\epsilon^{(\infty)}_{\Lambda}(s)=\varepsilon \quad\Longleftrightarrow\quad s=\eta_{\varepsilon}(\Lambda)=\eta^{(\infty)}_{\varepsilon}(\Lambda)\,,$$

which shows that the usual smoothing parameter is exactly the Rényi divergence of order ∞.

The following Theorem provides a lower bound of $\epsilon^{(\alpha)}_{\Lambda}(s)$.

Theorem 5. *For $\alpha>1$ and any n dimensional lattice Λ of unit volume, we have*

$$\epsilon^{(\alpha)}_{\Lambda}(s)\geq\frac{1}{\big(s^2\;\alpha^{\frac{1}{\alpha-1}}\big)^{\frac{n}{2}}}-1\to\infty,$$

given

$$s<\alpha^{-1/(2\alpha-2)}.$$

Proof. Since $\alpha>1$, we have

$$\begin{aligned}
\epsilon^{(\alpha)}_{\Lambda}(s) &= R_\alpha(\rho_{s,\Lambda}\|U)-1\\
&=\bigg(\int_{\mathcal{R}(\Lambda)}\Big(\sum_{\boldsymbol{\lambda}\in\Lambda}\rho_s(\boldsymbol{x}+\boldsymbol{\lambda})\Big)^{\alpha}d\boldsymbol{x}\bigg)^{\frac{1}{\alpha-1}}-1\\
&\geq\bigg(\int_{\mathcal{R}(\Lambda)}\sum_{\boldsymbol{\lambda}\in\Lambda}\rho_s(\boldsymbol{x}+\boldsymbol{\lambda})^{\alpha}d\boldsymbol{x}\bigg)^{\frac{1}{\alpha-1}}-1\\
&=\bigg(\sum_{\boldsymbol{\lambda}\in\Lambda}\int_{\mathcal{R}(\Lambda)+\boldsymbol{\lambda}}\rho_s(\boldsymbol{x})^{\alpha}d\boldsymbol{x}\bigg)^{\frac{1}{\alpha-1}}-1\\
&=\bigg(\int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})^{\alpha}d\boldsymbol{x}\bigg)^{\frac{1}{\alpha-1}}-1\\
&=\big(s^2\;\alpha^{\frac{1}{\alpha-1}}\big)^{-\frac{n}{2}}-1.
\end{aligned}$$

In particular, if

$$s^2\,\alpha^{\frac{1}{\alpha-1}} < 1,$$

then the factor $\left(s^2\,\alpha^{1/(\alpha-1)}\right)^{-n/2}$ grows without bound (exponentially) as $n \to \infty$. Hence the entire expression diverges in that regime. □

In the next section we will explore the regime of s that a lattice can have a vanishing Rényi flatness factor.

3 Proof Based on Minkowski-Hlawka Ensemble

In this section, we employ the Minkowski-Hlawka ensemble to examine the conditions under which $\epsilon_{\Lambda}^{(\alpha)}(s)$ approaches zero as n tends to infinity. Furthermore, we investigate the corresponding smoothing parameter $\eta_{\varepsilon}^{(\alpha)}(\Lambda)$ and its asymptotic behavior.

3.1 Minkowski–Hlawka Ensemble

The Minkowski–Hlawka ensemble is constructed by considering the space of unimodular lattices. Specifically, define

$$\mathfrak{L}_n \cong \mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z}),$$

where each lattice $\Lambda_n \in \mathfrak{L}_n$ is represented by a generator matrix $\boldsymbol{B}$ and the coset $\boldsymbol{B}\mathrm{SL}_n(\mathbb{Z})$ uniquely corresponds to the lattice $\boldsymbol{B}\mathbb{Z}^n$. In this ensemble, each lattice is identified up to scaling and basis transformation, which enables a uniform treatment of lattice properties. The space $\mathfrak{L}_n$ is endowed with the unique, normalized invariant Haar measure μ_n. A central result underlying this ensemble is Siegel's averaging formula [30]:

Lemma 6 ([16,30])**.** *For $n \geq 2$, let $f : \mathbb{R}^n \to \mathbb{R}$ be a Riemann integrable function such that $\|\boldsymbol{x}\|^{n+c} f(\boldsymbol{x})$ is bounded on $\mathbb{R}^n$ for some fixed $c > 0$. Then, for lattices Λ_n drawn according to the Haar measure μ_n, we have*

$$\mathbb{E}_{\Lambda_n \sim \mu_n}\left[\sum_{\boldsymbol{x}\in\Lambda_n\setminus\{0\}} f(\boldsymbol{x})\right] = \int_{\mathbb{R}^n} f(\boldsymbol{x})\,d\boldsymbol{x}.$$

Note that the lemma has been generalized to apply to essentially bounded Lebesgue measurable functions $f(\boldsymbol{x})$ that vanish outside a bounded region [22]. For convenience in subsequent proofs, we omit the notation $\Lambda_n \sim \mu_n$ and simply write Λ_n when the distribution is understood. With this ensemble, we are able to specify the average Rényi smoothing parameter.

3.2 Average Behavior for $\boldsymbol{\alpha} \in \mathbb{N}_{\geq \mathbf{2}}$

Fix an integer $\alpha \geq 2$. We first consider the average over the ensemble,

$$\mathbb{E}_{\Lambda}\left[\sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*} e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T}\right].$$

If $\Lambda = \boldsymbol{B}\mathbb{Z}^n$, with $\boldsymbol{B} \in \mathrm{SL}_n(\mathbb{R})$, then $\Lambda^* = \boldsymbol{B}^{-T}\mathbb{Z}^n$. Hence on cosets we have

$$\boldsymbol{B}\,\mathrm{SL}_n(\mathbb{Z}) \;\longmapsto\; \boldsymbol{B}^{-T}\,\mathrm{SL}_n(\mathbb{Z}).$$

It follows that averaging any summation over nonzero lattice points yields the same result whether one averages over Λ or over its dual Λ^* and it can be obtained that

$$\mathbb{E}_{\Lambda} = \mathbb{E}_{\Lambda^*}.$$

Therefore, in what follows we will no longer distinguish between $\mathbb{E}_{\Lambda}$ and $\mathbb{E}_{\Lambda^*}$, and denote them both simply by $\mathbb{E}_{\Lambda}$.

Rogers' mean-value formula then provides an explicit expression for this expectation.

Lemma 7. *[29, Theorem 4] Consider a Minkowski-Hlawka ensemble of random lattices Λ of dimension n and a Riemann-integrable function f. For $\alpha \in \mathbb{N}_{\geq 2}$, the average*

$$\mathbb{E}_{\Lambda}\left[\sum_{\boldsymbol{\lambda}_i\in\Lambda} f(\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\ldots,\boldsymbol{\lambda}_{\alpha-1})\right]$$

$$= f(\mathbf{0},\mathbf{0},\ldots,\mathbf{0}) + \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} f(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_{\alpha-1})d\boldsymbol{x}_1 d\boldsymbol{x}_2\ldots d\boldsymbol{x}_{\alpha-1} \tag{8}$$

$$+\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\cdots\frac{e_m}{q}\right)^n\int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} f\left(\sum_{i=1}^{m}\frac{\tilde{d}_{i1}}{q}\boldsymbol{x}_i,\ldots,\sum_{i=1}^{m}\frac{\tilde{d}_{i(\alpha-1)}}{q}\boldsymbol{x}_i\right)d\boldsymbol{x}_1\ldots d\boldsymbol{x}_m, \tag{9}$$

where the outer sum is over all divisions $(\boldsymbol{\nu};\boldsymbol{\mu}) = (\nu_1,\ldots,\nu_m;\mu_1,\ldots,\mu_{\alpha-1-m})$ of the numbers $1,2,\ldots,\alpha-1$ into two sequences,

$$\begin{aligned}
&1 \leq \nu_1 \leq \ldots \leq \nu_m \leq \alpha-1\\
&1 \leq \mu_1 \leq \ldots \leq \mu_{\alpha-1-m} \leq \alpha-1\\
&\nu_i \neq \mu_j, 1\leq i\leq m, 1\leq j\leq \alpha-1-m.
\end{aligned}$$

The inner sum is over all $m\times(\alpha-1)$ matrices $\tilde{\boldsymbol{D}}$, with integer elements having greatest common divisor (gcd) relatively prime to q, and with

$$\begin{aligned}
&\tilde{d}_{i\nu_j} = q\delta_{ij}, \quad i=1,\ldots,m, \quad j=1,\ldots,m,\\
&\tilde{d}_{i\mu_j} = 0, \quad \text{if} \quad \mu_j<\nu_i, i=1,\ldots,m, \quad j=1,\ldots,\alpha-1-m.
\end{aligned}$$

$\delta_{ij}=1$ only for $i=j$, $e_i=\gcd(\epsilon_i,q), i=1,\ldots,m$, where ϵ_i are the elementary divisors of matrix $\tilde{\boldsymbol{D}}$.

Proposition 1. *For order $\alpha \in \mathbb{N}_{\geq 2}$, the flatness factor of the Minkowski-Hlawka ensemble of random lattices satisfies*

$$\mathbb{E}_{\Lambda}\left[\epsilon_{\Lambda}^{(\alpha)}(s)\right] \leq \left(1 + \frac{1}{s^{(\alpha-1)n} \cdot \alpha^{n/2}} + \sum_{\boldsymbol{\nu},\boldsymbol{\mu}} \sum_{q=1}^{\infty} \sum_{\tilde{\boldsymbol{D}}} \left(\frac{e_1}{q}\frac{e_2}{q}\cdots\frac{e_m}{q}\right)^n \left(\frac{q}{s}\right)^{mn} |\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2}\right)^{1/(\alpha-1)} - 1 \tag{10}$$

following the notation of Lemma 7.

Proof. By Theorem 4, it can be obtained that

$$\mathbb{E}_{\Lambda}\left[\epsilon_{\Lambda}^{(\alpha)}(s)\right] = \mathbb{E}_{\Lambda}\left[\left(\sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*} e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T}\right)^{\frac{1}{\alpha-1}}\right] - 1$$

$$\leq \left(\mathbb{E}_{\Lambda} \sum_{\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_{\alpha-1}\in\Lambda^*} e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T}\right)^{\frac{1}{\alpha-1}} - 1$$

where Jensen's inequality is applied. To evaluate the expectation of the lattice-sum inside, define the function

$$f(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots, \boldsymbol{\lambda}_{\alpha-1}) = e^{-\pi s^2(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)(\boldsymbol{A}\otimes\boldsymbol{I}_n)(\boldsymbol{\lambda}_1^T,\ldots,\boldsymbol{\lambda}_{\alpha-1}^T)^T}$$

and apply Lemma 7 where the terms (8) and (9) are analyzed below.

The first term of (8) is

$$f(\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}) = 1,$$

while the second term of (8) can be derived as

$$\begin{aligned}\int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} f(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{\alpha-1}) d\boldsymbol{x}_1 d\boldsymbol{x}_2 \ldots d\boldsymbol{x}_{\alpha-1} &= \frac{1}{s^{(\alpha-1)n}}|(\boldsymbol{A}\otimes\boldsymbol{I}_n)|^{-1/2} \\ &= \frac{1}{s^{(\alpha-1)n}\cdot\alpha^{n/2}}\end{aligned}$$

For the infinite sum in (9), we derive

$$\begin{aligned}&\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\cdots\frac{e_m}{q}\right)^n \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} f\left(\sum_{i=1}^m \frac{\tilde{d}_{i1}}{q}\boldsymbol{x}_i, \ldots, \sum_{i=1}^m \frac{\tilde{d}_{i(\alpha-1)}}{q}\boldsymbol{x}_i\right) d\boldsymbol{x}_1\ldots d\boldsymbol{x}_m \\ &= \sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\cdots\frac{e_m}{q}\right)^n \int_{\mathbb{R}^n}\cdots\int_{\mathbb{R}^n} e^{-\frac{\pi}{q^2}s^2(\boldsymbol{x}_1^T,\cdots,\boldsymbol{x}_m^T)(\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T\otimes\boldsymbol{I}_n)(\boldsymbol{x}_1^T,\cdots,\boldsymbol{x}_m^T)^T} d\boldsymbol{x}_1\ldots d\boldsymbol{x}_m \\ &= \sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\cdots\frac{e_m}{q}\right)^n \left(\frac{q}{s}\right)^{mn}|\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2}.\end{aligned}$$

□

To bound the contribution of the infinite sum in (10), we use the following Lemma.

Lemma 8.

$$\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\frac{e_2}{q}\cdots\frac{e_m}{q}\right)^n\left(\frac{q}{s}\right)^{mn}|\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2}=O\left(\max_{1\le m\le\alpha-2}\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}}\right)$$

Proof. After permuting the columns, one may write

$$\widetilde{\boldsymbol{D}}=\left(q\boldsymbol{I}_m\mid\boldsymbol{C}\right),$$

where $\boldsymbol{I}_m$ is the $m\times m$ identity matrix and $\boldsymbol{C}$ is an $m\times(\alpha-1-m)$ integer matrix all of whose entries are coprime to q. Let $N(\boldsymbol{C},q)$ denote the number of vectors $\boldsymbol{x}\in(\mathbb{Z}/q\mathbb{Z})^m$ satisfying

$$\boldsymbol{C}^T\boldsymbol{x}\equiv\boldsymbol{0}\pmod q.$$

Then by [29, Lemma 1] we have

$$e_1e_2\cdots e_m\;=\;N(\boldsymbol{C},q)\;=\;\begin{cases}q^m, & \text{if } \boldsymbol{C}=\boldsymbol{0},\\ \le q^{m-1}, & \text{if } \boldsymbol{C}\neq\boldsymbol{0}.\end{cases}\tag{11}$$

Therefore

$$\begin{aligned}&\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\frac{e_2}{q}\cdots\frac{e_m}{q}\right)^n\left(\frac{q}{s}\right)^{mn}|\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2}\\ &\le\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\frac{1}{s^{mn}}\left(\frac{q^m}{\sqrt{|(q\boldsymbol{I},\boldsymbol{C})\boldsymbol{A}\left(q\boldsymbol{I},\boldsymbol{C}\right)^T|}}\right)^n\Bigg|_{\substack{q=1\\ \boldsymbol{C}=\boldsymbol{0}}}+\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|}}\right)^n\Bigg|_{\boldsymbol{C}\neq\boldsymbol{0}}\qquad(12)\\ &\le\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\frac{1}{s^{mn}}\left(\frac{q^m}{\sqrt{|(q\boldsymbol{I},\boldsymbol{C})\boldsymbol{A}\left(q\boldsymbol{I},\boldsymbol{C}\right)^T|}}\right)^n\Bigg|_{\substack{q=1\\ \boldsymbol{C}=\boldsymbol{0}}}+\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\boldsymbol{D}\in\mathcal{D}_q(\boldsymbol{P}_{\boldsymbol{\nu},\boldsymbol{\mu}})}\frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|}}\right)^n.\qquad(13)\end{aligned}$$

Here (12) is derived from (11). Case $\boldsymbol{C}=\boldsymbol{0}$ only appears when $q=1$, since 1 is the only positive integer coprime with 0. (13) is derived by relaxing matrices $\tilde{\boldsymbol{D}}$ that require elements of $\boldsymbol{C}$ coprime with q to matrices $\boldsymbol{D}$ that has nonzero $\boldsymbol{C}$. The notation $\mathcal{D}_q(\boldsymbol{P}_{\boldsymbol{\nu},\boldsymbol{\mu}})$ in (13) represents a set of matrices for permutation matrix $\boldsymbol{P}_{\boldsymbol{\nu},\boldsymbol{\mu}}$, which is defined as a $m\times m$ permutation matrix with $(\boldsymbol{P}_{\boldsymbol{\nu},\boldsymbol{\mu}})_{\nu_i,i}=1$ for $i=1,2,\ldots,m$. For any permutation matrix $\boldsymbol{P}$, $\mathcal{D}_q(\boldsymbol{P})$ is defined as

$$\mathcal{D}_q(\boldsymbol{P})\triangleq\{\boldsymbol{D}\in\mathbb{Z}^{m\times(\alpha-1)}:\boldsymbol{D}\boldsymbol{P}=(q\boldsymbol{I}_m,\boldsymbol{C}),\text{ where }\boldsymbol{C}\in\mathbb{Z}^{m\times(\alpha-1-m)},\boldsymbol{C}\neq\boldsymbol{0}\}.$$

Then it has been proved in Appendix A that

$$\sum_{q=1}^{\infty} \sum_{\boldsymbol{D} \in \mathcal{D}_q(\boldsymbol{P})} \frac{1}{s^{mn}} \left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|}} \right)^n = O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right). \tag{14}$$

Continuing from Eq. (13), we have

$$\begin{aligned} &\sum_{\boldsymbol{\nu},\boldsymbol{\mu}} \sum_{q=1}^{\infty} \sum_{\tilde{\boldsymbol{D}}} \left(\frac{e_1}{q} \frac{e_2}{q} \cdots \frac{e_m}{q} \right)^n \left(\frac{q}{s} \right)^{mn} |\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2} \\ &\leq \sum_{\boldsymbol{\nu},\boldsymbol{\mu}} \frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} + \sum_{\boldsymbol{\nu},\boldsymbol{\mu}} O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right) \qquad (15) \\ &= O\left(\max_{1 \leq m \leq \alpha-2} \frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right). \qquad (16) \end{aligned}$$

Here (15) is derived by computing the first term and applying (14) to the second term. (16) is derived by the fact that the sum is finite, and m is ranged from 1 to $\alpha - 2$ over all possible partitions $(\boldsymbol{\nu}, \boldsymbol{\mu})$. □

Combining results above, we are able to specify the average behavior of $\epsilon_{\Lambda}^{(\alpha)}(s)$.

Theorem 6. *For the Minkowski-Hlawka ensemble of random lattices, and order* $\alpha \in \mathbb{N}_{\geq 2}$*, if*

$$s > \alpha^{-1/2(\alpha-1)},$$

then

$$\mathbb{E}_{\Lambda}\left[\epsilon_{\Lambda}^{(\alpha)}(s) \right] = O\left(\max\left\{ \frac{1}{(\sqrt{2}s)^n}, \frac{1}{s^{(\alpha-1)n}\alpha^{\frac{n}{2}}} \right\} \right) \to 0.$$

Proof. Combining Proposition 1 and Lemma 8, it can be derived that

$$\begin{aligned} \mathbb{E}_{\Lambda}\left[\epsilon_{\Lambda}^{(\alpha)}(s) \right] &\leq \left(1 + \frac{1}{s^{(\alpha-1)n}\alpha^{\frac{n}{2}}} + \max_{1 \leq m \leq \alpha-2} O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right) \right)^{\frac{1}{\alpha-1}} - 1 \\ &\leq \frac{1}{\alpha-1} \frac{1}{s^{(\alpha-1)n}\alpha^{\frac{n}{2}}} + \frac{1}{\alpha-1} \max_{1 \leq m \leq \alpha-2} O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right) \\ &= O\left(\max_{1 \leq m \leq \alpha-1} \frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right). \qquad (17) \end{aligned}$$

where Bernoulli's inequality $(1+a)^{\frac{1}{\alpha-1}} - 1 \leq \frac{a}{\alpha-1}$ for $\alpha \geq 2$ is applied. For (17) to vanish as $n \to \infty$, it is required for each $1 \leq m \leq \alpha - 1$,

$$\frac{1}{s^m(m+1)^{\frac{1}{2}}} < 1 \Rightarrow \frac{1}{s^2} < (m+1)^{1/m}.$$

Since $g_1(m) = (m+1)^{1/m}$ is monotonically decreasing for $m \geq 1$, it suffices that

$$s > \alpha^{-1/2(\alpha-1)}.$$

To better estimate the convergence rate in (17), consider the function $g_2(m) = \frac{1}{s^m(m+1)^{1/2}}$. Setting $\frac{d}{dm}g_2(m) = 0$, we get the critical point $m^* = -\frac{1}{\log(s^2)} - 1$. If $m^* > 1$, $g_2(m)$ is decreasing for $[1, m^*)$, and then increasing for (m^*, ∞) . If $m^* \leq 1$, $g_2(m)$ is increasing over $[1, \infty)$. Combining these results, (17) can be expressed as

$$\mathbb{E}_\Lambda\left[\epsilon_\Lambda^{(\alpha)}(s)\right] = O\left(\max\left\{\frac{1}{(\sqrt{2}s)^n},\ \frac{1}{s^{(\alpha-1)n}\alpha^{\frac{n}{2}}}\right\}\right). \tag{18}$$

□

Thus we have obtained a full spectrum of smoothing parameters. The Rényi smoothing parameter decreases with the order α, from 1 for $\alpha \to \infty$ (the usual smoothing parameter) to $1/\sqrt{e}$ for $\alpha \to 1$ (KL divergence) [20].

We now present some examples for small orders α.

Example 1 ($\alpha = 2$)**.** *In this case, only the terms of* (8) *exist so that*

$$\mathbb{E}_\Lambda\left[\epsilon_\Lambda^{(2)}(s)\right] \leq \frac{1}{(\sqrt{2}s)^n} \to 0$$

if $s > \frac{1}{\sqrt{2}}$*. This agrees with the analysis using* (7)*, since*

$$\Theta_{\Lambda^*}(2s^2) = \frac{1}{(\sqrt{2}s)^n}\Theta_\Lambda\left(\frac{1}{2s^2}\right).$$

Example 2 ($\alpha = 3$)**.** *In this case,* $m = 1$ *and the only possible partitions are* $(\boldsymbol{\nu}, \boldsymbol{\mu}) = (1; 2)$ *and* $(2; 1)$*. Thus the* 1×2 *matrix* $\tilde{\boldsymbol{D}} = (q, p)$ *for* $q \geq 1, \gcd(p, q) = 1$ *or* $\tilde{\boldsymbol{D}} = (0, 1)$ [5]*; in either case it holds that* $\epsilon_1 = 1$*.*

The two terms of (8) *are given by*

$$\begin{aligned} &1 + \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{-\pi s^2(\boldsymbol{x}_1^T, \boldsymbol{x}_2^T)(\boldsymbol{I}_{2n} + \boldsymbol{J}_2 \otimes \boldsymbol{I}_n)(\boldsymbol{x}_1^T, \boldsymbol{x}_2^T)^T} d\boldsymbol{x}_1 d\boldsymbol{x}_2 \\ &= 1 + \frac{1}{(\sqrt{3}s^2)^n}. \end{aligned} \tag{19}$$

The infinite sum in (9) *is given by*

$$\begin{aligned} &\sum_{\boldsymbol{\nu},\boldsymbol{\mu}}\sum_{q=1}^{\infty}\sum_{\tilde{\boldsymbol{D}}}\left(\frac{e_1}{q}\frac{e_2}{q}\cdots\frac{e_m}{q}\right)^n\left(\frac{q}{s}\right)^{mn}|\tilde{\boldsymbol{D}}\boldsymbol{A}\tilde{\boldsymbol{D}}^T|^{-n/2} \\ &= \sum_{q=1}^{\infty}\sum_{p:\gcd(p,q)=1}\frac{1}{s^n}\left(\frac{1}{2(p^2+pq+q^2)}\right)^{n/2} + \left(\frac{1}{\sqrt{2}s}\right)^n \end{aligned}$$

[5] In general, the second case is $(0, q)$ for $q \geq 1$, but it is required that $\gcd(q, q) = 1$ which leaves $q = 1$ the only possibility.

$$\leq \left(\frac{1}{\sqrt{2}s}\right)^n \sum_{(p,q)\neq(0,0)} \left(\frac{1}{p^2+pq+q^2}\right)^{n/2}$$
$$= \left(\frac{1}{\sqrt{2}s}\right)^n \zeta_Q(n/2) \tag{20}$$

where $\zeta_Q(n/2)$ *is the Epstein zeta function associated with the positive-definite quadratic form* $Q(p,q) = p^2+pq+q^2$. *The Epstein zeta function corresponding to a binary quadratic form* Q *is defined by* $\zeta_Q(z) = \sum_{(p,q)\neq(0,0)} \frac{1}{Q(p,q)^z}$ *for* $\Re(z) > 1$. *The Epstein zeta function converges if* $n > 2$ *in this case. If we consider* (19) *and* (20) *both vanish, then it suffices that* (19) *vanishes, i.e.*

$$s > \frac{1}{\sqrt[4]{3}}.$$

But for the dominant term, we need to compare (19) *and* (20) *depending on* s. *In fact for* $n \geq 4$, $\zeta_Q(n/2) \leq \zeta_Q(2) \approx 2.873 < 3$. *Thus*

$$\mathbb{E}_\Lambda\left[\epsilon_\Lambda^{(3)}(s)\right] \leq \begin{cases} \dfrac{1}{2\cdot\left(\sqrt{3}\,s^2\right)^n}\,(1+o(1)), & s \in \left(\dfrac{1}{\sqrt[4]{3}}, \sqrt{\dfrac{2}{3}}\right], \\ \dfrac{3}{2\cdot\left(\sqrt{2}\,s\right)^n}\,(1+o(1)), & s \in \left(\sqrt{\dfrac{2}{3}}, +\infty\right). \end{cases}$$

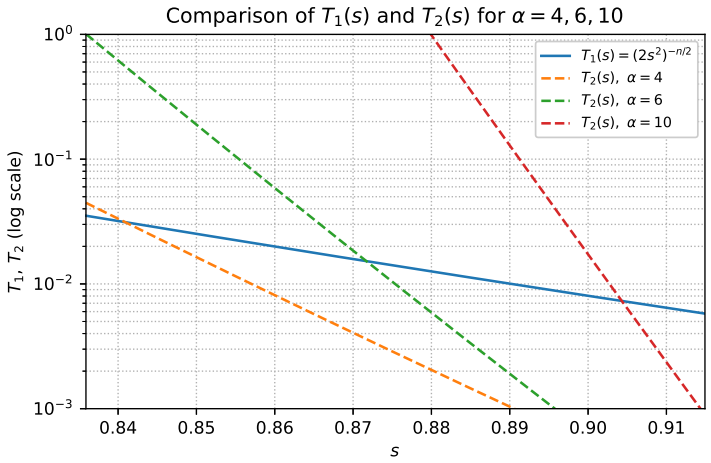

Fig. 1. Comparison of $T_1(s) = (2s^2)^{-n/2}$ and $T_2(s) = s^{-(\alpha-1)n}\,\alpha^{-n/2}$, with $n = 20$. The solid black line is $T_1(s)$; dashed lines are $T_2(s)$ for $\alpha = 4$, 6, and 10. The horizontal axis spans $s \approx [0.83, 0.92]$, and the vertical axis is logarithmic over $[10^{-3}, 1]$ to emphasize where each T_2 curve crosses T_1 (the critical s_c).

For large n and general α, one sees by comparing

$$\frac{1}{(\sqrt{2}s)^n} \quad \text{and} \quad \frac{1}{s^{(\alpha-1)n}\,\alpha^{n/2}}$$

that the sign of

$$\ln(2s^2) \;-\; \big[(\alpha-1)\ln(s^2)+\ln\alpha\big]$$

determines which term is exponentially larger: if s exceeds the unique critical value $s_c(\alpha)$ solving

$$\ln(2s^2) \;=\; (\alpha-1)\ln(s^2)+\ln\alpha,$$

then the Gaussian kernel term $(\sqrt{2}s)^{-n}$ dominates, whereas for $s < s_c(\alpha)$ the mixture term $\big(s^{(\alpha-1)}\,\alpha^{1/2}\big)^{-n}$ prevails. Figure 1 illustrates this crossover for $\alpha = 4, 6, 10$, showing how the threshold $s_c(\alpha)$ shifts to larger values as α increases.

4 Proof Based on Construction A

While the Minkowski-Hlawka ensemble encompasses a rich variety of lattices, it presents significant computational challenges for practical implementation. A more tractable alternative is Loeliger's Ensemble, which can be efficiently generated by extending linear codes through Construction A. In this section, we replace the Minkowski-Hlawka ensemble with the ensemble constructed from Construction A based on linear codes over $\mathbb{F}_p$.

4.1 Loeliger's Ensemble via Construction A

This ensemble is obtained by narrowing the Minkowski–Hlawka ensemble using Construction A, as described in [19]. Let p be a prime number and let $C \subseteq \mathbb{F}_p^n$ be an $[n,k]_p$-linear code over the finite field $\mathbb{F}_p$. The lattice associated with C is defined via Construction A as

$$\Lambda_C = \gamma \cdot \{\boldsymbol{x} \in \mathbb{Z}^n : \boldsymbol{x} \bmod p \in C\} = \gamma \cdot (C + p\mathbb{Z}^n)\,,$$

where γ is a scaling factor (typically chosen in $(0,1)$) to ensure that the lattice has the desired volume properties.

To analyze the average behavior of lattices constructed from linear codes, let $\mathcal{B}$ denote the ensemble of all $[n,k]_p$-linear codes over $\mathbb{F}_p^n$. The following averaging lemma, analogous to Siegel's formula, holds for the Construction A ensemble:

Lemma 9 (Averaging Lemma [5, 19])**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Riemann integrable function that is semi-admissible, i.e., $(1+\|\boldsymbol{x}\|)^{n+c}|f(\boldsymbol{x})|$ is bounded for some fixed $\boldsymbol{c} > 0$. Then, for any integer k satisfying $0 < k < n$ and any fixed volume V, the following holds:*

$$\frac{1}{|\mathcal{B}|}\sum_{C\in\mathcal{B}}\sum_{\boldsymbol{v}\in\Lambda_C} f(\boldsymbol{v}) = \sum_{\boldsymbol{v}\in p\mathbb{Z}^n} f(\gamma\boldsymbol{v}) + \frac{p^k-1}{p^n-1}\Bigg(\sum_{\boldsymbol{v}\in\mathbb{Z}^n} f(\gamma\boldsymbol{v}) - \sum_{\boldsymbol{v}'\in p\mathbb{Z}^n} f(\gamma\boldsymbol{v}')\Bigg),$$

and

$$\lim_{p\to\infty}\frac{1}{|\mathcal{B}|}\sum_{C\in\mathcal{B}}\sum_{\boldsymbol{v}\in\Lambda_C\setminus\{\mathbf{0}\}} f(\boldsymbol{v}) = \frac{1}{V}\int_{\mathbb{R}^n} f(\boldsymbol{v})\,d\boldsymbol{v},$$

with the constraint $\gamma^n p^{n-k} = V$.

Now we introduce some notations. For a pdf $f:\mathbb{R}^n\to\mathbb{R}^{\geq 0}$ and $j>0$, we define the capped L^j-norm integral as

$$\overline{\|f\|_j^j} \triangleq \min\left\{\|f\|_j^j,\,1\right\} = \begin{cases}\displaystyle\int_{\mathbb{R}^n} f(\boldsymbol{x})^j\,d\boldsymbol{x}, & j\geq 1,\\ 1, & j=0.\end{cases}$$

For notational convenience, we define $\|f\|_0^0 = 1$. Another notation is

$$\sum_{k_1+\cdots+k_i=i}$$

where the summation is over i nonnegative integers $k_1,\ldots,k_i$ satisfying $k_1+\cdots+k_i=i$

In the analysis that follows, we normalize the fundamental volume $V=1$ for simplicity. The key tool in our analysis is understanding the behavior of lattice point distributions. For any $\boldsymbol{m}\in\mathbb{F}_p^k$, let

$$\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}) \triangleq \rho_s(\boldsymbol{x}+\gamma\cdot\boldsymbol{G}\boldsymbol{m}+\gamma\cdot p\mathbb{Z}^n) = \sum_{\boldsymbol{v}\bmod p=\boldsymbol{G}\boldsymbol{m}}\rho_s(\boldsymbol{x}+\gamma\boldsymbol{v}),$$

where $\boldsymbol{G}\in\mathbb{F}_p^{n\times k}$ is the generator matrix for code $C\in\mathcal{B}$. For any non-zero codeword $\bar{\boldsymbol{v}}=\boldsymbol{G}\boldsymbol{m}\in\mathbb{F}_p^n\setminus\{\mathbf{0}\}$, we abbreviate $\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})$ as $\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})$. Leveraging the uniform distribution of codewords $\boldsymbol{G}\boldsymbol{m}$ over $\mathbb{F}_p^n$ for random $\boldsymbol{G}$, Lemma 9 simplifies to:

$$\begin{aligned}\mathbb{E}_{\boldsymbol{G}\sim\mathcal{B}}\left[\sum_{\boldsymbol{m}\in\mathbb{F}_p^k\setminus\{\mathbf{0}\}}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})\right] &= \frac{p^k-1}{p^n-1}\sum_{\bar{\boldsymbol{v}}\in\mathbb{F}_p^n\setminus\{\mathbf{0}\}}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})\\ &\sim \frac{1}{V}\int_{\mathbb{R}^n}\rho_s(\boldsymbol{x}+\boldsymbol{v})\,d\boldsymbol{v}\quad\big(p\to\infty,\ \gamma\to 0,\ \gamma^n p^{n-k}=V\big).\end{aligned}\tag{21}$$

For convenience, we shall write $\mathbb{E}_{\boldsymbol{G}}$ instead of $\mathbb{E}_{\boldsymbol{G}\sim\mathcal{B}}$ in what follows.

4.2 Average Behavior for $\alpha\in[2,\infty)$ with Gaussian Noise

To estimate $R_\alpha(\rho_{s,\Lambda_C}\|U)$ for order $\alpha\geq 2$, we need to extend (21) to higher moments and thus define the following,

$$S_{i,j}\triangleq\begin{cases}\mathbb{E}_{\boldsymbol{G}}\left[\left(\displaystyle\sum_{\boldsymbol{m}\in\mathbb{F}_p^k}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})\right)^i\right], & \text{if } i\geq 1 \text{ and } j=0.\\ \mathbb{E}_{\boldsymbol{G}}\left[\left(\displaystyle\sum_{\boldsymbol{m}\in\mathbb{F}_p^k}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})\right)^i\sum_{\boldsymbol{m}'\in\mathbb{F}_p^k}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}')^j\right], & \text{if } i\in\mathbb{N} \text{ and } j\geq 1.\end{cases}$$

The $S_{i,j}$ exhibits recursive structure. For instance, $S_{0,0} = 1$ and $S_{i,0} = S_{i-1,1}$ for $i \in \mathbb{N}^+$ establishes a recursive link between orders. In this way, all $S_{i,j}$ will be reduced to

$$S_{0,j} = \varphi_{\boldsymbol{x}}(\mathbf{0})^j + \frac{p^k - 1}{p^n - 1} \sum_{\bar{\boldsymbol{v}} \in \mathbb{F}_p^n \setminus \{\mathbf{0}\}} \varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j, \quad j \geq 1.$$

More specifically, the following recurrence relation and bounds can be established.

Proposition 2. *Based on the notations above, We have the following recurrence relation for $i \in \mathbb{N}^+,\ j \geq 1$:*

$$S_{i,j} \leq p^i S_{i-1,j+1} + S_{0,j} S_{i-1,1}, \tag{22}$$

Furthermore, based on this recurrence, we derive the following upper bound:

$$S_{i,j} \leq S_{0,j} S_{0,1}^i + p^{\frac{i(i+1)}{2}} \sum_{k_1 + \cdots + k_i = i} S_{0,j+k_1} \prod_{l=2}^{i} S_{0,k_l}, \tag{23}$$

Proof. The second claim follows by repeatedly calling for the first claim. For details, see Appendix B. □

Below we first provide an informal, intuitive outline of the proof to aid understanding, before presenting the full rigorous argument.

Given $\gamma^n p^{n-k} = V$ and for sufficiently large $p\gamma$ and small γ, it is easy to derive that

$$\lim_{p\to\infty} \varphi_{\boldsymbol{x}}(\mathbf{0})^j = \lim_{p\to\infty} \rho_s(\boldsymbol{x} + \gamma \cdot p\mathbb{Z}^n)^j = \rho_s(\boldsymbol{x})^j,$$

and combined with Lemma 9 we have

$$\lim_{p\to\infty} S_{0,j} = \rho_s(\boldsymbol{x})^j + \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x} + \boldsymbol{y})^j \, d\boldsymbol{y} = \rho_s(\boldsymbol{x})^j + \|\rho_s\|_j^j, \quad j \geq 1. \tag{24}$$

Applying Proposition 2 to the above, the bound of $S_{\alpha,0}$ with $\alpha \in \mathbb{N}^+$ can be obtained as follows,

$$\begin{aligned}\lim_{p\to\infty} S_{\alpha,0} = \lim_{p\to\infty} S_{\alpha-1,1} &\leq \lim_{p\to\infty} S_{0,1}^{\alpha} + \lim_{p\to\infty} p^{\frac{\alpha(\alpha-1)}{2}} \sum_{k_1 + \cdots + k_{\alpha-1} = \alpha} \prod_{l=1}^{\alpha-1} S_{0,k_l} \\ &\leq (\rho_s(\boldsymbol{x}) + 1)^{\alpha} + \lim_{p\to\infty} p^{\frac{\alpha(\alpha-1)}{2}} \sum_{k_1 + \cdots + k_{\alpha-1} = \alpha} \prod_{l=1}^{\alpha-1} \left(\rho_s(\boldsymbol{x})^{k_l} + \|\rho_s\|_{k_l}^{k_l} + o_p(1) \right).\end{aligned} \tag{25}$$

Meanwhile we have

$$\begin{aligned}\mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)}\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha}d\boldsymbol{x}\right] &= \mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)}\sum_{\boldsymbol{v}\in\Lambda_C}\rho_s(\boldsymbol{x}+\boldsymbol{v})\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha-1}d\boldsymbol{x}\right]\\ &= \mathbb{E}_{\Lambda_C}\left[\sum_{\boldsymbol{v}\in\Lambda_C}\int_{\mathcal{R}(\Lambda_C)}\rho_s(\boldsymbol{x}+\boldsymbol{v})\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha-1}d\boldsymbol{x}\right]\\ &= \mathbb{E}_{\Lambda_C}\left[\sum_{\boldsymbol{v}\in\Lambda_C}\int_{\mathcal{R}(\Lambda_C)+\boldsymbol{v}}\rho_s(\boldsymbol{x})\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha-1}d\boldsymbol{x}\right]\\ &= \int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})\mathbb{E}_{\Lambda_C}\left[\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha-1}d\boldsymbol{x}\right], \qquad (26)\end{aligned}$$

Thus combining (26) and (25), we are able to compute

$$\begin{aligned}&\lim_{p\to\infty}\mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)}\rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha+1}\,d\boldsymbol{x}\right]\\ &= \lim_{p\to\infty}\int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})S_{\alpha,0}\,d\boldsymbol{x} \qquad \text{by (26)}\\ &\le \underbrace{\int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})\big(\rho_s(\boldsymbol{x})+1\big)^{\alpha}\,d\boldsymbol{x}}_{(i)}\\ &\quad + \lim_{p\to\infty}\underbrace{p^{\frac{\alpha(\alpha-1)}{2}}\sum_{k_1+\cdots+k_{\alpha-1}=\alpha}\int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})\prod_{l=1}^{\alpha-1}\left(\rho_s(\boldsymbol{x})^{k_l}+\|\rho_s\|_{k_l}^{k_l}+o_p(1)\right)d\boldsymbol{x}}_{(ii)},\end{aligned}$$

where $o_p(1)$ tends to zero if $p\to\infty$. Our goal is to explore under which condition can we prove $(i)\to 1$ and $(ii)\to 0$.

Assume that

$$s > \alpha^{-\frac{1}{2(\alpha-1)}}.$$

Under this assumption, we have

$$\|\rho_s\|_{\beta}^{\beta} = \begin{cases} s^{-n(\beta-1)}\beta^{-\frac{n}{2}}\to 0, & \text{if } \alpha\ge\beta>1,\\ 1, & \text{if } \beta=1.\end{cases}$$

Thus for the first term (i), we expand:

$$(i) = \int_{\mathbb{R}^n}\rho_s(\boldsymbol{x})(\rho_s(\boldsymbol{x})+1)^{\alpha}\,d\boldsymbol{x} = \sum_{j=0}^{\alpha}\binom{\alpha}{j}\|\rho_s\|_{j+1}^{j+1} = 1+o_n(1).$$

Similarly, for the second term (ii) we note that the combinatorial contribution becomes for some constant c_α that

$$(ii) = \lim_{p\to\infty}p^{\frac{\alpha(\alpha-1)}{2}}\left(c_\alpha\sum_{k_1'+\cdots+k_{\alpha-1}'=\alpha}\|\rho_s\|_{k_1'+1}^{k_1'+1}\prod_{l=2}^{\alpha-1}\|\rho_s\|_{k_l'}^{k_l'}+o_p(1)\right),$$

which would diverge if $p \to \infty$ directly. In order to bound term (ii) and prove that it vanishes, we require p to be dependent on n, i.e. $p = p(n)$ as $n \to \infty$. Since each $\|f\|_{k_l'}^{k_l'}$ goes to 0 exponentially, the whole contribution will as well goes to 0 if we set $p(n)$ as a polynomial of n, for example $p(n) = O(n^2)$ or $O(n^3)$.

Collecting these results, we get the following bounds as $n \to \infty$:

$$\mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)} \rho_{s,\Lambda_C}(\boldsymbol{x})^{\alpha+1}\, d\boldsymbol{x}\right] \le \int_{\mathbb{R}^n} \rho_s(x) S_{\alpha,0}\, d\boldsymbol{x} \le (i)+(ii) \to 1$$
$$\Rightarrow \mathbb{E}_{\Lambda_C}\left[R_{\alpha+1}(\rho_{s,\Lambda_C}\|U)\right] \to 0.$$

The analysis looks nice, but requires a more careful way to deal with the Riemann integral approximation error in (24) and (25) since we assumed $p = p(n)$ is dependent with n. The following lemma helps to bound the approximation error in $e^{-\omega(n)}$.

Lemma 10. *For any $j > 0, j' \ge 1$, and under the conditions $p = \eta_n n^{1+2\varepsilon} \to \infty$, $\gamma = \eta_n^{-\frac{1}{2}} n^{-\frac{1}{2}-\varepsilon} \to 0$, $k = \frac{n}{2}$, and $\varepsilon > 0$, the following bounds are satisfied:*

$$S_{0,j} \le \varphi_{\boldsymbol{x}}(\mathbf{0})^j + \overline{\|\rho_s\|_j^j} + e^{-\omega(n)},$$
$$\int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j'} \varphi_{\boldsymbol{x}}(\mathbf{0})^j\, d\boldsymbol{x} \le \|\rho_s\|_{j+j'}^{j+j'} + e^{-\omega(n)}.$$

Here, $\eta_n \in \left(\frac{1}{2}, 1\right]$ is chosen to ensure that p remains prime, which is guaranteed by the fact that there exists a prime number between n' and $2n'$ for any integer n' [14].

Proof. See Appendix C. □

With this approximation error lemma, a rigorous proof of $\mathbb{E}_{\Lambda_C}\left[\epsilon_{\Lambda_C}^{(\alpha)}(s)\right] \to 0$ is provided below.

Theorem 7. *For an ensemble of random lattice based on Construction A of linear codes, $\alpha \in \mathbb{N}_{\ge 2}$, given*

$$s > \alpha^{-\frac{1}{2(\alpha-1)}},$$

then

$$\mathbb{E}_{\Lambda_C}\left[\epsilon_{\Lambda_C}^{(\alpha)}(s)\right] = O\left(\max\left\{\frac{1}{(\sqrt{2}s)^n},\ \frac{1}{s^{(\alpha-1)n}\alpha^{\frac{n}{2}}}\right\}\right).$$

Proof. We first estimate

$$\int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})(\varphi_{\boldsymbol{x}}(\mathbf{0})+1)^{\alpha-1} d\boldsymbol{x}$$
$$= \sum_{j=0}^{\alpha-1} \binom{\alpha-1}{j} \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) \varphi_{\boldsymbol{x}}(\mathbf{0})^j\, d\boldsymbol{x}$$
$$\le 1 + c'_\alpha \sum_{j=2}^{\alpha} \|\rho_s\|_j^j + e^{-\omega(n)}. \tag{27}$$

The last step follows by Lemma 10, and $c'_\alpha = \max_{j\in\{0,1,\ldots,\alpha-1\}} \binom{\alpha-1}{j}$. Based on this inequality, it can be obtained that

$$\mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)} \rho_{s,\Lambda_C}(\boldsymbol{x})^\alpha \, d\boldsymbol{x}\right] = \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) S_{\alpha-1,0}\, d\boldsymbol{x} \qquad \text{by Eq. (26)}$$

$$\leq p^{\frac{\alpha(\alpha-1)}{2}} \sum_{k_1+\cdots+k_{\alpha-1}=\alpha-1} \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) \prod_{l=1}^{\alpha-1} S_{0,k_l}\, d\boldsymbol{x} + \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) S_{0,1}^{\alpha-1} d\boldsymbol{x}$$

by Proposition 2

$$\leq p^{\frac{\alpha(\alpha-1)}{2}} \sum_{k_1+\cdots+k_{\alpha-1}=\alpha-1} \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) \prod_{j=1}^{\alpha-1} \left(\varphi_{\boldsymbol{x}}(\boldsymbol{0})^{k_j} + \overline{\|\rho_s\|_{k_j}^{k_j}}\right) d\boldsymbol{x} + \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})(\varphi_{\boldsymbol{x}}(\boldsymbol{0})+1)^{\alpha-1} d\boldsymbol{x} + e^{-\omega(n)}$$

by Lemma 10

$$\leq p^{\frac{\alpha(\alpha-1)}{2}} c_\alpha \sum_{k'_1+\cdots+k'_{\alpha-1}=\alpha-1} \prod_{j=2}^{\alpha-1} \overline{\|\rho_s\|_{k'_j}^{k'_j}} \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x}) \varphi_{\boldsymbol{x}}(\boldsymbol{0})^{k'_1} d\boldsymbol{x} + \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})(\varphi_{\boldsymbol{x}}(\boldsymbol{0})+1)^{\alpha-1} d\boldsymbol{x} + e^{-\omega(n)}$$

$$\leq p^{\frac{\alpha(\alpha-1)}{2}} c_\alpha \sum_{k'_1+\cdots+k'_{\alpha-1}=\alpha} \prod_{j=1}^{\alpha-1} \overline{\|\rho_s\|_{k'_j}^{k'_j}} + c'_\alpha \sum_{j=2}^{\alpha} \|\rho_s\|_j^j + 1 + e^{-\omega(n)}. \qquad (28)$$

The last step is derived by Eq. (27) and Lemma 10.

Noting that p is polynomial in n and

$$\|\rho_s\|_\beta^\beta = \frac{1}{s^{n(\beta-1)}\beta^{\frac{n}{2}}} \to 0, \quad \beta > 1,$$

which is exponentially small, we can retain only the exponential term with the dominant convergence bound in (28) as

$$\mathbb{E}_{\Lambda_C}\left[\int_{\mathcal{R}(\Lambda_C)} \rho_{s,\Lambda_C}(\boldsymbol{x})^\alpha d\boldsymbol{x}\right] = O\left(\max_{2\leq\beta\leq\alpha} \frac{1}{s^{n(\beta-1)}\beta^{\frac{n}{2}}}\right) + 1.$$

Based on this result, it can be derived that

$$\begin{aligned}
\mathbb{E}_{\Lambda_C}\left[\epsilon_{\Lambda_C}^{(\alpha)}(s)\right] &\leq \left(\mathbb{E}_{\Lambda_C} \int_{\mathcal{R}(\Lambda_C)} \rho_{s,\Lambda_C}(\boldsymbol{x})^\alpha d\boldsymbol{x}\right)^{\frac{1}{\alpha-1}} - 1 \\
&= \left(1 + O\left(\max_{2\leq\beta\leq\alpha} \frac{1}{s^{n(\beta-1)}\beta^{\frac{n}{2}}}\right)\right)^{\frac{1}{\alpha-1}} - 1. \\
&\leq O\left(\max_{2\leq\beta\leq\alpha} \frac{1}{s^{n(\beta-1)}\beta^{\frac{n}{2}}}\right). \qquad (29)
\end{aligned}$$

The proof is completed by following the same steps as those from (17) to (18) in Theorem 6. □

5 Cryptographic Applications

For cryptographic applications using integer lattices, we need the following adaption:

Definition 6 (Rényi Smoothing Parameter for Discrete Distributions). *For a p-ary lattice $p\mathbb{Z}^n \subset \Lambda \subset \mathbb{Z}^n$, and a noise distribution function $\rho(\boldsymbol{x})$ over $\mathbb{Z}^n$, denote Rényi divergence of order $\alpha \geq 0$ by* [6]

$$\mathbb{D}_\alpha(F_\Lambda \| U) = \left(\sum_{\boldsymbol{x} \in \mathbb{Z}^n \cap \mathcal{R}(\Lambda)} \frac{F_\Lambda(\boldsymbol{x})^\alpha}{1/|\mathbb{Z}^n/\Lambda|^{\alpha-1}} \right)^{\frac{1}{\alpha-1}},$$

where $F_\Lambda(\boldsymbol{x}) = \frac{\rho_\Lambda(\boldsymbol{x})}{\rho_{\mathbb{Z}^n}(\mathbf{0})}$. For a Gaussian noise distribution function $\rho_s(\boldsymbol{x})$ with parameter s, we write $\mathbb{D}_\alpha(F_\Lambda \| U)$ as $\mathbb{D}_\alpha(F_{s,\Lambda} \| U)$. The smoothing parameter $\tilde{\eta}_\varepsilon^{(\alpha)}(\Lambda)$ for $\varepsilon > 1$ is defined as the smallest $s > 0$ such that $\mathbb{D}_\alpha(F_{s,\Lambda} \| U) \leq \varepsilon$.

If $\alpha = \infty$, $\mathbb{D}_\alpha(F_{s,\Lambda} \| U)$ can be bounded using its continuous counterpart $R_\alpha(\rho_{s,\Lambda} \| U)$ in a straightforward manner [3, 13]. However, if $\alpha < \infty$, it becomes tricky to bound $\mathbb{D}_\alpha(F_{s,\Lambda} \| U)$. In the following, we show that $\mathbb{D}_\alpha(F_{s,\Lambda} \| U)$ and $\tilde{\eta}_\varepsilon^{(\alpha)}(\Lambda)$ are still upper-bounded by their continuous counterparts subject to appropriate scaling.

Lemma 11. *Let Λ be a p-ary lattice with volume p^{n-k} for $k < n$. Fix the volume $V(\Lambda')$ of $\Lambda' = \gamma\Lambda$ where $\gamma = p^{-(1-k/n)} V(\Lambda')^{1/n}$, and fix the noise parameter $s' = \gamma s$. Then $\mathbb{D}_\alpha(F_{s,\Lambda} \| U) \leq R_\alpha(\rho_{s',\Lambda'} \| U)$ and $\gamma \tilde{\eta}_\varepsilon^{(\alpha)}(\Lambda) \leq \eta_\varepsilon^{(\alpha)}(\Lambda')$ as $p \to \infty$.*

Proof. Firstly, we note that $\rho_{s,\mathbb{Z}^n}(\mathbf{0}) \geq 1$, thus $F_{s,\Lambda}(\boldsymbol{x}) \leq \rho_\Lambda(\boldsymbol{x})$. Secondly, it is easy to verify $\rho_{s,\Lambda}(\boldsymbol{x}) = \rho_{\gamma s,\gamma\Lambda}(\boldsymbol{x})^\alpha \gamma^{\alpha n}$. Then,

$$\begin{aligned}
\sum_{\boldsymbol{x} \in \mathbb{Z}^n \cap \mathcal{R}(\Lambda)} \frac{F_{s,\Lambda}(\boldsymbol{x})^\alpha}{1/|\mathbb{Z}^n/\Lambda|^{\alpha-1}} &\leq \sum_{\boldsymbol{x} \in \mathbb{Z}^n \cap \mathcal{R}(\Lambda)} \frac{\rho_{s,\Lambda}(\boldsymbol{x})^\alpha}{1/p^{(\alpha-1)(n-k)}} \\
&= \sum_{\boldsymbol{x} \in \gamma(\mathbb{Z}^n \cap \mathcal{R}(\Lambda))} \frac{\rho_{\gamma s,\gamma\Lambda}(\boldsymbol{x})^\alpha \gamma^{\alpha n}}{1/p^{(\alpha-1)(n-k)}} \\
&= \sum_{\boldsymbol{x} \in \gamma(\mathbb{Z}^n \cap \mathcal{R}(\Lambda))} \frac{\rho_{\gamma s,\gamma\Lambda}(\boldsymbol{x})^\alpha \gamma^{n}}{1/(p^{(n-k)}\gamma^n)^{\alpha-1}} \\
&\overset{p\to\infty}{\to} \int_{\mathcal{R}(\Lambda')} \frac{\rho_{s',\Lambda'}(\boldsymbol{x})^\alpha}{1/V(\Lambda')^{\alpha-1}} d\boldsymbol{x}.
\end{aligned}$$

The last step follows from the definition of Riemann integration since $\gamma \to 0$ as $p \to \infty$. Therefore, we have proven $\mathbb{D}_\alpha(F_{s,\Lambda} \| U) \leq R_\alpha(\rho_{s',\Lambda'} \| U)$; since the noise parameter is also a scaling, we have $\gamma\tilde{\eta}_\varepsilon^{(\alpha)}(\Lambda) \leq \eta_\varepsilon^{(\alpha)}(\Lambda')$. □

Since the tightness of security reductions is not the main focus of this paper, we concentrate on the savings in the sampling parameter. If needed, tight security reductions can still be achieved by selecting an appropriate order α [31].

[6] With overload of notation, here U denotes the uniform distribution over $\mathbb{Z}^n/\Lambda$.

5.1 GPV Signature Scheme

The sampling parameter of the GPV Signature Scheme [13] has been reduced in [3] by using Rényi smoothing of order ∞. Here we show it can be further reduced by Rényi smoothing of order $\alpha < \infty$.

The core technique underlying GPV's signature scheme is discrete Gaussian sampling over a trapdoor lattice [13]. Its security crucially relies on the property that the output distribution of discrete Gaussian sampling is oblivious to any particular basis used in the sampling process, therefore preventing leakage of the private key. We provide a high-level introduction to the GPV signature (see [13] for details). In key generation, one generates a hard public basis for a random lattice Λ, together with a short private basis of Λ. The public basis serves as the public key, while the private basis serves as the private key. Given a message $\boldsymbol{m}$, one uses the private basis to sample a point $\boldsymbol{x}$ from $D_{\Lambda+H(\boldsymbol{m},\boldsymbol{s}),r}$ with parameter r, where $H(\cdot)$ is a hash function and $\boldsymbol{s}$ is a random salt. The signature of $\boldsymbol{m}$ is $\boldsymbol{x}, \boldsymbol{s}$. The verifier checks that $\boldsymbol{x}$ is short and that $\boldsymbol{x} - H(\boldsymbol{m}, \boldsymbol{s}) \in \Lambda$ using the public basis.

It is shown in [13] that the security of GPV signing can be reduced to the hardness of the inhomogeneous short integer solution (ISIS) problem with approximation factor $\sqrt{n}r$. Therefore, the width r is the most important property of a discrete Gaussian sampler in this context. A key component in the security proof of the GPV signature scheme is the closeness of the distribution of $\boldsymbol{A}\boldsymbol{x} \mod q$, where $\boldsymbol{x} \leftarrow D_{\mathbb{Z}^m,r}$, to the uniform distribution $U(\mathbb{Z}_q^n)$. For statistical distance, this requires $r = \Omega(\sqrt{\lambda + \log q_s})$, where λ is the security parameter and q_s denotes the number of signing queries made by the attacker [3]. Using Rényi divergence of order ∞, [3] obtained a smaller bound, namely $r = \Omega(\sqrt{\log \lambda + \log q_s})$.

Notice that here the lattice Λ is identified with $\Lambda_{\boldsymbol{A}}^{\perp} = \{\boldsymbol{x} \in \mathbb{Z}^m : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \mod q\}$. This is actually Construction A defined with parity-check matrix $\boldsymbol{A}$, thus for prime q we can claim properties of the Loeliger ensemble thanks to Lemma 11. The Rényi divergence under consideration is given by $\mathbb{D}_\alpha(F_{r,\Lambda_{\boldsymbol{A}}^{\perp}} \| U)$. Now if we use Rényi smoothing of order $\alpha = 2$, by (7) the required parameter r can be further reduced by a constant $\frac{\sqrt{2}}{2}$, for the same bound ε on Rényi divergence. Although the constant may appear small, the security level is quite sensitive to the choice of r [26].

Assuming access to an oracle capable of sampling exactly at or slightly above the smoothing parameter [25, 34] (with potentially long running time), the improvement in the security level can be substantial. A small parameter s does not only give higher security, but also more compact signatures.

5.2 Dual Regev Encryption

Dual Regev encryption is a public-key encryption scheme based on the Learning With Errors (LWE) problem, formulated in the dual style [13]. In this version, $\boldsymbol{A} \in \mathbb{Z}_q^{n \times m}$ is a public matrix chosen uniformly at random, the secret key is $\boldsymbol{y} \leftarrow D_{\mathbb{Z}^m,r}$, and the public key is a vector $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{y} \in \mathbb{Z}_q^n$,.

To encrypt a bit $b \in \{0,1\}$, the sender samples $\boldsymbol{s} \leftarrow U(\mathbb{Z}_q^n)$, $\boldsymbol{e}_1 \leftarrow \chi^m$ and $e_2 \leftarrow \chi$, where χ an error distribution (typically a discrete Gaussian). The ciphertext is given by

$$(\boldsymbol{c}_1, c_2) = (\boldsymbol{A}^T\boldsymbol{s} + \boldsymbol{e}_1,\ \boldsymbol{u}^T\boldsymbol{s} + e_2 + \lfloor q/2 \rfloor \cdot b) \in \mathbb{Z}_q^m \times \mathbb{Z}_q.$$

The scheme is IND-CPA secure under the LWE assumption. A key step in the security proof is to show the public key $\boldsymbol{u} = \boldsymbol{A}\boldsymbol{y}$ is statistically close to uniform. Again, the relevant lattice here is $\Lambda_{\boldsymbol{A}}^{\perp} = \{\boldsymbol{x} \in \mathbb{Z}^m : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \mod q\}$. The Rényi divergence between the distribution of $\boldsymbol{u}$ and uniform is given by $\mathbb{D}_\alpha(F_{r,\Lambda_{\boldsymbol{A}}^{\perp}} \| U)$. Using Rényi smoothing of order ∞, [3] showed the sampling parameter $r = \Omega(\sqrt{\log \lambda})$, improving upon the bound $r = \Omega(\sqrt{\lambda})$ based on L^1 smoothing, for $\varepsilon = O(2^{-\lambda})$. If we use the Rényi smoothing parameter of order $\alpha < \infty$, the parameter r can be reduced further, resulting in more efficient implementation.

5.3 Cryptography Based on LIP

Given two lattices L and L' with bases $\boldsymbol{B}$ and $\boldsymbol{B'}$ respectively, the Lattice Isomorphism Problem (LIP) is to find an orthogonal matrix $\boldsymbol{O} \in O_n(\mathbb{R})$ and a unimodular matrix $\boldsymbol{U} \in GL_n(\mathbb{Z})$ such that $\boldsymbol{B'} = \boldsymbol{O}\boldsymbol{B}\boldsymbol{U}$. Lattices can also be represented using quadratic forms. Given a lattice basis $\boldsymbol{B}$, its Gram matrix is defined as $\boldsymbol{Q} = \boldsymbol{B}^T\boldsymbol{B}$. Two quadratic forms $\boldsymbol{Q}$ and $\boldsymbol{Q'}$ are equivalent if there exists a unimodular matrix $\boldsymbol{U}$ such that $\boldsymbol{Q'} = \boldsymbol{U}^T\boldsymbol{Q}\boldsymbol{U}$. Thus, the LIP in terms of quadratic forms can be formulated as follows: Given two lattices L and L' with Gram matrices $\boldsymbol{Q}$ and $\boldsymbol{Q'}$ respectively, the LIP asks to find a unimodular matrix $\boldsymbol{U} \in GL_n(\mathbb{Z})$ such that

$$\boldsymbol{Q'} = \boldsymbol{U}^T\boldsymbol{Q}\boldsymbol{U}$$

if it exists.

Recently, the LIP has found numerous applications in cryptography. In [8], an identification scheme, a KEM, and a signature scheme were given with security dependent on problems related to LIP. All these schemes require Gaussian sampling to generate a random form from the equivalent class of $\boldsymbol{Q}$, which, for the KEM and signature schemes, serve as the public key. Given a quadratic form $\boldsymbol{Q}$ and a parameter $s > \max\{\lambda_n(Q), \|B_Q^* \cdot \sqrt{\ln(2n + 4/\pi)}\|\}$, [8, Algorithm 1] returns a from $\boldsymbol{Q'} = \boldsymbol{U}^T\boldsymbol{Q}\boldsymbol{U}$ together with $\boldsymbol{U} \in GL_n(\mathbb{Z})$. This is achieved by running Gaussian sampling over the lattice represented by $\boldsymbol{Q}$ with parameter s. Reducing the parameter s would result in more efficient implementation.

Acknowledgments. The authors are grateful to Ling Liu for helpful discussions. This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) [grant numbers EP/X037010/1 and EP/Y004477/1].

References

1. Aggarwal, D., Dadush, D., Regev, O., Stephens-Davidowitz, N.: Solving the shortest vector problem in 2^n time using discrete Gaussian sampling: extended abstract.

In: Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing. STOC '15, Association for Computing Machinery, New York, NY, USA (2015)
2. Aggarwal, D., Dadush, D., Stephens-Davidowitz, N.: Solving the closest vector problem in 2^n time – the discrete Gaussian strikes again! In: 2015 IEEE 56th Annual Symposium on Foundations of Computer Science. pp. 563–582 (2015)
3. Bai, S., Langlois, A., Lepoint, T., Stehlé, D., Steinfeld, R.: Improved security proofs in lattice-based cryptography: Using the Rényi divergence rather than the statistical distance. In: Advances in Cryptology - ASIACRYPT 2015 - 21st International Conference on the Theory and Application of Cryptology and Information Security, Auckland, New Zealand, November 29 - December 3, 2015, Proceedings, Part I. pp. 3–24 (2015)
4. Bogdanov, A., Guo, S., Masny, D., Richelson, S., Rosen, A.: On the hardness of learning with rounding over small modulus. In: Kushilevitz, E., Malkin, T. (eds.) Theory of Cryptography. pp. 209–224. Springer Berlin Heidelberg, Berlin, Heidelberg (2016)
5. Campello, A.: Random ensembles of lattices from generalized reductions. IEEE Transactions on Information Theory **64**(7), 5231–5239 (2018)
6. Conway, J.H., Sloane, N.J.A.: Sphere Packings, Lattices, and Groups. Springer-Verlag, New York, 3 edn. (1998)
7. Debris-Alazard, T., Ducas, L., Resch, N., Tillich, J.P.: Smoothing codes and lattices: Systematic study and new bounds. IEEE Transactions on Information Theory **69**(9), 6006–6027 (2023)
8. Ducas, L., van Woerden, W.: On the lattice isomorphism problem, quadratic forms, remarkable lattices, and cryptography. In: Advances in Cryptology – EUROCRYPT 2022. pp. 643–673. Springer International Publishing, Cham (2022)
9. van Erven, T., Harremos, P.: Renyi Divergence and Kullback-Leibler Divergence **60**(7), 3797–3820 (July 2014). https://doi.org/10.1109/TIT.2014.2320500
10. Freitag, E.: Siegelsche modulfunktionen, vol. 254. Springer-Verlag (2013)
11. Gargava, N., Serban, V., Viazovska, M.: Moments of the number of points in a bounded set for number field lattices (2024), https://arxiv.org/abs/2308.15275
12. Gargava, N., Serban, V., Viazovska, M., Viglino, I.: Effective module lattices and their shortest vectors (2024), https://arxiv.org/abs/2402.10305
13. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: 40th Annual ACM Symposium on Theory of Computing. pp. 197–206. Victoria, Canada (2008)
14. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers. Oxford university press (1979)
15. Harville, D.A.: Matrix algebra from a statistician's perspective (1998)
16. Hlawka, E.: Zur geometrie der zahlen. Mathematische Zeitschrift **49**(1), 285–312 (1943)
17. Horn, R.A., Johnson, C.R.: Matrix analysis. Cambridge university press (2012)
18. Ling, C., Luzzi, L., Belfiore, J.C., Stehlé, D.: Semantically secure lattice codes for the Gaussian wiretap channel. IEEE Transactions on Information Theory **60**(10), 6399–6416 (Oct 2014)
19. Loeliger, H.A.: Averaging bounds for lattices and linear codes. IEEE Transactions on Information Theory **43**(6), 1767–1773 (1997)
20. Luzzi, L., Ling, C., Bloch, M.R.: Optimal rate-limited secret key generation from Gaussian sources using lattices. IEEE Transactions on Information Theory **69**(8), 4944–4960 (2023)

21. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. J. ACM **60**(6), 43:1–43:35 (Nov 2013)
22. Macbeath, A., Rogers, C.: A modified form of Siegel's mean value theorem. II. In: Mathematical Proceedings of the Cambridge Philosophical Society. vol. 54, pp. 322–326. Cambridge University Press (1958)
23. Micciancio, D., Regev, O.: Worst-case to average-case reductions based on Gaussian measures. In: Proc. Ann. Symp. Found. Computer Science. pp. 372–381. Rome, Italy (Oct 2004)
24. Pouly, A., Shen, Y.: Solving the shortest vector problem in $2^{0.63269n+o(n)}$ time on random lattices. Cryptology ePrint Archive, Paper 2024/1805 (2024), `https://eprint.iacr.org/2024/1805`
25. Pouly, A., Shen, Y.: Discrete Gaussian sampling for BKZ-reduced basis. In: Niederhagen, R., Saarinen, M.J.O. (eds.) Post-Quantum Cryptography. pp. 63–88. Springer Nature Switzerland, Cham (2025)
26. Prest, T.: Sharper bounds in lattice-based cryptography using the Rényi divergence. In: Advances in Cryptology - ASIACRYPT 2017 - 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I. pp. 347–374 (2017)
27. Prest, T., Kirchner, P., Killijian, M.O., Fouque, P.A., Lyubashevsky, V., Pornin, T.: FALCON: Fast-fourier lattice-based compact signatures over NTRU. `https://falcon-sign.info/falcon.pdf` (2020), NIST Post-Quantum Cryptography Standardization, Round 3 Submission
28. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. J. ACM **56**(6), 34:1–34:40 (2009)
29. Rogers, C.A.: Mean values over the space of lattices. Acta Math. **94**, 249–287 (1955), `https://doi.org/10.1007/BF02392493`
30. Siegel, C.L.: A mean value theorem in geometry of numbers. Annals of Mathematics **46**(2), 340–347 (1945)
31. Takashima, K., Takayasu, A.: Tighter security for efficient lattice cryptography via the rényi divergence of optimized orders. In: Au, M.H., Miyaji, A. (eds.) Provable Security. pp. 412–431. Springer International Publishing, Cham (2015)
32. Temkin, L.S.: Inequality. Oxford University Press (1993)
33. Terras, A.: Fourier coefficients of Eisenstein series of one complex variable for the special linear group. Transactions of the Americal Mathematical Society **205**, 97–114 (1975)
34. Wang, Z., Ling, C.: Lattice Gaussian sampling by Markov Chain Monte Carlo: Bounded distance decoding and trapdoor sampling. IEEE Transactions on Information Theory **65**(6), 3630–3645 (2019). `https://doi.org/10.1109/TIT.2019.2901497`

A Proof of Lemma 12

Lemma 12. *Let $m, \alpha \in \mathbb{N}$ be with $1 \le m < \alpha - 1$. For $q \in \mathbb{Z}^+$, define the set of matrices*

$$\mathcal{D}_q \triangleq \{\boldsymbol{D} \in \mathbb{Z}^{m\times(\alpha-1)} : \boldsymbol{D} = (q\boldsymbol{I}_m, \boldsymbol{C}),\ \textit{where } \boldsymbol{C} \in \mathbb{Z}^{m\times(\alpha-1-m)}, \boldsymbol{C} \neq \boldsymbol{0}\}$$

Let $\boldsymbol{A} = \boldsymbol{I}_{\alpha-1} + \boldsymbol{J}_{\alpha-1}$, where $\boldsymbol{J}_{\alpha-1} = \mathbf{1}_{\alpha-1}\mathbf{1}_{\alpha-1}^T$ is the all-one matrix. Then

$$\sum_{q=1}^{\infty} \sum_{\boldsymbol{D}\in\mathcal{D}_q} \frac{1}{s^{mn}} \left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|}} \right)^n = O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}} \right).$$

More generally, for any permutation matrix $\boldsymbol{P}$, the same result holds when summing over $\mathcal{D}_q(\boldsymbol{P}) \triangleq \{\boldsymbol{D} : \boldsymbol{DP} = (q\boldsymbol{I}_m, \boldsymbol{C}), \boldsymbol{C} \neq \boldsymbol{0}\}$ instead. Note that the case above corresponds to $\mathcal{D}_q = \mathcal{D}_q(\boldsymbol{I}_{\alpha-1})$.

Proof. Note that the general case can be reduced to the special case where $\boldsymbol{D} \in \mathcal{D}_q$, because

$$(\boldsymbol{DP})\boldsymbol{A}(\boldsymbol{DP})^{\boldsymbol{T}} = \boldsymbol{DP}(\boldsymbol{I}_{\alpha-1} + \boldsymbol{J}_{\alpha-1})\boldsymbol{P}^T\boldsymbol{D}^T = \boldsymbol{D}(\boldsymbol{I}_{\alpha-1} + \boldsymbol{J}_{\alpha-1})\boldsymbol{D}^T = \boldsymbol{DAD}^T.$$

Hence, we assume without loss of generality that $\boldsymbol{D} = (q\boldsymbol{I}_m, \boldsymbol{C})$.

Let $\boldsymbol{C} = (c_{ij}) \in \mathbb{Z}^{m\times(\alpha-1-m)} \setminus \boldsymbol{0}_{m\times(\alpha-1-m)}$. Define the composite radius R associated with the parameters q and $\boldsymbol{C}$ as

$$R(q, \boldsymbol{C}) \triangleq \sqrt{q^2 + \|\boldsymbol{C}\|_F^2} = \sqrt{q^2 + \sum_{i=1}^{m}\sum_{j=1}^{\alpha-1-m} c_{ij}^2}.$$

where $\|\boldsymbol{C}\|_F$ denotes the Frobenius norm of $\boldsymbol{C}$. Define

$$N(r) \triangleq \#\{(q, \boldsymbol{C}) \mid R(q, \boldsymbol{C}) = \sqrt{r}\},$$

i.e., $N(r)$ denotes the number of pairs $(q, \boldsymbol{C})$ for which $R(q, \boldsymbol{C}) = \sqrt{r}$. Then, we have the inequality

$$N(r) \leq (2\sqrt{r})^{m(\alpha-1-m)+1}, \tag{30}$$

which is derived by counting all possible values of $-\sqrt{r} \leq c_{ij}, q \leq \sqrt{r}$.

Then we partition the sum based on a fixed threshold R_0 dependent on m,

$$\begin{aligned}
&\sum_{q=1}^{\infty}\sum_{\boldsymbol{D}\in\mathcal{D}_q} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n \\
&= \sum_{q=1}^{\infty}\sum_{\substack{\boldsymbol{C}\in\mathbb{Z}^{m\times(\alpha-1-m)} \\ \boldsymbol{C}\neq\boldsymbol{0}}} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n \\
&= \sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})\geq R_0} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n + \sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})< R_0} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n,
\end{aligned} \tag{31}$$

where R_0 is a sufficiently large positive integer independent of n (to be determined later). In what follows, we estimate both sums for sufficiently large n.

1. Evaluate the infinite sum: $\sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})\geq R_0} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n$.

   We observe that the determinant $|\boldsymbol{DAD}^T|$ is a homogeneous polynomial of degree $2m$ with respect to the variables q and c_{ij}. Since $\boldsymbol{A} = \boldsymbol{I}_{\alpha-1}+\boldsymbol{J}_{\alpha-1} \succ 0$, hence $|\boldsymbol{DAD}^T| > 0$. Let us define the function

   $$g(q, c_{11}, \ldots, c_{ij}, \ldots) \triangleq |\boldsymbol{DAD}^T| > 0. \tag{32}$$

As a homogeneous polynomial of degree $2m$, this function satisfies the scaling property:

$$g(\lambda q, \lambda c_{11}, \ldots . \lambda c_{ij}, \ldots) = \lambda^{2m} g(q, c_{11}, \ldots, c_{ij}, \ldots),$$

for any scalar $\lambda > 0$. Specifically, let $\lambda = 1/R(q, \boldsymbol{C}) = 1/\sqrt{q^2 + \sum c_{ij}^2}$, then we have:

$$\begin{aligned} g(q, c_{11}, \ldots, c_{ij}, \ldots) &= R(q, \boldsymbol{C})^{2m} g\left(\frac{q}{\sqrt{q^2 + \sum c_{ij}^2}}, \frac{c_{00}}{\sqrt{q^2 + \sum c_{ij}^2}}, \ldots, \frac{c_{ij}}{\sqrt{q^2 + \sum c_{ij}^2}}, \ldots \right) \\ &= R(q, \boldsymbol{C})^{2m} g(q', c'_{11}, \ldots, c'_{ij}, \ldots), \end{aligned}$$

where $q' = \frac{q}{\sqrt{q^2 + \sum c_{ij}^2}}$ and $c'_{ij} = \frac{c_{ij}}{\sqrt{q^2 + \sum c_{ij}^2}}$.
This transformation normalizes our variables to the unit sphere, as can be verified by:

$$q'^2 + \sum c'^2_{ij} = \frac{q^2}{q^2 + \sum c_{ij}^2} + \frac{\sum c_{ij}^2}{q^2 + \sum c_{ij}^2} = 1.$$

Therefore, we can interpret $g(q', c'_{11}, \ldots, c'_{ij}, \ldots)$ as the evaluation of our polynomial $|\boldsymbol{DAD}^T|$ over the $m(\alpha - m - 1) + 1$ dimensional unit sphere. Thus let M be the minimum of the function over the unit sphere, i.e.

$$M \triangleq \min_{q'^2 + \sum c'^2_{ij} = 1} g(q', c'_{11}, \ldots, c'_{ij}, \ldots).$$

Here $M > 0$ since it is a normalization of Eq. (32).
Consequently, we have the lower bound:

$$|\boldsymbol{DAD}^T| = g(q, c_{11}, \ldots, c_{ij}, \ldots) = R(q, \boldsymbol{C})^{2m} g(q', c'_{11}, \ldots, c'_{ij}, \ldots) \geq R(q, \boldsymbol{C})^{2m} M.$$

Using this property, we proceed to bound the given sum:

$$\begin{aligned} &\sum_{(q, \boldsymbol{C}): R(q, \boldsymbol{C}) \geq R_0} \frac{1}{s^{mn}} \left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}} \right)^n \qquad (33) \\ &\leq \sum_{(q, \boldsymbol{C}): R(q, \boldsymbol{C}) \geq R_0} \frac{1}{s^{mn}} \left(\frac{R(q, \boldsymbol{C})^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}} \right)^n \leq \sum_{(q, \boldsymbol{C}): R(q, \boldsymbol{C}) \geq R_0} \frac{1}{s^{mn}} \frac{1}{(\sqrt{M} R(q, \boldsymbol{C}))^n} \\ &= \frac{1}{s^{mn}} \sum_{r = R_0^2}^{\infty} \frac{N(r)}{(Mr)^{\frac{n}{2}}} \\ &\leq \frac{1}{s^{mn}} \frac{2^{m(\alpha - 1 - m) + 1}}{M^{\frac{n}{2}}} \sum_{r = R_0^2}^{\infty} \frac{1}{r^{(n - m(\alpha - 1 - m) - 1)/2}} \qquad \text{by (30)} \end{aligned}$$

$$= \frac{1}{s^{mn}} \frac{2^{m(\alpha-1-m)+1}}{M^{\frac{n}{2}}} \sum_{r=R_0^2}^{\infty} \int_r^{r+1} \frac{dx}{r^{(n-m(\alpha-1-m)-1)/2}}$$

$$\leq \frac{1}{s^{mn}} \frac{2^{m(\alpha-1-m)+1}}{M^{\frac{n}{2}}} \int_{R_0^2}^{\infty} \frac{dx}{(x-1)^{(n-m(\alpha-1-m)-1)/2}}$$

$$= \frac{2^{m(\alpha-1-m)+2}}{n-m(\alpha-1-m)-3} \cdot \frac{1}{(s^2)^{\frac{mn}{2}} M^{\frac{n}{2}}} \cdot \frac{1}{(R_0^2-1)^{(n-m(\alpha-1-m)-3)/2}}$$

$$= O\left(\frac{1}{s^{mn}\big(M(R_0^2-1)\big)^{\frac{n}{2}}}\right) \quad \text{for sufficiently large } n. \tag{34}$$

2. Evaluate the finite sum $\sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})<R_0} \frac{1}{s^{mn}} \left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{DAD}^T|}}\right)^n$.
Note that the sum over all terms where $R(q,\boldsymbol{C}) < R_0$ is finite.
Let $s_1 \geq s_2 \geq \ldots \geq s_m$ be the m singular values of matrix $\boldsymbol{C}$. Since the singular values of $\boldsymbol{C}$ are the square roots of the eigenvalues of $\boldsymbol{CC}^T$, we find that s_i^2 for $i = 1, 2, \ldots, m$ are the eigenvalues of $\boldsymbol{CC}^T$ and satisfy the following properties:

$$\prod_{i=1}^{m} \left(q^2 + s_i^2\right) = \left|q^2 \boldsymbol{I}_m + \boldsymbol{CC}^T\right|,$$

$$\sum_{i=1}^{m} s_i^2 = \mathrm{Tr}(\boldsymbol{CC}^T) = \sum_{i,j} c_{ij}^2.$$

Denote $\boldsymbol{d}_m = \boldsymbol{D}\mathbf{1}_{\boldsymbol{\alpha-1}} = \left(q + \sum_j c_{1j},\ q + \sum_j c_{2j},\ \ldots,\ q + \sum_j c_{mj}\right)^T$ and the determinant of $\boldsymbol{DAD}^T$ is computed as follows,

$$\begin{aligned}
|\boldsymbol{DAD}^T| &= |q^2 \boldsymbol{I}_m + \boldsymbol{CC}^T + \boldsymbol{d}_m \boldsymbol{d}_m^T| \\
&= \left|q^2 \boldsymbol{I}_m + \boldsymbol{CC}^T\right| \cdot \left|\boldsymbol{I}_m + \left(q^2 \boldsymbol{I}_m + \boldsymbol{CC}^T\right)^{-1} \boldsymbol{d}_m \boldsymbol{d}_m^T\right| && (35) \\
&= \prod_{i=1}^{m} \left(q^2 + s_i^2\right) \cdot \left(1 + \boldsymbol{d}_m^T \left(q^2 \boldsymbol{I}_m + \boldsymbol{CC}^T\right)^{-1} \boldsymbol{d}_m\right) && (36) \\
&\geq \prod_{i=1}^{m} \left(q^2 + s_i^2\right) \cdot \left(1 + \frac{\boldsymbol{d}_m^T \boldsymbol{d}_m}{q^2 + s_1^2}\right) && (37) \\
&= \prod_{i=1}^{m} \left(q^2 + s_i^2\right) + \prod_{i=2}^{m} \left(q^2 + s_i^2\right) \boldsymbol{d}_m^T \boldsymbol{d}_m \\
&\geq q^{2m} + q^{2m-2} \sum_{i=1}^{m} s_i^2 + q^{2m-2} \boldsymbol{d}_m^T \boldsymbol{d}_m && (38) \\
&= q^{2m-2} \left(q^2 + \sum_{i=1}^{m} \left((q + \sum_j c_{ij})^2 + \sum_j c_{ij}^2\right)\right)
\end{aligned}$$

$$\begin{aligned} &\geq q^{2m-2}(q^2+m) \\ &\geq q^{2m-2}(1+m). \end{aligned} \tag{39}$$

Here (35) is derived by identity $\left|\boldsymbol{A}+\boldsymbol{B}\right| = \left|\boldsymbol{A}\right| \cdot \left|\boldsymbol{I}+\boldsymbol{A}^{-1}\boldsymbol{B}\right|$. (36) is derived by noting that $\left(q^2\boldsymbol{I}_m+\boldsymbol{C}\boldsymbol{C}^T\right)^{-1}\boldsymbol{d}_m\boldsymbol{d}_m^T$ is a dyadic product matrix of a column vector $\left(q^2\boldsymbol{I}_m+\boldsymbol{C}\boldsymbol{C}^T\right)^{-1}\boldsymbol{d}_m$ and a row vector $\boldsymbol{d}_m^T$, and apply the matrix-determinant lemma [15, p. 446], i.e. $|\boldsymbol{I}+\boldsymbol{u}\boldsymbol{v}^T| = 1+\boldsymbol{v}^T\boldsymbol{u}$. (37) is derived by Rayleigh-Ritz theorem [17, p. 176]. To establish (38), we begin by applying the inequality $\prod_{i=1}^m\left(q^2+s_i^2\right) \geq q^{2m}+q^{2m-2}\sum_{i=1}^m s_i^2$ which provides a lower bound for the first term, and subsequently employ the inequality $\prod_{i=2}^m\left(q^2+s_i^2\right) \geq q^{2m-2}$ on the last term. (39) is obtained from the fact that $q \geq 1$, and for each i we have $(q+\sum_j c_{ij})^2+\sum_j c_{ij}^2 \geq 1$ holds. This is because the result is always a non-zero integer for any $c_{ij} \in \mathbb{Z}$ and $q \in \mathbb{Z}^+$.

Then based on (39), it can be obtained that

$$\begin{aligned} \sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})<R_0} \frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|}}\right)^n &\leq \sum_{(q,\boldsymbol{C}):R(q,\boldsymbol{C})<R_0} \frac{1}{s^{mn}}\frac{1}{(m+1)^{n/2}} \\ &= O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}}\right), \end{aligned} \tag{40}$$

where the last step is derived since it is a finite sum, and the number of terms does not depend on n.

By choosing

$$R_0 \geq \sqrt{1+\frac{m+1}{M}},$$

and plugging the infinite sum bound (34) and finite sum bound (40) into (31), the proof of the lemma will be finished,

$$\begin{aligned} &\sum_{q=1}^{\infty}\sum_{\boldsymbol{D}\in\mathcal{D}_q}\frac{1}{s^{mn}}\left(\frac{q^{m-1}}{\sqrt{|\boldsymbol{D}\boldsymbol{A}\boldsymbol{D}^T|}}\right)^n \\ &\leq O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}}\right)+O\left(\frac{1}{s^{mn}\left(M(R_0^2-1)\right)^{\frac{n}{2}}}\right) \\ &= O\left(\frac{1}{s^{mn}(m+1)^{\frac{n}{2}}}\right). \end{aligned}$$

□

B Proof of Proposition 2

Proof. 1. We begin with proof of the recurrence relation claim of $S_{i,j} \le p^i S_{i-1,j+1} + S_{0,j} S_{i-1,1}$ for $i \in \mathbb{N}^+$ and $j \ge 1$.

$$
\begin{aligned}
S_{i,j} &= \mathbb{E}_{\boldsymbol{G}}\left[\Big(\sum_{\boldsymbol{m}\in\mathbb{F}_p^k} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})\Big)^i \sum_{\boldsymbol{m}'\in\mathbb{F}_p^k} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}')^j\right] \\
&= \mathbb{E}_{\boldsymbol{G}}\left[\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^k} \prod_{l=1}^{i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_l) \sum_{\boldsymbol{m}'\in\mathbb{F}_p^k} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}')^j\right] \\
&= \underbrace{\mathbb{E}_{\boldsymbol{G}}\left[\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^k} \prod_{l=1}^{i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_l) \sum_{\boldsymbol{m}'\in U_i\setminus\{\boldsymbol{0}\}} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}')^j\right]}_{(i)} \\
&\quad + \underbrace{\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^k} \mathbb{E}_{\boldsymbol{G}\boldsymbol{m}_1,\ldots,\boldsymbol{G}\boldsymbol{m}_i}\left[\prod_{l=1}^{i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_l)\Big(\varphi_{\boldsymbol{x}}(\boldsymbol{0})^j + \mathbb{E}_{\boldsymbol{G}|\boldsymbol{G}\boldsymbol{m}_1,\ldots,\boldsymbol{G}\boldsymbol{m}_i}\Big[\sum_{\boldsymbol{m}'\notin U_i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}')^j\Big]\Big)\right]}_{(ii)}.
\end{aligned}
\tag{41}
$$

Here we denote $U_i = \mathrm{span}\{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\}$, and divide the vectors of $\boldsymbol{m}'$'s into two parts depending on whether they belong to $U_i \setminus \{\boldsymbol{0}\}$ or not. For vectors not in U_i, we split the expectation into conditional fixed codewords $\{\boldsymbol{G}\boldsymbol{m}_1,\ldots,\boldsymbol{G}\boldsymbol{m}_i\}$ in the last step. Now we estimate (i) and (ii) separately. To estimate (i), we need the following inequality.

Lemma 13 (Rearrangement Inequality, [32])**.** *Let $a_1 \le a_2 \le \cdots \le a_n$ and $b_1 \le b_2 \le \cdots \le b_n$ be sequences of real numbers. Then,*

$$a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \ge a_{s(1)} b_1 + a_{s(2)} b_2 + \cdots + a_{s(n)} b_n,$$

for any permutation s of $\{1,2,\ldots,n\}$.

Now we begin with

$$\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^*} \prod_{\ell=1}^{i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_\ell) \sum_{\boldsymbol{m}'\in U_i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})^j.$$

Since any $\boldsymbol{m}'$ can be written as $\boldsymbol{m} = \sum_{\ell=1}^{i} d_\ell\, \boldsymbol{m}_\ell$ with $d = (d_1,\ldots,d_i) \in \mathbb{F}_p^i \setminus \{0\}$, this becomes

$$\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^*} \prod_{\ell=1}^{i} \varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_\ell) \sum_{\substack{d=(d_1,\ldots,d_i)\in\mathbb{F}_p^i \\ d\neq 0}} \varphi_{\boldsymbol{x}}\Big(\boldsymbol{G},\sum_{\ell=1}^{i} d_\ell\, \boldsymbol{m}_\ell\Big)^j \quad \text{(some linear combinations may coincide).}$$

Re-ordering the sums,

$$\sum_{d\in\mathbb{F}_p^i\setminus\{0\}}\ \sum_{\boldsymbol{m}_1,\dots,\boldsymbol{m}_i\in\mathbb{F}_p^*}\Big(\prod_{\ell=1}^{i}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_\ell)\Big)\ \varphi_{\boldsymbol{x}}\Big(\boldsymbol{G},\sum_{\ell=1}^{i}d_\ell\,\boldsymbol{m}_\ell\Big)^j.$$

Now set

$$i_0(d)\ =\ \min\{\ell : d_\ell\neq 0\},$$

and group by which coordinate is first nonzero:

$$(*)=\sum_{d\neq 0}\ \sum_{\substack{\boldsymbol{m}_1,\dots,\boldsymbol{m}_{i_0-1},\\ \boldsymbol{m}_{i_0+1},\dots,\boldsymbol{m}_i\in\mathbb{F}_p^*}}\Big(\prod_{\substack{\ell=1\\ \ell\neq i_0}}^{i}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_\ell)\Big)\sum_{\boldsymbol{m}_{i_0}\in\mathbb{F}_p^*}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_{i_0})\ \varphi_{\boldsymbol{x}}\Big(\boldsymbol{G},\sum_{\ell=1}^{i}d_\ell\,\boldsymbol{m}_\ell\Big)^j.$$

Note that as $\boldsymbol{m}_{i_0}$ runs through $\mathbb{F}_p^*$, since $d_{i_0}\neq 0$, the linear combination $\sum_\ell d_\ell\,\boldsymbol{m}_\ell$ also runs through $\mathbb{F}_p^*$. Hence by the "crypto-lemma" $\varphi_{\boldsymbol{x}}\big(\boldsymbol{G},\sum_\ell d_\ell\boldsymbol{m}_\ell\big)$ is just a permutation of $\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_{i_0})$.
Applying Lemma 13 to the marked term $(*)$ gives

$$(*)\ \le\ \sum_{d\neq 0}\ \sum_{\boldsymbol{m}_1,\dots,\boldsymbol{m}_i\in\mathbb{F}_p^*}\Big(\prod_{\ell\neq i_0}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_\ell)\Big)\sum_{\boldsymbol{m}_{i_0}\in\mathbb{F}_p^*}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_{i_0})^{\,j+1}.$$

Since $\sum_{d\neq 0}1=p^i-1$ and the remaining products factor,

$$(i)\le(p^i-1)\mathbb{E}_{\boldsymbol{G}}\Big(\sum_{\boldsymbol{m}\in\mathbb{F}_p^*}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})\Big)^{i-1}\ \sum_{\boldsymbol{m}\in\mathbb{F}_p^*}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})^{\,j+1}\ <\ p^{\,i-1}\,S_{i-1,j+1},\tag{42}$$

Now we estimate (ii). The Gaussian binomial coefficient, denoted

$$\begin{bmatrix}n\\k\end{bmatrix}_p=\frac{(1-p^n)(1-p^{n-1})\cdots(1-p^{n-k+1})}{(1-p^k)(1-p^{k-1})\cdots(1-p)},$$

counts the number of k-dimensional subspaces of an n-dimensional vector space over $\mathbb{F}_p$. Denote $V=\mathrm{span}\{\boldsymbol{G}\boldsymbol{m}_1,\dots,\boldsymbol{G}\boldsymbol{m}_i\}$ and then $V^\perp$ as the complement linear space of V with $\dim(V^\perp)=n-\dim(V)$. Thus

$$\mathbb{F}_p^n=V\oplus V^\perp.$$

If V is fixed, the number of linear codes in $\mathcal{B}$ containing V is $\begin{bmatrix}n-\dim(V)\\k-\dim(V)\end{bmatrix}_p$ and the number of linear codes in $\mathcal{B}$ both containing V and a given $\bar{\boldsymbol{v}}\notin V$ is $\begin{bmatrix}n-\dim(V)-1\\k-\dim(V)-1\end{bmatrix}_p$. Note that $\bar{\boldsymbol{v}}$ is a combination of vectors from V and $V^\perp$. Thus we have

$$\mathbb{E}_{\boldsymbol{G}|\boldsymbol{G}\boldsymbol{m}_1,\ldots,\boldsymbol{G}\boldsymbol{m}_i}\left[\sum_{\boldsymbol{m}\notin U_i}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m})^j\right] = \frac{\begin{bmatrix}n-\dim(V)-1\\ k-\dim(V)-1\end{bmatrix}_p}{\begin{bmatrix}n-\dim(V)\\ k-\dim(V)\end{bmatrix}_p}\sum_{\bar{\boldsymbol{v}}\notin V}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j$$
$$= \frac{p^{k-\dim(V)}-1}{p^{n-\dim(V)}-1}\sum_{\bar{\boldsymbol{v}}\notin V}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j \le \frac{p^k-1}{p^n-1}\sum_{\bar{\boldsymbol{v}}\in\mathbb{F}_p^n\setminus\{\boldsymbol{0}\}}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j$$

Thus

$$(ii) \le \sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^k}\mathbb{E}_{\boldsymbol{G}\boldsymbol{m}_1,\ldots,\boldsymbol{G}\boldsymbol{m}_i}\left[\prod_{l=1}^{i}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_l)\right]\left(\varphi_{\boldsymbol{x}}(\boldsymbol{0})^j + \frac{p^k-1}{p^n-1}\sum_{\bar{\boldsymbol{v}}\in\mathbb{F}_p^n\setminus\{\boldsymbol{0}\}}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j\right)$$
$$= S_{0,j}\mathbb{E}_{\boldsymbol{G}}\left[\sum_{\boldsymbol{m}_1,\ldots,\boldsymbol{m}_i\in\mathbb{F}_p^k}\prod_{l=1}^{i}\varphi_{\boldsymbol{x}}(\boldsymbol{G},\boldsymbol{m}_l)\right]$$
$$= S_{0,j}S_{i-1,1}. \tag{43}$$

Substituting inequalities (42) and (43) into (41) yields the proof of the first claim of (22).

2. Now we prove the first claim of (23) by induction on i.
With the case $i=1$ settled in part 1, suppose the desired inequality is true for all integers $i' \le i$. We then turn to deriving an upper bound for $S_{i+1,j}$

$$S_{i+1,j} \le p^{i+1}S_{i,j+1} + S_{0,j}S_{i,1}$$
$$\le p^{i+1}S_{0,j+1}S_{0,1}^i + p^{i+1}\cdot p^{\frac{i(i+1)}{2}}\sum_{k_1+\cdots+k_i=i}S_{0,j+1+k_1}\prod_{l=2}^{i}S_{0,k_l}$$
$$+ S_{0,j}S_{0,1}^{i+1} + p^{\frac{i(i+1)}{2}}\sum_{k_1+\cdots+k_i=i}S_{0,j}\cdot S_{0,1+k_1}\prod_{l=2}^{i}S_{0,k_l} \tag{44}$$
$$\le S_{0,j}S_{0,1}^{i+1} + p^{\frac{(i+1)(i+2)}{2}}\sum_{k_1+\cdots+k_{i+1}=i+1}S_{0,j+k_1}\prod_{l=2}^{i+1}S_{0,k_l}. \tag{45}$$

(44) is derived by induction. Thus the claim of upper bound for $S_{i,j}$ for $i\in\mathbb{N}^+$ and $j\ge 1$ follows.

□

C Proof of Lemma 10

Proof. First it is obvious that for $r\ge 1$,

$$N(r) \;=\; \#\big\{\,\boldsymbol{u}\in\mathbb{Z}^n : \|\boldsymbol{u}\|^2 = r\big\} \le (2\sqrt{r}+1)^n \le (3\sqrt{r})^n.$$

Then for some fixed β, as $p = \eta_n n^{1+2\varepsilon} \to \infty, \gamma = 1/\eta_n^{\frac{1}{2}} n^{\frac{1}{2}+\varepsilon} \to 0, \gamma^n p^{n-k} = 1, k = \frac{n}{2}, \varepsilon > 0$

$$\begin{aligned}\sum_{\xi \in \frac{1}{\gamma}\mathbb{Z}^n \setminus \{\mathbf{0}\}} e^{-\beta^2 \|\xi\|^2} &= \sum_{\boldsymbol{u} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}} e^{-\frac{\beta^2}{\gamma^2}\|\boldsymbol{u}\|^2} \le \sum_{r=1}^{\infty} \sum_{\substack{\boldsymbol{u} \in \mathbb{Z}^n \\ \|\boldsymbol{u}\|^2 = r}} e^{-\frac{\beta^2}{\gamma^2} r} \\ &= \sum_{r=1}^{\infty} N(r) e^{-\frac{\beta^2}{\gamma^2} r} \le \sum_{r=1}^{\infty} (3\sqrt{r})^n e^{-\frac{\beta^2}{\gamma^2} r}. \end{aligned} \tag{46}$$

Denote $g(r) = (3\sqrt{r})^n e^{-\frac{\beta^2}{\gamma^2} r}$, The critical point $r_0(n)$ is determined by

$$\frac{d \ln g(r)}{dr} = \tfrac{n}{2r} - \tfrac{\beta^2}{\gamma^2} = 0 \Rightarrow r_0(n) = \frac{n\,\gamma^2}{2\,\beta^2}.$$

Thus

$$\begin{aligned}\sum_{r \le r_0(n)} (3\sqrt{r})^n \exp\Big(-\tfrac{\beta^2}{\gamma^2} r\Big) \;&\le\; r_0(n)\,(3\sqrt{r_0(n)})^n \exp\Big(-\tfrac{\beta^2}{\gamma^2} r_0(n)\Big) \\ &=\; 3^n \Big(\tfrac{\gamma\sqrt{n}}{\sqrt{2}\,\beta}\Big)^{n+2} e^{-\,n/2} = O\Big(3^n\, e^{-n/2}\, n^{-\varepsilon(n+2)}\Big) \\ &= e^{-\omega(n)}. \end{aligned} \tag{47}$$

Continuing from Eq. (46), by (47) we have

$$\begin{aligned}\sum_{\xi \in \frac{1}{\gamma}\mathbb{Z}^n \setminus \{\mathbf{0}\}} e^{-\beta^2 \|\xi\|^2} &= \sum_{r \le r_0(n)} (3\sqrt{r})^n \exp\Big(-\tfrac{\beta^2}{\gamma^2} r\Big) + \sum_{r > r_0(n)} (3\sqrt{r})^n \exp\Big(-\tfrac{\beta^2}{\gamma^2} r\Big) \\ &\le e^{-\omega(n)} + \int_0^{\infty} (3\sqrt{r})^n e^{-\frac{\beta^2}{\gamma^2} r} dr \\ &= e^{-\omega(n)} + \frac{3^n \gamma^{n+2}}{\beta^{n+2}} \Gamma\big(\tfrac{n}{2}+1\big) = e^{-\omega(n)} + \frac{3^n}{\beta^{n+2}} \cdot \frac{e^{\frac{n+2}{2}\log\frac{n+2}{2} - \frac{n}{2} + O(1)}}{e^{\frac{n+2}{2}\log n + \epsilon(n+2)\log n}} \\ &= e^{-\omega(n)}. \end{aligned} \tag{48}$$

Similarly, we have

$$\sum_{\boldsymbol{u} \in p\gamma\mathbb{Z}^n \setminus \{\mathbf{0}\}} e^{-\frac{1}{j}\|\boldsymbol{u}\|^2} \le e^{-\omega(n)}. \tag{49}$$

Given $\rho_s(\boldsymbol{x}) = \frac{1}{s^n} e^{-\pi \|\boldsymbol{x}\|^2 / s^2}$ and assuming $\boldsymbol{u}_0 = \mathbf{0}$, we compute the Fourier transform in $\boldsymbol{v}$ of $\prod_{r=0}^{j-1} \rho_s(\boldsymbol{x} + \boldsymbol{v} + \boldsymbol{u}_r)$ as follows:

$$\mathcal{F}\left\{\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\boldsymbol{v}+\boldsymbol{u}_r)\right\}(\boldsymbol{\xi})$$
$$=\exp\left(\frac{\frac{1}{j}\left\|\sum_{r=1}^{j-1}\boldsymbol{u}_r\right\|^2-\sum_{r=1}^{j-1}\|\boldsymbol{u}_r\|^2}{s^2}+2\pi\mathrm{i}\left\langle\boldsymbol{x}+\boldsymbol{u}',\boldsymbol{\xi}\right\rangle\right)\cdot\mathcal{F}\left\{\frac{1}{s^{jn}}\exp\left(-\frac{\pi j\|\boldsymbol{v}\|^2}{s^2}\right)\right\}(\boldsymbol{\xi}) \tag{50}$$

$$\Rightarrow\left|\mathcal{F}\left\{\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\boldsymbol{v}+\boldsymbol{u}_r)\right\}(\boldsymbol{\xi})\right|\le\exp\left(-\frac{1}{s^2j}\sum_{r=1}^{j-1}\|\boldsymbol{u}_r\|^2\right)\cdot\frac{1}{s^{n(j-1)}j^{n/2}}e^{-\frac{\pi s^2}{j}\|\boldsymbol{\xi}\|^2}. \tag{51}$$

Here in (50) phase shift property is applied and $\boldsymbol{u}'=-\frac{1}{j}\sum_{r=1}^{j-1}\boldsymbol{u}_r$. (51) is obtained by applying Cauchy-Schwarz inequality $\left\|\sum_{r=1}^{j-1}\boldsymbol{u}_r\right\|^2\le(j-1)\sum_{r=1}^{j-1}\|\boldsymbol{u}_r\|^2$, and $|e^{-2\pi\mathrm{i}\left\langle\boldsymbol{x}+\boldsymbol{u}',\boldsymbol{\xi}\right\rangle}|=1$.

With the tools and inequalities above, we are able to present the proof of the lemma. When $j\in\mathbb{N}^+$, it can be derived that

$$\begin{aligned}
S_{0,j}&=\varphi_x(\mathbf{0})^j+\frac{p^k-1}{p^n-1}\sum_{\bar{\boldsymbol{v}}\in\mathbb{F}_p^n\setminus\{\mathbf{0}\}}\varphi_{\boldsymbol{x}}(\bar{\boldsymbol{v}})^j\\
&\le\varphi_x(\mathbf{0})^j+p^{k-n}\sum_{\bar{\boldsymbol{v}}\in\mathbb{F}_p^n\setminus\{\mathbf{0}\}}\left(\sum_{\boldsymbol{u}\in p\gamma\mathbb{Z}^n}\rho_s(\boldsymbol{x}+\gamma\bar{\boldsymbol{v}}+\boldsymbol{u})\right)^j\\
&=\varphi_x(\mathbf{0})^j+p^{k-n}\sum_{\bar{\boldsymbol{v}}\in\gamma\mathbb{F}_p^n\setminus\{\mathbf{0}\}}\sum_{\boldsymbol{u}\in p\gamma\mathbb{Z}^n}\sum_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1}\in p\gamma\mathbb{Z}^n}\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\bar{\boldsymbol{v}}+\boldsymbol{u}+\boldsymbol{u}_r)\\
&=\varphi_x(\mathbf{0})^j+\sum_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1}\in p\gamma\mathbb{Z}^n}p^{k-n}\sum_{\bar{\boldsymbol{v}}\in\gamma\mathbb{F}_p^n\setminus\{\mathbf{0}\}}\sum_{\boldsymbol{u}\in p\gamma\mathbb{Z}^n}\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\bar{\boldsymbol{v}}+\boldsymbol{u}+\boldsymbol{u}_r)\\
&\le\varphi_x(\mathbf{0})^j+\sum_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1}\in p\gamma\mathbb{Z}^n}p^{k-n}\sum_{\boldsymbol{v}\in\gamma\mathbb{Z}^n}\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\boldsymbol{v}+\boldsymbol{u}_r)\\
&=\varphi_x(\mathbf{0})^j+\sum_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1}\in p\gamma\mathbb{Z}^n}\frac{1}{p^{n-k}\gamma^n}\sum_{\xi\in\frac{1}{\gamma}\mathbb{Z}^n}\mathcal{F}\left\{\prod_{r=0}^{j-1}\rho_s(\boldsymbol{x}+\boldsymbol{v}+\boldsymbol{u}_r)\right\}(\boldsymbol{\xi})\quad\text{by Poisson summation formula}\\
&\le\varphi_x(\mathbf{0})^j+\sum_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_{j-1}\in p\gamma\mathbb{Z}^n}\sum_{\xi\in\in\frac{1}{\gamma}\mathbb{Z}^n}e^{-\frac{1}{s^2j}\sum_{r=1}^{j-1}\|\boldsymbol{u}_r\|^2}\cdot\frac{1}{s^{n(j-1)}j^{n/2}}e^{-\frac{\pi s^2}{j}\|\boldsymbol{\xi}\|^2}\ \text{by Eq. (51)}
\end{aligned}$$

$$
\begin{aligned}
&= \varphi_x(\mathbf{0})^j + \frac{1}{s^{n(j-1)} j^{n/2}} \Bigg(\sum_{\xi \in \in \frac{1}{\gamma}\mathbb{Z}^n} e^{-\frac{\pi s^2}{j}\|\boldsymbol{\xi}\|^2} \Bigg) \Bigg(\sum_{\boldsymbol{u} \in p\gamma\mathbb{Z}^n} e^{-\frac{1}{s^2 j}\|\boldsymbol{u}\|^2} \Bigg)^{j-1} \\
&\leq \varphi_x(\mathbf{0})^j + \frac{1}{s^{n(j-1)} j^{n/2}} (1 + e^{-\omega(n)})(1 + e^{-\omega(n)})^{j-1} \text{ by Eq. (48) and (49)} \\
&= \varphi_x(\mathbf{0})^j + \frac{1}{s^{n(j-1)} j^{n/2}} + e^{-\omega(n)} \\
&= \varphi_{\boldsymbol{x}}(\mathbf{0})^j + \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^j \, d\boldsymbol{x} + e^{-\omega(n)}. \qquad (52)
\end{aligned}
$$

Meanwhile, it can be obtained that

$$
\begin{aligned}
\int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j'} \varphi(\mathbf{0})^j d\boldsymbol{x} &= \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j'} \Bigg(\sum_{\boldsymbol{u} \in p\gamma\mathbb{Z}^n} \rho_s(\boldsymbol{x} + \boldsymbol{u}) \Bigg)^j \\
&= \sum_{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j \in p\gamma\mathbb{Z}^n} \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j'} \rho_s(\boldsymbol{x} + \boldsymbol{u}_1) \ldots \rho_s(\boldsymbol{x} + \boldsymbol{u}_j) d\boldsymbol{x} \\
&\leq \sum_{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_j \in p\gamma\mathbb{Z}^n} \exp\left(\frac{-\frac{j'}{j+j'} \pi \sum_{r=1}^{j} \|\boldsymbol{u}_r\|^2}{s^2} \right) \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j+j'} \, d\boldsymbol{x} \text{ same in Eq. (51)} \\
&= \Bigg(\sum_{\boldsymbol{u} \in p\gamma\mathbb{Z}^n} e^{\frac{-\frac{j'}{j+j'} \pi \|\boldsymbol{u}\|^2}{s^2}} \Bigg)^j \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j+j'} \, d\boldsymbol{x} \\
&\leq (1 + e^{-\omega(n)}) \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j+j'} d\boldsymbol{x} \qquad \text{by Eq. (49)} \\
&= \int_{\mathbb{R}^n} \rho_s(\boldsymbol{x})^{j+j'} d\boldsymbol{x} + e^{-\omega(n)}. \qquad (53)
\end{aligned}
$$

Combining results above the proof is finished. □