Refined Evaluation Methods of Decryption Failure Rate in Lattice-Based Public-Key Encryption with Message Encoding

Guoqing Zhou[1,2], Lin Wang[1], Yue Cao[1], Baosheng Huang[1], Sen Hou[1], and Dawu Gu[2(✉)]

[1] National Key Lboratory of Security Communication, Chengdu, 610041, China
{mathzhou,wanglin4math}@outlook.com,
{2670505187,1196274553,2236468760}@qq.com

[2] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China
dwgu@sjtu.edu.cn

Abstract. Lattice-based cryptography is one of the most promising candidates for post-quantum cryptography. Decryption Failure Rate (DFR) is a critical metric for the correctness and security of lattice-based public-key encryption (PKE) schemes. At present, most DFR evaluation methods for lattice-based PKE with message encoding rely on oversimplified assumptions, rough approximations, and fail to fully exploit the geometric structure of special encoding lattices, resulting in loose or inaccurate bounds.

To address these limitations, this paper proposes a refined DFR evaluation framework for lattice-based PKE with message encoding, focusing on two mainstream decoding paradigms: Maximum Likelihood Decoding (MLD) and Bounded Distance Decoding (BDD). For MLD-based schemes, we precisely characterize the minimal vectors of Barnes-Wall (BW) lattices and derive tighter union bounds by leveraging the encoding lattice structure. For BDD-based schemes, we introduce a noncentral chi-squared distribution based method to model mixed noise composed of Gaussian and discrete components, avoiding the inaccuracy of conventional pure Gaussian approximations. Furthermore, we extend the framework to algebraic-lattice-based PKE by analyzing the variance and correlation of polynomial product coefficients, and propose a weighted chi-squared distribution with saddlepoint approximation to handle correlated coefficients.

We validate the proposed methods on representative lattice-based PKEs including CNTR, Scloud$^+$, and CNTR-Prime, resulting in the DFR upper bound approximate 15 bits decreasing for CNTR, 1 bit decreasing for Scloud$^+$, and at least 84 bits increasing for CNTR-Prime. The experimental results show that our refined analysis yields significantly tighter and more accurate DFR bounds compared with traditional approaches, providing a reliable theoretical basis for parameter selection and security certification of practical lattice-based cryptosystems.

Keywords: Lattice-based Cryptography · Cryptanalysis · Decryption Failure Rate · Lattice Code · Algebraic Structure

1 Introduction

With the rapid advancement of quantum computing, the traditional pubic key cryptosystems, such as RSA and ECC, are increasingly unable to guarantee digital security [23,24]. In response, the National Institute of Standards and Technology (NIST) initiated a dedicated effort to standardize post-quantum cryptographic algorithms, with lattice-based cryptography standing out as one of the most significant and promising candidates. Algorithms such as Kyber [22], Dilithium [18], and Falcon [20] have been selected as finalists in NIST's PQC standardization project, underscoring the importance of lattice-based systems in the post-quantum era. Due to the characteristics of lattice-based encryption algorithms[16], it introduces concepts of decryption failures and decryption failure rate (DFR). DFR refers to the probability that the decryption process fails to correctly recover the original message due to the accumulation of noises during encryption. During encryption, small noise vectors are added to ensure the security of the system. These noise vectors, typically sampled from narrow discrete distributions, ensure that the ciphertext does not leak information about the plaintext. However, this intentional noise can accumulate during the decryption process, potentially leading to decoding failures if the noises exceed the correction capabilities of the system.

While DFRs are typically very low in well-designed systems[6,10,22], they are a critical factor in determining the overall functionality and security of lattice-based encryption schemes. Obviously a failed decryption can result in data loss or even system downtime. Beyond mere functionality issues, DFR is a critical security metric: on the one hand, DFR impacts the tightness of constructing IND-CCA encryption/KEMs in the (quantum) random oracle model[12,13]. On the other hand, decryption failures leak correlated information about the secret key, enabling powerful attacks such as failure-boosting [7] and key-recovery attacks [5,8] that exploit repeated failure events to compromise the scheme. As a result, tightly bounding and minimizing DFR is essential for both correctness and provable security in practical lattice cryptosystems.

To mitigate decryption failures, modern lattice PKE schemes [14,25,21] increasingly integrate message encoding based on lattice codes [17], which enhance noise tolerance by mapping messages to well-spaced lattice points within a fundamental region of a nested lattice structure. Lattice codes provide natural error-correcting capability by maximizing the minimum distance between codewords, expanding the effective decoding region and reducing the probability that noise causes decryption failure. Within the Lindner-Peikert framework [16], the common method for DFR evaluation typically follows two paradigms: for Bounded Distance Decoding (BDD), DFR is upper-bounded by the probability that the Euclidean norm of the total noise exceeds a fixed decoding radius; for Maximum Likelihood Decoding (MLD), bounds rely on containment within the Voronoi

region of the origin, often using union bounds over minimal lattice vectors. Existing analyses frequently use Gaussian tail bounds, infinity-norm constraints, or heuristic approximations to compute these probabilities efficiently, supporting parameter selection for schemes like Scloud$^+$ [25], and CNTR [14].

Despite substantial progress, current DFR evaluation methods suffer from following notable limitations that compromise accuracy and tightness.

(1). MLD-based bounds depend on precise characterization of Voronoi regions, which are computationally intractable for high-dimensional or structured lattices like Barnes-Wall (BW) lattices, forcing researchers to use rough approximations or incomplete union bounds. Existing methods often do not exploit the full structure of specialized encoding lattices, missing opportunities to derive tighter, more precise DFR bounds.
(2). Conventional BDD analyses using Euclidean norm bounds fail to account for mixed noise distributions (e.g., combinations of Gaussian and discrete uniform noise), leading to overestimation or underestimation of DFR in schemes like Scloud$^+$.
(3). Finally, most DFR evaluations ignore the algebraic structure of polynomial rings in algebraic-lattice-based PKE, where different modulus polynomials induce distinct variance and correlation patterns in polynomial product coefficients, leading to unaccounted errors and imprecise bounds for schemes such as CNTR-Prime [15].

These gaps motivate the need for refined, structure-aware DFR analysis that accurately models real noise distributions, leverages encoding lattice geometry, and accounts for algebraic ring effects to produce tight and reliable bounds for modern post-quantum lattice cryptosystems.

1.1 Contributions.

In this work, we consider refined evaluation methods of DFR in lattice-based PKEs with message encoding, and make the following contributions.

(1). We establish a precise structure-aware DFR evaluation method for MLD solvers by deeply analyzing the minimal vectors and kissing numbers of BW_N lattices. We derive a compact union bound formula tailored for scaled BW_N lattices, which eliminates redundant terms and delivers tighter bounds than conventional approximations.
(2). We propose a novel DFR evaluation approach for BDD solvers to handle mixed noise (Gaussian + discrete). By modeling the squared Euclidean norm of total noise as a noncentral chi-squared distribution and designing a trimming strategy for efficient computation, we resolve the inaccuracy of traditional Euclidean-norm bounds in schemes with non-Gaussian noise.
(3). We extend the DFR evaluation to algebraic-lattice-based PKE by systematically analyzing the variance and coefficient correlation of polynomial products under three typical modulus polynomials(x^d+1, $x^d-x^{\frac{d}{2}}+1$, x^d-x-1). For rings with correlated coefficients (e.g., $\mathbb{Z}[x]/(x^d-x-1)$), we propose

a weighted chi-squared distribution method combined with saddlepoint approximation to accurately compute tail probabilities for correlated noise vectors.

(4). We apply the refined methods to three representative post-quantum schemes: CNTR, Scloud$^+$, and CNTR-Prime. our refined approach reduces the DFR upper bound by approximately 15 bits for CNTR, 1 bit for Scloud$^+$, and yields at least 84 bits increased bound for CNTR-Prime [1]. The results demonstrate that our framework provides more accurate and tighter DFR upper bounds than original analyses, supporting possible optimized parameter design for security, efficiency, and correctness.

1.2 Organization.

This paper is organized as follows. In Sect. 2, we introduce preliminary information related to basic knowledge. Section 3 presents a refined evaluation method of DFR with MLD solver. Section 4 presents a refined evaluation method of DFR with BDD solver, and Section 5 extends the DFR evaluation to algebraic-lattice-based PKE by analyzing coefficient variance and correlation. We conclude this paper in Sect. 6.

2 Preliminaries

We denote vectors by lower-case bold letters, e.g., $\mathbf{a}, \mathbf{b}, \mathbf{c}, \cdots$, and matrices by upper-case bold letters, e.g., $\mathbf{A}, \mathbf{B}, \mathbf{C}, \cdots$. For a vector $\mathbf{v} \in \mathbb{R}^n$, the Eucilean norm is denoted by $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$, the infinity norm is denoted by $\|\mathbf{v}\|_\infty = \max_{1 \le i \le n} |v_i|$. For positive integer $n < m$, $[n]$ denotes the set $\{0, 1, \cdots, n-1\}$ and $[n : m]$ denotes the set $\{n, n+1, \cdots, m-1\}$.

2.1 Basic definition of lattice and distribution

Definition 1 (Lattice). *A lattice $\Lambda \in \mathbb{R}^n$ is a discrete subset defined as*

$$\Lambda = \mathcal{L}(\mathbf{B}_{n\times d}) = \mathbf{B}_{n\times d} \cdot \mathbb{Z}^d = \{c_1\mathbf{b}_1 + c_2\mathbf{b}_2 + \cdots + c_d\mathbf{b}_d \mid c_i \in \mathbb{Z}, 1 \le i \le d\},$$

where $\mathbf{B}_{n\times d} = (\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_d)$, called a basis matrix of Λ, is a matrix of rank d. If $d = n$, we call Λ a full-rank lattice.

The *scaled* lattice $c \cdot \Lambda = \{c \cdot \mathbf{v} \mid \mathbf{v} \in \Lambda\}$. The minimal norm of vectors in $\Lambda \backslash \{\mathbf{0}\}$ is denoted by $\lambda_1(\Lambda)$. Lattice can be partitioned as $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \Lambda_2 \cup \cdots$, where Λ_i, the i-th *layer* from the origin, is chosen so that the norm of any vector in Λ_i is a increasing constant by i. The number of vectors in Λ_1, denoted by $\tau(\Lambda)$, is called *kissing number*.

Definition 2 (Basic region). *A basic region of the lattice $\Lambda \in \mathbb{R}^n$ is a set $\mathcal{P}_\Lambda$ that satisfies the following properties:*

[1] The code is available at https://github.com/zhou-gq/RefinedDFR.

1. *Covering property:* $\cup_{\mathbf{v}\in\Lambda}\{\mathbf{v}+\mathcal{P}_\Lambda\}=\mathbb{R}^n$*;*
2. *Partitioning property:* $\forall\ \mathbf{u},\mathbf{v}\in\Lambda$*, if* $\{\mathbf{u}+\mathcal{P}_\Lambda\}\cap\{\mathbf{v}+\mathcal{P}_\Lambda\}\neq\emptyset$*, then* $\mathbf{u}=\mathbf{v}$*.*

For a point $\mathbf{u}\in\Lambda$, the *Voronoi region* $\mathcal{V}(\mathbf{u})$, a special case of basic region, is the set of all points in $\mathbb{R}^n$ which the closest lattice point is $\mathbf{u}$, i.e.,

$$\mathcal{V}(\mathbf{u})=\{\mathbf{x}\in\mathbb{R}^n\mid\|\mathbf{x}-\mathbf{u}\|\leq\ \|\mathbf{x}-\mathbf{v}\|,\ \forall\ \mathbf{v}\in\Lambda\}.$$

Write half space $\Pi_{\mathbf{u}}(\mathbf{v})=\{\mathbf{x}\in\mathbb{R}^n\mid\|\mathbf{x}-\mathbf{u}\|\leq\ \|\mathbf{x}-\mathbf{v}\|\}$, then $\mathcal{V}(\mathbf{u})=\cap_{\mathbf{v}\in\Lambda\setminus\{\mathbf{u}\}}\Pi_{\mathbf{u}}(\mathbf{v})$. [4, Sec. 8] implies that Voronoi region is a finite intersection, i.e., $\mathcal{V}(\mathbf{u})=\cap_{\mathbf{v}\in R(\mathbf{u})}\Pi_{\mathbf{u}}(\mathbf{v})$, where $R(\mathbf{u})$ is a minimal finite set of *relative points* in lattice defining $\mathcal{V}(\mathbf{u})$. Note that $\mathbf{u}+\Lambda_1\subseteq R(\mathbf{u})$.

Next, we give the definitions on lattice hard problems CVP and BDD.

Definition 3 (CVP). *Given a lattice* Λ *and* $\mathbf{t}\in\mathbb{R}^n$*, the closest vector problem (CVP) asks to find a vector* $\mathbf{u}\in\Lambda$ *that is closest to* $\mathbf{t}$*, i.e.,* $\|\mathbf{t}-\mathbf{u}\|\leq\|\mathbf{t}-\mathbf{v}\|,\ \forall\ \mathbf{v}\in\Lambda$*.*

Definition 4 (BDD). *Given a lattice* Λ *and* $\mathbf{t}\in\mathbb{R}^n$ *such that dist*$(\mathbf{t},\Lambda)\leq r$ *for a parameter* r*, the bounded distance decode (BDD) problem asks to find a vector* $\mathbf{u}\in\Lambda$ *that is closest to* $\mathbf{t}$*, i.e.,* $\|\mathbf{t}-\mathbf{u}\|\leq\|\mathbf{t}-\mathbf{v}\|,\ \forall\ \mathbf{v}\in\Lambda$*.*

Algorithms solve the above two problems are referred as maximum likelihood decoding (MLD) solver and BDD solver respectively.

A random variable X follows the *central binomial distribution* of parameter 1, denoted by $X\leftarrow\mathrm{CBD}_1$, refers to the probability distribution where $\Pr(X=\pm1)=0.25$ and $\Pr(X=0)=0.5$. The *convolution* of two distributions refers to the distribution of the sum of two random variables following these distributions. The convolution of n independent and identically distribution D is denoted by $[n]D$. For example, the central binomial distribution of parameter η, denoted by CBD_η, is the convolution of η independent CBD_1, i.e., $\mathrm{CBD}_\eta=[\eta]\mathrm{CBD}_1$. The *constant hamming distribution* of parameter (n,h), denoted by $\mathcal{H}_{n,h}$, refers to the set of vectors containing exactly $(n-2h)$ zeros, h ones and h negative ones.

Definition 5 (Noncentral chi-squared distribution). *A variable* Y *follows noncentral chi-squared distribution* $\chi^2_{k,\lambda}$*, if* $Y=\sum_{i=1}^k(X_i+\mu_i)^2$*, where* X_i *follows standard normal distribution and noncentrality parameter* $\lambda=\sum_{i=1}^k\mu_i^2$*. The cumulative distribution function of* $\chi^2_{k,\lambda}$ *is*

$$F(x;k,\lambda):=\Pr(Y\leq x\mid Y\leftarrow\chi^2_{k,\lambda})=\sum_{j=0}^{\infty}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}\cdot\Pr(Z\leq x\mid Z\leftarrow\chi^2_{k+2j}),$$

where χ^2_{k+2j} *is the central chi-squared distribution of* $k+2j$ *degrees of freedom.*

2.2 Lattice code and BW_N lattice

In this section, we give the concepts of lattice code and corresponding message encoding and decoding function. Then we introduce BW_N lattice, a widely used encoding lattice family in lattice-based cryptosystems.

Definition 6 (Lattice code). *Two lattices Λ_f and Λ_c are nested if $\Lambda_c \subseteq \Lambda_f$. A lattice code $\mathcal{C}(\Lambda_f, \Lambda_c)$ is the set of points in Λ_f that lie within a basic region of Λ_c, i.e.,*

$$\mathcal{C}(\Lambda_f, \Lambda_c) = \Lambda_f \cap \mathcal{P}_{\Lambda_c}.$$

Lemma 1 (Diagonal nesting [26]). *Let $\Lambda_c \subseteq \Lambda_f$ be nested lattices. There exists basis matrices $\mathbf{B}_f$ and $\mathbf{B}_c$ such that*

$$\mathbf{B}_c = \mathbf{B}_f \cdot \mathrm{diag}(p_1, p_2, \cdots, p_d),$$

where p_i are positive integers for all $1 \leq i \leq d$.

From Lemma 1, there always exists a labeling function, mapping the message into lattice code.

Lemma 2 (Labeling function [17]). *Let $\Lambda_c \subseteq \Lambda_f$ be nested lattices. Basis matrices $\mathbf{B}_f$ and $\mathbf{B}_c$ satisfy $\mathbf{B}_c = \mathbf{B}_f \cdot \mathrm{diag}(p_1, p_2, \cdots, p_d)$. Let the message space be the*

$$\mathcal{M} = [p_1] \times [p_2] \cdots, \times [p_d],$$

The labeling function $f : \mathcal{M} \to \mathcal{C}(\Lambda_f, \Lambda_c),\ \mathbf{m} \mapsto (\mathbf{B}_f \mathbf{m} \bmod \Lambda_c)$ is bijective.

Based on the labeling function above, we can construct message encode and decode function as follows:

- $\mathsf{Encode} = f : \mathcal{M} \to \mathcal{C}(\Lambda_f, \Lambda_c),\ \mathbf{m} \mapsto (\mathbf{B}_f \mathbf{m} \bmod \Lambda_c)$,
- $\mathsf{Decode} : \mathbb{R}^n \to \mathcal{M}, \mathbf{v} \mapsto f^{-1}(Q_{\Lambda_f}(\mathbf{v}) \bmod \Lambda_c)$,

where function Q denotes the MLD solver or BDD solver.

The following lemma describes the correct decoding condition when involved noise.

Lemma 3 (Correct decoding condition [17]). *Let $\mathbf{e} \in \mathbb{R}^n$ be a noise, then for any $\mathbf{m} \in \mathcal{M}$, $\mathsf{Decode}(\mathsf{Encode}(\mathbf{m}) + \mathbf{e}) = \mathbf{m}$ if and only if $Q_{\Lambda_f}(\mathbf{e}) \in \Lambda_c$.*

Therefore, the probability of decoding failure can be evaluated as

$$P_e = \Pr(Q_{\Lambda_f}(\mathbf{e}) \notin \Lambda_c) \leq \Pr(Q_{\Lambda_f}(\mathbf{e}) \neq \mathbf{0}). \tag{1}$$

Now, we introduce BW_N lattices. They were first defined by Barnes and Wall [1] as follows:

Definition 7 (Barnes-Wall lattice). *Let* $\{\lambda_i\}_{i=0}^{n}$ *be integers satisfying* $\lambda_i = \lfloor \frac{i}{2} \rfloor$. *Write* $N = 2^n$, *denote* BW_N *by the lattice generated by all vectors* $2^{\lambda_{n-r}}[C_r]$, *where* C_r *runs over all cosets of dimension* r *in* $\mathbb{F}_2^n$. *For any* $W \subseteq \mathbb{F}_2^n$, $[W]$ *denotes all the vectors* $\mathbf{x} \in \{0,1\}^N$ *such that* $x_{\mathbf{a}} = \mathbf{1}_{\mathbf{a} \in W}$.

BW_N lattices can also be constructed from Reed-Muller codes.

Lemma 4 ([9]). *Let* $N = 2^n$, *then*

$$\mathrm{BW}_N = \sum_{\substack{1 \le r \le n-1 \\ r \text{ is odd}}} 2^{\frac{r-1}{2}} \cdot \mathrm{RM}(r,n) + 2^{\lfloor \frac{n}{2} \rfloor} \cdot \mathbb{Z}^N$$

where $\mathrm{RM}(r,n)$ *is the* $(2^n, \sum_{i=0}^{r} \binom{n}{i}, 2^{n-r})$ *Reed-Muller code.*

It is clear from above Lemma 4 that $2^{\lfloor \frac{n}{2} \rfloor} \cdot \mathbb{Z}^N \subseteq \mathrm{BW}_N$, which means that a lattice code can be constructed from scaled nested lattices $\Lambda_f = c \cdot \mathrm{BW}_N$ and $\Lambda_c = c \cdot 2^{\lfloor \frac{n}{2} \rfloor} \cdot \mathbb{Z}^N$.

For example, $\mathrm{BW}_8 = (8,4,4) + 2 \cdot \mathbb{Z}^8$, where $(8,4,4)$ Reed-Muller code is just the $(8,4,4)$ extended Hamming code.

2.3 General method of DFR evaluation

In this section, we introduce two general methods of DFR evaluation of lattice-based PKE. The difference between evaluation methods depends on the type of decoding function.

A PKE scheme consists of three algorithms.

- Key generation: $(pk, sk) = \mathsf{Keygen}(1^\lambda)$;
- Encryption: $c = \mathsf{Enc}(m, pk)$;
- Decryption: $\widetilde{m} = \mathsf{Dec}(c, sk)$.

Lindner-Peikert framework is the most widely used framework for lattice-based PKE. It can be described as follows:

Write the noises caused by compressing $\mathbf{u}$ and v by $\mathbf{r_u}$ and r_v respectively. From the LWE assumption, the compressing noise $\mathbf{r_u}$ and r_v of $\mathbf{u}$ and v can be simulated as

$$\mathbf{r_u} \leftarrow \left\{ \mathsf{Decompress}_1(\mathsf{Compress}_1(\mathbf{u})) - \mathbf{u} \middle| \mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^n \right\},$$
$$r_v \leftarrow \left\{ \mathsf{Decompress}_1(\mathsf{Compress}_1(v)) - v \middle| v \xleftarrow{\$} \mathbb{Z}_q \right\},$$

with standard deviation $\sigma_{\mathbf{u}}$ and σ_v respectively. In decryption algorithm,

$$\begin{aligned} \tilde{v} - \mathbf{s}^T \tilde{\mathbf{u}} &= v + r_v - \mathbf{s}^T(\mathbf{u} + \mathbf{r_u}) \\ &= (\mathbf{A}\mathbf{s} + \mathbf{e})^T \mathbf{r} + e_2 + \mathsf{Encode}(m) + r_v - \mathbf{s}^T(\mathbf{A}^T\mathbf{r} + \mathbf{e}_1 + \mathbf{r_u}) \\ &= \mathsf{Encode}(m) + \left(\mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u}) + e_2 + r_v\right). \end{aligned}$$

$\mathsf{Keygen}(1^\lambda)$	$\mathsf{Enc}(m, pk)$	$\mathsf{Dec}(c, sk)$
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$	$\mathbf{r} \xleftarrow{\$} D_{\mathbf{r}}$	$(\mathbf{c_1}, c_2) = c$
$\mathbf{s} \xleftarrow{\$} D_{\mathbf{s}}$	$\mathbf{e}_1 \xleftarrow{\$} D_{\mathbf{e}_1}$	$\tilde{\mathbf{u}} = \mathsf{Decompress}_1(\mathbf{c_1})$
$\mathbf{e} \xleftarrow{\$} D_{\mathbf{e}}$	$e_2 \xleftarrow{\$} D_{e_2}$	$\tilde{v} = \mathsf{Decompress}_2(c_2)$
$\mathbf{b} = \mathbf{As} + \mathbf{e}$	$\mathbf{u} = \mathbf{A}^T\mathbf{r} + \mathbf{e}_1$	$\widetilde{m} = \mathsf{Decode}(\tilde{v} - \mathbf{s}^T\tilde{\mathbf{u}})$
$pk = (\mathbf{A}, \mathbf{b})$	$v = \mathbf{b}^T\mathbf{r} + e_2 + \mathsf{Encode}(m)$	
$sk = \mathbf{s}$	$\mathbf{c_1} = \mathsf{Compress}_1(\mathbf{u})$	
	$c_2 = \mathsf{Compress}_2(v)$	
	$c = (\mathbf{c_1}, c_2)$	

Fig. 1. Lindner-Peikert framework of lattice-based PKE. $D_{\mathbf{s}}, D_{\mathbf{e}}, D_{\mathbf{r}}, D_{\mathbf{e}_1}$ and D_{e_2} are distributions of polynomials with small coefficients. $\mathsf{Encode}(\cdot)$ (resp. $\mathsf{Decode}(\cdot)$) is the encoding (resp. decoding) function of message (resp. ciphertext). $\mathsf{Compress}(\cdot)$ (resp. $\mathsf{Decompress}(\cdot)$) is the compressing (resp. decompressing) function.

Therefore, $\widetilde{m} = m$ if and only if $\mathsf{Decode}(\mathsf{Encode}(m)+E_{\text{total}}) = m$, where $E_{\text{total}} = \mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u}) + e_2 + r_v$ is the total noise.

For lattice-based PKE with lattice code, encoding function $\mathsf{Encode}(\cdot)$ is constructed from a lattice code $\mathcal{C}(\Lambda_f, \Lambda_c)$ as showing in Section 2.2. m is a vector in $\mathbb{R}^d$. If d is larger than the dimension N of Λ_f, then m can be split into $\lceil d/N \rceil$ vectors of dimension N, with encoding and decoding operations performed on each, resulting in the final DFR being increased by a factor of $\lceil d/N \rceil$. In this case, we say the PKE uses $\Lambda_f^{\lceil d/N \rceil}$ for encoding. For simplicity, the subsequent theoretical analysis only considers the case where $d = N$.

There are two common types of decoding function corresponding to MLD solver or BDD solver.

$$\mathsf{Decode} : \mathbb{R}^d \to \mathcal{M},$$
$$\mathbf{v} \mapsto \mathsf{Encode}^{-1}(Q_{\Lambda_f}(\mathbf{v}) \bmod \Lambda_c),\ Q = \text{ MLD solver or BDD solver.}$$

By formula (1), $\delta = \Pr(Q_{\Lambda_f}(E_{\text{total}}) \neq \mathbf{0})$ can be regarded as the DFR result. Here we discuss two general methods of DFR evaluation.

Case 1. The decoding function uses BDD solver with parameter r. In this case, $Q_{\Lambda_f}(E_{\text{total}}) = \mathbf{0}$ if $\|E_{\text{total}}\| \leq r$, thus,

$$\delta \leq \Pr\left(\|E_{\text{total}}\| > r\right). \tag{2}$$

Case 2. The decoding function uses MLD solver. In this case, $Q_{\Lambda_f}(E_{\text{total}}) = \mathbf{0}$ if and only if $E_{\text{total}} \in \mathcal{V}(\mathbf{0})$, where $\mathcal{V}(\mathbf{0})$ is the Voronoi region of $\mathbf{0} \in \Lambda_f$. It means that $\delta = \Pr(E_{\text{total}} \notin \mathcal{V}(\mathbf{0}))$.

Obviously, $B(\mathbf{0}, \frac{\lambda_1(\Lambda_f)}{2}) \subseteq \mathcal{V}(\mathbf{0})$, so there is a rough evaluation like case 1,

$$\delta = \Pr(E_{\text{total}} \notin \mathcal{V}(\mathbf{0})) \leq \Pr\left(\|E_{\text{total}}\| > \frac{\lambda_1(\Lambda_f)}{2}\right). \tag{3}$$

There is another evaluation using the structure of Voronoi region.

$$\begin{aligned}\mathcal{V}(\mathbf{0}) &= \cap_{\mathbf{v}\in R(\mathbf{0})}\{\mathbf{x} \in \mathbb{R}^d \mid \ \|\mathbf{x}\| \leq \|\mathbf{x}-\mathbf{v}\|\} \\ &= \cap_{\mathbf{v}\in R(\mathbf{0})}\{\mathbf{x} \in \mathbb{R}^d \mid \langle \mathbf{x}, \mathbf{v}\rangle \leq \frac{1}{2}\|\mathbf{v}\|^2\},\end{aligned}$$

so considering union bound,

$$\begin{aligned}\delta &= \Pr(E_{\text{total}} \notin \mathcal{V}(\mathbf{0})) \\ &= \Pr(E_{\text{total}} \in \cup_{\mathbf{v}\in R(\mathbf{0})}\{\mathbf{x} \in \mathbb{R}^d \mid \langle \mathbf{x}, \mathbf{v}\rangle > \frac{1}{2}\|\mathbf{v}\|^2\}) \\ &\leq \sum_{\mathbf{v}\in R(\mathbf{0})} \Pr\left(\langle E_{\text{total}}, \mathbf{v}\rangle > \frac{1}{2}\|\mathbf{v}\|^2\right).\end{aligned} \tag{4}$$

Remark 1. The trivial message encoding used in Kyber [22] and Frodo [10] can also be integrated into this framework. Let nested lattices $\Lambda_f = C\mathbb{Z}^d$, $\Lambda_c = kC\mathbb{Z}^d$. Then the relative points of $\mathbf{0} \in \Lambda_f$ are of the form $(0, \cdots, 0, \pm C, 0, \cdots, 0)$. Hence from formula (4),

$$\delta \leq \sum_{\mathbf{v}\in R(\mathbf{0})} \Pr\left(\langle E_{\text{total}}, \mathbf{v}\rangle > \frac{1}{2}\|\mathbf{v}\|^2\right) = \Pr\left(\|E_{\text{total}}\|_\infty > \frac{C}{2}\right),$$

which is coincide with the original evaluation method in Kyber and Frodo.

Since it is always difficult to describe Voronoi region, the following formula gives an approximate method [17] when E follows a distribution of standard deviation $\sigma \ll \lambda_1(\Lambda_f)$,

$$\delta = \Pr(E_{\text{total}} \notin \mathcal{V}(\mathbf{0})) \lesssim \sum_{\mathbf{v}\in \Lambda_1} \Pr\left(\langle E_{\text{total}}, \mathbf{v}\rangle > \frac{1}{2}\|\mathbf{v}\|^2\right), \tag{5}$$

where Λ_1 is the set of all minimal vectors in $\Lambda_f \backslash \{\mathbf{0}\}$.

3 Refined Evaluation with MLD-based Decoding

In this section, we describe the pattern of nonzero minimal vectors in BW_N lattice. Using the symmetry of these vectors, we give a refined evaluation method of DFR in PKEs with BW_N lattice encoding and MLD-based decoding. Based on the refined method, we re-evaluated the DFR of CNTR scheme [14].

3.1 Nonzero Minimal Vectors in BW_N Lattice

Following the concepts in [1], the rank of vector $\mathbf{x} \neq \mathbf{0}$ of BW_N to be the largest $r \in [n+1]$ for which all coordinates $\mathbf{x}_\alpha$ are divisible by 2^{λ_r}, and the minimum M of BW_N is $\min_{r\in[n+1]}(n-r+2\lambda_r)$. From Theorem 3.2 in [1], a vector $\mathbf{x} \neq \mathbf{0}$ is a minimal vector in BW_N if and only if its rank R satisfies

$$n - R + 2\lambda_R = M.$$

Recall that $\lambda_i = \left\lfloor \frac{i}{2} \right\rfloor$ for $0 \leq i \leq n$, it is easy to show that $M = n-1$ and every odd $R \in [n+1]$ is the solution of above equation. From the section 5 in [1], the number of minimal vectors of every odd rank R is $\mathrm{Num}_{n,R} = 2^{n+1+\binom{n-R}{2}} \cdot K_{n,R}$, where

$$K_{n,R} = \frac{(2^n-1)(2^{n-1}-1)\cdots(2^{n-R+1}-1)}{(2^R-1)(2^{R-1}-1)\cdots(2-1)}.$$

The coordinate of minimal vector $\mathbf{x} \in \mathrm{BW}_N$ of rank odd R is

$$x_{\mathbf{a}} = \begin{cases} 2^{\lambda_R}(\mathbf{1} - 2z_{\mathbf{a}}) \ , & \text{if } \mathbf{a} \in C_{n-R}, \\ \mathbf{0} \qquad\qquad\quad , & \text{otherwise} \ , \end{cases}$$

where $z_{\mathbf{a}}$ is defined on some coset C_{n-R} and has the value 0 or 1. In short, every minimal vector of odd rank R has 2^{n-R} coordinates equal to $\pm 2^{\lambda_R}$ and other coordinates equal to 0. As a result, the norm of minimal vectors

$$\lambda_1(\mathrm{BW}_N) = \sqrt{2^{n-R} \cdot 2^{2\lambda_R}} = \sqrt{2^M} = \sqrt{\frac{N}{2}}.$$

For example $N = 8$, the number and shape of minimal vectors in BW_8 is summarized in following Table 1.

Table 1. number and shape of minimal vectors in BW_8

rank R	number $\mathrm{Num}_{n,R}$	shape
1	224	$(\pm1, \pm1, \pm1, \pm1, 0, 0, 0, 0)$
3	16	$(\pm2, 0, 0, 0, 0, 0, 0, 0)$

3.2 Refined Evaluation Method of DFR

It is known that $R(\mathbf{0}) = \Lambda_1$ for (scaled) BW_4 and BW_8 lattices [3, Sec. H]. To evaluate the DFR using (scaled) BW_N lattice code, we should consider exact formula (4) when $N \leq 8$ and approximate formula (5) when $N \geq 16$. In general, we should evaluate the upper bound of DFR as

$$\delta = \sum_{\mathbf{v} \in \Lambda_1(\mathrm{BW}_N)} \Pr\left(\langle E_{\mathrm{total}}, \mathbf{v} \rangle > \frac{1}{2}\|\mathbf{v}\|^2 \right).$$

Suppose encoding lattice $\Lambda_f = c \cdot \mathrm{BW}_N$ and noise vector $E_{\text{total}} = (E_1, E_2, \cdots, E_N) \leftarrow D_E^N$, where D_E is a central distribution of real numbers. From the analysis in section 3.1, the vectors in Λ_1 of lattice Λ_f belong to different odd rank, and every minimal vector of odd rank R has 2^{n-R} coordinates equal to $\pm c \cdot 2^{\lambda_R}$ and other coordinates equal to 0. Therefore,

$$\delta = \sum_{\text{odd } R \in [n]} \mathrm{Num}_{n,R} \cdot \Pr\left(x > c \cdot 2^{n-2-\frac{R-1}{2}} \middle| x \leftarrow [2^{n-R}]D_E\right).$$

From the analysis above, we give the following Algorithm 1 to evaluate DFR of lattice-based PKE with BW_N lattice encoding and MLD-based decoding.

Algorithm 1: refined DFR evaluation method of lattice-based PKE with BW_N lattice encoding and MLD-based decoding

Input: lattice-based PKE framework and parameters.

Output: DFR result δ.

1 Extract the encoding lattice $\Lambda_f = c \cdot \mathrm{BW}_N$ and total noise vector $E_{\text{total}} = (E_1, E_2, \cdots, E_d) \leftarrow D_E^d$;

2 Calculate the part result

$$\delta_0 = \sum_{\text{odd } R \in [n]} \mathrm{Num}_{n,R} \cdot \Pr\left(x > c \cdot 2^{n-2-\frac{R-1}{2}} \middle| x \leftarrow [2^{n-R}]D_E\right),$$

where $\mathrm{Num}_{n,R} = 2^{n+1+\binom{n-R}{2}} \cdot \frac{(2^n-1)(2^{n-1}-1)\cdots(2^{n-R+1}-1)}{(2^R-1)(2^{R-1}-1)\cdots(2-1)}$;

3 **return** $\delta = \frac{d}{N} \cdot \delta_0$.

3.3 Impact on CNTR Scheme

CNTR [14] is an encryption scheme based on NTRU problem [11] and RLWE problem [19]. It can be described as following Fig. 2.

In Fig. 2, $\mathsf{Encode}(\cdot)$ function performs encoding operations on every 4 bits of the message. For any $m \in \{0,1\}^4$, $\mathsf{Encode}(m) = m \cdot \frac{q}{2}\mathbf{H} \bmod q$, where matrix

$$\mathbf{H} = \begin{pmatrix} 1&1&1&1&0&0&0&0 \\ 0&0&1&1&1&1&0&0 \\ 0&0&0&0&1&1&1&1 \\ 0&1&0&1&0&1&0&1 \end{pmatrix}$$

is a generator matrix of $(8,4,4)$ extanded hamming code, equivalently $\mathrm{RM}(1,3)$ Reed-Muller code. $\mathsf{Decode}(\cdot)$ function, given a vector in $\mathbb{R}^8$, finds the closest point in the q-ary lattice $\{v \in \mathbb{Z}^8 \mid (v \bmod q) \in \frac{q}{2} \cdot \mathrm{RM}(1,3)\}$ and then recovers the message. From the analysis in [14], the noise polynomial $E_{\text{total}} = gr + \frac{q}{q_2}\varepsilon f$, where $g \xleftarrow{\$} D_1$, $r \xleftarrow{\$} D_2$, $\varepsilon \leftarrow \{\lfloor \frac{q_2}{q} u \rceil - \frac{q_2}{q} u \mid u \xleftarrow{\$} R_q\}$, $f = 2f' + 1$, $f' \xleftarrow{\$} D_1$.

$\mathsf{Keygen}(1^\lambda)$	$\mathsf{Enc}(m, pk)$
1 : $f' \xleftarrow{\$} D_1 \subseteq R_q$	1 : $r \xleftarrow{\$} D_2$
2 : If $f \notin R_q^\times$, restart	2 : $c = \lfloor \frac{q_2}{q}(hr + \mathsf{Encode}(m)) \rceil$
3 : $f = 2f' + 1$	
4 : $g \xleftarrow{\$} D_1 \subseteq R_q$	$\mathsf{Dec}(c, sk)$
5 : $h = g/f \in R_q$	1 : $\widetilde{m} = \mathsf{Decode}(\frac{q}{q_2} \cdot fc \bmod q)$
6 : $pk = h,\ sk = f$	

Fig. 2. CNTR scheme.

The polynomial ring $R_q = \mathbb{Z}_q/(x^d - x^{\frac{d}{2}} + 1)$. The correct decryption condition in original paper is shown in the following lemma.

Lemma 5 (Correct decryption condition [14]). *Let E_i be the i-th octet of $gr + \frac{q}{q_2}\varepsilon f \in \mathbb{Z}_q/(x^d - x^{\frac{d}{2}} + 1)$, then the decryption in CNTR is correct if $\|E_i\| < \frac{q}{2}$ for any $i \in [\frac{d}{8}]$.*

In the context of lattice code, let $\Lambda_c = q\mathbb{Z}^8$ and $\Lambda_f = \frac{q}{2} \cdot \mathrm{BW}_8 = \frac{q}{2} \cdot (\mathrm{RM}(8,4,4) + 2\mathbb{Z}^8)$, then the function $\mathsf{Encode}(\cdot)$ can be regarded as the labeling function of lattice code $\mathcal{C}(\Lambda_f, \Lambda_c)$. The function $\mathsf{Decode}(x) = \mathsf{Encode}^{-1}(Q_{\Lambda_f}(x) \bmod \Lambda_c)$ is the corresponding decoding function using MLD solver.

Suppose the noise vector $E_{\text{total}} = (E_1, E_2, \cdots, E_N) \leftarrow D_E^N$. From the shape of minimal vectors in BW_8 in Table 1, the rank 3 vectors in $\Lambda_f = \frac{q}{2} \cdot \mathrm{BW}_8$ are of the form $(\pm q, 0, 0, 0, 0, 0, 0, 0)$, which fall into Λ_c and have no impact on decryption failure. Hence there is another refined evaluation formula of DFR as follows.

$$\begin{aligned}\delta_{D_E} &= \Pr(Q_{\Lambda_f}(E_{\text{total}}) \notin \Lambda_c) \\ &\le \mathrm{Num}_{3,1} \cdot \Pr\left(x > q \mid x \leftarrow [4]D_E\right) \\ &= 224 \cdot \Pr\left(x > q \mid x \leftarrow [4]D_E\right)\end{aligned}$$

To describe the noise expansion caused by multiplication in the algebraic ring, we introduce the following propositions.

Proposition 1. *Suppose $F = \sum_{u \in I_F} f_{i_u} g_{j_u}$ and $G = \sum_{v \in I_G} f_{i_v} g_{j_v}$, where independent $f_{i_u}, f_{i_v} \leftarrow D_1$, $g_{j_u}, g_{j_v} \leftarrow D_2$ for $i \in [d]$. D_1 and D_2 are central distribution with variance σ_1^2 and σ_2^2 respectively. (i_u, j_u) (resp. (i_v, j_v)) is distinct for different u (resp. v). If the index set $S = \#\{(u,v) \mid (i_u, j_u) = (i_v, j_v), u \in I_F, v \in I_G\}$ has cardinality of N, then*

$$\mathrm{Cov}(F, G) = N \cdot \sigma_1^2 \sigma_2^2.$$

Specially, $\sigma^2(F) = \mathrm{Cov}(F, F) = \#I_F \cdot \sigma_1^2 \sigma_2^2$.

Proof. Obviously $\mathbb{E}(F) = \mathbb{E}(G) = 0$. It follows that

$$\begin{aligned}\mathrm{Cov}(F,G) &= \mathbb{E}(F \cdot G)\\ &= \sum_{(u,v)\in S} \mathbb{E}(f_{i_u} g_{j_u} f_{i_v} g_{j_v}) + \sum_{(u,v)\notin S} \mathbb{E}(f_{i_u} g_{j_u} f_{i_v} g_{j_v})\\ &= \sum_{(u,v)\in S} \mathbb{E}(f_{i_u}^2 g_{j_u}^2)\\ &= N \cdot \sigma_1^2 \sigma_2^2.\end{aligned}$$

Proposition 2. *Let* $R_q = \mathbb{Z}_q/(x^d - x^{\frac{d}{2}} + 1)$, $f = \sum_{i=0}^{d-1} f_i x^i \in R_q$, $g = \sum_{i=0}^{d-1} g_i x^i \in R_q$ *,where independent* $f_i \leftarrow D_1$, $g_i \leftarrow D_2$ *for* $i \in [d]$. D_1 *and* D_2 *are central distribution with variance* σ_1^2 *and* σ_2^2 *respectively. The multiplication* $h := fg = \sum_{i=0}^{d-1} h_i x^i \in R_q$. *Then for all* $k \in [d]$,

$$\sigma^2(h_k) = N_k \cdot \sigma_1^2 \sigma_2^2,$$

where N_k *is the number of* $f_i g_j$ *in the sum of* h_k. *Especially for* $\frac{n}{2} \le k \le n-1$,

$$h_k = \sum_{0\le i\le k-\frac{d}{2}} (f_i g_{k-i} + f_{i+\frac{d}{2}} g_{k-i} + f_{i+\frac{d}{2}} g_{k-\frac{d}{2}-i}) + \sum_{k-\frac{d}{2}<i<\frac{d}{2}} (f_i g_{k-i} + f_i g_{k+\frac{d}{2}-i} + f_{i+\frac{d}{2}} g_{k-i})$$

Proof. From simple calculation, for all $k \in [d]$,

$$h_k = \sum_{\substack{i+j=k\\ 0\le k\le d-1}} f_i g_j - \sum_{\substack{i+j=k+d\\ 0\le k\le d-2}} f_i g_j + \sum_{\substack{i+j=k+\frac{d}{2}\\ \frac{d}{2}\le k\le d-1}} f_i g_j + \sum_{\substack{i+j=k+d\\ \frac{d}{2}\le k\le d-2}} f_i g_j - \sum_{\substack{i+j=k+\frac{3d}{2}\\ 0\le k\le \frac{d}{2}-2}} f_i g_j.$$

It means that h_k shares a form $f_{i_1} g_{j_1} + f_{i_2} g_{j_2} + \cdots f_{i_{N_k}} g_{j_{N_k}}$, where $(i_u, j_u) \neq (i_v, j_v)$ for all different u, v. All f_{i_u} and g_{j_u} follow independent central distributions. From Proposition 1, it is hold that $\sigma^2(h_k) = N_k \cdot \sigma_1^2 \sigma_2^2$ where N_k is the number of $f_i g_j$ in the sum of h_k. Moreover,

$$N_k = \begin{cases} \frac{3d}{2} - k - 1\ , & 0 \le k \le \frac{d}{2} - 1,\\ \frac{3d}{2} & ,\ \frac{d}{2} \le k \le d-1. \end{cases} \tag{6}$$

For $\frac{n}{2} \le k \le n-1$, $h_k = \sum_{i+j=k} f_i g_j + \sum_{i+j=k+\frac{d}{2}} f_i g_j$. Suppose $i + j = k$.

- If $i \in [0, k - \frac{d}{2}]$, then $j = k - i \in [\frac{d}{2}, k]$, $k + \frac{d}{2} - i \in [d, k + \frac{d}{2}]$, $k + \frac{d}{2} - j = \frac{d}{2} + i \in [\frac{d}{2}, k]$. All possible $f_j g_j$ are $f_i g_{k-i}$, $f_{i+\frac{d}{2}} g_{k-i}$ and $f_{i+\frac{d}{2}} g_{k-\frac{d}{2}-i}$, so there are terms of $f_i g_{k-i} + f_{i+\frac{d}{2}} g_{k-i} + f_{i+\frac{d}{2}} g_{k-\frac{d}{2}-i}$.
- If $i \in (k - \frac{d}{2}, \frac{d}{2})$, then $j = k - i \in (k - \frac{d}{2}, \frac{d}{2})$, $k + \frac{d}{2} - i \in (k, d)$, $k + \frac{d}{2} - j = \frac{d}{2} + i \in (k, d)$. All possible $f_j g_j$ are $f_i g_{k-i}$, $f_i g_{k+\frac{d}{2}-i}$, $f_{i+\frac{d}{2}} g_{k-i}$, so there are terms of $f_i g_{k-i} + f_i g_{k+\frac{d}{2}-i} + f_{i+\frac{d}{2}} g_{k-i}$.
- If $i \ge \frac{d}{2}$, then $i - \frac{d}{2} \in [0, \frac{d}{2})$ falls into two cases above, which is counted.

Remark 2. It should be noted that the variances of first $\frac{d}{2}$ coefficients in h are smaller than that of last $\frac{d}{2}$ ones. When $\frac{d}{2} \leq k \leq d-1$, h_k is the sum of $\frac{d}{2}$ independent terms of $(a_1b_1 + a_1b_2 + a_2b_2)$.

From Proposition 2 and formula (6) , the last $\frac{d}{2N}$ coefficients of total noise polynomial in CNTR scheme share the same distrition D_{last}. The first $\frac{d}{2N}$ coefficients are narrower than the last $\frac{d}{2N}$ ones, which implies the failure probabilities of first $\frac{d}{2N}$ coefficients are smaller than those of last $\frac{d}{2N}$ ones. Using union bound, it holds that

$$\delta_{\text{CNTR}} \leq \delta_1 + \delta_2 + \cdots + \delta_{\frac{d}{2N}} + \frac{d}{2N} \cdot \delta_{D_{\text{last}}} \in \left[\frac{d}{2N} \cdot \delta_{D_{\text{last}}}, \frac{d}{N} \cdot \delta_{D_{\text{last}}} \right],$$

which means that we can use $\frac{d}{N} \cdot \delta_{D_{\text{last}}}$ as the upper bound of DFR of CNTR scheme with one bit loss.

Here, we compare the DFR results using two evaluation methods in Table 2. The first method follows original analysis of CNTR scheme [14], equivalently, analysis in Lemma 5. Since there is no detail to calculate the Eucilean norm of noise vector, we model the noise distribution D_{last} as a central Gaussian distribution with variance $\sigma^2 = \frac{3d}{2} \cdot \sigma_1^2 \left(\sigma_2^2 + \left(\frac{2q}{q_2}\right)^2 \sigma_\varepsilon^2 \right) + \left(\frac{q}{q_2}\right)^2 \sigma_\varepsilon^2$. The second method uses our refined analysis, which uses BW_8 lattice code and models the noise distribution D_{last} as the same distribution in first method.

Table 2. compare of DFR results in CNTR scheme

	CNTR-512	CNTR-768	CNTR-1024
d	512	768	1024
(q, q_2)	$(3457, 2^{10})$	$(3457, 2^{10})$	$(3457, 2^{10})$
(D_1, D_2)	$(\text{CBD}_5, \text{CBD}_5)$	$(\text{CBD}_3, \text{CBD}_3)$	$(\text{CBD}_2, \text{CBD}_2)$
original method [14]	$2^{-153.89}$	$2^{-209.26}$	$2^{-264.91}$
our refined method	$\mathbf{2^{-169.68}}$	$\mathbf{2^{-226.45}}$	$\mathbf{2^{-283.19}}$

Remark 3. It can be noticed that the DFR result of our refined method is smaller than that of original method, which means that the original method is not tight enough. Besides, the result of original method is different from the result in [14, Tab. 3]. It may be caused by the different method of calculating the Eucilean norm of noise vector.

4 Refined Evaluation with BDD-based Decoding

In this section, we point out the drawback of current method of DFR evaluation in PKEs using lattice encoding and BDD-based decoding. We then present

two methods to overcome this drawback, and finally re-evaluate the DFR of Scloud$^+$scheme using the more efficient method.

4.1 Drawback of Current Methods

In LPR framework as shown in Fig. 1, the noise vector $E_{\text{total}} = \mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u}) + e_2 + r_v$. Every coefficient in $\mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u})$ is the convolution of at least d distributions, which can be reasonably modeled as a Gaussian distribution. However, $e_2 + r_v$ is not the case. e_2 is usually sampled from a narrow distribution and has little impact on DFR evaluation. In contrast, r_v follows typically a wide uniform distribution, and modeling it as a Gaussian distribution introduces significant approximation difference.

For example, In Kyber [22] settings, $q = 3457$ and $R_q = \mathbb{Z}_q[x]/(x^{256} + 1)$, there is a huge difference between three methods as shown in following Table 3. The Gaussian simulation method refers to simulating E_{total} as a central Gaussian distribution with variance $\sigma^2(E_{\text{total}})$, the entirely discrete method refers to characterizing E_{total} completely using a discrete probability distribution table, and the partial Gaussian method refers to simulating $\mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u})$ as a Gaussian distribution while characterizing $e_2 + r_v$ using a discrete probability distribution table.

Table 3. the DFR result of Kyber scheme using different evaluation methods

	Kyber-512	Kyber-768	Kyber-1024
k	2	3	4
$(D_{\mathbf{s}}, D_{\mathbf{e}})$	$(\text{CBD}_3, \text{CBD}_3)$	$(\text{CBD}_2, \text{CBD}_2)$	$(\text{CBD}_2, \text{CBD}_2)$
$(D_{\mathbf{r}}, D_{\mathbf{e_1}}, D_{e_2})$	$(\text{CBD}_3, \text{CBD}_2, \text{CBD}_2)$	$(\text{CBD}_2, \text{CBD}_2, \text{CBD}_2)$	$(\text{CBD}_2, \text{CBD}_2, \text{CBD}_2)$
(du, dv)	$(10, 4)$	$(10, 4)$	$(11, 5)$
Gaussian simulation	$2^{-75.83}$	$2^{-81.13}$	$2^{-145.54}$
entirely discrete	$2^{-139.14}$	$2^{-165.24}$	$2^{-175.20}$
partial Gaussian	$2^{-147.43}$	$2^{-172.62}$	$2^{-181.11}$

The entirely discrete method is the most accurate among the three after ensuring computational accuracy. It can be seen in Table 3 that the Gaussian simulation method and the entirely discrete method give significantly different evaluations of DFR. The main reason for this large difference is that noise item r_v almost uniformly distributed in $\mathbb{Z}_{\lfloor p/2^{dv} \rceil}$, which is instantiated as $\mathbb{Z}_{208}$ and $\mathbb{Z}_{104}$, making it difficult to simulate as a Gaussian distribution. In comparison, the difference in DFR results assessed by the partial Gaussian method and the entirely discrete method are less than 10 bits, which are acceptable.

Kyber [22] actually uses the entirely discrete method for calculating its DFR, because trivial message encoding analyzed in Remark 1, requiring only calculation of the infinity norm of noise vector, which can be efficiently implemented.

However, the PKE using non-trivial message encoding and BDD-based decoding typically requires the computation of the Euclidean norm to estimate DFR. There exist following drawbacks in current methods.

1. Entirely discrete method. The square of Euclidean norm of a vector $\mathbf{e} = (e_1, e_2, \cdots, e_n)$ is $\|\mathbf{e}\|^2 = \sum_{i=1}^{n} e_i^2$. Suppose $e_i \leftarrow D_i$, then the distribution of $\|\mathbf{e}\|^2$ is the convolution of all square of D_i for $1 \leq i \leq n$. In most cryptographic instantiations, the squared range of distribution D_i becomes very large, making their convolution difficult to compute in practice.
2. Gaussian simulation method. When the rounding noise follows a almost uniform distribution with large range, the inaccuracy of Gaussian simulation method will be unacceptable like the case in Table 3.

Therefore, it is necessary to find a compromise method for calculating the Euclidean norm in the context of lattice-based PKEs to balance the accuracy and feasibility of the evaluation.

4.2 Compromise Evaluation Method of DFR

When the decoding function uses BDD solver, $Q_{\Lambda_f}(E_{\text{total}}) = \mathbf{0}$ if $\|E_{\text{total}}\| \leq r$ and $\text{DFR} \leq \Pr(\|E_{\text{total}}\| > r)$. Suppose the noise vector $E_{\text{total}} = (E_1 + F_1, E_2 + F_2, \cdots, E_N + F_N)$, where $E_i \leftarrow \mathcal{N}(0, \sigma^2)$, $F_i \leftarrow D_{\text{dis}}$ for $1 \leq i \leq N$. D_{dis} is a discrete distribution which can be described by a probability distribution table. For every fixed $f = (f_1, f_2, \cdots, f_N)$, $\|(E_1 + f_1, E_2 + f_2, \cdots, E_N + f_N)\|^2$ naturally follows a noncentral chi-squared distribution. The probability of $\|(E_1 + f_1, E_2 + f_2, \cdots, E_N + f_N)\|$ smaller than a number x is

$$\begin{aligned}
&\Pr(\|(E_1 + f_1, E_2 + f_2, \cdots, E_N + f_N)\| \leq x) \\
=&\Pr(\sum_{i=1}^{N}(E_i + f_i)^2 \leq x^2 \mid E_i \leftarrow \mathcal{N}(0, \sigma^2)) \\
=&\Pr\left(\sum_{i=1}^{N}\left(\frac{E_i}{\sigma} + \frac{f_i}{\sigma}\right)^2 \leq \left(\frac{x}{\sigma}\right)^2 \middle| E_i \leftarrow \mathcal{N}(0, \sigma^2)\right) \\
=&F\left(\left(\frac{x}{\sigma}\right)^2; N, \frac{\|f\|^2}{\sigma^2}\right).
\end{aligned}$$

Using the probability table of distribution D_{dis}, it is easy to get distribution $D_{\text{norm}} = \{\|(F_1, F_2, \cdots, F_N)\|^2 \mid F_i \leftarrow D_{\text{dis}}\}$ and its probability table. Therefore, the probability of $\|E_{\text{total}}\|$ smaller than a number x is

$$\Pr(\|E_{\text{total}}\| \leq x) = \mathbb{E}\left(F\left(\left(\frac{x}{\sigma}\right)^2; N, \frac{Y}{\sigma^2}\right) \middle| Y \leftarrow D_{\text{norm}}\right).$$

The cumulative distribution function $F(x; k, \lambda)$ is a infinite sum. To ensure the efficiency and precision of calculation, we consider the following trimming

strategy. First, preset the precision to *prec*. In order to control the inaccuracy within 2^{-prec}, it suffices to ensure to compute an expression $T(x;k,\lambda)$ satisfying $|F(x;k,\lambda)-T(x;k,\lambda)| \leq 2^{-prec}$. It is known that

$$\begin{aligned}
&\left|F(x;k,\lambda)-\sum_{j=0}^{n}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}\cdot\Pr(Z\leq x \mid Z\leftarrow\chi^2_{k+2j})\right| \\
=&\sum_{j=n+1}^{\infty}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}\cdot\Pr(Z\leq x \mid Z\leftarrow\chi^2_{k+2j}) \\
\leq&\sum_{j=n+1}^{\infty}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!} \\
=&e^{-\lambda/2}\cdot\frac{e^{\xi}(\lambda/2)^{n+1}}{(n+1)!} \\
\leq&\frac{(\lambda/2)^{n+1}}{(n+1)!}.
\end{aligned}$$

Denote $M_{\lambda,prec}=\min\left\{n\middle|\frac{(\lambda/2)^{n+1}}{(n+1)!}\leq 2^{-prec}\right\}$ and select $T(x;k,\lambda)=\sum_{j=0}^{M_{\lambda,prec}}\frac{e^{-\lambda/2}(\lambda/2)^j}{j!}\cdot\Pr(Z\leq x \mid Z\leftarrow\chi^2_{k+2j})$, then

$$\begin{aligned}
&\left|\Pr(\|E_{\text{total}}\|\leq r)-\mathbb{E}\left(T\left(\left(\frac{r}{\sigma}\right)^2;N,\frac{Y}{\sigma^2}\right)\middle|Y\leftarrow D_{\text{norm}}\right)\right| \\
\leq&\mathbb{E}\left(\left|F\left(\left(\frac{r}{\sigma}\right)^2;N,\frac{Y}{\sigma^2}\right)-T\left(\left(\frac{r}{\sigma}\right)^2;N,\frac{Y}{\sigma^2}\right)\right|\middle|Y\leftarrow D_{\text{norm}}\right) \\
\leq&\mathbb{E}\left(\frac{(\lambda/2)^{M_{\lambda,prec}+1}}{(M_{\lambda,prec}+1)!}\middle|Y\leftarrow D_{\text{norm}}\right) \\
\leq&2^{-prec},
\end{aligned}$$

From the analysis above, we give the following Algorithm 2 to evaluate DFR of lattice-based PKE using lattice encoding and BDD-based decoding.

4.3 Impact on Scloud$^+$ Scheme

Scloud$^+$ [25] is a unstructured lattice-based encryption scheme using scaled BW_{32} as encoding lattice. In Scloud$^+$ setting, $D_{\mathbf{s}}=\mathcal{H}_{n,h_1}$, $D_{\mathbf{e}}=\text{CBD}^m_{\eta_1}$, $D_{\mathbf{r}}=\mathcal{H}_{n,h_2}$, $D_{\mathbf{e}_1}=\text{CBD}^n_{\eta_2}$, $D_{e_2}=\text{CBD}_{\eta_2}$. $\mathsf{Compress}_1(x)=\lfloor\frac{q_1}{q}\cdot x\rceil$ and $\mathsf{Decompress}_1(x)=\frac{q}{q_1}\cdot x$. $\mathsf{Compress}_2(x)=\lfloor\frac{q_2}{q}\cdot x\rceil_{\text{odd}}$ and $\mathsf{Decompress}_2(x)=\frac{q}{q_2}\cdot x$, where $\lfloor x\rceil_{\text{odd}}$ denotes the nearest integer to x, where any half-integer is rounded to the closest odd integer. Scloud$^+$ uses $\left(\frac{q}{2^\tau}\cdot\text{BW}_{32}\right)^{\frac{l_m}{\mu}}$ lattice for encoding and decoding, where $\lambda_1(\frac{q}{2^\tau}\cdot\text{BW}_{32})=\frac{4q}{2^\tau}$. The correct decryption condition in original paper is shown in the following lemma.

Algorithm 2: refined DFR evaluation method of lattice-based PKE using lattice encoding and BDD-based decoding

Input: lattice-based PKE framework and parameters, precision *prec*.
Output: DFR result δ.

1 Extract the correct decryption condition $\Pr(\|E_{\text{total},i}\| < r)$, where $E_{\text{total},i} = (E_1 + F_1, E_2 + F_2, \cdots, E_N + F_N)$ is i-th N-dimensional part of the d-dimensional vector E_{total}, and $E_i \leftarrow \mathcal{N}(0, \sigma^2)$, $F_i \leftarrow D_{\text{dis}}$ for $1 \le i \le N$;
2 Generate distribution $D_{\text{norm}} = \{\|(F_1, F_2, \cdots, F_N)\|^2 \mid F_i \leftarrow D_{\text{dis}}\}$;
3 Set $\delta_0 = 0$;
4 **for** $(a, pr_a) \in D_{\text{norm}}$ **do**
5 Calculate $M = \min\left\{n \middle| \frac{(\frac{a}{2\sigma^2})^{n+1}}{(n+1)!} \le 2^{-prec}\right\}$;
6 Calculate part result $\delta_0 = \delta_0 + pr_a \cdot \left(1 - T\left(\left(\frac{r}{\sigma}\right)^2; N, \frac{x}{\sigma^2}\right)\right)$, where

$$T(x; k, \lambda) = \sum_{j=0}^{M} \frac{e^{-\lambda/2}(\lambda/2)^j}{j!} \cdot \Pr(Z \le x \mid Z \leftarrow \chi^2_{k+2j});$$

7 **return** $\delta = \frac{d}{N} \cdot \delta_0$.

Lemma 6 (Correct decryption condition [14]). *Let $E_{\text{total},i}$ be i-th part of vector $E_{\text{total}} = \mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u}) + e_2 + r_v$ then the decryption in Scloud$^+$ Scheme is correct if $\|E_{\text{total},i}\| < \frac{2q}{2^\tau}$ for any $i \in [\frac{l_m}{\mu}]$.*

The first part $E = \mathbf{e}^T\mathbf{r} - \mathbf{s}^T(\mathbf{e}_1 + \mathbf{r_u})$ in noise vector E_{total} can be simulated as a central Gaussian distribution with standard deviation σ_E, where

$$\sigma_E^2 = 2h_2 \cdot \frac{\eta_1}{2} + 2h_1 \cdot (\frac{\eta_1}{2} + \sigma_{\mathbf{u}}^2).$$

The second part $F = e_2 + r_v$ can be characterized by a probability table D_{dis}, with $D_{\text{norm}} = \{\|(F_1, F_2, \cdots, F_{32})\|^2 \mid F_i \leftarrow D_{\text{dis}}\}$. Hence there is another refined evaluation formula of DFR as follows.

$$\begin{aligned}\delta &= \frac{l_m}{\mu} \cdot \Pr\left(\|E_{\text{total}}\| > \frac{\lambda_1(\frac{q}{2^\tau} \cdot \text{BW}_{32})}{2}\right) \\ &= \frac{l_m}{\mu} \cdot \left(1 - \mathbb{E}\left(T\left(\left(\frac{2q}{2^\tau \sigma_E}\right)^2; 32, \frac{Y}{\sigma_E^2}\right) \middle| Y \leftarrow D_{\text{norm}}\right)\right).\end{aligned}$$

Here, we compare the DFR results using two evaluation methods in Table 4. The first method follows original analysis of Scloud$^+$scheme [25], and models total noise distribution as a central Gaussian distribution with variance $\sigma_{\text{total}}^2 = 2h_2 \cdot \frac{\eta_1}{2} + 2h_1 \cdot (\frac{\eta_1}{2} + \sigma_{\mathbf{u}}^2) + \frac{\eta_2}{2} + \sigma_v^2$. The second method uses our refined analysis, and split the total noise to a sum of Gaussian distribution and discrete distribution.

[2] The original DFR result in [25] forgot the repetition coefficient $\frac{l_m}{\mu}$. Here we make corrections.

Table 4. compare of DFR results in Scloud$^+$scheme

	Scloud$^+$-128	Scloud$^+$-192	Scloud$^+$-256
(q, q_1, q_2)	$(2^{12}, 2^9, 2^7)$	$(2^{12}, 2^{12}, 2^{10})$	$(2^{12}, 2^{10}, 2^7)$
(m, n)	$(600, 600)$	$(928, 896)$	$(1136, 1120)$
(h_1, h_2)	$(150, 150)$	$(224, 232)$	$(280, 284)$
(η_1, η_2)	$(7, 7)$	$(2, 1)$	$(3, 2)$
τ	3	4	3
repetition $\frac{l_m}{\mu}$	2	2	4
original method [2]	$2^{-133.21}$	$2^{-199.64}$	$2^{-263.74}$
our refined method	$\mathbf{2^{-133.39}}$	$\mathbf{2^{-199.65}}$	$\mathbf{2^{-265.11}}$

Remark 4. It can be noticed that the DFR result of our refined method is smaller than that of original method, which means that the original method is not tight enough. DFR reduction in result does not exceed 2 bits, which is because the noise caused by ciphertext compression in Scloud$^+$is relatively small. Therefore, for general ciphertext-compressed, message-encoded, and BDD-decoded lattice-based PKEs, we still suggest using our refined method for evaluating DFR.

5 Refined Evaluation with Special Algebraic Structure

In this section, we present the impact of algebraic structure on the polynomial products, and study the method of DFR evaluation in the algebraic-lattice-based PKE using message encoding in polynomial ring $R_q = \mathbb{Z}_q/(x^d - x - 1)$. Based on the refined method, we re-evaluated the DFR of CNTR-Prime scheme [15].

5.1 Variance and Correlation of Coefficients of Product Polynomial

Different modulus polynomials result in distinct behaviors of products of ring elements. On the one hand, the summation terms for each coefficient of the product polynomial are affected by the modulus polynomial, which in turn leads to differences in the variance of coefficient distributions. On the other hand, the independence between coefficients of the product polynomial is closely related to the modulus polynomial. Although the independence among coefficients is difficult to characterize, in latticebased cryptography settings, coefficients of the product polynomial can be modeled as a Gaussian distribution, making it feasible to consider the correlation induced by the modulus polynomial.

Here, we give the variance and correlation of coefficients of product polynomial in three common algebraic structure.

Proposition 3. *Let* $R_q = \mathbb{Z}_q/(\phi(x))$, $f = \sum_{i=0}^{d-1} f_i x^i \in R_q$, $g = \sum_{i=0}^{d-1} g_i x^i \in R_q$ *,where independent* $f_i \leftarrow D_1$, $g_i \leftarrow D_2$ *for* $i \in [d]$. D_1 *and* D_2 *are central distribution with variance* σ_1^2 *and* σ_2^2 *respectively. The multiplication* $h := fg = \sum_{i=0}^{d-1} h_i x^i \in R_q$.

- *If* $\phi(x) = x^d+1$, *then* $\sigma^2(h_k) = d\cdot\sigma_1^2\sigma_2^2$ *for any* $k \in [d]$, *and* $\mathrm{Cov}(h_{k_1}, h_{k_2}) = 0$ *for any* $k_1 \neq k_2$.
- *If* $\phi(x) = x^d - x^{\frac{d}{2}} + 1$, *then* $\sigma^2(h_k) = \begin{cases} (\frac{3d}{2} - k - 1)\cdot\sigma_1^2\sigma_2^2 \ , & 0 \leq k \leq \frac{d}{2} - 1, \\ \frac{3d}{2}\cdot\sigma_1^2\sigma_2^2 \ , & \frac{d}{2} \leq k \leq d-1, \end{cases}$ *and* $\mathrm{Cov}(h_{k_1}, h_{k_2}) = \begin{cases} 0 \ , & k_2 - k_1 \neq 0, \pm\frac{d}{2}, \\ (d - \min\{k_1, k_2\} - 1)\cdot\sigma_1^2\sigma_2^2 \ , & k_2 - k_1 = \pm\frac{d}{2}. \end{cases}$
- *If* $\phi(x) = x^d - x - 1$, *then* $\sigma^2(h_k) = \begin{cases} d\cdot\sigma_1^2\sigma_2^2 \ , & k = 0, \\ (2d-k)\cdot\sigma_1^2\sigma_2^2 \ , & 1 \leq k \leq d-1, \end{cases}$ *and* $\mathrm{Cov}(h_{k_1}, h_{k_2}) = \begin{cases} 0 \ , & k_2 - k_1 \neq 0, \pm 1, \\ (d - \min\{k_1, k_2\} - 1)\cdot\sigma_1^2\sigma_2^2 \ , & k_2 - k_1 = \pm 1. \end{cases}$

Proof. Let the set $I_k = \{(i,j) \mid f_i g_j$ is a summation term of $h_k\}$. From Proposition 1, it suffices to find $\#I_k$ and $\#(I_{k_1} \cap I_{k_2})$.

- If $\phi(x) = x^d + 1$, then $h_k = \sum_{i+j=k} f_i g_j - \sum_{i+j=k+d} f_i g_j$ with $\#I_k = d$ for any $k \in [d]$. The index set I_k consists of all (i,j) satisfying $i + j = k, n + k$. Thus for any possible $k_1 \neq k_2$, $I_{k_1} \cap I_{k_2} = \emptyset$.
- If $\phi(x) = x^d - x^{\frac{d}{2}} + 1$, then from the proof of Proposition 2,

$$\#I_k = \begin{cases} (\frac{3d}{2} - k - 1) \ , & 0 \leq k \leq \frac{d}{2} - 1, \\ \frac{3d}{2} \ , & \frac{d}{2} \leq k \leq d-1. \end{cases}$$

  The index set I_k is a subset of (i,j) satisfying $i + j = k, k + \frac{d}{2}, k + d, k + \frac{3d}{2}$. Suppose $k_1 \leq k_2$. If $k_2 \neq k_1, k_1 + \frac{d}{2}$, then obviously $I_{k_1} \cap I_{k_2} = \emptyset$. If $k_2 = k_1 + \frac{d}{2}$, then $I_{k_1} \cap I_{k_2} = \{(i,j) \mid i + j = k_1 + d, i \in [d], j \in [d]\}$ with cardinality

$$\#(I_{k_1} \cap I_{k_2}) = (d-1) - (k_1 + d - (d-1)) + 1 = d - k_1 - 1.$$

- If $\phi(x) = x^d - x - 1$, then for any $k \in [d]$,

$$h_k = \sum_{\substack{i+j=k \\ 0\leq k\leq d-1}} f_i g_j + \sum_{\substack{i+j=k+d \\ 0\leq k\leq d-2}} f_i g_j + \sum_{\substack{i+j=k+d-1 \\ 1\leq k\leq d-1}} f_i g_j,$$

  with $\#I_k = \begin{cases} d \ , & k = 0, \\ (2d-k) \ , & 1 \leq k \leq d-1. \end{cases}$ The index set I_k is a subset of (i,j) satisfying $i + j = k, k + d, k + d - 1$. Suppose $k_1 \leq k_2$. If $k_2 \neq k_1, k_1 + 1$, then $I_{k_1} \cap I_{k_2} = \emptyset$. If $k_2 = k_1 + 1$, then $I_{k_1} \cap I_{k_2} = \{(i,j) \mid i + j = k_1 + d, i \in [d], j \in [d]\}$ with cardinality

$$\#(I_{k_1} \cap I_{k_2}) = (d-1) - (k_1 + d - (d-1)) + 1 = d - k_1 - 1.$$

Since message encoding usually takes several adjacent coefficients of the message polynomial, the adjacent coefficients of noise polynomial are consequently incorporated into the decryption correctness condition. Therefore from Proposition 3, the case of the modulus polynomial $\phi(x) = x^d - x - 1$ will introduce correlation into the evaluation of DFR.

5.2 Refined Evaluation Method of DFR

Suppose the DFR of a algebraic-lattice-based PKE with modulus polynomial $\phi(x) = x^d - x - 1$ with can be evaluated by

$$\delta \leq \Pr\left(\|\mathbf{e}\| > r\right),$$

where the d-dimensional noise vector $\mathbf{e}$ follows the multivariate Gaussian distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{M}$. Since covariance matrix is always real, symmetric and semi-positive, there exists an orthogonal transformation $\mathbf{A}$ such that $\mathbf{A}^T\mathbf{M}\mathbf{A} = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots \sigma_d^2)$. Therefore,

$$\begin{aligned}\delta &\leq \Pr\left(\|\mathbf{e}\| > r\right) = \Pr\left(\|\mathbf{A}\mathbf{e}\| > r\right) \\ &= \Pr\left(\sum_{i=1}^{d} \sigma_i^2 \cdot X_i^2 > r^2 \mid \text{ independent } X_i \leftarrow \mathcal{N}(0,1)\right).\end{aligned}$$

The random variable $\sum_{i=1}^{d} \sigma_i^2 \cdot X_i^2$ follows the weighted chi-squared distribution. Saddlepoint approximation [2] is a efficient method to compute small tail probability, the detail is shown in following lemma.

Lemma 7 ([2]). *Suppose random variable $Y = \sum_{i=1}^{d} \sigma_i^2 \cdot X_i^2$, where independent $X_i \leftarrow \mathcal{N}(0,1)$. Write the cumulant generating function $K(x) = -\frac{1}{2}\sum_{i=1}^{d} \log(1 - 2\sigma_i^2 x)$ and its derivative function $K'(x)$, $K''(x)$. Let $s \in [0, \frac{1}{2 \cdot \max \sigma_i^2}]$ be the root of $K'(x) = B$. Then*

$$\Pr\left(Y > B\right) \approx 1 - \Phi\left(w + \frac{\log(v/w)}{w}\right),$$

where $w = \sqrt{2(s \cdot B - K(s))}$, $v = s\sqrt{K''(s)}$, $\Phi(x)$ is the cumulative distribution function of $\mathcal{N}(0,1)$.

From the analysis above, we give the following Algorithm 3 to evaluate DFR of algebraic-lattice-based PKE using message encoding with polynomial ring $R_q = \mathbb{Z}_q/(x^d - x - 1)$.

5.3 Impact on CNTR-Prime Scheme

CNTR-Prime scheme [15] is a variant of CNTR scheme. The main difference between CNTR-Prime and CNTR is the algebraic structure and parameters. In CNTR-Prime scheme, the modulus polynomial is selected as $x^d - x - 1$ instead of $x^d - x^{\frac{d}{2}} - 1$.

Here, we compare the DFR results using two evaluation methods in Table 5. The first method follows original analysis of CNTR-Prime scheme [15]. Since there is no detail to calculate the Eucilean norm of noise vector $gr + \frac{q}{q_2}\varepsilon f = gr + \frac{2q}{q_2}\varepsilon f' + \frac{q}{q_2}\varepsilon$, we use another sufficient correct decryption condition

$$\left\| \left(gr + \frac{2q}{q_2}\varepsilon f' \right)_i \right\| < \frac{q}{2} - \sqrt{2} \cdot \frac{q}{q_2},$$

Algorithm 3: DFR evaluation method of algebraic-lattice-based PKE using message encoding with polynomial ring $R_q = \mathbb{Z}_q/(x^d - x - 1)$

Input: lattice-based PKE framework and parameters.
Output: DFR result δ.

Extract the correct decryption condition $\Pr(\|\mathbf{e}_i\| < r)$, where $\mathbf{e}_i$ is i-th N-dimensional part of the d-dimensional vector $\mathbf{e}$, and $\mathbf{e}$ is the sum of several polynomial products with single-layer variance σ_e^2;

Generate the covariance matrix $\sigma_e^2\mathbf{M}$ with $\mathbf{M}_{k,k} = \begin{cases} d & , \ k = 0, \\ (2d-k) & , \ 1 \le k \le d-1, \end{cases}$ $\mathbf{M}_{k,k+1} = \mathbf{M}_{k+1,k} = d - k - 1$ for $k \in [d-1]$, other entries 0;

for $i = 0 \to \frac{d}{N} - 1$ **do**

  Generate the principal minor $\mathbf{M}_i := \mathbf{M}[i \cdot N : (i+1)N]$;

  Calculate the eigenvalues $\sigma_1^2, \sigma_2^2, \cdots \sigma_N^2$ of the matrix $\mathbf{M}_i$;

  Evaluate the part result $\delta_i = 1 - \Phi\left(w + \frac{\log(v/w)}{w}\right)$, where $\Phi(x)$ is the cumulative distribution function of $\mathcal{N}(0,1)$, $w = \sqrt{2\left(s \cdot \left(\frac{r}{\sigma_e}\right)^2 - K(s)\right)}$, $v = s\sqrt{K''(s)}$, $K(x) = -\frac{1}{2}\sum_{i=1}^{N} \log(1 - 2\sigma_i^2 x)$ with its derivative function $K'(x)$, $K''(x)$, $s \in [0, \frac{1}{2 \cdot \max \sigma_i^2}]$ the root of $K'(x) = \left(\frac{r}{\sigma_e}\right)^2$;

return $\delta = \sum_{i=0}^{\frac{d}{N}-1} \delta_i$.

and model every coefficient of $gr + \frac{2q}{q_2}\varepsilon f'$ as an independent central Gaussian distribution. The second method uses our refined analysis, which considers the correlation between coefficients.

Table 5. compare of DFR results in CNTR-Prime scheme

	CNTR-512	CNTR-768	CNTR-1024
d	653	761	1277
(q, q_2)	$(4621, 2^{11})$	$(4591, 2^{10})$	$(7879, 2^{10})$
(D_1, D_2)	$(\mathrm{CBD}_3, \mathrm{CBD}_3)$	$(\mathrm{CBD}_2, \mathrm{CBD}_2)$	$(\mathrm{CBD}_2, \mathrm{CBD}_2)$
original method [15]	$2^{-590.07}$	$2^{-300.16}$	$2^{-187.91}$
our refined method	$\mathbf{2^{-314.48}}$	$\mathbf{2^{-162.62}}$	$\mathbf{2^{-103.03}}$

Remark 5. It can be noticed that the DFR result of our refined method is significantly larger than that of original method, which means that the original method is incorrect. Besides, the result of original method is different from the result in [15, Tab. 2]. It may be caused by the different method of calculating the Eucilean norm of noise vector.

6 Conclusion

This work proposes a refined and comprehensive framework to evaluate the decryption failure rate (DFR) for lattice-based public-key encryption with message encoding. We fix flaws in traditional DFR analysis by using Barnes-Wall lattice structure for tighter MLD bounds, noncentral chi-squared modeling for mixed-noise BDD scenarios, and weighted chi-squared with saddlepoint approximation to handle coefficient correlation induced by algebraic ring structures. We systematically characterize the variance and correlation of polynomial product coefficients under common ring moduli, filling the gap of ignoring algebraic effects in existing DFR evaluations. Experiments on CNTR, Scloud$^+$, and CNTR-Prime show our method gives much tighter and more accurate DFR bounds than conventional approaches. This work provides a structure-aware theoretical foundation for DFR evaluation of both generic and algebraic lattice-based cryptosystems, supporting possible better parameter design and more reliable security proofs for practical post-quantum cryptosystems.

References

1. Barnes, E.S., Wall, G.E.: Some extreme forms defined in terms of abelian groups. Journal of the Australian Mathematical Society **1**(1), 4763 (1959). https://doi.org/10.1017/S1446788700025064
2. Chen, T., Lumley, T.: Numerical evaluation of methods approximating the distribution of a large quadratic form in normal variables. Comput. Stat. Data Anal. **139**(C), 7581 (Nov 2019). https://doi.org/10.1016/j.csda.2019.05.002
3. Conway, J., Sloane, N.: Voronoi regions of lattices, second moments of polytopes, and quantization. IEEE Transactions on Information Theory **28**(2), 211–226 (1982). https://doi.org/10.1109/TIT.1982.1056483
4. Conway, J.H., Sloane, N.J.A.: On the voronoi regions of certain lattices. SIAM Journal on Algebraic Discrete Methods **5**(3), 294–305 (1984). https://doi.org/10.1137/0605031
5. D'Anvers, J.P., Guo, Q., Johansson, T., Nilsson, A., Vercauteren, F., Verbauwhede, I.: Decryption failure attacks on IND-CCA secure lattice-based schemes. In: Lin, D., Sako, K. (eds.) PKC 2019: 22nd International Conference on Theory and Practice of Public Key Cryptography, Part II. Lecture Notes in Computer Science, vol. 11443, pp. 565–598. Springer, Cham (Apr 2019). https://doi.org/10.1007/978-3-030-17259-6_19
6. D'Anvers, J.P., Karmakar, A., Roy, S.S., Vercauteren, F.: Saber: Module-LWR based key exchange, CPA-secure encryption and CCA-secure KEM. In: Joux, A., Nitaj, A., Rachidi, T. (eds.) AFRICACRYPT 18: 10th International Conference on Cryptology in Africa. Lecture Notes in Computer Science, vol. 10831, pp. 282–305. Springer, Cham (May 2018). https://doi.org/10.1007/978-3-319-89339-6_16
7. D'Anvers, J.P., Rossi, M., Virdia, F.: (One) failure is not an option: Bootstrapping the search for failures in lattice-based encryption schemes. In: Canteaut, A., Ishai, Y. (eds.) Advances in Cryptology – EUROCRYPT 2020, Part III. Lecture Notes in Computer Science, vol. 12107, pp. 3–33. Springer, Cham (May 2020). https://doi.org/10.1007/978-3-030-45727-3_1

8. D'Anvers, J.P., Vercauteren, F., Verbauwhede, I.: On the impact of decryption failures on the security of LWE/LWR based schemes. Cryptology ePrint Archive, Report 2018/1089 (2018), https://eprint.iacr.org/2018/1089
9. Forney, G.: Coset codes. ii. binary lattices and related codes. IEEE Transactions on Information Theory **34**(5), 1152–1187 (1988). https://doi.org/10.1109/18.21246
10. Glabush, L., Longa, P., Naehrig, M., Peikert, C., Stebila, D., Virdia, F.: FrodoKEM: A CCA-secure learning with errors key encapsulation mechanism. IACR Communications in Cryptology (CiC) **2**(3), 25 (2025). https://doi.org/10.62056/ayivom2hd
11. Hoffstein, J., Pipher, J., Silverman, J.H.: Ntru: A ring-based public key cryptosystem. In: Buhler, J.P. (ed.) Algorithmic Number Theory. pp. 267–288. Springer Berlin Heidelberg, Berlin, Heidelberg (1998)
12. Hofheinz, D., Hövelmanns, K., Kiltz, E.: A modular analysis of the Fujisaki-Okamoto transformation. In: Kalai, Y., Reyzin, L. (eds.) TCC 2017: 15th Theory of Cryptography Conference, Part I. Lecture Notes in Computer Science, vol. 10677, pp. 341–371. Springer, Cham (Nov 2017). https://doi.org/10.1007/978-3-319-70500-2_12
13. Jiang, H., Zhang, Z., Chen, L., Wang, H., Ma, Z.: IND-CCA-secure key encapsulation mechanism in the quantum random oracle model, revisited. In: Shacham, H., Boldyreva, A. (eds.) Advances in Cryptology – CRYPTO 2018, Part III. Lecture Notes in Computer Science, vol. 10993, pp. 96–125. Springer, Cham (Aug 2018). https://doi.org/10.1007/978-3-319-96878-0_4
14. Liang, Z., Fang, B., Zheng, J., Zhao, Y.: Compact and efficient KEMs over NTRU lattices. Comput. Stand. Interfaces **89**(C) (2024). https://doi.org/10.1016/j.csi.2023.103828
15. Liang, Z., Zhao, X., Fang, B., Zhao, Y.: Efficient andcompact ntru-based key encapsulation mechanism in large-galois-group prime-degree prime-ideal number field. Jounal of Software **36(2)**, 747–775 (2025), http://www.jos.org.cn/1000-9825/7161.htm
16. Lindner, R., Peikert, C.: Better key sizes (and attacks) for LWE-based encryption. In: Kiayias, A. (ed.) Topics in Cryptology – CT-RSA 2011. Lecture Notes in Computer Science, vol. 6558, pp. 319–339. Springer, Berlin, Heidelberg (Feb 2011). https://doi.org/10.1007/978-3-642-19074-2_21
17. Lyu, S., Liu, L., Ling, C., Lai, J., Chen, H.: Lattice codes for lattice-based PKE. Designs, Codes and Cryptography **92**(4), 917–939 (2024). https://doi.org/10.1007/s10623-023-01321-6
18. Lyubashevsky, V., Ducas, L., Kiltz, E., Lepoint, T., Schwabe, P., Seiler, G., Stehlé, D., Bai, S.: CRYSTALS-DILITHIUM. Tech. rep., National Institute of Standards and Technology (2022), available at https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022
19. Lyubashevsky, V., Peikert, C., Regev, O.: On ideal lattices and learning with errors over rings. In: Gilbert, H. (ed.) Advances in Cryptology – EUROCRYPT 2010. Lecture Notes in Computer Science, vol. 6110, pp. 1–23. Springer, Berlin, Heidelberg (May / Jun 2010). https://doi.org/10.1007/978-3-642-13190-5_1
20. Prest, T., Fouque, P.A., Hoffstein, J., Kirchner, P., Lyubashevsky, V., Pornin, T., Ricosset, T., Seiler, G., Whyte, W., Zhang, Z.: FALCON. Tech. rep., National Institute of Standards and Technology (2022), available at https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022

21. Saliba, C., Luzzi, L., Ling, C.: Error correction for FrodoKEM using the gosset lattice. In: International Zurich Seminar on Information and Communication(IZS 2022), Zurich, Switzerland. ETH, Zurich (2022). https://doi.org/10.3929/ethz-b-000535279
22. Schwabe, P., Avanzi, R., Bos, J., Ducas, L., Kiltz, E., Lepoint, T., Lyubashevsky, V., Schanck, J.M., Seiler, G., Stehlé, D., Ding, J.: CRYSTALS-KYBER. Tech. rep., National Institute of Standards and Technology (2022), available at https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022
23. Shor, P.W.: Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. SIAM Review **41**(2), 303–332 (1999). https://doi.org/10.1137/S0036144598347011
24. Tani, S.: Claw finding algorithms using quantum walk. Theoretical Computer Science **410**(50), 5285–5297 (2009). https://doi.org/10.1016/j.tcs.2009.08.030
25. Wang, A., Zheng, Z., Zhao, C., Qiu, Z., Zeng, G., Wang, X.: Scloud+: An efficient LWE-based KEM without ring/module structure. In: Security Standardisation Research. pp. 147–174. springer Nature Switzerland, Cham (2025). https://doi.org/10.1007/978-3-031-87541-0_7
26. Zamir, R., Nazer, B., Kochman, Y., Bistritz, I.: Nested lattices. Cambridge University Press (2014)