Rejection Sampling Revisit: How to Choose Parameters in Lattice-Based Signature

Zhongxiang Zheng, Anyu Wang, and Lingyue Qin

Institute for Advanced Study, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Zhongxiang Zheng; zhengzx13@tsinghua.org.cn

Received 30 March 2021; Accepted 19 May 2021; Published 7 June 2021

Academic Editor: Xiao Chen

Copyright © 2021 Zhongxiang Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rejection sampling technology is a core tool in the design of lattice-based signatures with 'Fiat–Shamir with Aborts' structure, and it is related to signing efficiency and signature size as well as security. In the rejection sampling theorem proposed by Lyubashevsky, the masking vector of rejection sampling is chosen from discrete Gaussian distribution. However, in practical designs, the masking vector is more likely to be chosen from bounded uniform distribution due to better efficiency and simpler implementation. Besides, as one of the third-round candidate signatures in the NIST postquantum cryptography standardization process, the 3rd round version of CRYSTALS-Dilithium has proposed a new method to decrease the rejection probability in order to achieve better efficiency and smaller signature size by decreasing the number of nonzero coefficients of the challenge polynomial according to the security levels. However, it is seen that small entropies in this new method may lead to higher risk of forgery attack compared with former schemes proposed in its 2nd version. Thus, in this paper, we first analyze the complexity of forgery attack for small entropies and then introduce a new method to decrease the rejection probability without loss of security including the security against forgery attack. This method is achieved by introducing a new rejection sampling theorem with tighter bound by utilizing Rényi divergence where masking vector follows uniform distribution. By observing large gaps between the security claim and actual security bound in CRYSTALS-Dilithium, we propose two series of adapted parameters for CRYSTALS-Dilithium. The first set can improve the efficiency of the signing process in CRYSTALS-Dilithium by factors of 61.7% and 41.7%, according to the security levels, and ensure the security against known attacks, including forgery attack. And, the second set can reduce the signature size by a factor of 14.09% with small improvements in efficiency at the same security level.

1. Introduction

With the rapid developments in quantum algorithms and computations, research in lattice-based cryptography has attracted considerable attention because lattice-based cryptosystems are likely to be effective against quantum computing attacks in the future. The first lattice-based cryptosystem is proposed by Ajtai and Dwork [1] in 1997 which is also known as the first cryptosystem that achieves worst case to average case reduction. Since then, many well-known lattice-based cryptosystems have been designed, including GGH [2] by Goldreich et al. and NTRU [3] by Hoffstein et al., as well as LWE by Regev [4]. Nowadays, schemes with various features, such as digital signatures [5, 6], identity-based and attribute-based encryption [7, 8], zero-knowledge proof [9], and fully homomorphic schemes [10], can be realized based on these basic designs. On the contrary, the developments of methods in solving lattice problem including enumeration [11, 12] and lattice reduction algorithms [13, 14], as well as sieving algorithms [15, 16], also contribute to the selection of parameters in these schemes. As a result, lattice-based cryptosystems are now regarded as promising candidates for the NIST postquantum cryptography standardization process.

Most lattice-based signatures are designed based on three general structures, namely, GGH structure, Fiat–Shamir structure, and GPV trapdoor structure. The GGH signature is the first practical lattice-based signature scheme which is proposed in [2] and known as the source of signatures following the GGH structure. This scheme is based
on the closest vector problem (CVP) and enjoys the advantages such as high efficiency, small signature size, and simple verification. However, analysis [17] shows that signatures of the scheme leak the information of the secret key; thus, the secret key can be recovered by collecting enough number of signatures. As a result, many variants based on GGH structure concentrate on improving the security against the attack proposed in [17]. As for another basic type, Fiat–Shamir structure is first used to design practical lattice-based signature scheme in [18]. This work combines the Fiat–Shamir structure with rejection sampling technology to avoid the risk of secret leakage. Due to its high security, high efficiency, and small signature size, many variants have been proposed based on this work including [19–21]. Among the signature schemes based on Fiat–Shamir structure, two schemes named CRYSTALS-Dilithium [22] and qTESLA [23] have been widely studied because they are known as the 2nd round NIST postquantum cryptography standardization candidates. Moreover, recently CRYSTALS-Dilithium has become one of the 3rd round NIST candidates of signatures. The other type of lattice-based signature schemes is based on GPV trapdoor structure [24], such as [25, 26]. Compared with those based on Fiat–Shamir structure, the schemes have smaller signature size but lower efficiency. Furthermore, it should be noted that the scheme FALCON [27] is a 3rd round NIST postquantum cryptography standardization candidate with the GPV trapdoor structure.

As an important subroutine in Fiat–Shamir structure, rejection sampling technology is widely used in the design of signatures schemes. The idea of this process is simple but effective, where it demands the signer selectively outputs signatures to ensure that the secret key should not be leaked by signatures. To achieve this goal, the rejection sampling process will choose to output a signature or reject it according to a fixed condition. This technology is first introduced in [18] and then further improved in [19]. When first introduced in [18], the vector for masking the secret is chosen from a bounded uniform distribution and then changed to be chosen from discrete Gaussian distribution in [19]. Besides, a theoretical analysis is also provided in [19] to prove that, under properly chosen parameters, a masking vector sampled from discrete Gaussian distribution can be used to protect the secret key from leakage by ensuring the output distribution of the rejection sampling process is statistically close to a certain discrete Gaussian distribution. In other words, the upper bound of statistical distance between the output distribution and the ideal one is small.

Although discrete Gaussian distribution enjoys the property of high security, sampling from it demands more time and space complexity than from a uniform distribution. As a result, many practical schemes choose to sample from a bounded uniform distribution for the masking vector including the two NIST candidates, CRYSTALS-Dilithium [22] and qTESLA [23]. Besides, the method to increase the success probability of rejection sampling without loss of securities is an important issue in the design of these schemes. For example, in the third-round version of CRYSTALS-Dilithium, a new technique is used to decrease the rejection probability to achieve better efficiency and smaller signature size by decreasing the number of nonzero coefficients of the challenge polynomial according to the security levels. However, it is seen that small entropies in this new method may lead to higher risk of forgery attack compared with former schemes proposed in the 2nd version which will be described in Section 2.

In this paper, we propose another way to increase the success probability of rejection sampling without loss of security. This idea is obtained by firstly proposing a more practical rejection sampling theorem with masking vector sampled from bounded uniform distribution, where a tighter bound is achieved by using Rényi divergence rather than statistical distance. Secondly, we use the proposed theorem to analyze the parameters used in CRYSTALS-Dilithium and observe that more accurate security estimation can be obtained due to the new rejection sampling theorem which allows us to adjust the parameters by balancing the securities and optimize their efficiencies as well as sizes. Our result shows that, by choosing proper parameters, the efficiency of sign algorithm in CRYSTALS-Dilithium can be further improved depending on the security levels. As lattice-based signatures with the Fiat–Shamir structure usually have higher efficiency but larger size compared with other types of lattice-based signatures, how to minimize the size of public key and signature is a core issue in the designs. So, we further propose a variant of the scheme with optimized size by utilizing our rejection sampling theorem which reduces the size of signature at the same security level. This is the third contribution of our paper.

The rest of the paper is organized as follows. In Section 2, we introduce some background about lattice, discrete Gaussian sampling, LWE problem, divergences, and rejection sampling technology. Our analysis of the rejection sampling theorem for the uniform distribution and its proof are presented in Section 3. In Section 4, some applications of the above theorem are described, including a security analysis of CRYSTALS-Dilithium parameters and several variants of CRYSTALS-Dilithium which can provide higher efficiency of signing and smaller signature size. Finally, we give our conclusion in Section 5.

2 Preliminaries

For \( x \in \mathbb{R} \), let \( \lfloor x \rfloor \) be the maximum integer that is no more than \( x \) and let \( \lfloor x \rfloor \) be the nearest integer to \( x \). Let \( \mathbb{Z}_q \) denote the set of integers in \([0, q - 1]\).

2.1 Lattice. An \( m \)-dimensional lattice is a discrete additive subgroup in \( \mathbb{R}^m \) which can be represented as the set of integer linear combination of \( n \) linearly independent vectors \([b_1, \ldots, b_n]\), i.e.,

\[
\mathcal{L}(B) = \left\{ \sum_{i=1}^{n} x_i b_i \mid x_i \in \mathbb{Z}, \forall i \in [1, n] \right\},
\]

where \( B = [b_1, \ldots, b_n] \) is called a basis of \( \mathcal{L} \) which is not unique, \( n (n \leq m) \) is the rank of the lattice, a lattice is called full-rank if \( m = n \). The determinant of \( \mathcal{L} \) is defined as...
2. Gaussian Distribution over Lattices. For \( s > 0 \), the Gaussian function is defined as
\[
\rho_{s,c}(y) = e^{-\|y-c\|^2/s^2},
\]
for \( y \in \mathbb{R}^n \), where \( s \) is called the width. When \( s = 1 \) or \( c = 0 \), the corresponding subscript is usually omitted for simplicity.

Definition 1 (discrete Gaussian distribution). For \( s > 0 \) and \( c \in \mathbb{R}^n \), the discrete Gaussian distribution \( D_{L+c,s} \) over \( L + c \) is defined as
\[
D_{L+c,s}(x) = \frac{\rho_s(x)}{\rho_s(L+c)}
\]
where \( x \in L + c \) and \( \rho_s(L+c) = \sum_{x \in L+c} \rho_s(x) \). We call \( \sigma = s/\sqrt{2\pi} \) the standard deviation for \( D_{L+c,s} \).

It is difficult to calculate the sum \( \rho_s(L) \) directly, but it is related to the sum of values of a Gaussian function over the dual lattice according to the celebrated Poisson summation formula.

Lemma 1 (Poisson summation formula, see [28]). For an \( n \)-dimensional lattice \( L \), let \( s > 0 \) and \( t \in \mathbb{R}^n \), and the following hold:
\[
\begin{align*}
(1) \quad & \rho_s(L) = (s^n/\det(L))\rho_{1/s}(L^*) \\
(2) \quad & \rho_s(L + t) = (s^n/\det(L))\sum_{w \in L^*} e^{2\pi i \langle w, t \rangle} \rho_{1/s}(w)
\end{align*}
\]

There is a tail bound for the continuous Gaussian distribution and the discrete Gaussian distribution also has a similar property which was first proven by Banaszczyk [28]. The following is a refinement to the bound of Banaszczyk given in [29].

Lemma 2 (tail bound, see [29]). For the variable \( X \sim N(0,1) \), then \( \Pr(\|X\| > t) \leq 2e^{-\pi t^2} \). For an \( n \)-dimensional lattice \( L \) and a vector \( t \in \mathbb{R}^n \), let \( s > 0 \) and \( c \geq (1/\sqrt{2\pi}) \), and we have
\[
\Pr_{x \sim D_{L+c,s}}(\|X\| > cs\sqrt{n}) \leq \left(\frac{2\pi e^n}{c^2} \right) e^{-mc^2} \rho_s(L) / \rho_s(L + t).
\]
\[\Delta_\alpha(P|Q) = \left( \sum_{k \in \text{supp}(P)} \frac{P(k)^\alpha}{Q(k)^{\alpha-\tau}} \right)^{1/(\alpha-1)}. \tag{9}\]

According to the research of [31], using Rényi divergence to estimate security can provide smaller parameters in designing lattice-based schemes than using statistical distance.

2.5. Rejection Sampling. Rejection sampling is an important tool which is widely used in designing lattice-based signature [19, 22, 23]. It is first proposed in [18] and can be used to produce a distribution that is statistically close to another one. In this way, one can output a distribution without leaking information of the secret keys and the lower bound of the complexity against attacks which use the information of signatures and has been given in Theorem 1.

**Theorem 1** (rejection sampling theorem, see [19]). Let \( V \) be a subset of \( \mathbb{Z}^m \) in which all elements have norms less than \( T \), \( \sigma \) be some element in \( \mathbb{R} \) such that \( \sigma = \omega \left( T \sqrt{\log m} \right) \), and \( h: V \rightarrow R \) be a probability distribution. Then, there exists a constant \( M = O(1) \) such that the distribution of the following algorithm \( \mathcal{S} \) is as follows:

1. \( v \leftarrow h \)
2. \( z \leftarrow D_{v, \sigma}^m \)
3. \( \text{output } (z, v) \text{ with probability } \min \left( \rho_{v, \sigma}^m(z) / M \rho_{v, \sigma}^m(z) \right) \).

Within statistical distance \( 2^{-\omega(\log m)} / M \) of the distribution, algorithm \( \mathcal{S} \) is as follows:

1. \( v \leftarrow h \)
2. \( z \leftarrow D_{v, \sigma}^m \)
3. \( \text{output } (z, v) \text{ with probability } 1/M. \)

Moreover, the probability that \( \mathcal{S} \) outputs something is at least \( (1 - 2^{-\omega(\log m)}) / M. \)

2.6. Overview of Signatures Based on Fiat–Shamir with Aborts. Fiat–Shamir with Aborts approach is an LWE-based signature framework that is firstly introduced in [18]. Based on this framework, many improvements have been proposed for better security and efficiency in [19, 22, 23]. The overview of the ‘Fiat–Shamir with Aborts’ framework can be summarized in Figure 1. Note that the original scheme proposed in [18] is based on the LWE problem, and further improvements [19, 22, 23] are mainly designed based on ring-LWE or module-LWE to achieve better efficiency. In this paper, we concentrate on practical designs; thus, all elements as well as computations in the following paper are in the polynomial ring \( \mathbb{Z}_q[X]/(X^n + 1) \).

2.7. The Entropy of Challenge Polynomial. In the updated version of the 3rd-round submission of CRYSTALS-Dilithium, a new method is introduced to decrease the rejection probability in order to achieve better efficiency where the number of nonzero coefficients of the challenge polynomial \( c \) varies according to the security levels. For example, in the 3rd round version, the challenge polynomial \( c \) are chosen from \( B_{324}, B_{350}, B_{39}, B_{60}, B_{60}, \) and \( B_{60} \), respectively, for the parameters with the security claims of 55, 89, 112, 165, 229, 298, and 343, while in the 2nd version, \( c \) is chosen from \( B_{60} \) for all parameters. Since all coefficients of the challenge polynomial are in the set of \( \{-1, 0, 1\} \), the entropy of a challenge polynomial chosen from \( B_4 \) is log \( \left( \frac{256}{\tau} \right) + \tau \) bit. However, small entropy leads to a direct forgery attack without any valid message for ‘Fiat–Shamir with Aborts’ structure. Recall the verification process of the schemes; given the public key \( pk = (A, t) \), the adversary can be performed by finding a set of signature \( \sigma = (\sigma', c') \) which satisfies \( \|z\|_{\infty} < \gamma_1 - \beta \) and \( c' = H ((M[w]) = H (M [\text{HighBits}(A z' - c' t, 2y_2)). For any fixed \( c' \in B_4 \), an adversary can forge a signature by picking some \( z' \) that satisfies \( \|z'\|_{\infty} < \gamma_1 - \beta \) and check if \( c' = H (M [\text{HighBits}(A z' - c' t, 2y_2)). Since the entropy of \( z' \) is much larger than the entropy of \( c' \), the adversary shall succeed with the complexity of \( O \left( \log \left( \frac{256}{\tau} \right) + \tau \right) \) in the classical model. Furthermore, by regarding the formula \( c' = H (M [\text{HighBits}(A z' - c' t, 2y_2)) \) as a function of \( z' \), the Grover algorithm can be used to achieve quadratic speedup with the complexity of \( O \left( \sqrt{\log \left( \frac{256}{\tau} \right) + \tau} \right) \) in the quantum model. As a result, the securities of these sets of parameters against forgery attack shall be 67, 80, 96, 112, 128, 128, and 128 in the quantum model, where the corresponding security claims are 55, 89, 112, 165, 229, 298, and 343. For these seven sets of parameters, the last six of them suffer from this quantum forgery attack and the last two sets of parameters...
even may not be secure in the classical model, as shown in Table 1. Since the proposed idea of decreasing the rejection probability may also decrease the security of the scheme, in Section 3, we will provide another method to achieve this goal without loss of security and use it to introduce variants of CRYSTALS-Dilithium which achieve better efficiency and smaller signature size. Since the security claims of the third round parameters in CRYSTALS-Dilithium have large gaps with the complexities of the forgery attack, our comparisons shall be conducted based on the second version of parameters in CRYSTALS-Dilithium rather than the third-round ones (note that the practical verification process of the 3rd version CRYSTALS-Dilithium is more complex than the framework shown in Figure 1 due to the application of hint vector as well as two stages’ sampling process; however, it is easy to check that the proposed forgery attack also applies to this practical scheme).

3. Rejection Sampling Theorem for Uniform Distribution

Rejection sampling theorem proposed in [19] can be used to estimate the security of the rejection sampling process against secret recovery attacks by computing the upper bound of the statistical distance of the output distribution and the target one where the two distributions follow discrete Gaussian distribution with distinct centers. However, in practical designs, uniform distributions are often used rather than discrete Gaussian distributions. This makes it more efficient and more convenient to sample elements, but the complexity of recovering secret key from the output of such samples remains unknown. Besides, by utilizing Rényi divergence instead of statistical distance used in the rejection sampling theorem of [19], a tighter security bound which leads to smaller parameter size can be obtained. So, let us start with clearly defining the problem and then providing a theorem about solving the problem.

**Definition 7** (distinguish problem for rejection sampling with bounded uniform masking vector). Let \( S_\gamma \) be a uniform distribution with elements in \([-\gamma, \ldots, 0, \ldots, \gamma]\), \( D_\eta \) be an arbitrary distribution with the support \([-\eta, \ldots, 0, \ldots, \eta]\), and \( \beta < \gamma \) be a positive constant. Given a number of samples, the goal is to decide which of the two algorithms the samples follow.

Algorithm \( \mathcal{A} \):
1. \( y \xleftarrow{} S_\gamma^m \),
2. \( x \xleftarrow{} D_\eta^m \),
3. if \( \|y + x\|_\infty \geq \gamma - \beta \), restart,
4. output \( z = y + x \).

Algorithm \( \mathcal{F} \):
1. \( y \xleftarrow{} S_\gamma^m \),
2. if \( \|y\|_\infty \geq \gamma - \beta \), restart,
3. output \( z = y \).

**Theorem 2**. Given a distinguish problem for rejection sampling with bounded uniform masking vector defined by the probability distribution \( S_\gamma \) and \( D_\eta \) and integers \( m > 0, \eta > 0, \eta \geq \beta > 0 \), the Rényi divergence between the output distributions, denoted as \( \Delta_{\alpha} \), is

\[
\Delta_{\alpha}(P_{z_i}(k)\|Q(k)) = \left( \frac{\sum_{k \in [-\gamma+\beta, \gamma-\beta-1]} P_{z_i}(k)^{\alpha}}{Q(k)^{\alpha-1}} \right)^{1/(\alpha-1)},
\]

where \( Q(k) = 1/(2(\gamma - \beta) - 1) \),

\[
P_{z_i}(k) = \begin{cases} \frac{1}{2\gamma + 1} \times \frac{1}{1 - S} & \text{if } k \in [-\gamma + \eta + 1, \gamma - \eta - 1]; \\ \frac{1}{2\gamma + 1} \times \frac{\text{CDF}_D(k + \gamma)}{1 - S}, & \text{if } k \in [-\gamma + \beta + 1, -\gamma + \eta]; \\ \frac{1}{2\gamma + 1} \times \frac{(1 - \text{CDF}_D(k - \gamma - 1))}{1 - S}, & \text{if } k \in [\gamma - \eta, \gamma - \beta - 1], \end{cases}
\]

and \( S = \sum_{k=-\eta-\gamma}^{-\gamma+\beta} P_{z_i}(k) + \sum_{k=\gamma-\beta}^{\gamma+\eta} P_{z_i}(k), \ \text{CDF}_D(k) = \sum_{i=\eta}^{k} P_{x \leftarrow D_\eta}(x = i), \text{ and } P_{z_i}(k) \text{ is defined as follows:} \)
Theorem. Let us first study the distribution of \( z = y + x \) without rejection. As the support of \( y_i \) (\( i \in [1, m] \)) and \( x_i \) (\( i \in [1, m] \)) are separately sampled from \( \{-y, \ldots, 0, \ldots, y\} \) and \( \{-\eta, \ldots, 0, \ldots, \eta\} \), it is clear that \( z_i \) (\( i \in [1, m] \)) has the output distribution as follows:

\[
P_z(k) = \begin{cases} 
\frac{1}{2y+1}, & \text{if } k \in [-y + \eta + 1, y - \eta - 1], \\
\frac{1}{2y+1} \times \text{CDF}_D(k + \gamma), & \text{if } k \in [-y - \eta, -y + \eta], \\
\frac{1}{2y+1} \times (1 - \text{CDF}_D(k - \gamma - 1)), & \text{if } k \in [y - \eta, y + \eta], 
\end{cases} 
\]

(12)

where \( \text{CDF}_D \) denotes the cumulative distribution function of the distribution \( D_{\eta} \) and \( \text{CDF}_D(k) = \sum_{m=0}^{\infty} P_{x \sim D_{\eta}}(x = i) \).

When applying rejection sampling with the condition \( |z_i| \geq y - \beta \) to \( z_i \), whether \( \beta > \eta \) or not shall influence the output distribution, if \( \beta > \eta \), we have \( P_z'(k) \) (\( k \in [-y + \beta + 1, y - \beta - 1] \)) as follows:

\[
P_z'(k) = \frac{1}{2y+1} \times \frac{1}{(2(y - \beta) - 1)/2(y + 1)} = \frac{1}{2(y - \beta) - 1}, \quad \text{if } k \in [-y + \beta, y - \beta]. 
\]

(14)

And, for \( 0 < \beta \leq \eta \), we have \( P_z'(k) \) (\( k \in [-y + \beta + 1, y - \beta - 1] \)) as follows:

\[
P_z'(k) = \begin{cases} 
\frac{1}{2y+1} \times \frac{1}{1 - s}, & \text{if } k \in [-y + \eta + 1, y - \eta - 1], \\
\frac{1}{2y+1} \times \frac{\text{CDF}_D(k + \gamma)}{1 - S}, & \text{if } k \in [-y + \beta + 1, -y + \eta], \\
\frac{1}{2y+1} \times \frac{(1 - \text{CDF}_D(k - \gamma - 1))}{1 - S}, & \text{if } k \in [y - \eta, y - \beta - 1], 
\end{cases} 
\]

(15)
where $S = \sum_{k=\eta}^{-\eta} P_{z_i}(k) + \sum_{k=\gamma}^{-\gamma} P_{z_i}(k)$. As a result, we now have the output distribution of algorithm $\mathcal{A}'$ denoted as $P_{\mathcal{A}'}(k)$, and the output distribution $P^*$ of algorithm $\mathcal{A}$ can be derived in a similar way. For $\beta > 0$, we have $P_{z_i}(k) (k \in [-\gamma + \beta, 1, \gamma - \beta - 1])$ as follows:

$$P_{z_i}(k) = \frac{1}{2(\gamma - \beta)-1}, \quad \text{if } k \in [-\gamma + \beta + 1, \gamma - \beta - 1].$$

(16)

Now, we have clear descriptions of the output distributions of the two algorithms, and it is seen that the two distributions are exactly the same when $\beta \leq \eta$. To measure the distances between the two probability distributions and evaluate the security, we shall recall the definition of Rényi divergence.

For any two discrete probability distributions $P$ and $Q$ such that $\text{Supp}(P) \subset \text{Supp}(Q)$ and $\alpha \in (1, +\infty)$, the Rényi divergence of order $\alpha$, denoted as $\Delta_{\alpha}$, is defined by

$$\Delta_{\alpha}(P|Q) = \left( \sum_{k \in \text{Supp}(P)} \frac{P(k)^{\alpha}}{(Q(k))^{1/\alpha}} \right)^{1/(\alpha-1)}. \quad (17)$$

Combining the result of $P_{\mathcal{A}'}(k)$ and the definition of $Q(k)$, this finishes the proof. $\square$

4. Applications of the Rejection Sampling

4.1. Distinguish Analysis for Rejection Sampling. With the help of Theorem 2 and equation (18), we can evaluate the security against attacks utilizing information of signatures by Rényi divergence for practical lattice-based signature schemes, including the NIST candidates. We shall take the parameters used in CRYSTALS-Dilithium as examples to show how to analyze the lower bound of the complexities of these attacks by utilizing Theorem 2. Besides, it should be noted that the compress technology proposed in [21] is commonly used to reduce the length of signature (which corresponds the rejection condition about the infinite norm of low bits shown in line 11 of Figure 1), this process can be viewed as another distinguish problem of rejection sampling for uniform distribution with different parameters. Take the parameters used in CRYSTALS-Dilithium-II and III which separately correspond to the first and the second level of NIST’s categories as examples, we can normalize them into the following distinguish problems in Table 2, where we use ‘R’ to denote the distinguish problem of rejection sampling process and ‘C’ to denote the distinguish problem of compress process.

Since the upper bound of advantage for attacks utilizing the outputs of signatures is given by equation (18), say when advantage is no more than $2^{-\alpha}$, the lower bound of complexity for attacks is at least $2^{\alpha}$. Thus, applying Theorem 2 to the two distinguish problems, we can get the following results of security analysis, as shown in Table 3. Besides, we also take the proposed forgery attack into consideration because the security analysis, as shown in Table 3, the security of these attacks utilizing the information of signatures are much larger than other types of attacks especially for the forgery attack brought by the small entropy of $c$. By observing the large gaps between complexities of different type attacks, refined parameters which can provide better efficiency, smaller sizes, and higher security can be obtained by balancing these complexities. We will discuss how to choose parameters and what can be achieved in Section 4.2.

4.2. Choosing New Parameters for Rejection Sampling of Dilithium. In Section 4.1, we estimate the complexities of corresponding distinguish problems and observe there exist large gaps between the complexities of different types of attacks. As the parameters of rejection sampling relate to the efficiency, the security, and the signature size of the schemes at the same time, we can balance the gaps in order to achieve better efficiency, higher security, and smaller size of the schemes. The balancing shall be used based on parameters of CRYSTALS-Dilithium, and it should be noted that this technology can be naturally applied to other signature schemes.
using rejection sampling for uniform distributed masking vectors. Our approaches contain the following steps:

1. Utilizing the method in Section 2.7 to balance the complexities of forgery attack by adjusting proper entropy of $c$ according to the security levels

2. Utilizing Theorem 2 to balance the complexities of the two distinguish problems by setting $\beta_1$ and $\beta_2$ separately for rejection sampling process and compress process rather than using the same $\beta$ for the two processes

3. Utilizing the methods of primal attack, dual attack, and SIS attack used in [22] to balance the complexities of various types of attacks by choosing proper $\beta$, $\eta$, and $c$

To apply these modifications, new parameter $\beta_1$ and $\beta_2$ should be introduced to replace $\beta$, and the revised framework is shown in Figure 2. Besides, since the choices of parameters for $\beta_1$ and $\beta_2$ are very close, the hardness reduction of the framework in Figure 2 follows the one in Figure 1 naturally.

The success probability of rejection sampling and compress process relates to the efficiency of signature because the sign process will be continuously repeated until a proper signature is outputted, and the success probability is computed as

$$P_{\text{succ}} \approx e^{-256 \left( \left( \frac{\beta_1}{\beta_2} \right)^{\eta_0} \left( \frac{\beta_1}{\beta_2} \right)^{\eta_1} \right)}.$$  \hspace{1cm} (19)

Based on these analyses, we choose parameters by designing a program which contains the algorithms of success probabilities, primal attack estimation, dual attack estimation, SIS attack estimation, and the distinguish attack estimation given by Theorem 2. With the input of parameters, the program outputs these complexities and properties. And, the final results are obtained by testing different values iteratively and make a balance of these complexities and the efficiency.

The comparisons of the parameters in this work (separately denoted as This Work-I and This Work-II corresponding to different security levels) and those in CRYSTALS-Dilithium are shown in Table 4. The implementations can be found in https://github.com/Anonymous496/Digital-signatures. And, the experiments of efficiency are conducted with the environment of Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz.

As the signing procedure will repeat several times until a signature is outputted, the success probabilities influence the efficiency of signing process directly. In other words, the efficiency of the signing process with our technique are 61.7% and 41.7% faster than that in CRYSTALS-Dilithium according to the security levels. Furthermore, since the sizes of public key and signature are considered as more important factors than their efficiency for signature schemes based on Fiat–Shamir structure, we can use the proposed technology to introduce a new set of parameters with smaller signature size and keep the same security level with small improvement in efficiency. We denote the adapted scheme as this Work-III, and the comparisons can be found in Table 5.

From the comparison, it is seen that the signature size of the proposed scheme is 14.09% smaller than the original one with small improvement in signing efficiency and keeps the same security level compared with CRYSTALS-Dilithium-II. It should be noted that similar optimizations can also be applied to other sets of parameters in CRYSTALS-Dilithium.
Table 4: Parameters of the adapted scheme I.

| Parameters          | Dilithium-II | This Work-I | Dilithium-III | This Work-II |
|---------------------|--------------|-------------|---------------|--------------|
| \( q \)             | 8380417      | 8380417     | 8380417       | 8380417      |
| \( d \)             | 14           | 14          | 14            | 14           |
| Weight of \( c \)   | 60           | 60          | 60            | 60           |
| Entropy of \( c \)  | 257          | 257         | 257           | 257          |
| \( \gamma_1 \)      | 523776       | 523776      | 523776        | 523776       |
| \( \gamma_2 \)      | 261888       | 261888      | 261888        | 261888       |
| \((k, l)\)          | (4, 3)       | (4, 3)      | (5, 4)        | (5, 4)       |
| \( \eta \)          | 6            | 6           | 5             | 5            |
| \( \beta \)         | 325          | —           | 275           | —            |
| \( \beta_1 \)       | —            | 170         | —             | 180          |
| \( \beta_2 \)       | —            | 170         | —             | 180          |
| \( \omega \)        | 80           | 80          | 96            | 96           |
| pk size (bytes)     | 1184         | 1184        | 1472          | 1472         |
| sig size (bytes)    | 2044         | 2044        | 2701          | 2701         |
| Expectation of repeats | 5.74   | 2.49        | 6.56          | 3.42         |
| keygen (ms)         | 0.10         | 0.10        | 0.15          | 0.15         |
| Sign (ms)           | **0.55**     | **0.34**    | **0.85**      | **0.60**     |
| Verify (ms)         | 0.12         | 0.12        | 0.17          | 0.17         |
| SIS block size      | 340          | 340         | 488           | 488          |
| Classical/quantum SIS security | 99.28/90.1 | 99.28/90.1 | 142.49/129.59 | 142.49/129.59 |
| LWE block size      | 255          | 255         | 377           | 377          |
| Classical/quantum LWE security | 96.24/90.1 | 96.24/90.1 | 142.28/130.91 | 142.28/130.91 |
| \( 1/\text{adv}(R) \) | 300.43       | 95.19       | 317.58        | 129.77       |
| \( 1/\text{adv}(C) \) | 298.34       | 93.19       | 315.58        | 128.0        |
| Classical/quantum forgery attack | **257/128** | **196/98**  | **257/128**   | **257/128**  |
| Security against known attacks | 90.1         | 90.1        | 128           | 128          |

KeyGen ():  
01. A ← \( R_{k \times l}^{q} \),  
02. \((s_1, s_2) ← S_{\gamma_1} \times S_{\gamma_2}^{d} \),  
03. t := As_1 + s_2,  
04. return \( pk = (A, t), sk = (A, t, s_1, s_2) \).

Sign (sk, M):  
05. z := \( \perp \),  
06. while \( z := \perp \) do,  
07. \( y ← S_{\gamma_1} \),  
08. \( w_1 := \text{HighBits} (Ay, 2_{\gamma_2}) \),  
09. \( c ∈ B_{60} := H(M || w_1) \),  
10. \( z := y - cs_1 \),  
11. if \( ||z||_{\infty} ≥ \gamma_1 - \beta_1 \) or \( ||\text{LowBits} (Ay - cs_2, 2_{\gamma_2})||_{\infty} ≥ \gamma_2 - \beta_2 \) or \( \text{HighBits} (Ay - cs_2, 2_{\gamma_2}) \neq w_1 \), then \( z := \perp \),  
12. return \( σ = (z, c) \),  
Verify (pk, M, o):  
13. \( w'_1 := \text{HighBits} (Az - ct, 2_{\gamma_2}) \),  
14. if \( ||z||_{\infty} < \gamma_1 - \beta \) and \( c = H(M || w'_1) \) then return ‘verified successfully’.

Figure 2: Revised framework of signatures based on Fiat–Shamir with Aborts.
5. Conclusion

In this paper, we study rejection sampling technology for lattice-based signatures and concentrate on the conditions for practical designs. We first introduce a new rejection sampling theorem for bounded uniform distributed masking vectors which is widely used in current designs where a tighter result is obtained due to the usage of Rényi divergence, and then, we use the proposed theorem to analyze the complexities against attacks utilizing information of signatures for the parameters in CRYSTALS-Dilithium and observe that there exist large gaps between complexities of different types of attacks, e.g., forgery attack and key recovery attack. Thirdly, we propose two series of adapted parameters for CRYSTALS-Dilithium. The first set can improve the efficiency of the signing process in CRYSTALS-Dilithium by factors of 61.7% and 41.7% according to the security levels and ensure the same signature size as well as security claims including forgery attack. And, the second set can reduce the signature size by a factor of 14.09% with small improvement in signing efficiency and keep the same security level.

Data Availability

The data used to support the findings of the study are available at https://github.com/Anonymous496/Digital-signatures.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by National Key Research and Development Program of China (Grant nos. 2017YFA0303903 and 2018YFA0704701), Major Program of Guangdong Basic and Applied Research (Grant no. 2019B030302008), and Major Scientific and Technological Innovation Project of Shandong Province (Grant no. 2019JZZY010133).

References

[1] M. Ajtai and C. Dwork, “A public-key cryptosystem with worst-case/average-case equivalence,” in Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing, pp. 284–293, ACM, El Paso, TX, USA, May 1997.

[2] O. Goldreich, S. Goldwasser, and S. Halevi, “Public-key cryptosystems from lattice reduction problems,” in Advances in Cryptology - CRYPTO ’97, 17th Annual International Cryptology Conference Proceedings, vol. 1294, Lecture Notes in Computer Science, pp. 112–131, Springer, Santa Barbara, CA, USA, August 1997.

[3] J. Hoffstein, J. Pipher, and J. H. Silverman, “NTRU: a ring-based public key cryptosystem,” in Proceedings of the Algorithmic Number Theory, Third International Symposium, ANTS-III, vol. 1423, Lecture Notes in Computer Science, pp. 267–288, Springer, Portland, OR, USA, June 1998.

[4] O. Regev, “On lattices, learning with errors, random linear codes, and cryptography,” in Proceedings of the 37th Annual ACM Symposium on Theory of Computing, pp. 84–93, ACM, Baltimore, MD, USA, May 2005.

[5] R. Hiromasa, “Digital signatures from the middle-product LWE,” in Proceedings of the provable security - 12th
[6] V. Lyubashevsky, “Digital signatures based on the hardness of ideal lattice problems in all rings,” in *Proceedings of the Advances in Cryptology - ASIACRYPT 2016 - 22nd International Conference on the Theory and Application of Cryptology and Information Security, Part II*, vol. 10032, Lecture Notes in Computer Science, pp. 196–214, Hanoi, Vietnam, December 2016.

[7] C. Xie, J. Weng, W. Lu, and L. Hou, “General construction of revocable identity-based fully homomorphic signature,” *Science China Information Sciences*, vol. 63, no. 3, 2020.

[8] F. Luo, S. Al-Kuwari, W. Susilo, and D. H. Duong, “Attribute-based proxy re-signature from standard lattices and its applications,” *Computer Standards & Interfaces*, vol. 75, p. 103499, 2021.

[9] Y. Zhang, X. Liu, Y. Yin, Q. Zhang, and H. Jia, “On new zero-knowledge proofs for fully anonymous lattice-based group signature scheme with verifier-local revocation,” in *Proceedings of the Applied Cryptography and Network Security Workshops - ACNS 2020*, vol. 12418, Lecture Notes in Computer Science, pp. 381–399, Springer, Rome, Italy, June 2020.

[10] F. Luo, F. Wang, K. Wang, and K. Chen, “Fully homomorphic encryption based on the ring learning with rounding problem,” *IET Information Security*, vol. 13, no. 6, pp. 639–648, 2019.

[11] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis,” *Mathematics of Computation*, vol. 44, no. 170, p. 463, 1985.

[12] N. Gama, P. Q. Nguyen, and O. Regev, “Lattice enumeration using extreme pruning,” in *Proceedings of the Advances in Cryptology - EUROCRYPT 2010*, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, vol. 6110, Lecture Notes in Computer Science, pp. 257–278, Springer, Monaco/French Riviera, May 2010.

[13] A. K. Lenstra, H. W. Lenstra, and L. Lovász, “Factoring polynomials with rational coefficients,” *Mathematische Annalen*, vol. 261, no. 6, pp. 515–534, 1982.

[14] Y. Chen and P. Q. Nguyen, “Bkz 2.0: better lattice security estimates,” in *Proceedings of the Advances in Cryptology - ASIACRYPT 2011 - 17th International Conference on the Theory and Application of Cryptology and Information Security*, vol. 7073, Lecture Notes in Computer Science, pp. 1–20, Springer, Seoul, South Korea, December 2011.

[15] P. Q. Nguyen and T. Vidick, “Sieve algorithms for the shortest vector problem are practical,” *Journal of Mathematical Cryptology*, vol. 2, no. 2, pp. 181–207, 2008.

[16] A. Becker, L. Ducas, N. Gama, and T. Laarhoven, “New directions in nearest neighbor searching with applications to lattice sieving,” in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016*, pp. 10–24, SIAM, Arlington, VA, USA, January 2016.

[17] P. Q. Nguyen and O. Regev, “Learning a paralelepiped: cryptanalysis of GGH and NTRU signatures,” in *Proceedings of the Advances in Cryptology - EUROCRYPT 2006*, 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques, vol. 4004, Lecture Notes in Computer Science, pp. 271–288, Springer, St. Petersburg, Russia, May 2006.

[18] V. Lyubashevsky, “Fiat-shamir with aborts: applications to lattice and factoring-based signatures,” in *Proceedings of the Advances in Cryptology - ASIACRYPT 2009*, 15th International Conference on the Theory and Application of Cryptology and Information Security, vol. 5912, Lecture Notes in Computer Science, pp. 598–616, Springer, Tokyo, Japan, December 2009.

[19] V. Lyubashevsky, “Lattice signatures without trapdoors,” in *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques Proceedings*, vol. 7237, Lecture Notes in Computer Science, pp. 738–755, Springer, Cambridge, UK, April 2012.

[20] T. Güneysu, V. Lyubashevsky, and T. Pöppelmann, “Practical lattice-based cryptography: a signature scheme for embedded systems,” in *Cryptographic Hardware and Embedded Systems - CHES 2012 - 14th International Workshop Proceedings*, vol. 7428, Lecture Notes in Computer Science, pp. 530–547, Springer, Leuven, Belgium, September 2012.

[21] S. Bai and S. D. Galbraith, “An improved compression technique for signatures based on learning with errors,” in *Topics in Cryptology - CT-RSA 2014 - The Cryptographer’s Track at the RSA Conference 2014 Proceedings*, vol. 8366, Lecture Notes in Computer Science, pp. 28–47, Springer, San Francisco, CA, USA, February 2014.

[22] V. Lyubashevsky, L. Ducas, E. Kiltz et al., “Crystals-dilithium,” *Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2018, no. 1, pp. 238–268, 2018.

[23] E. Alkim, S. L. M. Paulo, Barreto, and N. Bindel, “The lattice-based digital signature scheme qTESL,” in *Proceedings of the Applied Cryptography and Network Security-18th International Conference [ACNS] 2020*, Rome, Italy, October 2020.

[24] C. Gentry, C. Peikert, and V. Vaikuntanathan, “Trapdoors for hard lattices and new cryptographic constructions,” in *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pp. 197–206, ACM, Victoria, British Columbia, Canada, May 2008.

[25] J. Alwen and C. Peikert, “Generating shorter bases for hard random lattices,” *Theory of Computing Systems*, vol. 48, no. 3, pp. 535–553, 2011.

[26] D. Micciancio and C. Peikert, “Trapdoors for lattices: simpler, tighter, faster, smaller,” in *Advances in Cryptology - EUROCRYPT 2012 - 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques Proceedings*, vol. 7237, Lecture Notes in Computer Science, pp. 700–718, Springer, Cambridge, UK, April 2012.

[27] T. Prest, P. A. Fouque, J. Hoffstein et al., “Falcon,” National Institute of Standards and Technology, Gaithersburg, MD, USA, 2018.

[28] W. Banaszczyk, “New bounds in some transference theorems in the geometry of numbers,” *Mathematische Annalen*, vol. 296, no. 1, pp. 625–635, 1993.

[29] C. Tian, M. Liu, and G. Xu, “Measure inequalities and the transference theorem in the geometry of numbers,” *Proceedings of the American Mathematical Society*, vol. 142, no. 1, pp. 47–57, 2014.

[30] V. Lyubashevsky, C. Peikert, and O. Regev, “On ideal lattices and learning with errors over rings,” in *Proceedings of the Advances in Cryptology - EUROCRYPT 2010*, 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques, vol. 6110, Lecture Notes in Computer Science, pp. 257–282, Springer, St. Petersburg, Russia, May 2006.
[31] S. Bai, T. Lepoint, A. Roux-Langlois, A. Sakzad, D. Stehlé, and R. Steinfeld, "Improved security proofs in lattice-based cryptography: using the Rényi divergence rather than the statistical distance," *Journal of Cryptology*, vol. 31, no. 2, pp. 610–640, 2018.