文章编号:1671-4229(2024)02-0057-08

# 基于 NTRU 格上非球型离散高斯采样的优化

柴惠哲，唐春明*，贾惠文
(广州大学 数学与信息科学学院，广东 广州　510006)

**摘　要**：随着量子计算机的飞速发展,后量子密码成为研究热点。格密码因性能均衡、安全基础牢靠,以及功能丰富等特点成为后量子密码中的主流。原像采样是格密码中的核心算法,被广泛应用于诸多高级密码方案的构造,格上 Hash-and-Sign 数字签名是最简单、最直接的应用。从技术上原像采样算法分为 GPV 型和 Peikert 型,前者的特点是输出质量高,但算法通常只能串行执行;后者支持并行运算,但输出质量较差。文章将非球面高斯技术应用于 NTRU 格上的 Peikert 型采样算法,旨在提升其效率。具体选取了两种参数模式,和原始 NTRU 格上的 Peikert 型采样算法相比,模式 1 可以提高基于该采样算法数字签名的安全强度并降低签名尺寸;模式 2 在不降低安全性的前提下,可以进一步降低签名尺寸。实验结果表明,在模式 1 中,安全性提升约 18% ~20%,签名尺寸降低约 15%;模式 2 保持安全性不变,但是签名尺寸降低约 30% ~35%。

**关键词**：格密码；NTRU 格；非球面高斯采样
**中图分类号**：TN918.4　　**文献标志码**：A

# Optimization of non-spherical discrete Gaussian sampling based on NTRU lattice

*CHAI Hui-zhe*, *TANG Chun-ming**, *JIA Hui-wen*
(School of Mathematics and Information Science, Guangzhou University, Guangzhou 510006, China)

**Abstract**: With the rapid development of quantum computers, post-quantum cryptography has become a research hotspot. Lattice cryptography has become the mainstream in post-quantum cryptography due to its balanced performance, solid security foundation, and rich functions. Pre-image sampling is the core algorithm in lattice cryptography and is widely used in the construction of many advanced cryptography schemes. Hash-and-Sign digital signature on lattice is its simplest and most direct application. Technically, pre-image sampling algorithms are divided into GPV and Peikert. The former is characterized by high output quality, but the algorithm can usually only be executed serially; the latter supports parallel operations, but the output quality is poor. This article applies non-spherical Gaussian technology to the Peikert sampling algorithm on the NTRU lattice, aiming to improve its efficiency. Specifically, two parameter modes were selected. Compared with the Peikert sampling algorithm on the original NTRU lattice, mode 1 can improve the security strength of digital signatures based on this sampling algorithm and reduce the size of the signature; mode 2 does not reduce security. Under the premise, the signature size can be further reduced. Experimental results show that in mode 1, the security is improved by about 18% ~20% and the signature size is reduced by about 15%; in mode 2,

**收稿日期**：2023 - 09 - 20；**修回日期**：2023 - 12 - 08
**基金项目**：国家重点研发计划资助项目(2021YFB3100200)；国家自然科学基金资助项目(12171114)
**作者简介**：柴惠哲(1999—),男,硕士研究生. E-mail:1290501583@qq.com
* 通信作者. E-mail:ctang@gzhu.edu.cn
引文格式：柴惠哲，唐春明，贾惠文. 基于 NTRU 格上非球型离散高斯采样的优化[J]. 广州大学学报(自然科学版),2024,23(2):57-64.

the security remains unchanged, but the signature size is reduced by about 30% ~35%.

**Key words**: lattice cryptography; NTRU; non-spherical Gaussian

## 0 引 言

随着时代的发展,传统的基于大整数分解和离散对数求解困难问题的公钥密码体制如RSA,ECC等不再安全[1]。因此,"后量子密码"体系的建立迫在眉睫。而格上困难问题存在最坏情况到平均情况的归约特点,具有抗量子攻击性,为未来的签名方案提供了理论基础。

格密码首次成功应用追溯至2008年由Gentry等[2]提出,利用高斯采样算法构造了陷门单向函数,应用于签名方案中,并给出了安全性证明。其主要思想是Klein[3]的最近平面算法随机化版本,在任意的格基上离散采样一个向量,输出的结果与基的几何形状无关,从而隐藏了陷门信息。随后,2010年由Peikert[4]提出利用卷积乘法,将输出向量分布的协方差矩阵卷积成球形高斯分布,隐藏了陷门信息,并且这种算法在q模格上生成的采样更高效。随后,2012年Micciancio等[5]对Pei10中的算法进行了更简易、更高效的优化,与GPV成为现在两种重要的高斯采样方案。

高斯采样框架的提出,引起了学者们的广泛兴趣,越来越多的研究者投入到该领域的研究工作中。例如Ducas等[6]的(DDLL13)对输出的分布与目标分布有更小的误差,提高采样的质量;Ducas等[7]的DLP14中,将它们的内容与通过随机化Babai最近平面算法获得的高斯采样算法相结合,被纳入了GPV框架。他们分析了构造的安全性,并提供了GPV框架中签名方案的第一个合理有效的实现;然而,这样做是以缓慢的签名算法为代价的,对此Ducas等[8]在DP16文章中提出利用快速傅里叶方法加快正交基的计算,并且在NTRU格上利用快速傅里叶算法加速采样过程,最终计算时间下降到O(N),同时也利用了底层数域的域结构来实现更好的复杂性。最终提出了一套成熟的签名方案:FALCON[9],现已入围后量子密码第四轮的选拔。然而FALCON方案虽然很快,但其签名算法非常复杂,难以实现,不适合并行化,也难以抵御侧信道攻击,因此,针对这些缺点,MITAKA[10]方案对FALCON方案进行优化,使用了一种新的技术来生成更高质量的陷门,可以在很大程度上减轻安全损失并提高效率,新的方案可以更简单、在线/离线、更容易并行化地运行,同时防止侧信道攻击;Espitau等[11]又针对FALCON和MITAKA框架通过椭圆采样,降低了模数q,并通过适当的编码技术改善了生成签名和密钥所需空间较大的问题。

FALCON等方案的实现需要高斯采样算法,常用GPV和Peikert算法。Peikert算法适用于与基$\boldsymbol{B}$的最大奇异值$s_1(\boldsymbol{B})$一样小的高斯参数$s$,在密码学应用中,高斯参数是控制具体安全性的主要量,也是潜在最坏情况格问题的近似因子。这是为了保证方案安全,对手很难在高斯的可能半径$s\sqrt{n}$内找到格点。很容易证明$s_1(\boldsymbol{B})$至少是$\max\|b_i\|$,因此,无法从比GPV算法更窄的高斯分布中采样。同时,任何基$\boldsymbol{B}$总是可以被有效地处理以得到$s_1(\boldsymbol{B})\leqslant n\cdot\|\acute{\mathrm{B}}\|$,所以,Peikert算法至多比GPV算法宽松$n$倍,安全性远低于GPV算法。故本文提出一种优化方案,可以在保证算法运行时间的情况下,改进Peikert算法中卷积的协方差矩阵,使采样分布的高斯参数更小,即得到更高安全性的签名以及减少所需的存储空间。

具体而言,本研究将原方案采样所使用的离散球面高斯分布推广到非球面高斯分布上,而不会泄露任何关于陷门的信息。由于采样安全性和所需存储空间大小与算法的高斯参数密切相关,优化方案允许设置参数以提高安全级别或减小签名大小。因此,为满足不同的需求设置了两种参数模式。在模式1中,为了提高伪造签名攻击的安全性,即降低原像采样的上界$l_2=\sqrt{\gamma_1^2+\gamma_2^2}$,在保证算法正确性的条件下,调整$\gamma_1,\gamma_2$的取值使得$l_2$最小,最终原像采样的上界缩小至原方案的53%,签名算法中的安全性提高18%,同时,采样所需的存储空间也有所减少。在模式2中,为了减少存储空间$storage=n(1+\log(\gamma_1))$的大小,并且保证算法的正确性和伪造签名攻击的安全性。

存储空间的大小只与 $\gamma_2$ 相关,这是因为实际利用采样算法的签名方案是基于 NTRU 格的,只需一半高斯向量作为签名输出,而另一半则在验证过程中恢复。选择满足条件的最小 $\gamma_2$,最终可以减少采样的存储大小至原方案的 66%。最后,对优化后的算法进行安全性分析,利用 BKZ 算法评估签名伪造和密钥恢复的复杂度。本文算法的优化仅涉及降低原像采样的上界,故在伪造签名攻击下的安全性得到了提高,而密钥恢复攻击下的安全性与高斯采样的参数有关,故此安全性保持不变。

## 1　预备知识

### 1.1　线性代数

一组线性无关的向量集 $\boldsymbol{B}=\{b_1,b_2,\cdots,b_k\}$ 的 Gram-Schmidt 化是 $\tilde{\boldsymbol{B}}=\{\tilde{b}_1,\tilde{b}_2,\cdots,\tilde{b}_k\}$,其中,任意一个向量 $b_i$ 都正交于 $\mathrm{span}\{b_1,b_2,\cdots,b_{i-1}\}$。一个向量 $p=(a_0,a_1,\cdots,a_{n-1})$ 生成的反循环矩阵 $\boldsymbol{A}(p)$形式如下:

$$\boldsymbol{A}(\boldsymbol{P})=\begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ -a_{n-1} & a_0 & \cdots & a_{n-2} \\ \cdots & \cdots & & \cdots \\ -a_1 & -a_2 & \cdots & a_0 \end{bmatrix}。$$

一个对称矩阵 $\boldsymbol{\Sigma}\in\mathbb{R}^{m\times m}$是正定的,记作 $\boldsymbol{\Sigma}>0$(或 $\boldsymbol{\Sigma}\geq 0$),如果对所有的非零 $x\in\mathbb{R}^m$,有 $x^t\boldsymbol{\Sigma}x>0$(或 $x^t\boldsymbol{\Sigma}x\geqslant 0$)。如果 $\boldsymbol{\Sigma}=\boldsymbol{B}\boldsymbol{B}^t$,$\boldsymbol{B}$ 为 $\boldsymbol{\Sigma}\geqslant 0$ 的平方根,记作$\sqrt{\boldsymbol{\Sigma}}=\boldsymbol{B}$。对于任意标量 $s$,$\boldsymbol{\Sigma}>s$ 表示 $\boldsymbol{\Sigma}$ 的所有特征值都严格大于 $s$,用 $s_1(\boldsymbol{B})=\max_u\|\boldsymbol{B}_u\|$ 表示 $\boldsymbol{B}$ 的最大奇异值,其中,最大值取决于单位向量 $u\in\mathbb{R}^m$。

任意 $n\times n$ 矩阵 $S$ 和非空索引集合 $\boldsymbol{I},\boldsymbol{J}\subset\{1,\cdots,n\}$,对于 $\boldsymbol{S}$ 中选择位置 $\boldsymbol{S}[I,j]\subset\boldsymbol{I}\times\boldsymbol{J}$ 的元素所得到的方阵,记作 $\boldsymbol{S}[\boldsymbol{I},\ \boldsymbol{J}]$,如果 $\boldsymbol{I}=\boldsymbol{J}$,简写为 $\boldsymbol{S}[\boldsymbol{I}]$。对于任意非奇异矩阵 $\boldsymbol{S}\in\mathbb{R}^{n*n}$,并且索引集合满足 $\boldsymbol{I}\cup\bar{\boldsymbol{I}}=\{1,\cdots,n\}$,$\boldsymbol{I}\cap\bar{\boldsymbol{I}}=\varnothing$,称

$$\boldsymbol{S}/\boldsymbol{I}=\boldsymbol{S}[\boldsymbol{I}]-\boldsymbol{S}[\boldsymbol{I},\bar{\boldsymbol{I}}]\boldsymbol{S}[\bar{\boldsymbol{I}}]^{-1}\boldsymbol{S}[\bar{\boldsymbol{I}},\boldsymbol{I}]$$

是 $\boldsymbol{S}[\bar{\boldsymbol{I}}]$的舒尔补。特别地,如果 $\boldsymbol{S}=\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B}^t & \boldsymbol{D}\end{bmatrix}$,那么 $\boldsymbol{A}$ 的舒尔补就是矩阵就是 $\boldsymbol{S}/\boldsymbol{A}=\boldsymbol{D}-\boldsymbol{B}^t\boldsymbol{A}^{-1}\boldsymbol{B}$。对于任意的索引集 $\boldsymbol{I}$,对称矩阵 $\boldsymbol{S}$ 是正定的当且仅当 $\boldsymbol{S}[\boldsymbol{I}]$和 $\boldsymbol{S}/\boldsymbol{S}[\boldsymbol{I}]$都是正定的。

### 1.2　格

**定义 1(格)**　设 $H=R_m$,格是 $H$ 的离散加法子群。对于基 $\boldsymbol{B}=\{b_1,b_2,\cdots,b_n\}\in H_n$,记作 $\mathcal{L}(\boldsymbol{B})$并且称 $\boldsymbol{B}$ 生成的格为向量集

$$\left\{\sum_{i=1}^{n}x_ib_i\mid x_i\in\mathbb{Z}\right\}$$

一个格通常记作 $\Lambda$,或者$\mathcal{L}(\boldsymbol{B})$,当提供基 $\boldsymbol{B}$ 时。

**定义 2(格的行列式)**　设 $\Lambda=\mathcal{L}(\boldsymbol{B})$是秩为 $n$ 的格。记 $\det(\Lambda)$ 为 $\Lambda$ 的行列式,其值为 $\sqrt{\det(\boldsymbol{B}\boldsymbol{B}^t)}$。当 $\boldsymbol{B}$ 是一个方阵,直接的,$\det(\Lambda)=|\det(\boldsymbol{B})|$。

可以证明格与基的选取无关。设 $\Lambda$ 上的任意两组基 $\boldsymbol{B}_1,\boldsymbol{B}_2$,则存在一个单模矩阵 $\boldsymbol{U}$,使等式 $\boldsymbol{B}_1=\boldsymbol{U}\boldsymbol{B}_2$ 成立。

#### 1.2.1　格上困难问题

为了格理论的完整性,首先回顾格上两类重要的困难性问题。

**定义 3(SVP-Shortest Vector Problem)**　给定一个 $n$ 维格 $\Lambda$,找到一个格向量使得 $\|v\|=\lambda_1(\Lambda)$

**定义 4(CVP-Closest Vector Problem)**　给定一个 $n$ 维格 $\Lambda$ 和一个点 $c\in H$,找到一个格向量使得 $\|c-v\|=\mathrm{dist}(c,\lambda)=\min_{z\in\Lambda}(\|c-z\|)$

上述的两类格上困难问题,在目前的现代计算机中很难解决,目前还不清楚在量子计算机中是否存在能显著提高效率的算法,与现在主流的密码体系基于大整数分解和离散对数问题形成鲜明对比,后者已知在量子计算机下存在(概率)多项式时间内解决。

#### 1.2.2　NTRU 格

NTRU 格是密码学中广泛使用的一类特殊格,因为它们的环形结构允许在对格执行常数时间上获得一个因子 $n$,从而生成高效且紧凑的密码系统。

**定义 5(NTRU 格)**　令 $f,g,F,G\in\mathbb{Z}^m$使得,

$$fG-gF=q\ \mathrm{mod}\ (x^N+1),$$

那么 $f,g,F,G$ 生成的 NTRU 矩阵就是行生成的分块矩阵,

$$\boldsymbol{B}_{f,g,F,G}=\begin{pmatrix}\mathcal{A}(f) & \mathcal{A}(g)\\ \mathcal{A}(F) & \mathcal{A}(G)\end{pmatrix},$$

**引理 1**　令 $f,g,F,G$ 满足定,$h=gf^{-1}\bmod q$,那

么 $\boldsymbol{A}_{h,q}=\begin{pmatrix}-\mathcal{A}(h) & \boldsymbol{I}_N\\ q\boldsymbol{I}_N & \boldsymbol{O}_N\end{pmatrix}$,$\boldsymbol{B}_{f,g}=\begin{pmatrix}\mathcal{A}(f) & \mathcal{A}(g)\\ \mathcal{A}(F) & \mathcal{A}(G)\end{pmatrix}$生成相同的格

$$\Lambda_{h,q}=\{u,v\in\mathbb{Z}^N \mid u+vh=0 \bmod q\}。$$

如果 $N,q$ 足够大并且 $f,g$ 的生成满足条件,则 $h$ 与随机向量计算不可区分,并且 $f,g$ 难以通过 $h$ 恢复,故 $\boldsymbol{B}_{f,g}$ 和 $\boldsymbol{A}_{h,q}$ 成为了公钥密码体系的基本原语(作为私钥和公钥)。

**引理2** 令 $\boldsymbol{B}_{f,g,F,G}$ 是 NTRU 格,则 $\boldsymbol{B}_{f,g,F,G}$ 的最大奇异值为

$$s_1(\boldsymbol{B}_{f,g})=\sqrt{\max_{\omega\in\Omega_m}(\lambda_\omega)},$$

其中,$C(x)\triangleq(ff^*+gg^*+FF^*+GG^*)(x)$ 并且 $\lambda_\omega\triangleq\frac{1}{2}(C(\omega)+(-1)^i\sqrt{C^2(\omega)-4q^2})$。

### 1.3 高斯分布

$m$ 维高斯函数 $\rho:\mathbb{R}^m\to(0,1)$,定义为 $\rho(x)=\exp(-\pi\cdot(x,x))$ 由一个可逆矩阵 $\boldsymbol{B}$ 得到一个线性变换得到

$$\rho_B(x)=\rho(\boldsymbol{B}^{-1}\cdot x)=\exp(-\pi\cdot x^t\Sigma x),$$

其中,$\Sigma=\boldsymbol{B}\boldsymbol{B}^t$,对任意的 $c\in\operatorname{span}\{B\}$,$\rho_{\sqrt{\Sigma}}$ 以 $c$ 为中心的高斯函数定义为 $\rho_{\sqrt{\Sigma},c}(x)=\rho_{\sqrt{\Sigma}}(x-c)$,格上归一化后,得到协方差为 $\frac{\Sigma}{2\pi}$ 的连续高斯分布 $\mathcal{D}_{\Lambda,\sqrt{\Sigma},c}$。将分布限制在离散的结构上,得到离散的高斯分布,对任意的 $x\in\Lambda$,

$$\mathcal{D}_{\Lambda,\sqrt{\Sigma},c}=\frac{\rho_{\sqrt{\Sigma},c}(x)}{\rho_{\sqrt{\Sigma},c}(\Lambda)}。$$

**定义6(光滑参数)** 对任意 $\varepsilon>0$,格 $\Lambda$ 的光滑参数为最小的 $s>0$,使得 $\rho(s\cdot\Lambda^*)\leqslant(1+\varepsilon)$。

**引理3** 令 $\Sigma>0$,格 $\Lambda\subset\operatorname{span}(\Sigma)$,记 $\sqrt{\Sigma}\geqslant\eta_\epsilon(\Lambda)$,如果 $\rho_{\sqrt{\Sigma^{-1}}}(\Lambda^*)\leqslant(1+\epsilon)$。

**引理4** 令 $\Lambda\subset\mathbb{R}^m$ 是基为 $\boldsymbol{B}$ 的格,$\epsilon>0$,则

$$\eta_\epsilon\leqslant\|\widetilde{\boldsymbol{B}}\|\cdot\sqrt{\frac{\ln\left(2m\left(1+\frac{1}{\epsilon}\right)\right)}{\pi}}。$$

**引理5** $\boldsymbol{B}$ 为 $m$ 维格的基,对可忽略的函数 $\varepsilon(m)$,令 $s\geqslant\eta_\varepsilon(\Lambda)$,则 $\Pr_{x\leftarrow D_{\Lambda,s}}[\|x\|\geqslant s\cdot\sqrt{m}]\leqslant negl(m)$。

**引理6** 令 $v\in\mathbb{Z}^m$ 是一个向量,其中,$\|v\|^m\leqslant s\sqrt{m}$,存储该向量所需的最大位数的上界为 $m(1+\lceil\log s\rceil)$。

### 1.4 基于 Peikert 的高斯采样算法

Peikert 算法由 Peikert 在 2010 年提出,其主要的思想是利用 Roundoff 算法求解 CVP 问题进行采样,但此方法产生的采样结果会产生一个和基相关的协方差矩阵 $\boldsymbol{\Sigma}_1=\sigma^2\boldsymbol{B}^t\boldsymbol{B}$,从而泄露陷门信息。为了保证安全性,Peikert 提出了预采样一个非球型高斯分布拥有协方差 $\boldsymbol{\Sigma}_2=s^2\boldsymbol{I}-\sigma^2\boldsymbol{B}^t\boldsymbol{B}$,再和 Roundoff 算法输出的结果相卷积最后得到不泄露陷门信息的球形高斯分布 $\boldsymbol{\Sigma}=s^2\boldsymbol{I}$。卷积的正确性,简单来说,如果 $X$ 和 $Y$ 是两个独立的随机变量,则它们之和 $X+Y$ 的概率分布是它们各自分布的卷积。此外,对于连续(不一定是球型)高斯分布,协方差矩阵在卷积下是可加的。剩下的问题是对参数 $s$ 的选择,对任何(非退化)高斯的协方差矩阵都是对称正定的,即它的所有特征值都是正的。相反,每个正定矩阵都是某个高斯分布的协方差,可以通过计算协方差矩阵的"平方根"对其进行有效采样。因此,充分必要条件是 $\boldsymbol{\Sigma}_1$ 的所有特征值都小于 $s^2$。同时,该算法适用于超过给定基 $\boldsymbol{B}$ 的最大奇异值的任何 $s$。

Algorithm 1: PeiSample($\boldsymbol{B}_1,\sigma,E,c$)

Input: $\boldsymbol{B}_1\in\mathbb{Z}_q^{2N\times 2N}$ 是格 $\Lambda(\boldsymbol{B}_1)$ 的基矩阵,舍入参数 $\sigma=\omega(\sqrt{\log n})$,正定协方差矩阵

$\boldsymbol{\Sigma}=s^2\boldsymbol{I}>\boldsymbol{\Sigma}_1=\sigma^2\boldsymbol{B}_1\boldsymbol{B}_1^t$,向量 $c\in\mathbb{R}^n$

Output: 向量 $x\in\Lambda+c$,与在 $\mathcal{D}_{\Lambda+c,\sqrt{\Sigma}}$ 分布有 $negl(n)$ 的误差

1 令 $\boldsymbol{\Sigma}_2=\boldsymbol{\Sigma}-\boldsymbol{\Sigma}_1>0$ 并且计算 $\boldsymbol{B}_2=\sqrt{\boldsymbol{\Sigma}_2}$;

2 选择一个向量 $x_2\leftarrow\mathcal{D}_{\sqrt{\Sigma_2}}$,即 $x_2\leftarrow\boldsymbol{B}_2\cdot\mathcal{D}_1$;

3 返回 $x\leftarrow c-\boldsymbol{B}_1\lfloor\boldsymbol{B}_1^{-1}(c-x_2)\rceil$;

算法的正确性由下面的定理保证。

**定理1** 令 $\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2>0$ 都是正定矩阵,$\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_1+\boldsymbol{\Sigma}_2>0$ 并且 $\boldsymbol{\Sigma}_3^{-1}=\boldsymbol{\Sigma}_1^{-1}+\boldsymbol{\Sigma}_2^{-1}>0$。令 $\Lambda_1$,$\Lambda_2$ 为两个格使得对任意的正整数 $\epsilon\geqslant\frac{1}{2}$,$\sqrt{\boldsymbol{\Sigma}_1}\geqslant\eta_\epsilon(\Lambda_1)$,$\sqrt{\boldsymbol{\Sigma}_2}\geqslant\eta_\epsilon(\Lambda_2)$,并且令任意的 $c_1,c_2\in R_n$,对以下的概率实验

$$x \leftarrow \mathcal{D}_{\Lambda_2+c_2,\sqrt{\Sigma_2}},$$

$$x_1 \leftarrow x_2 + \mathcal{D}_{\Lambda_2+c_1-x_2,\sqrt{\Sigma_1}},$$

$x_1$ 的边缘分布与$\mathcal{D}_{\Lambda_1+c_1,\sqrt{\Sigma}}$有 $8\epsilon$ 距离,此外,对任意的 $x_1 \in \Lambda_1 + c_1$,对给定 $x_1 = \bar{x}_1$ 的 $x_2 \in \Lambda_2 + c_2$条件分布与 $c_3 + \mathcal{D}_{\Lambda_2+c_2-x_3,\sqrt{\Sigma_3}}$有 $2\epsilon$ 的统计距离,其中,$\Sigma_3^{-1}c_3 = \Sigma_1^{-1}\bar{x}_1$。

## 2　基于非球面的 Peikert 采样优化

### 2.1　非球面采样优化后的 Peikert 算法

2022 年,Jia 等[12]提出了一种针对 MP12 采样框架的优化,此研究专注于(ring－)LWE 和 (ring－)SIS 假设下改进随机预言机模型中的哈希和签名。本文运用非球型高斯技术对 Peikert 采样算法进行优化。具体来说,将原始的离散球形高斯分布推广到非球型高斯分布,且不会泄漏有关陷门的任何信息。该思路适用于精确的陷门设置和近似的陷门设置。Peikert 算法同样需要陷门设置,对陷门矩阵进行改造,使其在满足算法正确性的同时,最大程度地提高安全性或者最大程度降低存储空间的大小。Algorithm 2 是优化后的非球型离散高斯采样算法框架。

Algorithm 2: PeiSample($B_1, \sigma, E, c$)

Input: $\boldsymbol{B}_1 \in \mathbb{Z}_q^{2N\times 2N}$是格 $\Lambda(B_1)$ 的基矩阵,舍入参数 $\sigma = \omega(\sqrt{logn})$,正定协方差矩阵

$$\Sigma = \begin{pmatrix} \gamma_1^2 & 0 \\ 0 & \gamma_2^2 \end{pmatrix} > \Sigma_1 = \sigma^2 \boldsymbol{B}_1 \boldsymbol{B}_1^t,\text{向量 } c \in \mathbb{R}^n$$

Output:向量 $x \in \Lambda + c$,与在$\mathcal{D}_{\Lambda+c,\sqrt{\Sigma}}$分布有 $negl(n)$的误差

1 令 $\Sigma_2 = \Sigma - \Sigma_1 > 0$ 并且计算 $\boldsymbol{B}_2 = \sqrt{\Sigma_2}$;

2 选择一个向量 $x_2 \leftarrow \mathcal{D}_{\sqrt{\Sigma_2}}$,即 $x_2 \leftarrow \boldsymbol{B}_2 \cdot \mathcal{D}_1$;

3 返回 $x \leftarrow c - \boldsymbol{B}_1 \lfloor \boldsymbol{B}_1^{-1}(c - x_2) \rceil$;

为了满足 Algorithm 2 的正确性 $\Sigma_1, \Sigma_2$ 需是正定矩阵。

$$\Sigma_2 = \begin{pmatrix} \gamma_1^2 - \sigma^2(ff^t + gg^t) & -\sigma^2(fF^t + gG^t) \\ -\sigma^2(Ff^t + Gg^t) & \gamma_2^2 - \sigma^2(FF^t + GG^t) \end{pmatrix}。$$

根据引理 $\Sigma_2$ 是正定矩阵当且仅当 $\Sigma[I]$和 $\Sigma/\Sigma[I]$都是正定矩阵。即

$$\Sigma[I] = \gamma_2^2 - \sigma^2(FF^t + GG^t) > 0,$$

$$\gamma_2 > \sigma s_1(F,G),$$

$$\frac{\Sigma}{\Sigma[I]} = \gamma_1^2 I - \sigma^2(ff^t + gg^t) - H^t \cdot (\gamma_2^2 I - \sigma^2(FF^t + GG^t))^{-1} \cdot H > 0 =$$

$$\gamma_1^2 I - \frac{\sigma^2(ff^t + gg^t)(\gamma_2^2 - \sigma^2(FF^t + GG^t)) + \sigma^4(fF^t + gG^t)(Ff^t + Gg^t)}{\gamma_2^2 I - \sigma^2(FF^t + GG^t)} =$$

$$\gamma_1^2 I - \frac{\sigma^2\gamma_2^2(ff^t + gg^t) - (fG - gF)(fG - gF)^t\sigma^4}{\gamma_2^2 I - \sigma^2(FF^t + GG^t)},$$

$$\gamma_1 \geqslant \frac{\sqrt{\sigma^2\gamma_2^2 s_1^2(f,g) - \sigma^4 q^2}}{\sqrt{\gamma_2^2 - \sigma^2 s_1^2(F,G)}}。$$

### 2.2　签名算法

利用优化后的原像采用算法,可以构造 Algorithm3－5 的签名方案。首先,令$\mathcal{D}$是$\mathbb{Z}$上的某种分布,$f,g$ 的每一项都是在$\mathcal{D}$中的采样,再计算满足条件 $fG - gF = q \mod (x^N + 1)$ 的 $F, G$。构造公钥为 $h = gf^{-1} \bmod q$,构造私钥为

$$\boldsymbol{B}_{f,g,F,G} = \begin{pmatrix} \mathcal{A}(f) & \mathcal{A}(g) \\ \mathcal{A}(F) & \mathcal{A}(G) \end{pmatrix}。$$

签名时,对明文进行一次哈希运算生成 $t$,再将 $t$ 嵌入到格中和采样算法的结果进行减法运算,最终生成的 $s_1$ 和 $s_2$满足 $s_1 + s_2\, h = t$。而在传输过程中只需要 $s_2$,在验证过程中恢复 $s_1$,在满足 $\|(s_1, s_2)\| \leqslant 1.04\sigma\sqrt{N}$时接受,反之拒绝。

Algorithm 3: KeyGen($N, q$)

Input: $N$, $q$

*Output*: 私钥 $\boldsymbol{B}_{f,g,F,G} \in \mathbb{Z}_q^{2N\times 2N}$,和公钥 $\boldsymbol{h} \in \mathbb{Z}_q$

1 $f$, $g \leftarrow \mathcal{D}$其中$\mathcal{D}$是$\mathbb{Z}$上某种分布; 计算 $F, G \in \mathbb{Z}$使得 $fG - gF = q\ mod\ (x^N + I)$;

2 $\boldsymbol{h} = gf^{-1} \bmod q$;

3 $\boldsymbol{B}_{f,g,F,G} = \begin{pmatrix} \mathcal{A}(f) & \mathcal{A}(g) \\ \mathcal{A}(F) & \mathcal{A}(G) \end{pmatrix}$;

4 返回 $SK \leftarrow \boldsymbol{B}_{f,g,F,G}$, PK$\leftarrow \boldsymbol{h}$;

Algorithm 4: Signature($\boldsymbol{B}$,$m$)

Input: 私钥 $\boldsymbol{B}\in\mathbb{Z}_q^{2N\times 2N}$,哈希函数:$H:\{0,1\}^*\to\mathcal{Z}_q$,标准差 $\sigma$,明文 $\boldsymbol{m}$

Output: 签名 $s_2\in\mathbb{Z}_q$

1 $t\leftarrow H(m)\in Z_q$;

2 $(s_1,s_2)=(t,0)-\text{PeiSample}(\boldsymbol{B},\sigma,Q(t,0))$;

//$s_1+s_2h=t$

3 返回 $s_2$;

## 2.3 优化方案的参数

模式1采用椭圆采样优化后,可以提高安全性。采样算法的输出质量和生成的高斯参数有关,使得敌手很难在更小的范围内找到满足条件的格点,原方案基于NTRU格上的原像采样上界为 $s\sqrt{2n}$,而椭圆采样优化后的原像采样上界缩小至 $\sqrt{n(\gamma_1^2+\gamma_2^2)}$,调整 $\gamma_1$, $\gamma_2$ 的取值在保证算法正确性的同时,最大程度上提高采样算法的安全性;模式2采用椭圆采样优化后,可以降低预采样的存储大小。由于本文采用NTRU格结构,它是对短向量$(s_1,s_2)$进行采样的陷门,使得对于任何 $t$,满足$s_1+s_2h=t$,而不泄露关于其自身的任何信息。在已知私钥 $h$ 和 $t$ 下,只需要存储较小的 $s_2$在验证的时候可以通过计算恢复向量$(s_1,s_2)$。而 $s_2$ 的存储大小和预采样的基矩阵相关,对原方案存储大小为 $n(1+\log(s))$,优化后方案的存储大小为 $n(1+\log(\gamma_2))$。

Algorithm 5: Verify($\boldsymbol{h}$,$m$,$s_2$)

Input: 私钥 $\boldsymbol{h}\in\mathbb{Z}_q$,哈希函数:$H:\{0,1\}^*\to\mathbb{Z}_q$,明文 $\boldsymbol{m}$

Output: 签名 or 拒绝

1 $t\leftarrow H(m)\in Z_q$;

2 $s_1=t-s_2h$;

3 if $\|(s_1,s_2)\|\leqslant 1.04\sigma\sqrt{N}$, then

4 | 接受;

5 else

6 | 拒绝;

7 end

**模式1** 为了满足算法的正确性,令 $\gamma_1\geqslant\frac{\sqrt{\sigma^2\gamma_2^2s_1^2(f,g)-\sigma^4q^2}}{\sqrt{\gamma_2^2-\sigma^2s_1^2(F,G)}}$,使得 $\|l_2\|=\sqrt{\gamma_1^2n+\gamma_2^2n}$,最小可得

$$l_2\geqslant\sqrt{n}\cdot\sqrt{\frac{a\gamma_2^2-c}{\gamma_2^2-b}+\gamma_2^2}=\sqrt{n}\cdot\sqrt{a+b+\frac{a\cdot b-c}{\gamma_2^2-b}+(\gamma_2^2-b)},$$

其中, $a=\sigma^2s_1^2(f,g)$,$b=\sigma^2s_1^2(F,G)$,$c=\sigma^4q^2$

$l_2$ 取得最小值

$$l_2=\sqrt{n(\sqrt{\sigma^4s_1^2(F,G)s_1^2(f,g)-\sigma^4q^2}+\sigma^2s_1^2(f,g)+\sqrt{\sigma^4s_1^2(F,G)s_1^2(f,g)-\sigma^4q^2}+\sigma^2s_1^2(F,G))}$$

当且仅当 $\frac{a\cdot b-c}{\gamma_2^2-b}=\gamma_2^2-b$,故

$$\gamma_1^2=\sqrt{\sigma^4s_1^2(F,G)s_1^2(f,g)-\sigma^4q^2}+\sigma^2s_1^2(f,g),$$

$$\gamma_2^2=\sqrt{\sigma^4s_1^2(F,G)s_1^2(f,g)-\sigma^4q^2}+\sigma^2s_1^2(F,G)。$$

**模式2** 令签名的安全性不变,即 $\sqrt{2n\cdot s^2}=\sqrt{n\cdot(\gamma_1^2+\gamma_2^2)}$,并且满足算法的正确性 $\gamma_1\geqslant\frac{\sqrt{\sigma^2\gamma_2^2s_1^2(f,g)-\sigma^4q^2}}{\sqrt{\gamma_2^2-\sigma^2s_1^2(F,G)}}$,使得存储空间 $n(1+\log(\gamma_1))$最小,故

$$\gamma_1^2\geqslant\frac{\sigma^2(2s^2-\gamma_1^2)s_1^2(f,g)-\sigma^4q^2}{2s^2-\gamma_1^2-\sigma^2s_1^2(F,G)},$$

即

$$\gamma_1^4+(s_1^2(F,G)\sigma^2-sZ_1^2(f,g)\sigma^2-2s^2)*\gamma_1^2+2s^2\sigma^2s_1^2(f,g)-\sigma^4q^2\leqslant 0,$$

则存储空间最小需满足:

$$\gamma_1^2=\frac{1}{2}(-b-\sqrt{b^2-8s^2\sigma^2s_1^2(f,g)+4\sigma^4q^2})。$$

此时的存储空间为

$$n\left(\log\left(\sqrt{\frac{1}{2}[-b-\sqrt{b^2-8s^2\sigma^2s_1^2(f,g)+4\sigma^4q^2}]}\right)+1\right),$$

其中, $b=s_1^2(F,G)\sigma^2-s_1^2(f,g)\sigma^2-2s^2$。

## 2.4 优化后方案的结果与安全性分析

表1是模式1和模式2优化后与原方案在原像采样上界和存储空间的比较。

**表 1　模式 1、模式 2 与原方案的比较**

Table 1　Comparison of mode 1, mode 2 and the original

| 参数 | $q=2\,247, \sigma=0.707, n=512$ | | | $q=2\,247, \sigma=0.707, n=1\,024$ | | |
|---|---|---|---|---|---|---|
| | 模式 1 | 模式 2 | 原方案 | 模式 1 | 模式 2 | 原方案 |
| $\gamma_1$ | 126.56 | 37.63 | – | 193.98 | 54.97 | – |
| $\gamma_2$ | 491.02 | 949.96 | – | 770.62 | 1 495.43 | – |
| $s$ | – | – | 410.43 | – | – | 743.62 |
| $\|x_i\|$ | 11 485.39 | 21 513.88 | 21 513.88 | 25 433.65 | 47 887.01 | 47 887.01 |
| storage | 4.07 | 3.19 | 4.82 | 8.80 | 6.94 | 9.48 |
| Forg. | 74.91 | 62.19 | 62.19 | 158.19 | 133.51 | 133.51 |
| Key Recov. | 94.87 | 94.87 | 94.87 | 218.62 | 218.62 | 218.62 |

2.4.1　参数选取

本文利用 pycharm 平台模拟这种离散高斯采样表现本方案优化的效果。利用均值为 0,标准差为$\frac{1}{\sqrt{2}}$的分布对 $n$ 维 $f,g$ 的每项进行采样,实验表明,此时采样范围更小,安全性更高。再利用 NTRU 格的结构,求解出 $F,G$,满足 $fG-gF=q(\bmod x^N+1)$,为了保证算法的正确性,$f,g$ 的采样结果应令 $s_1(\boldsymbol{B}_{f,g,F,G})<1.15N^{0.25}\sqrt{\log N}\sqrt{q}$,否则将重新选取,保证特征值的大小是为了防止采样的结果过大,伪造签名的概率提高,使得安全性降低。最后代入上述计算最大奇异值的结果得出 $\gamma_1$、$\gamma_2$、原像采样的上界 $l_2$ 和存储空间 storage。本文选择参数 $q=2\,477$、$\sigma=0.707$ 来模拟实际结果,当 $q$ 的选取过大时,计算复杂度提升的同时,采样空间也会变大,敌手更容易伪造签名,安全性会降低;过小的 $q$ 取值会令敌手更容易破解密钥。实验表明,当 $q=2\,477$ 时,优化的算法有最好的结果。$\sigma$ 的选取也会影响采样的结果,选取 $\sigma=0.707$,令采样的结果大概率落入正确的范围内,降低了敌手攻击成功的可能性。

2.4.2　算法优化结果

在模式 1 中,希望得到最小的原像采样空间,以提高伪造签名攻击的安全性,当 $n=512$ 时,原方案的 $s=410.43$,原像采样大小的上界为 21 513.88,优化后 $\gamma_1=126.56$、$\gamma_2=491.02$,上界减少至 11 485.39,上界缩小至原来的 53%;当 $n=1\,024$ 时,原方案的 $s=743.62$,原像采样大小上界为 47 887.01,优化后 $\gamma_1=193.98$、$\gamma_2=770.62$,上界减少至 25 433.65,上界缩小至原来的 53%。除此之外,存储空间相较于原方案也有所降低。在模式 2 中,希望在保证算法正确性的同时占用最小的存储空间。当 $n=512$ 时,原方案的 $s=410.43$,存储空间为 4.82 kb,优化后,$\gamma_1=37.63$,$\gamma_2=949.96$,存储空间降低至 3.19 kb,而伪造签名攻击下的安全性几乎没有变化;当 $n=1\,024$ 时,原方案的 $s=743.62$,存储空间为 9.48 kb,优化后 $\gamma_1=54.97$,$\gamma_2=1\,495.43$,存储空间降低至 6.94 kb,而伪造签名攻击下的安全性几乎没有变化。

2.4.3　安全性分析

最后需要评估优化后算法的安全性,本文使用常用的密码分析法,即借助密钥恢复和签名伪造最佳攻击的复杂性验证算法的安全性。遵循几何级数假设(GSA),断言一个优化基的 Gram-Schmidt 向量的范数随着几何衰减而减小。进一步说,利用自对偶 BKZ 归约算法评估本方案的安全性,第一种是评估求解 RSIS 问题的困难性,即伪造签名问题的安全性;第二种是评估公钥的随机性,即评估密钥恢复的安全性。伪造签名问题的安全性是由原像采样大小的上界决定,而密钥恢复攻击的安全性是由参数 $n$、$q$,以及标准差 $\tau$ 决定。BKZ 算法是通过估计 SVP-Oracle 的时间复杂度来评估安全性的,在问题 $ISIS_{n,q,\beta}$ 中,求解一个向量,其 $l_2$ 范数为 $2^{2\sqrt{n\log q\log\delta}}$,因此令采样 log 向量 $\beta=\sqrt{\gamma_1^2\,n+{\gamma_2}^2 n}=2^{2\sqrt{n\log q\log\delta}}$,求解 $\delta$。再根据 $\delta=\left(\frac{\kappa}{2\pi\mathrm{e}}(\pi\kappa)^{\frac{1}{\kappa}}\right)^{\frac{1}{2(\kappa-1)}}$ 求 $\kappa$,最后算 $2^{0.292\kappa}$ 的值,就是伪造签名攻击下的安全性参数结果,值越高代表安

全性越高。利用计算机模拟结果,选择之前的参数,当 $n=512$ 时,原方案的伪造签名攻击(见表1中的 Forg.)下安全性为62.19,模式1安全性提高至74.91,模式2安全性保持62.19;当 $n=1\ 024$ 时,原方案的伪造签名攻击下安全性为136.42,模式1安全性提高至158.19,模式2的安全性保持136.42;在密钥恢复攻击(见表1中的 Key Recov.)下,当 $n=512$ 时,安全性为94.87;当 $n=1\ 024$ 时,安全性为218.62,密钥恢复攻击的安全性没有变化。本文对算法的优化仅涉及降低原像采样的上界,故在伪造签名攻击下的安全性得到提高,而密钥恢复攻击下的安全性与高斯采样的参数有关,故此安全性保持不变。

## 3 总 结

本文优化了 Peikert 采样算法,通过将球型高斯采样修改为非球型高斯采样,而同时不泄露陷门信息。为满足不同需求,提出了两种优化模式,模式1为提高伪造签名攻击的安全性,最终将高斯采样的上界缩小至原来的53%,签名算法的安全性提高18%;模式2为减少存储空间大小,最终将采样的存储空间减少至原来的66%。此外,本文中对采样的优化方法可以应用至所有以 Peikert 采样算法为基础的密码方案中,以提高签名算法的安全性或采用的减少存储空间。

**参考文献:**


[1] Shor P. 35th Annual Symposium on Foundations of Computer Science, November 20-22, 1994[C]. Washington: IEEE Computer Society Santa Fe,1994.

[2] Gentry C,Peikert C,Vaikuntanathan V. STOC '08: Symposium on Theory of Computing, May 17-20, 2008[C]. Victoria British Columbia Canada: Association for Computing Machinery, 2008.

[3] Klein P. Symposium on Discrete Algorithms, January 09-11 2000[C]. San Francisco:Society for Industrial and Applied Mathematics, 2000.

[4] Peikert C. 30th Annual Cryptology Conference, August 15-19, 2010[C]. Santa Barbara:Springer, 2010.

[5] Micciancio D,Peikert C. 31st Annual International Conference on the Theory and Applications of Cryptographic Techniques, April 15-19, 2012[C]. Cambridge:Springer,2012.

[6] Ducas L, Durmus A,Lepoint T,et al. 33rd Annual Cryptology Conference, August 18-22,2013[C]. Santa Barbara:Springer, 2013.

[7] Ducas L,Lyubashevsky V,Prest T. 20th International Conference on the Theory and Application of Cryptology and Information Security, December 07-11,2014[C]. Kaoshiung:Springer,2014.

[8] Ducas L,Prest T. ISSAC'16: International Symposium on Symbolic and Algebraic Computation, July 20-22, 2016[C]. Waterloo: Association for Computing Machinery,2016.

[9] Pierre-Alain F,Jeffrey H, Paul K, et al. Falcon: Fast-fourier lattice-based compact signatures over ntru[J]. Submission to the NIST's Post-quantum Cryptography Standardization Process, 2018, 36(5):31-41.

[10] Espitau T. ACM Asia Conference on Computer and Communications Security, June 7, 2021[C]. Hong Kong: Association for Computing Machinery,2021.

[11] Espitau T,Tibouchi M,Wallet A,et al. 42nd Annual International Cryptology Conference, August 15-18, 2022[C]. Santa Barbara:Springer,2022.

[12] Jia H, Hu Y, Tang C. Lattice-based hash-and-sign signatures using approximate trapdoor, revisited[J]. IET Information Security, 2021,16(1):41-50.


【责任编辑:陈 钢】