

半均匀 LWE 问题的紧致归约^{*}

王洋^{1,2}, 王明强¹

¹(山东大学 数学学院, 山东 济南 250100)

²(密码科学技术技术全国重点实验室, 北京市 5159 信箱, 100878)

通讯作者: 王明强, E-mail: wangmingqiang@sdu.edu.cn



摘要: 在部分实用化的格密码协议设计和应用中, 需要用到公开矩阵服从特定分布的、非均匀的 LWE 问题的困难性来证明相应密码体制的安全性. 最近, 贾文娟等人给出了半均匀 LWE 问题的具体定义, 并采用类似证明熵 LWE 问题困难性的归约路线证明了欧式格/理想格/模格上半均匀 LWE 问题的困难性. 但是, 贾文娟等人的归约方法 (在维数和误差分布的高斯参数等方面) 会引入较大的归约损失. 同时需要引入额外的、非标准的困难性假设来证明环上的半均匀 LWE 问题的困难性. 利用 Hint-LWE 问题困难性的归约技巧, 给出了半均匀 LWE 问题困难性更紧致的归约. 采用的归约方法几乎不受代数结构的影响, 可以统一地应用到欧式格/理想格/模格上定义的非均匀 LWE 问题. 可以基于标准的 LWE 假设证明对应欧式格/理想格/模格上的半均匀 LWE 问题的困难性而无需引入任何额外的非标准困难性假设. 归约结果保持相应 LWE 问题的维数不变, 且归约过程中对应 LWE 问题的误差高斯参数的归约损失较小.

关键词: 基于格的密码学; 格中困难问题的归约; 半均匀 LWE 问题; Hint-LWE 问题; 离散高斯分布
中图法分类号: TP309

中文引用格式: 王洋, 王明强. 半均匀 LWE 问题的紧致归约. 软件学报. <http://www.jos.org.cn/1000-9825/7388.htm>

英文引用格式: Wang Y, Wang MQ. Tighter Reductions of LWE Problems with Semi-Uniform Seeds. Ruan Jian Xue Bao/Journal of Software (in Chinese). <http://www.jos.org.cn/1000-9825/7388.htm>

Tighter Reductions of LWE Problems with Semi-uniform Seeds

WANG Yang^{1,2}, WANG Ming-Qiang¹,

¹(School of Mathematics, Shandong University, Jinan 250100, China)

²(State Key Laboratory of Cryptology, P.O. Box 5159, Beijing, 100878, China)

Abstract: In certain designs and applications of practical lattice-based cryptography, one may need to use a special kind of LWE problems, in which the public matrix is sampled from some non-uniform distribution, to argue securities of corresponding schemes. Recently, Jia *et al.* formalized definitions of LWE problems with semi-uniform seeds, and proved corresponding hardness of Euclidean/ideal/module lattice-based LWE problems with semi-uniform seeds by using similar reduction roadmaps as those used to show the hardness of the so-called entropic LWE problems. However, Jia *et al.*'s reduction will introduce great losses in Gaussian parameters of errors and dimensions. Also, additional non-standard assumptions are needed to argue the hardness of semi-uniform LWE problems over rings. In this paper, we present a tighter reduction for LWE problems with semi-uniform seeds by modifying techniques used in proving the hardness of Hint-LWE problems. Our reductions are almost not affected by algebraic structures of corresponding problems, and could be applied to Euclidean/ideal/module lattice-based LWE problems with semi-uniform seeds uniformly. By using our methods, it is possible to prove the hardness of corresponding LWE problems with semi-uniform seeds based on standard LWE assumptions without introducing any additional non-standard assumption. Our reductions also keep the dimension of corresponding LWE problems unchanged. Meanwhile, the reduction loss of Gaussian parameters of corresponding LWE problems is relatively small.

* 基金项目: 国家重点研发计划项目(2021YFB3100200); 保密通信全国重点实验室稳定支持计划项目(2024, WD202402); 密码科学技术全国重点实验室开放课题(MMKFKT202207); 山东省自然科学基金(ZR2022QF039)

收稿时间: 2024-06-25; 修改时间: 2024-09-05; 采用时间: 2024-12-30; jos 在线出版时间: 2025-01-20

Key words: Lattice-Based Cryptography; Reductions of Lattice-Based Problems; Semi-Uniform LWE Problems; Hint-LWE Problems; Discrete Gaussian Distributions

欧式格 LWE (Learning with Errors Problems) 问题^[1]及其 (代数) 变体环/模 LWE 问题^[2,3]是目前格密码中应用最广泛的一类平均情形的数学困难问题, 基于欧式格/环/模 LWE 问题及其适当形式的变体问题几乎可以设计已知的所有密码原语. 到目前为止, 关于 LWE 问题实用化变体问题的研究可以粗略地分为三类 (更具体的分类讨论可参考文献 [4]). 第一类是研究 LWE 问题的秘密或者误差分布为更一般的分布 (例如 $\{0,1\}$ 上的二元分布, 小区间 $[-\alpha, \alpha]$ 上的均匀分布, 熵 LWE 问题等) 时, 对应 LWE 问题的困难性^[5-9]; 第二类是研究 LWE 问题的秘密或者误差存在部分泄露信息时, 对应的 LWE 问题 (即 Hint-LWE 问题) 的困难性^[5, 10-15]; 第三类是研究 LWE 问题的公开矩阵为非均匀分布时, 对应的 LWE 问题 (即非均匀 LWE 问题) 的困难性^[4, 16-18].

最近, 贾文娟等人^[4, 17-18]系统地研究了一类特殊的非均匀 LWE 问题, 即半均匀 LWE 问题, 的困难性, 给出了半均匀 LWE 问题的形式化定义, 并采用类似证明熵 LWE 问题困难性的归约路线证明了欧式格/理想格/模格上对应半均匀 LWE 问题的困难性. 以欧式格为例, 称 $\mathbb{Z}_q^{m \times d}$ 上的分布 \mathcal{D} 为 η -半均匀分布, 如果存在 $\mathbb{Z}^{m \times d}$ 上的一族可以有效取样的分布 $\{\phi_U\}_{U \in \mathbb{Z}_q^{m \times d}}$, 使得取样过程 $\{U + E_U: U \leftarrow U(\mathbb{Z}_q^{m \times d}), E_U \leftarrow \phi_U\}$ 所输出的分布与 \mathcal{D} 统计不可区分, 且以接近于 1 的概率有 E_U 的谱范数不超过 η 成立. 对应地, 可以定义判定版本的非均匀 LWE 问题 $DLWE_{\mathbb{Z}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 为区分如下分布:

- (1) $\{(A, A \cdot s + e \bmod q\mathbb{Z}): A \leftarrow \mathcal{D}, s \leftarrow \chi_2, e \leftarrow \chi_1\}$;
- (2) $\{(A, \mathbf{u}): A \leftarrow \mathcal{D}, \mathbf{u} \leftarrow U(\mathbb{Z}_q^m)\}$.

值得注意的是, 对任意的 $U \in \mathbb{Z}_q^{m \times d}$, 当选取 ϕ_U 为特定的固定函数 $E_U = \left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot U \right\rfloor \right\rfloor - U$ 时, 对应的半均匀 LWE 问题即为区分如下分布:

- (1) $\left\{ \left(\left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot A \right\rfloor \right\rfloor, \left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot A \right\rfloor \right\rfloor \cdot s + e \bmod q\mathbb{Z} \right) : A \leftarrow U(\mathbb{Z}_q^{m \times d}), s \leftarrow \chi_2, e \leftarrow \chi_1 \right\}$;
- (2) $\left\{ \left(\left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot A \right\rfloor \right\rfloor, \mathbf{u} \right) : A \leftarrow U(\mathbb{Z}_q^{m \times d}), \mathbf{u} \leftarrow U(\mathbb{Z}_q^m) \right\}$.

值得指出的是, 由 A 可以很容易地计算 $\left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot A \right\rfloor \right\rfloor$. 利用这里一类特殊的半均匀 LWE 问题可以证明 NIST (National Institute of Standards and Technology) 后量子竞赛第一轮的首选算法 Kyber 采用的基础公钥加密方案的 IND-CPA (indistinguishability under chosen plaintext attack) 安全性^[19-20]. 也就是说, 可以证明采用 LP 框架^[21]和压缩公钥的方式来设计的高效公钥加密方案也是满足 IND-CPA 安全的. 早期, 由于缺乏对应非均匀 LWE 问题到标准 LWE 问题的严格归约, Kyber 设计团队从 NIST 第二轮竞选开始取消了方案中对公钥进行压缩的设计方式^[22-24], 这导致对应方案的通信带宽的增加. 在实践应用中, 这一类特殊的半均匀 LWE 问题可以结合密钥共识等技术来设计非常高效的格密码方案^[25-29].

虽然已知的半均匀 LWE 问题困难性的归约结果适用于一般的秘密分布 (即归约中仅要求秘密 s 的分布在特定条件下拥有足够的熵即可, s 所服从分布的具体形式可以不定). 但是由于采用的是类似证明熵 LWE 问题困难性的归约路线, 已知的半均匀 LWE 问题的归约在维数和高斯分布参数等方面的归约损失较大^[4, 17-18]. 具体来讲, 证明 d 维的半均匀 LWE 问题的困难性要用的 \tilde{d} ($d = O(\tilde{d} \cdot \log q)$) 维的标准 LWE 问题的困难性. 同时, 对于一般的 η -半均匀分布 \mathcal{D} , 归约过程中高斯误差参数的损失与样本数目有关 (特别是样本数为 $\text{poly}(\lambda)$ 时, 误差参数的膨胀倍数较大). 另外, 在证明环中 (求解版本) 半均匀 LWE 问题的困难性时, 需要引入额外的、相对而言非标准的困难性假设, 即某种判定版本的 NTRU 问题对应的假设.

在实用中,一般采用的秘密/误差分布为离散高斯分布的LWE问题来进行密码协议的设计或者困难问题的归约讨论.一个自然的问题是针对特定的离散高斯分布的情形,能否给出非均匀LWE问题的更紧致的归约.这也是本研究的主要出发点和研究动机所在.

在本文中,我们推广了Hint-LWE问题困难性研究中用到的技巧并将其应用于半均匀LWE问题困难性的研究中,给出了半均匀LWE问题困难性更紧致的归约方法.本文采用的归约方法基本不受代数结构的限制,可以统一地应用于欧式格/理想格/模格上定义的半均匀LWE问题.具体来讲,本文的归约方式具有以下优势:

- (1) 归约结果是保持维数的:即 d 维的半均匀LWE问题的困难性仅需要用到 d 维的标准LWE问题的困难性来保证;
- (2) 归约过程中的高斯参数损失较少,且对于一般的 η -半均匀分布 \mathcal{D} ,归约中的高斯参数与样本数无关:即高斯参数为 σ 的 d 维的半均匀LWE问题的困难性可以由高斯参数为 $O(\eta \cdot \sigma)$ 的 d 维的标准LWE问题的困难性来保证.

本文第1节将介绍非均匀LWE问题的简单研究现状.第2节将介绍本文所需的基础知识,包括(理想/模)格,高斯分布,(半均匀)LWE问题的定义和一些有用的引理等.第3节将介绍本文的主要归约,并将我们的归约结果与已知结果做简单比较.最后,在第4节总结全文.

1 非均匀LWE问题的相关工作

在初始LWE问题的样本中,公开矩阵一般选自随机均匀分布,即 $A \leftarrow U(\mathbb{Z}_q^{m \times d})$.在已知的部分LWE问题的变体问题研究中,经常需要将 A 换成所谓的Lossy形式(即 $A = B \cdot C + E$,使用多秘密的LWE样本来替换均匀矩阵)来进行讨论^[6-8].一般而言, $A = B \cdot C + E$ 对应的分布不是与均匀分布统计不可区分的.但是在对应LWE问题困难这个计算假设下,可以证明对应的“非均匀”LWE问题也是困难的.在本文中,非均匀LWE问题指的是公开矩阵 A 服从的分布 \mathcal{D} 与 $U(\mathbb{Z}_q^{m \times d})$ 不是统计或者计算不可区分的情况下对应的LWE问题.

据我们所知,在不考虑某些 A 为特定的循环矩阵的情况下对应的LWE问题^[30](此类问题对应 A 的系数在某些环上均匀选择)时,Bonch等人第一次讨论了非均匀LWE问题的困难性^[16],并证明了当公开矩阵 A 服从的分布为(1) \mathbb{Z}^d 上适当参数的离散高斯分布;(2) k 足够大时对应的 $\{0,1\}^k$ 上的均匀分布;(3) \mathbb{Z}^d 的某些足够大的线性子空间上的均匀分布时,对应的非均匀LWE问题的困难性可以由标准LWE问题的困难性来保证.此外,Bruna等人提出的连续LWE问题(continuous LWE problems)对应的公开矩阵 A 服从适当的连续高斯分布^[31],Gupte等人也证明了适当参数条件下,此类问题的困难性可以由标准的LWE问题来保证^[32].

但是,上述几类非均匀LWE问题不适用于某些格密码应用(例如压缩公钥的方式来设计的加密/密钥封装方案^[25-29]).近期,贾文娟等人形式化定义了半均匀LWE问题^[4, 17-18],并采用类似证明熵LWE问题困难性的归约路线,结合Renyi散度、格中正则性结论(leftover hash lemma)等工具证明了欧式格/理想格/模格中适当参数的半均匀LWE问题是困难的.

2 基础知识

在本节中,我们将给出符号说明,并介绍在本文讨论过程中要经常用到的概念、定义和部分引理.

使用符号 $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ 来依次表示全体整数/有理数/实数组成的集合.对于给定的正整数 $q \in \mathbb{Z}$,定义符号 $\mathbb{Z}_q := \mathbb{Z}/q \cdot \mathbb{Z}$ 并使用符号 $[q]$ 来表示集合 $\{1, 2, \dots, q\}$.如无特殊说明,我们默认 \mathbb{Z}_q 中陪集的代表元取自集合 $[-\frac{q}{2}, \frac{q}{2}) \cap \mathbb{Z}$.本文使用粗体小写字母 \mathbf{x} 来表示向量,使用斜体大写字母 A 来表示矩阵或者集合.如无特殊说明,在本文中默认使用列向量,并使用符号 A^T 来表示矩阵 A 的转置(向量亦看为特殊的矩阵);符号

I_d 表示 $d \times d$ 的单位矩阵. 对于向量 $\mathbf{x} = (x_1, \dots, x_d)^T \in \mathbb{R}^d$, 定义 $\|\mathbf{x}\| := \sum_{i=1}^d x_i^2$, $\|\mathbf{x}\|_1 := \sum_{i=1}^d |x_i|$, $\|\mathbf{x}\|_\infty := \max_{i \in [d]} |x_i|$. 对于矩阵 $A \in \mathbb{R}^{d \times d}$, 使用符号 $s_k(A), k \in [d]$ 来表示矩阵 A 的 d 个不同的奇异值. 我们默认对奇异值按照大小排序为 $s_d(A) \leq \dots \leq s_2(A) \leq s_1(A)$, 此时 $s_1(A)$ 即为 A 的谱范数. 对于有限集合 S , 符号 $U(S)$ 表示集合 S 上的均匀分布. 对于概率分布 \mathcal{D} , 使用符号 $x \leftarrow \mathcal{D}$ 来表示随机变量 x 服从分布 \mathcal{D} . 集合 S 上的离散概率分布 \mathcal{D}_1 和 \mathcal{D}_2 的统计距离 $\Delta(\mathcal{D}_1, \mathcal{D}_2)$ 定义为 $\Delta(\mathcal{D}_1, \mathcal{D}_2) := \frac{1}{2} \cdot \sum_{a \in S} |\Pr_{x \leftarrow \mathcal{D}_1}[x = a] - \Pr_{y \leftarrow \mathcal{D}_2}[y = a]|$. 我们使用符号 λ 来表示安全参数, 并称一个函数 $f(\lambda)$

是可忽略的, 如果对任意的 $c > 0$, 存在 λ_0 , 使得对任意的 $\lambda \geq \lambda_0$, 不等式 $f(\lambda) \leq \frac{1}{\lambda^c}$ 恒成立. 当具体的函数形式对我们的讨论没有影响时, 一般使用符号 $\text{negl}(\lambda)$ 来表示 (某个或者某些不同的) 可忽略的函数. 如果 $\Delta(\mathcal{D}_1, \mathcal{D}_2) \leq \text{negl}(\lambda)$, 则称概率分布 \mathcal{D}_1 和 \mathcal{D}_2 是统计接近的/统计不可区分的.

2.1 格与离散高斯分布

在本文中, 我们仅讨论满秩的整数格. 更确切地说, 本文中讨论的 d 维欧式格 $\Lambda := \mathcal{L}(B)$ 可以写成形如 $\mathbb{Z} \cdot \mathbf{b}_1 + \dots + \mathbb{Z} \cdot \mathbf{b}_d$ 的形式, 这里 $B := (\mathbf{b}_1 | \dots | \mathbf{b}_d) \in \mathbb{Z}^{d \times d}$, d 为任意正整数. 格 Λ 的对偶格 Λ^* 的定义为 $\Lambda^* := \{\mathbf{x} \in \mathbb{R}^d : \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}, \forall \mathbf{y} \in \Lambda\}$. 我们也关心特殊的理想/模格. 为了简化讨论, 本文固定正整数 $n = 2^n$, 并考虑特殊的分圆域 $\mathcal{K} := \mathbb{Q}[x]/(x^n + 1)$ 及其代数整数环 $\mathcal{R} := \mathbb{Z}[x]/(x^n + 1)$ (值得注意的是, 采用与本文类似的分析方法, 本文的结论可以推广到一般的代数数域中). 对于 \mathcal{R} 中的多项式 $a = a_0 + a_1 \cdot x + \dots + a_{n-1} \cdot x^{n-1}$, 我们采用系数嵌入 σ_c 将其嵌入到 \mathbb{Z}^n , 即 $\sigma_c(a) := (a_0, \dots, a_{n-1})^T$. 注意到 \mathcal{R} 中两个元素的多项式乘法可以表示为矩阵乘以向量的形式. 对于给定的 $a = a_0 + a_1 \cdot x + \dots + a_{n-1} \cdot x^{n-1}$ 和 $b = b_0 + b_1 \cdot x + \dots + b_{n-1} \cdot x^{n-1}$, 经简单计算可知 $\sigma_c(a \cdot b) = M_c(a) \cdot \sigma_c(b) = M_c(b) \cdot \sigma_c(a)$. 其中,

$$M_c(a) := \begin{pmatrix} a_0 & -a_{n-1} & -a_{n-2} & \dots & -a_1 \\ a_1 & a_0 & -a_{n-1} & \dots & -a_2 \\ a_2 & a_1 & a_0 & \dots & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & \dots & a_0 \end{pmatrix}.$$

易知对于 \mathcal{R} 的任意理想 J , $\sigma_c(J)$ 为一个满秩格.

对于任意给定的正实数 $s \in \mathbb{R}$ 和向量 $\mathbf{c} \in \mathbb{R}^d$, 定义以 s 为参数、以 \mathbf{c} 为中心的 d 维高斯函数为 $\rho_{s,\mathbf{c}}(\mathbf{x}) := e^{-\pi \frac{\|\mathbf{x}-\mathbf{c}\|^2}{s^2}}$. 对应地, 以 s 为参数、以 \mathbf{c} 为中心的 d 维连续高斯分布 $D_{s,\mathbf{c}}$ 对应的概率密度函数为

$\frac{1}{s^d} \cdot \rho_{s,\mathbf{c}}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^d$. 可以将对应的高斯分布限制在任意的格 (或者离散集合) Λ 上, 对应的离散高斯分布

$D_{\Lambda,s,\mathbf{c}}$ 的概率分布函数为 $\frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{\rho_{s,\mathbf{c}}(\Lambda)}, \mathbf{x} \in \Lambda$. 这里, $\rho_{s,\mathbf{c}}(\Lambda) := \sum_{\mathbf{x} \in \Lambda} \rho_{s,\mathbf{c}}(\mathbf{x})$. 当 $\mathbf{c} = \mathbf{0}$ 或者 $s = 1$ 时, 我们

省略相关符号中对应的下标, 例如使用 $\rho_s(\mathbf{x}), \rho(\mathbf{x})$ 等. 给定任意非奇异的矩阵 $B \in \mathbb{R}^{d \times d}$, 矩阵 $\Sigma := B \cdot B^T$ 是一个实对称正定矩阵 (记为 $\Sigma > 0$). 定义 $\rho_B(\mathbf{x}) := \rho(B^{-1} \cdot \mathbf{x}) = e^{-\pi \cdot \mathbf{x}^T \cdot \Sigma^{-1} \cdot \mathbf{x}}$. 注意到 $\rho_B(\mathbf{x})$ 的取值仅仅与 Σ 和 \mathbf{x} 相关, 因此我们记 $\rho_{\sqrt{\Sigma}}(\mathbf{x}) := \rho_B(\mathbf{x})$. 为了简化讨论, 这里的 $\sqrt{\Sigma}$ 选择为 Σ 的 Cholesky 分解 (即满足条件 $\Sigma = \sqrt{\Sigma} \cdot \sqrt{\Sigma}^T, \sqrt{\Sigma}$ 为下三角矩阵). 事实上, 取 $\sqrt{\Sigma}$ 为任意满足条件 $\Sigma = B \cdot B^T$ 的矩阵 B 并不会影响后续的讨论. 对于两个实对称正定矩阵 A, B , 如果 $A - B$ 仍为实对称正定矩阵, 则记 $A > B$.

当我们在 \mathcal{R} 中讨论时, 我们使用符号 $f \leftarrow D_{\mathcal{R},s}$ 来表示多项式 f 的系数独立地取自参数为 s 的 1 维离散高斯分布 $D_{\mathbb{Z},s}$. 由于 $(D_{\mathbb{Z},s})^d = D_{\mathbb{Z}^d,s}$, 因此 $f \leftarrow D_{\mathcal{R},s}$ 等价于 f 的系数组成的向量服从分布 $D_{\mathbb{Z}^n,s}$, 即 $\sigma_c(f) \leftarrow D_{\mathbb{Z}^n,s}$. 特别地, 我们有 $D_{\mathbb{R}^d,s} = D_{\mathbb{Z}^n,d,s}$. 假设 $\varepsilon > 0$ 为某正实数. 对任意的 d 维格 Λ , 定义格 Λ 的

(与 ε 相关的) 光滑参数 $\eta_\varepsilon(A)$ 为 $\min_{s>0}\{s: \rho_{1/s}(A^*\setminus\{\mathbf{0}\}) \leq \varepsilon\}$. 类似地, 对于任意的实对称正定矩阵 $\Sigma > \mathbf{0}$, 如果条件 $\eta_\varepsilon(\sqrt{\Sigma}^{-1} \cdot A) \leq 1$ 成立, 则记 $\sqrt{\Sigma} \geq \eta_\varepsilon(A)$. 根据定义, 条件 $\eta_\varepsilon(\sqrt{\Sigma}^{-1} \cdot A) \leq 1$ 等价于 $\rho(\sqrt{\Sigma}^T \cdot A^*\setminus\{\mathbf{0}\}) = \rho_{\sqrt{\Sigma}^{-1}}(A^*\setminus\{\mathbf{0}\}) \leq 0$. 在本文后面的讨论中, 我们需要用到如下引理^[1, 8, 10, 33-35].

引理 1: 以下事实成立.

- (1) 假设 Σ 是实对称正定矩阵, A 为某个格, 实数 $\varepsilon \in (0, \frac{1}{2})$. 如果 $s_1(\Sigma) \leq \eta_\varepsilon(A)^{-2}$, 则有 $\sqrt{\Sigma} \geq \eta_\varepsilon(A)$.
- (2) 对于 d 维格 \mathbb{Z}^d , 存在可忽略的 $\varepsilon = \varepsilon(d)$ 使得 $\eta_\varepsilon(\mathbb{Z}^d) \leq \omega(\sqrt{\log d})$.
- (3) 假设已知 d 维格 $\Lambda = \mathcal{L}(B)$ 的一组格基 B , 实数 $r = \omega(\sqrt{\log d})$. 则对任意的向量 $\mathbf{c} \in \mathbb{R}^d$ 和满足条件 $\Sigma > r^2 \cdot B \cdot B^T$ 的实对称正定矩阵 Σ , 存在利用 B 的多项式时间的取样算法从一个与分布 $D_{\Lambda, \sqrt{\Sigma}, \mathbf{c}}$ 统计接近的分布中抽取样本.

注意到在 σ_c 的作用下, \mathcal{R} 对应 \mathbb{Z}^n . 根据引理 1, 对任意的向量 $\mathbf{c} \in \mathbb{R}^{n \cdot d}$ 和满足条件 $\Sigma > \omega(\sqrt{\log n \cdot d})$ 的实对称正定矩阵 Σ , 我们可以有效地从一个与 $D_{\mathcal{R}^d, \sqrt{\Sigma}, \mathbf{c}}$ 统计不可区分的分布中进行抽样.

引理 2: 假设实数 $\varepsilon \in (0, \frac{1}{2})$, Σ_1, Σ_2 为实对称正定矩阵, $\Sigma_3^{-1} := \Sigma_2^{-1} + \Sigma_1^{-1}$ 且满足条件 $\sqrt{\Sigma_3} \geq \eta_\varepsilon(\mathbb{Z}^d)$. 则对任意的向量 $\mathbf{c} \in \mathbb{R}^d$, 分布 $\{\mathbf{x}_1 + \mathbf{x}_2: \mathbf{x}_1 \leftarrow D_{\mathbb{Z}^d, \sqrt{\Sigma_1}, \mathbf{c}}, \mathbf{x}_2 \leftarrow D_{\mathbb{Z}^d, \sqrt{\Sigma_2}, \mathbf{c}}\}$ 与 $D_{\mathbb{Z}^d, \sqrt{\Sigma_1 + \Sigma_2}, \mathbf{c}}$ 的统计距离不超过 $2 \cdot \varepsilon$.

利用引理 2, 我们可以将某个离散高斯分布拆分为某两个适当参数的离散高斯分布的和.

2.2 LWE 问题

LWE 问题最早由 Regev 提出^[1], 是目前格密码设计中应用最广泛的一类平均情形的格中困难问题.

本文用符号 \mathfrak{R} 来表示环 \mathbb{Z} 或者 \mathcal{R} , 并定义 $\mathfrak{R}_q := \mathfrak{R}/q \cdot \mathfrak{R}$. 在系数嵌入 σ_c 的作用下, \mathcal{R} 中元素的加法和乘法可以使用 \mathbb{Z}^n 中向量的加法, 以及对应矩阵和向量的乘法来表示. 注意到 σ_c 可以平凡地扩展到以 \mathcal{R} 中的元素为元素的向量或者矩阵上. 为了讨论方便, 定义 \mathbb{Z} 上的 σ_c 作用为恒等映射. 这样, σ_c 可以视为 \mathfrak{R} 上的 \mathbb{Z} -模同态, 或者视为 \mathfrak{R}_q 上的 \mathbb{Z}_q -模同态.

对于正整数 m, d, q 以及 \mathfrak{R}^m 上的概率分布 χ_1 , \mathfrak{R}^d 上的概率分布 χ_2 , $\mathfrak{R}_q^{m \times d}$ 上的概率分布 \mathcal{D} , 定义环 \mathfrak{R} 相关的判定版本的 LWE 问题 (记为 $DLWE_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$) 为**如下问题**: 区分分布 $(A, A \cdot \mathbf{s} + \mathbf{e} \bmod q \cdot \mathfrak{R})$ 与分布 (A, \mathbf{u}) . 这里, $A \leftarrow \mathcal{D}$, $\mathbf{s} \leftarrow \chi_2$, $\mathbf{e} \leftarrow \chi_1$, $\mathbf{u} \leftarrow U(\mathfrak{R}_q^m)$. 注意到在上述定义中, 参数 m 对应的是 LWE 问题的样本数; \mathcal{D} 即为 LWE 问题公开矩阵服从的分布; χ_1 和 χ_2 分别为 LWE 问题的误差向量和秘密向量服从的分布. 为了下文讨论方便, 当样本形式为 $(A, A \cdot \mathbf{s} + \mathbf{e} \bmod q \cdot \mathfrak{R})$ 时, 我们称对应的样本取自 $DLWE_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题的分布. 对于任意的 (量子) 概率多项式时间的敌手 \mathfrak{A} , 定义其解决 $DLWE_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题的优势为

$$\text{Adv}_{\mathfrak{A}}(DLWE_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})) := |\text{Pr}[\mathfrak{A}(A, A \cdot \mathbf{s} + \mathbf{e} \bmod q \cdot \mathfrak{R}) = 1] - \text{Pr}[\mathfrak{A}(A, \mathbf{u}) = 1]|.$$

当 $\mathfrak{R} = \mathbb{Z}$ 时, 上述判定版本 LWE 问题对应欧式格上的 LWE 问题. 此时, 参数 d 为对应的 (格的) 维数. 而当 $\mathfrak{R} = \mathcal{R}$ 时, 上述判定版本 LWE 问题对应现实中最常用的、扩张次数为 2 的幂次的分圆域上定义的相应 LWE 问题. 此时, 对应的 (格的) 维数为 $n \cdot d$. 当 $d = 1$ 时对应环 LWE 问题, 而当 $d \geq 2$ 时对应模 LWE 问题.

一般而言, LWE 问题的公开矩阵的分布为 $\mathcal{D} = U(\mathfrak{R}_q^{m \times d})$. 当 $\mathcal{D} \neq U(\mathfrak{R}_q^{m \times d})$ 时, 我们称对应的 LWE 问题为非均匀 LWE 问题. 本文关心的是非均匀 LWE 问题中的一类特殊的半均匀 LWE 问题, 即 \mathcal{D} 为下面定义的 η -半均匀分布^[4, 17-18].

定义 3: 假设 m, d 为正整数. 对于任意的正实数 $\eta > 0$, 我们称分布 \mathcal{D} 为 $\mathfrak{R}_q^{m \times d}$ 上的 η -半均匀分布, 如

果存在 $\mathfrak{R}^{m \times d}$ 上的一族可以有效取样的概率分布 $\{\phi_U\}_{U \in \mathfrak{R}^{m \times d}}$ 使得对任意随机取样的 $A \leftarrow \mathcal{D}, U \leftarrow \mathcal{U}(\mathfrak{R}_q^{m \times d})$ 和 $E_U \leftarrow \phi_U$, 下面两个条件成立:

- (1) 随机变量 A 与 $U + E_U \bmod q \cdot \mathfrak{R}$ 统计不可区分;
- (2) $\Pr_{U \leftarrow \mathcal{U}(\mathfrak{R}_q^{m \times d}), E_U \leftarrow \phi_U}[\mathfrak{s}_1(\sigma_c(E_U)) \leq \eta] \geq 1 - \text{negl}(\lambda)$.

3 半均匀 LWE 问题的困难性研究

本节将给出半均匀 LWE 问题相关的更紧的归约方法. 在给出具体的归约之前, 我们需要用到下面这个关键引理. 引理 4 的证明所采用到的方法与文献 [10] 中的方法类似.

引理 4: 假设 σ_1, σ_2 为正实数, \tilde{m}, \tilde{n} 为正整数, $F \in \mathbb{Z}^{\tilde{m} \times \tilde{n}}$ 为任意一个使得矩阵 $\Sigma_0 := (\frac{1}{\sigma_1^2} \cdot I_{\tilde{n}} + \frac{1}{\sigma_2^2} \cdot F^T \cdot F)^{-1}$

存在的矩阵. 则以下两个取样过程输出的分布相同:

- (1) 先取样 $\mathbf{s} \leftarrow D_{\mathbb{Z}^{\tilde{n}}, \sigma_1}$ 和 $\mathbf{e} \leftarrow D_{\mathbb{Z}^{\tilde{m}}, \sigma_2}$, 再计算 $\mathbf{z} = F \cdot \mathbf{s} + \mathbf{e}$, 最后输出 $(\mathbf{s}, \mathbf{z}) \in \mathbb{Z}^{\tilde{n}} \times \mathbb{Z}^{\tilde{m}}$;
- (2) 先取样 $\mathbf{s} \leftarrow D_{\mathbb{Z}^{\tilde{n}}, \sigma_1}$ 和 $\mathbf{e} \leftarrow D_{\mathbb{Z}^{\tilde{m}}, \sigma_2}$, 再计算 $\mathbf{z} = F \cdot \mathbf{s} + \mathbf{e}$ 和 $\mathbf{c} = \frac{1}{\sigma_2^2} \cdot \Sigma_0 \cdot F^T \cdot \mathbf{z}$; 随后采样 $\tilde{\mathbf{s}} \leftarrow D_{\mathbb{Z}^{\tilde{n}}, \sqrt{\Sigma_0}, \mathbf{c}}$;

最后输出 $(\tilde{\mathbf{s}}, \mathbf{z}) \in \mathbb{Z}^{\tilde{n}} \times \mathbb{Z}^{\tilde{m}}$.

证明: 记输出的第一个随机变量为 \mathbf{x} , 第二个随机变量为 \mathbf{y} . 注意到对任意的 $(\mathbf{v}, \mathbf{w}) \in \mathbb{Z}^{\tilde{n}} \times \mathbb{Z}^{\tilde{m}}$,

$$\Pr[\mathbf{x} = \mathbf{v} \wedge \mathbf{y} = \mathbf{w}] = \Pr[\mathbf{y} = \mathbf{w}] \cdot \Pr[\mathbf{x} = \mathbf{v} | \mathbf{y} = \mathbf{w}].$$

同时, 在两个取样过程中, $\Pr[\mathbf{y} = \mathbf{w}] = \Pr[\mathbf{z} = F \cdot \mathbf{s} + \mathbf{e} = \mathbf{w}]$ 相同. 因此, 为了证明这两个取样过程最后输出的分布相同, 只需证明在取样过程 (1) 中, 在已知 $F \cdot \mathbf{s} + \mathbf{e}$ 的情况下, \mathbf{s} 服从的条件分布为 $D_{\mathbb{Z}^{\tilde{n}}, \sqrt{\Sigma_0}, \mathbf{c}}$ 即可.

为此, 任取 $(\mathbf{v}, \mathbf{w}) \in \mathbb{Z}^{\tilde{n}} \times \mathbb{Z}^{\tilde{m}}$. 在取样过程 (1) 中, 我们有

$$\begin{aligned} \Pr[\mathbf{x} = \mathbf{v} \wedge \mathbf{y} = \mathbf{w}] &= \Pr[\mathbf{s} = \mathbf{v} \wedge F \cdot \mathbf{s} + \mathbf{e} = \mathbf{w}] \\ &= \Pr[\mathbf{s} = \mathbf{v} \wedge \mathbf{e} = \mathbf{w} - F \cdot \mathbf{v}] \\ &= \frac{1}{\rho_{\sigma_1}(\mathbb{Z}^{\tilde{n}}) \cdot \rho_{\sigma_2}(\mathbb{Z}^{\tilde{m}})} \cdot e^{-\pi \left(\frac{1}{\sigma_1^2} \mathbf{v}^T \cdot \mathbf{v} + \frac{1}{\sigma_2^2} (\mathbf{w} - F \cdot \mathbf{v})^T \cdot (\mathbf{w} - F \cdot \mathbf{v}) \right)} \\ &= \frac{1}{\rho_{\sigma_1}(\mathbb{Z}^{\tilde{n}}) \cdot \rho_{\sigma_2}(\mathbb{Z}^{\tilde{m}})} \cdot e^{-\pi \left((\mathbf{v} - \mathbf{c})^T \cdot \Sigma_0^{-1} \cdot (\mathbf{v} - \mathbf{c}) - \mathbf{c}^T \cdot \Sigma_0^{-1} \cdot \mathbf{c} + \frac{1}{\sigma_2^2} \mathbf{w}^T \cdot \mathbf{w} \right)}. \end{aligned}$$

这里, 我们用到了条件 $\Sigma_0^{-1} = \Sigma_0^{-T}$. 注意到

$$\begin{aligned} \Pr[\mathbf{y} = \mathbf{w}] &= \Pr[\mathbf{z} = \mathbf{w}] = \sum_{\mathbf{a} \in \mathbb{Z}^{\tilde{n}}} \Pr[\mathbf{s} = \mathbf{a}] \cdot \Pr[\mathbf{e} = \mathbf{w} - F \cdot \mathbf{s} | \mathbf{s} = \mathbf{a}] \\ &= \frac{1}{\rho_{\sigma_1}(\mathbb{Z}^{\tilde{n}}) \cdot \rho_{\sigma_2}(\mathbb{Z}^{\tilde{m}})} \cdot \sum_{\mathbf{a} \in \mathbb{Z}^{\tilde{n}}} e^{-\pi \frac{\|\mathbf{a}\|^2}{\sigma_1^2}} \cdot e^{-\pi \frac{\|\mathbf{w} - F \cdot \mathbf{a}\|^2}{\sigma_2^2}}, \end{aligned}$$

我们可以得到

$$\begin{aligned} \Pr[\mathbf{x} = \mathbf{v} | \mathbf{y} = \mathbf{w}] &= \Pr[\mathbf{s} = \mathbf{v} | \mathbf{z} = \mathbf{w}] \\ &= \Pr[\mathbf{s} = \mathbf{v} \wedge \mathbf{z} = \mathbf{w}] / \Pr[\mathbf{z} = \mathbf{w}] \\ &= \frac{e^{-\pi \left(-\mathbf{c}^T \cdot \Sigma_0^{-1} \cdot \mathbf{c} + \frac{1}{\sigma_2^2} \mathbf{w}^T \cdot \mathbf{w} \right)}}{\sum_{\mathbf{a} \in \mathbb{Z}^{\tilde{n}}} e^{-\pi \frac{\|\mathbf{a}\|^2}{\sigma_1^2}} \cdot e^{-\pi \frac{\|\mathbf{w} - F \cdot \mathbf{a}\|^2}{\sigma_2^2}}} \cdot e^{-\pi \cdot (\mathbf{v} - \mathbf{c})^T \cdot \Sigma_0^{-1} \cdot (\mathbf{v} - \mathbf{c})}. \end{aligned}$$

因为式子 $\frac{e^{-\pi(-c^T \cdot \Sigma_0^{-1} \cdot c + \frac{1}{\sigma_2^2} w^T \cdot w)}}{\sum_{a \in \mathbb{Z}^{\tilde{n}}} e^{-\pi \frac{\|a\|^2}{\sigma_1^2}} \cdot e^{-\pi \frac{\|w-F \cdot a\|^2}{\sigma_2^2}}}$ 与 v 无关, 且有概率等式 $\sum_{v \in \mathbb{Z}^{\tilde{n}}} \Pr[s = v | z = w] = 1$ 成立, 所以我们有

$$\frac{e^{-\pi(-c^T \cdot \Sigma_0^{-1} \cdot c + \frac{1}{\sigma_2^2} w^T \cdot w)}}{\sum_{a \in \mathbb{Z}^{\tilde{n}}} e^{-\pi \frac{\|a\|^2}{\sigma_1^2}} \cdot e^{-\pi \frac{\|w-F \cdot a\|^2}{\sigma_2^2}}} = \left(\sum_{v \in \mathbb{Z}^{\tilde{n}}} e^{-\pi(v-c)^T \cdot \Sigma_0^{-1} \cdot (v-c)} \right)^{-1} = \frac{1}{\rho_{\sqrt{\Sigma_0}, c}(\mathbb{Z}^{\tilde{n}})}.$$

因此,

$$\Pr[s = v | z = w] = \frac{1}{\rho_{\sqrt{\Sigma_0}, c}(\mathbb{Z}^{\tilde{n}})} \cdot e^{-\pi(v-c)^T \cdot \Sigma_0^{-1} \cdot (v-c)}.$$

也就是说, 在已知 $z = F \cdot s + e = w$ 的情况下, s 服从的条件分布为 $D_{\mathbb{Z}^{\tilde{n}}, \sqrt{\Sigma_0}, c}$.

证毕.

3.1 半均匀 LWE 问题的归约

本文的主要定理如下.

定理 5: 假设 $\varepsilon = \text{negl}(\lambda)$ 为某可忽略的函数, m, q, d 为正整数, \mathcal{D} 为 $\mathfrak{R}_q^{m \times d}$ 上的 η -半均匀分布, 正实数 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \omega(\max\{\sqrt{\log m \cdot n}, \sqrt{\log d \cdot n}\})$ 且满足条件 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathfrak{R}^m)$, $\sigma_1 = \sqrt{\gamma^2 + \sigma_3^2}$, $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$.

记 $\chi_1 = D_{\mathfrak{R}^m, \sigma_1}$, $\chi_2 = D_{\mathfrak{R}^d, \sigma_2}$, $\chi_3 = D_{\mathfrak{R}^m, \sigma_3}$, $\chi_4 = D_{\mathfrak{R}^d, \sigma_4}$. 如果存在一个 (量子) 概率多项式时间的敌手 \mathfrak{A} 可以以 δ 的概率解决 $\text{DLWE}_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题, 则存在一个 (量子) 概率多项式时间的敌手 \mathfrak{B} 可以以 $\delta - \text{negl}(\lambda)$ 的概率解决 $\text{DLWE}_{\mathfrak{R}, m, q, \chi_3}^d(\chi_4; U(\mathfrak{R}_q^{m \times d}))$ 问题.

证明: 根据假设, \mathcal{D} 为 $\mathfrak{R}_q^{m \times d}$ 上的 η -半均匀分布. 由定义 3, 存在 $\mathfrak{R}^{m \times d}$ 上的一族可以有效取样的概率分布 $\{\phi_U\}_{U \in \mathfrak{R}^{m \times d}}$ 使得对任意随机取样的 $A \leftarrow \mathcal{D}$, $U \leftarrow U(\mathfrak{R}_q^{m \times d})$ 和 $E_U \leftarrow \phi_U$, 随机变量 A 与 $U + E_U \bmod q \cdot \mathfrak{R}$ 统计不可区分. 同时, 有 $\Pr_{U \leftarrow U(\mathfrak{R}_q^{m \times d}), E_U \leftarrow \phi_U}[\mathfrak{s}_1(\sigma_c(E_U)) \leq \eta] \geq 1 - \text{negl}(\lambda)$ 成立.

对任意 (量子) 概率多项式时间的敌手 \mathfrak{A} , 我们定义一系列区分游戏 Game_i , 并定义 $p_i := \Pr_{\text{Game}_i}[\mathfrak{A}(A, b) = 1]$. 这里, $i \in [6]$, $(A, b) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$.

Game₁: 取样 $A \leftarrow \mathcal{D}$, $s \leftarrow \chi_2$, $e \leftarrow \chi_1$; 计算 $b = A \cdot s + e \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, b) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

注意到 Game_1 中输出的样本 (A, b) 即为 $\text{DLWE}_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题的原始分布.

Game₂: 取样 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $E_U \leftarrow \phi_U$, $s \leftarrow \chi_2$, $e \leftarrow \chi_1$; 如果 $\mathfrak{s}_1(\sigma_c(E_U)) > \eta$, 则输出 \perp (代表采样出错, 游戏终止); 否则, 计算 $A = U + E_U \bmod q \cdot \mathfrak{R}$, $b = A \cdot s + e \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, b) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

根据定义 3, 在 Game_2 中输出 \perp 的概率是可忽略的. 当 Game_2 正常输出 (A, b) 时, Game_2 中输出的 A 与 Game_1 中输出的 A 统计不可区分. 因此, 我们有 $|p_1 - p_2| \leq \text{negl}(\lambda)$.

Game₃: 取样 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $E_U \leftarrow \phi_U$, $s \leftarrow \chi_2$; 如果 $\mathfrak{s}_1(\sigma_c(E_U)) > \eta$, 则输出 \perp ; 否则继续取样 $e_1 \leftarrow D_{\mathfrak{R}^m, \gamma}$, $e_2 \leftarrow D_{\mathfrak{R}^m, \sigma_3}$; 再计算 $z = E_U \cdot s + e_1$, $A = U + E_U \bmod q \cdot \mathfrak{R}$, $b = U \cdot s + z + e_2 \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, b) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

下面当 $\mathfrak{R} = \mathbb{Z}$ 时取 $N = m$, $\tilde{N} = d$; 而当 $\mathfrak{R} = \mathcal{R}$ 时取 $N = m \cdot n$, $\tilde{N} = n \cdot d$. 如果令 $\Sigma_1 := \gamma^2 \cdot I_N$, $\Sigma_2 := \sigma_3^2 \cdot I_N$ 并取 $\Sigma_3 := \frac{\gamma^2 \cdot \sigma_3^2}{\gamma^2 + \sigma_3^2} \cdot I_N$, 则根据假设 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathfrak{R}^m) = \eta_\varepsilon(\mathbb{Z}^N)$ 以及引理 2 可知, 概率分布

$D_{\mathfrak{R}^m, \gamma} + D_{\mathfrak{R}^m, \sigma_3}$ 与概率分布 $D_{\mathfrak{R}^m, \sqrt{\gamma^2 + \sigma_3^2}} = \chi_1$ 的统计距离不超过 $2 \cdot \varepsilon$. 在 Game_3 中, 我们有

$$\begin{aligned} b &= U \cdot s + z + e_2 \bmod q \cdot \mathfrak{R} = U \cdot s + E_U \cdot s + e_1 + e_2 \bmod q \cdot \mathfrak{R} \\ &= (U + E_U) \cdot s + (e_1 + e_2) \bmod q \cdot \mathfrak{R} \\ &= A \cdot s + (e_1 + e_2) \bmod q \cdot \mathfrak{R}. \end{aligned}$$

所以, 根据上述分析, 我们可以得到 $|p_2 - p_3| \leq \text{negl}(\lambda)$.

Game₄: 取样 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $E_U \leftarrow \phi_U$, $\mathbf{s} \leftarrow \chi_2$; 如果 $s_1(\sigma_c(E_U)) > \eta$, 则输出 \perp ; 否则继续取样 $\mathbf{e}_1 \leftarrow D_{\mathfrak{R}^m, \gamma}$, $\mathbf{e}_2 \leftarrow D_{\mathfrak{R}^m, \sigma_3}$; 再计算 $A = U + E_U \bmod q \cdot \mathfrak{R}$, $\mathbf{z} = E_U \cdot \mathbf{s} + \mathbf{e}_1$, 并取 $\Sigma_0 := (\frac{1}{\sigma_2^2} \cdot I_N + \frac{1}{\gamma^2} \cdot \sigma_c(E_U)^T \cdot \sigma_c(E_U))^{-1}$, $\mathbf{c} = \frac{1}{\gamma^2} \cdot \Sigma_0 \cdot \sigma_c(E_U)^T \cdot \mathbf{z}$; 随后取样 $\tilde{\mathbf{s}} \leftarrow D_{\mathfrak{R}^d, \sqrt{\Sigma_0, \mathbf{c}}}$, 并利用二元组 $(\tilde{\mathbf{s}}, \mathbf{z})$ 来计算 $\mathbf{b} = U \cdot \tilde{\mathbf{s}} + \mathbf{z} + \mathbf{e}_2 \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, \mathbf{b}) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

我们先来分析 **Game₄** 中所需要的分布是否可以有效地取样. 注意到 $\Sigma_0^{-1} = \frac{1}{\sigma_2^2} \cdot I_N + \frac{1}{\gamma^2} \cdot \sigma_c(E_U)^T \cdot \sigma_c(E_U)$, 根据参数选择, 我们有 $s_1(\Sigma_0^{-1}) \leq \frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$. 所以, 可以得到 $s_N(\Sigma_0) > 2 \cdot \sigma_4^2$, 进而有 $\Sigma_0 > \sigma_4^2 \cdot I_N$ 成立. 由引理 1 以及我们选择的参数条件 $\sigma_4 \geq \omega(\sqrt{\log n \cdot d})$ 可知, 可以有效地从一个与 $D_{\mathfrak{R}^d, \sqrt{\Sigma_0, \mathbf{c}}}$ 统计不可区分的分布中采样 $\tilde{\mathbf{s}}$. 当所有的分布均可以有效采样时, **Game₃** 与 **Game₄** 的区别仅仅在于用于计算 \mathbf{b} 的 (\mathbf{s}, \mathbf{z}) 与 $(\tilde{\mathbf{s}}, \mathbf{z})$ 的不同. 根据引理 4, 在我们的参数选择下, (\mathbf{s}, \mathbf{z}) 与 $(\tilde{\mathbf{s}}, \mathbf{z})$ 的分布统计不可区分. 因此, 我们有 $|p_3 - p_4| \leq \text{negl}(\lambda)$.

Game₅: 取样 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $E_U \leftarrow \phi_U$, $\mathbf{s} \leftarrow \chi_2$; 如果 $s_1(\sigma_c(E_U)) > \eta$, 则输出 \perp ; 否则继续取样 $\mathbf{e}_1 \leftarrow D_{\mathfrak{R}^m, \gamma}$, $\mathbf{e}_2 \leftarrow D_{\mathfrak{R}^m, \sigma_3}$; 再计算 $A = U + E_U \bmod q \cdot \mathfrak{R}$, $\mathbf{z} = E_U \cdot \mathbf{s} + \mathbf{e}_1$, 并取 $\Sigma_0 := (\frac{1}{\sigma_2^2} \cdot I_N + \frac{1}{\gamma^2} \cdot \sigma_c(E_U)^T \cdot \sigma_c(E_U))^{-1}$, $\mathbf{c} = \frac{1}{\gamma^2} \cdot \Sigma_0 \cdot \sigma_c(E_U)^T \cdot \mathbf{z}$; 随后取样 $\mathbf{s}_1 \leftarrow D_{\mathfrak{R}^d, \sqrt{\Sigma_0 - \sigma_4^2 \cdot I_N, \mathbf{c}}}$ 和 $\mathbf{s}_2 \leftarrow D_{\mathfrak{R}^d, \sigma_4}$ 并计算 $\tilde{\mathbf{s}} = \mathbf{s}_1 + \mathbf{s}_2 \bmod q \cdot \mathfrak{R}$; 再利用二元组 $(\tilde{\mathbf{s}}, \mathbf{z})$ 来计算 $\mathbf{b} = U \cdot \tilde{\mathbf{s}} + \mathbf{z} + \mathbf{e}_2 \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, \mathbf{b}) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

我们已经知道 $s_N(\Sigma_0) > 2 \cdot \sigma_4^2$, 所以 $s_N(\Sigma_0 - \sigma_4^2 \cdot I_N) > \sigma_4^2$. 根据引理 1 和我们的参数选择, 可以有效地从一个与 $D_{\mathfrak{R}^d, \sqrt{\Sigma_0 - \sigma_4^2 \cdot I_N, \mathbf{c}}}$ 统计接近的分布进行采样. 如果现在取 $\Sigma_1^{-1} = \frac{1}{\sigma_4^2} \cdot I_N$, $\Sigma_2^{-1} = (\Sigma_0 - \sigma_4^2 \cdot I_N)^{-1}$ 并令 $\Sigma_3^{-1} = \Sigma_1^{-1} + \Sigma_2^{-1}$, 简单计算即知 $s_1(\Sigma_3^{-1}) \leq \frac{1}{\sigma_4^2} + s_1(\Sigma_2^{-1}) < \frac{2}{\sigma_4^2} \leq (\eta_\varepsilon(\mathfrak{R}^d))^{-2}$ 成立. 由引理 1 可知, $\sqrt{\Sigma_3} \geq \eta_\varepsilon(\mathfrak{R}^d)$. 进而根据引理 2, 分布 $D_{\mathfrak{R}^d, \sqrt{\Sigma_0 - \sigma_4^2 \cdot I_N, \mathbf{c}}} + D_{\mathfrak{R}^d, \sigma_4}$ 与分布 $D_{\mathfrak{R}^d, \sqrt{\Sigma_0, \mathbf{c}}}$ 的统计距离不超过 $2 \cdot \varepsilon$, 即

Game₄ 和 **Game₅** 中对应的 $\tilde{\mathbf{s}}$ 的分布统计不可区分. 所以, 我们可以得到 $|p_4 - p_5| \leq \text{negl}(\lambda)$.

Game₆: 取样 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $E_U \leftarrow \phi_U$, $\mathbf{s} \leftarrow \chi_2$ 和 $\mathbf{e}_1 \leftarrow D_{\mathfrak{R}^m, \gamma}$; 如果 $s_1(\sigma_c(E_U)) > \eta$, 则输出 \perp ; 否则计算 $A = U + E_U \bmod q \cdot \mathfrak{R}$ 和 $\mathbf{z} = E_U \cdot \mathbf{s} + \mathbf{e}_1$; 令 $\Sigma_0 := (\frac{1}{\sigma_2^2} \cdot I_N + \frac{1}{\gamma^2} \cdot \sigma_c(E_U)^T \cdot \sigma_c(E_U))^{-1}$, $\mathbf{c} = \frac{1}{\gamma^2} \cdot \Sigma_0 \cdot \sigma_c(E_U)^T \cdot \mathbf{z}$; 随后取样 $\mathbf{s}_1 \leftarrow D_{\mathfrak{R}^d, \sqrt{\Sigma_0 - \sigma_4^2 \cdot I_N, \mathbf{c}}}$ 和 $\mathbf{u} \leftarrow U(\mathfrak{R}_q^m)$; 计算 $\mathbf{b} = U \cdot \mathbf{s}_1 + \mathbf{z} + \mathbf{u} \bmod q \cdot \mathfrak{R}$; 最后输出 $(A, \mathbf{b}) \in \mathfrak{R}_q^{m \times d} \times \mathfrak{R}_q^m$ 给 \mathfrak{A} .

在 **Game₆** 中, 由于 \mathbf{u} 是独立随机地取自分布 $U(\mathfrak{R}_q^m)$, 因此有 $\mathbf{b} \leftarrow U(\mathfrak{R}_q^m)$ 成立. 根据 η -半均匀分布, A 服从的分布与 \mathcal{D} 统计不可区分. 根据假设, \mathfrak{A} 可以以 δ 的概率解决 $\text{DLWE}_{\mathfrak{R}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题. 所以我们可以得到 $|p_1 - p_6| \geq \delta - \text{negl}(\lambda)$. 综合 **Game₁** 到 **Game₅** 的分析, 我们有 $|p_5 - p_6| \geq \delta - \text{negl}(\lambda)$.

注意到在 **Game₅** 中,

$$\mathbf{b} = U \cdot \tilde{\mathbf{s}} + \mathbf{z} + \mathbf{e}_2 \bmod q \cdot \mathfrak{R} = U \cdot \mathbf{s}_1 + \mathbf{z} + U \cdot \mathbf{s}_2 + \mathbf{e}_2 \bmod q \cdot \mathfrak{R}.$$

现在, 我们可以利用敌手 \mathfrak{A} 的能力来构造敌手 \mathfrak{B} 以解决 $\text{DLWE}_{\mathfrak{R}, m, q, \chi_3}^d(\chi_4; U(\mathfrak{R}_q^{m \times d}))$ 问题. 假设敌手 \mathfrak{B} 获得的样本为 $(U, \tilde{\mathbf{b}})$, 其中 $U \leftarrow U(\mathfrak{R}_q^{m \times d})$, $\mathbf{s}_2 \leftarrow D_{\mathfrak{R}^d, \sigma_4}$, $\mathbf{e}_2 \leftarrow D_{\mathfrak{R}^m, \sigma_3}$, $\tilde{\mathbf{b}} = U \cdot \mathbf{s}_2 + \mathbf{e}_2 \bmod q \cdot \mathfrak{R}$ 或者

$U \leftarrow U(\mathbb{R}_q^{m \times d}), \tilde{\mathbf{b}} \leftarrow U(\mathbb{R}_q^m)$. 则 \mathfrak{B} 进行如下操作:

- (1) 取样 $E_U \leftarrow \phi_U, \mathbf{s} \leftarrow \chi_2$; 如果 $s_1(\sigma_c(E_U)) > \eta$, 则输出 \perp ; 否则继续取样 $\mathbf{e}_1 \leftarrow D_{\mathbb{R}^m, \gamma}$;
- (2) 计算 $A = U + E_U \bmod q \cdot \mathbb{R}$ 和 $\mathbf{z} = E_U \cdot \mathbf{s} + \mathbf{e}_1$, 随后取 $\Sigma_0 := (\frac{1}{\sigma_2^2} \cdot I_N + \frac{1}{\gamma^2} \cdot \sigma_c(E_U)^T \cdot \sigma_c(E_U))^{-1}$,
 $\mathbf{c} = \frac{1}{\gamma^2} \cdot \Sigma_0 \cdot \sigma_c(E_U)^T \cdot \mathbf{z}$;
- (3) 取样 $\mathbf{s}_1 \leftarrow D_{\mathbb{R}^d, \sqrt{\Sigma_0 - \sigma_2^2 \cdot I_N} \cdot \mathbf{c}}$, 随后计算 $\mathbf{b} = U \cdot \mathbf{s}_1 + \mathbf{z} + \tilde{\mathbf{b}} \bmod q \cdot \mathbb{R}$;
- (4) 输出 $(A, \mathbf{b}) \in \mathbb{R}_q^{m \times d} \times \mathbb{R}_q^m$ 给 \mathfrak{A} .

最后 \mathfrak{B} 输出 \mathfrak{A} 的输出即可.

显然, 根据我们的参数选择, 上述 \mathfrak{B} 的操作可以在概率多项式时间内完成. 下面来分析所构造的敌手 \mathfrak{B} 可以成功解决 $\text{DLWE}_{\mathbb{R}, m, q, \chi_3}^d(\chi_4; U(\mathbb{R}_q^{m \times d}))$ 问题的概率. 易知当 $(U, \tilde{\mathbf{b}})$ 来自均匀分布时, \mathfrak{B} 给 \mathfrak{A} 的样本 (A, \mathbf{b}) 所服从的分布与 Game_6 的输出所服从的分布相同. 否则, \mathfrak{B} 给 \mathfrak{A} 的样本 (A, \mathbf{b}) 所服从的分布与 Game_5 的输出所服从的分布相同. 因此,

$$\text{Adv}_{\mathfrak{B}}(\text{DLWE}_{\mathbb{R}, m, q, \chi_3}^d(\chi_4; U(\mathbb{R}_q^{m \times d}))) = |p_5 - p_6| \geq \delta - \text{negl}(\lambda).$$

证毕.

注意到定理 5 中的条件 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \omega(\max\{\sqrt{\log m \cdot n}, \sqrt{\log d \cdot n}\})$ 是为了对 \mathbb{R} 进行统一分析. 当 $\mathbb{R} = \mathbb{Z}$ 时, 相应的条件可以改为 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \omega(\max\{\sqrt{\log m}, \sqrt{\log d}\})$. 我们对定理 5 中所选择的参数条件进行一些简单的说明. 条件 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathbb{R}^m)$ 和 $\sigma_1 = \sqrt{\gamma^2 + \sigma_3^2}$ 是为了保证从 Game_2 到 Game_3 的过

程中, 我们可以对 Game_2 中的原始误差 \mathbf{e} 进行拆分; 条件 $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$ 是为了保证我们可以应用关键引理 4; 本质上讲, 条件 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \omega(\max\{\sqrt{\log m \cdot n}, \sqrt{\log d \cdot n}\})$ 保证了证明过程中对应的离散高斯分布均可以有效地取样. 同时, 此条件连同条件 $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$ 也保证了从 Game_4 到 Game_5 的过渡中可以对 Game_4 中的原始误差 \mathbf{s} 进行拆分. 最后得到的 Game_5 (以及 Game_6) 即为所期望的分布形式, 非常便于我们来构造概率多项式时间的算法 (归约) 来嵌入 $\text{DLWE}_{\mathbb{R}, m, q, \chi_3}^d(\chi_4; U(\mathbb{R}_q^{m \times d}))$ 问题的实例.

3.2 与已知结果的对比分析

利用定理 5, 可以得到半均匀的(欧式格/环/模)LWE 问题的困难性结果. 在本小节中, 对任意的整数 $x \in \mathbb{Z}$, 使用符号 $\lfloor x \rfloor$ 来表示不超过 x 的最大整数, 并定义 $\lfloor x \rfloor := \lfloor x + \frac{1}{2} \rfloor$. 相应的取整符号可以平凡的推广到向量 (多项式) 或者矩阵上.

当 $\mathbb{R} = \mathbb{Z}$ 时, 可以得到 d 维欧式格对应的半均匀 LWE 问题的困难性归约.

推论 6: 假设 $\varepsilon = \text{negl}(\lambda)$ 为某可忽略的函数, m, q, d 为正整数, \mathcal{D} 为 $\mathbb{Z}_q^{m \times d}$ 上的 η -半均匀分布, 正实数 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \sqrt{2} \cdot \omega(\max\{\sqrt{\log m}, \sqrt{\log d}\})$ 且满足条件 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathbb{Z}^m)$, $\sigma_1 = \sqrt{\gamma^2 + \sigma_3^2}$, $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$. 记

$\chi_1 = D_{\mathbb{Z}^m, \sigma_1}$, $\chi_2 = D_{\mathbb{Z}^d, \sigma_2}$, $\chi_3 = D_{\mathbb{Z}^m, \sigma_3}$, $\chi_4 = D_{\mathbb{Z}^d, \sigma_4}$. 如果存在一个 (量子) 概率多项式时间的敌手 \mathfrak{A} 可以以 δ 的概率解决 $\text{DLWE}_{\mathbb{Z}, m, q, \chi_1}^d(\chi_2; \mathcal{D})$ 问题, 则存在一个 (量子) 概率多项式时间的敌手 \mathfrak{B} 可以以 $\delta - \text{negl}(\lambda)$ 的概率解决 $\text{DLWE}_{\mathbb{Z}, m, q, \chi_3}^d(\chi_4; U(\mathbb{Z}_q^{m \times d}))$ 问题.

如果选择 $\sigma_2 = 2\sqrt{2} \cdot \sigma_4$, $\gamma = 2\sqrt{2} \cdot \sigma_4 \cdot \eta$, 则 $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} = \frac{1}{4 \cdot \sigma_4^2} < \frac{1}{2 \cdot \sigma_4^2}$. 由引理 1, $\eta_\varepsilon(\mathbb{Z}^m) \leq \omega(\sqrt{\log m})$. 而 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} = \frac{\gamma}{\sqrt{1 + (\frac{\gamma}{\sigma_3})^2}} = \frac{\sigma_3}{\sqrt{1 + (\frac{\sigma_3}{\gamma})^2}}$, 所以在上述参数设置下, 当 $\sigma_3, \gamma \geq \sqrt{2} \cdot \omega(\sqrt{\log m})$ 时, 不等式 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathbb{Z}^m)$ 恒成立. 注意到 $\sigma_1 = \sqrt{\gamma^2 + \sigma_3^2}$, 因此,

- 当选取参数 $\sigma_3 = \sigma_4 =: \sigma$ 时,可以将标准 (即秘密与误差服从相同参数的离散高斯分布) 的 $DLWE_{\mathbb{Z},m,q,D_{\mathbb{Z}^m,\sigma}}^d(D_{\mathbb{Z}^d,\sigma}; U(\mathbb{Z}_q^{m \times d}))$ 问题归约到公开矩阵服从任意 η -半均匀分布 \mathcal{D} 的半均匀 $DLWE_{\mathbb{Z},m,q,\chi_1}^d(\chi_2; \mathcal{D})$ 问题. 其中, $\chi_1 = D_{\mathbb{Z}^m,\sigma_1}$, $\sigma_1 = O(\eta \cdot \sigma)$; $\chi_2 = D_{\mathbb{Z}^d,\sigma_2}$, $\sigma_2 = O(\sigma)$.

值得指出的是,在不计多项式规模样本数的情况下,误差 e 所服从的高斯分布的参数大小决定了对应 LWE 问题的困难程度 (例如,采用文献 [36] 的方法可以将秘密取自均匀分布的 LWE 问题归约到秘密取自与误差相同参数的离散高斯分布的 LWE 问题,即所谓的标准形式 (normal form) 的 LWE 问题;而秘密取自均匀分布的 LWE 问题是同等条件下最难的一类问题). 因此,我们在这里仅举例说明 $\sigma_3 = \sigma_4$ 的情形.

与文献 [4] 的归约结果相比,本文的归约结果是保持格的维数 d 的,但是文献 [4] 中的归约结果需要依赖于相应的低维 LWE 问题的困难性. 同时,在我们的归约结果中,在渐进意义下两个问题的误差参数大小的归约损失为 η ,与样本数目无关. 但是文献 [4] 中的归约结果在渐进意义下两个问题的误差参数大小的归约损失与样本数相关 (约为 $\sigma_1 = \tilde{O}(\sqrt{m} \cdot \sigma + \eta)$). 不过值得注意的是,本文的归约结果仅仅局限于秘密/误差服从离散高斯分布的情形,且仅适用于判定版本 LWE 问题的归约. 而文献 [4] 的归约同时适用于求解和判定版本问题的归约,但是对于判定版本的对应问题,需要额外限制 q 为素数或者要求秘密 s 取自 $\{0,1\}$ 上的、特定条件下熵足够大的二元分布.

假设 $0 < p \leq q$ 为正整数,对任意的矩阵 $U \in \mathbb{Z}_q^{m \times d}$,可以定义固定函数 $E_U = \left\lfloor \frac{q}{p} \cdot \left\lfloor \frac{p}{q} \cdot U \right\rfloor \right\rfloor - U$. 对任意的 $x \in \mathbb{Z}_q$,简单计算知 $|x - \lfloor \frac{q}{p} \cdot \lfloor \frac{p}{q} \cdot x \rfloor | \leq \lfloor \frac{q}{2p} \rfloor$. 因此,我们有 $\varepsilon_1(E_U) \leq \sqrt{m \cdot d} \cdot \lfloor \frac{q}{2p} \rfloor$. 定义分布 $\mathcal{D}_{p,q,m,d}$ 为: 取样 $U \in \mathbb{Z}_q^{m \times d}$, 输出 $\lfloor \frac{q}{p} \cdot \lfloor \frac{p}{q} \cdot U \rfloor \rfloor$, 则 $\mathcal{D}_{p,q,m,d}$ 为特殊的 $\sqrt{m \cdot d} \cdot \lfloor \frac{q}{2p} \rfloor$ -半均匀分布. 上述讨论的结论可以自然地应用到 $\mathcal{D}_{p,q,m,d}$ 对应的半均匀 LWE 问题.

当 $\mathfrak{R} = \mathcal{R}$ 时,可以得到对应的判定版本的环/模半均匀 LWE 问题的困难性归约.

推论 7: 假设 $\varepsilon = \text{negl}(\lambda)$ 为某可忽略的函数, m, q, d 为正整数, \mathcal{D} 为 $\mathcal{R}_q^{m \times d}$ 上的 η -半均匀分布, 正实数 $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \gamma \geq \sqrt{2} \cdot \omega(\max\{\sqrt{\log m \cdot n}, \sqrt{\log d \cdot n}\})$ 且满足条件 $\frac{\gamma \cdot \sigma_3}{\sqrt{\gamma^2 + \sigma_3^2}} \geq \eta_\varepsilon(\mathcal{R}^m)$, $\sigma_1 = \sqrt{\gamma^2 + \sigma_3^2}$, $\frac{1}{\sigma_2^2} + \frac{\eta^2}{\gamma^2} < \frac{1}{2 \cdot \sigma_4^2}$. 记 $\chi_1 = D_{\mathcal{R}^m,\sigma_1}$, $\chi_2 = D_{\mathcal{R}^d,\sigma_2}$, $\chi_3 = D_{\mathcal{R}^m,\sigma_3}$, $\chi_4 = D_{\mathcal{R}^d,\sigma_4}$. 如果存在一个 (量子) 概率多项式时间的敌手 \mathfrak{A} 可以以 δ 的概率解决 $DLWE_{\mathcal{R},m,q,\chi_1}^d(\chi_2; \mathcal{D})$ 问题, 则存在一个 (量子) 概率多项式时间的敌手 \mathfrak{B} 可以以 $\delta - \text{negl}(\lambda)$ 的概率解决 $DLWE_{\mathcal{R},m,q,\chi_3}^d(\chi_4; U(\mathcal{R}_q^{m \times d}))$ 问题.

类似于欧式格中的半均匀 LWE 问题的参数分析, 可以推出

- 当选取参数 $\sigma_3 = \sigma_4 =: \sigma$ 时,可以将标准的 $DLWE_{\mathcal{R},m,q,D_{\mathcal{R}^m,\sigma}}^d(D_{\mathcal{R}^d,\sigma}; U(\mathcal{R}_q^{m \times d}))$ 问题归约到公开矩阵服从任意 η -半均匀分布 \mathcal{D} 的半均匀 $DLWE_{\mathcal{R},m,q,\chi_1}^d(\chi_2; \mathcal{D})$ 问题. 其中, $\chi_1 = D_{\mathcal{R}^m,\sigma_1}$, $\sigma_1 = O(\eta \cdot \sigma)$; $\chi_2 = D_{\mathcal{R}^d,\sigma_2}$, $\sigma_2 = O(\sigma)$.

注意到当 $d = 1$ 时,上述结果即为对应的环上的半均匀 LWE 问题的困难性归约结果. 与文献 [17] 的结论相比,我们可以将标准的环 LWE 问题归约到对应的半均匀环 LWE 问题,而不需要使用额外的困难性假设 (例如某些特定形式的 DSPR (decisional small polynomial ratio) 假设,即对应为判定版本的 NTRU 假设). 当 $d \geq 2$ 时,与文献 [4] 的结果相比,在渐近的意义下,我们的误差参数的归约损失仅为 η ,与其余参数无关. 特别地,我们的归约也是保持格的维数的,不需要像文献 [4, 17-18] 那样基于低维度 (更确切地说,低秩数) 的模 LWE 问题的困难性.

当定义分布 $\mathcal{D}_{p,q,m,d}$ 为: 取样 $U \in \mathcal{R}_q^{m \times d}$, 输出 $\lfloor \frac{q}{p} \cdot \lfloor \frac{p}{q} \cdot U \rfloor \rfloor$ 时,简单计算易知 $\mathcal{D}_{p,q,m,d}$ 为特殊的 $n \cdot \sqrt{m \cdot d} \cdot \lfloor \frac{q}{2p} \rfloor$ -半均匀分布. 利用分布 $\mathcal{D}_{p,q,m,d}$ 对应的半均匀模 LWE 问题可以证明,当采用类似 NIST 第一轮候选算法 Kyber 那种压缩公钥的设计方式时,对应的底层加密方案也满足 IND-CPA 安全性. 具体方案的构造如下. 由于相关方案的 IND-CPA 安全性的证明是标准的,我们仅给出大概解释,证明的具体细节可以参考文献 [4, 25-29].

KeyGen(1^λ): 给定安全参数 λ , 密钥生成算法采样 $A \leftarrow U(\mathcal{R}_q^{d \times d})$, $s \leftarrow D_{\mathcal{R}^d,\sigma_2}$, $e \leftarrow D_{\mathcal{R}^d,\sigma_1}$. 计算 $b = A \cdot s + e$, 并返回公钥 $pk = (A, \lfloor \frac{p}{q} \cdot b \rfloor)$ 和私钥 $sk = s$.

Enc(pk, m): 假设 $pk = (A, \lfloor \frac{p}{q} \cdot b \rfloor)$, 要加密消息 $\sigma_c(m) \in \{0,1\}^n$, 加密算法先计算 $\tilde{b} = \lfloor \frac{q}{p} \cdot \lfloor \frac{p}{q} \cdot b \rfloor \rfloor$. 随后, 采样 $r \leftarrow D_{\mathcal{R}^d,\sigma_2}$, $e_1 \leftarrow D_{\mathcal{R}^d,\sigma_2}$, $e_2 \leftarrow D_{\mathcal{R},\sigma_1}$, 并计算 $c_1 = A^T \cdot r + e_1$ 和 $c_2 = \tilde{b}^T \cdot r + e_2$. 最后返回密文 $ct = (c_1, c_2)$.

Dec(sk, ct): 假设 $ct = (c_1, c_2)$, $sk = s$. 解密算法计算并返回 $m = \lfloor \frac{2}{q} \cdot (c_2 - s^T \cdot c_1) \rfloor$.

简单计算可知,对于适当选择的参数,上述方案可以正确解密. 为证明方案的 IND-CPA 安全性,需要用

到两次相应判定版本的 LWE 假设. 第一次是将 (A, \mathbf{b}) 替换为均匀随机的 (A, \mathbf{u}) , 用到的假设是 $\text{DLWE}_{\mathcal{R}, d, q, D_{\mathcal{R}, \sigma_1}}^d(D_{\mathcal{R}, \sigma_2}; \mathbf{U}(\mathcal{R}_q^{d \times d}))$ 问题是困难的; 第二次是将密文替换为随机均匀的元素, 其中替换密文 \mathbf{c}_1 用到的假设仍是 $\text{DLWE}_{\mathcal{R}, d, q, D_{\mathcal{R}, \sigma_1}}^d(D_{\mathcal{R}, \sigma_2}; \mathbf{U}(\mathcal{R}_q^{d \times d}))$ 问题是困难的. 但是, 替换密文 \mathbf{c}_2 需要用到的困难性假设是 $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma_1}}^d(D_{\mathcal{R}, \sigma_2}; \mathcal{D}_{p, q, 1, d})$. 这里, $\mathcal{D}_{p, q, 1, d}$ 对应的分布为: 取样 $\mathbf{u} \leftarrow \mathbf{U}(\mathcal{R}_q^d)$, 输出 $\lfloor \frac{q}{p} \cdot \lfloor \frac{p}{q} \cdot \mathbf{u} \rfloor \rfloor$. 注意到分布 $\mathcal{D}_{p, q, 1, d}$ 是 $n \cdot \sqrt{d} \cdot \lfloor \frac{q}{2p} \rfloor$ -半均匀分布, 采用本文的归约结果可知, $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma_1}}^d(D_{\mathcal{R}, \sigma_2}; \mathcal{D}_{p, q, 1, d})$ 问题的困难性可以由 $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma}}^d(D_{\mathcal{R}, \sigma}; \mathbf{U}(\mathcal{R}_q^d))$ 问题的困难性来保证. 对应的参数关系是 $\sigma_1 = O(n \cdot \sqrt{d} \cdot \lfloor \frac{q}{2p} \rfloor \cdot \sigma)$, $\sigma_2 = O(\sigma)$. 为兼顾解密错误率和安全强度, 参数 p, q 相差的比特数一般为 $O(1)$. 采用文献 [27] 的经验估计, 非对称 LWE 问题的困难性与误差参数约为 $\sqrt{\sigma_1 \cdot \sigma_2} = O(n^{\frac{1}{2}} \cdot d^{\frac{1}{4}}) \cdot \sigma$ 的对称 LWE 问题的困难性相当. 因此, 当将本文的归约结果应用于 Kyber 类采用压缩公钥的设计方式来设计的公钥加密方案时, 对应的理论归约损失约为 $O(n^{\frac{1}{2}} \cdot d^{\frac{1}{4}})$. 值得指出的是, 本文的归约方式对于参数 d (和 q) 的限制很小. 可以选择 $d = O(1)$, 此时对应的理论归约损失约为 $O(n^{\frac{1}{2}})$.

我们采用文献 [4] 中的 5.2 小节的定理来估计采用上述设计方式的公钥加密方案 IND-CPA 安全性的理论归约损失. 此时, $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma_1}}^d(D_{\mathcal{R}, \sigma_2}; \mathcal{D}_{p, q, 1, d})$ 问题的困难性可以由 $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma}}^k(\mathbf{U}(\mathcal{R}_q^k); \mathbf{U}(\mathcal{R}_q^k))$ 问题的困难性和 $\text{DLWE}_{\mathcal{R}, 1, q, D_{\mathcal{R}, \sigma_1}}^d(\mathbf{U}(\mathcal{R}_q^d); \mathbf{U}(\mathcal{R}_q^d))$ 问题的困难性来保证. 这里, $\sigma_1 = O(n \cdot \sqrt{d} \cdot \lfloor \frac{q}{2p} \rfloor \cdot \sigma + \sigma^2)$. 如果我们也选择 $\sigma_2 = O(\sigma)$ 的话, 根据离散高斯的性质大致估计文献 [4] 中的熵不等式条件可以得到对应模 LWE 问题的秩至少应该满足 $d \geq O(k \cdot \log q + \log n + \omega(\log \lambda))$ (注意到即使是将 $D_{\mathcal{R}, \sigma_2}$ 替换为随机均匀分布, 这个条件也应该满足). 因此严格来讲, 文献 [4] 的归约结果不适用于 $d = O(1)$ 的情形. 此外, 文献 [4] 的归约结果对于参数 q 以及秘密分布也有额外的限制. 例如要求 q 为分裂性质良好的素数, 且要求秘密的分布模 $q\mathcal{R}$ 的每一个素理想均具有足够大的熵. 如果加上这些限制, 文献 [4] 的归约结果将更加受限.

4 总结

在本文中, 我们将 Hint-LWE 相关问题研究中的方法进行改进并应用到半均匀 LWE 问题的研究之中, 给出了非均匀 LWE 问题更简单、更紧致的困难性归约. 具体来讲, 本文的归约方法几乎不受代数结构的影响, 可以一致地应用到对应的欧式格/环/模上的半均匀 LWE 问题的归约. 相比于已知的半均匀 LWE 问题的归约结果, 本文给出的归约是保持对应 LWE 问题的维数的, 归约损失小且归约损失 (对应的 LWE 问题的误差所服从的离散高斯参数) 与样本数等条件无关. 当将本文的归约应用到环 LWE 问题时, 对应的归约方法可以基于标准的环 LWE 问题给出非均匀环 LWE 问题的困难性证明, 归约结论无需用到 DSPR 等额外的 (相对而言非标准的) 困难性假设.

References:

- [1] Regev O. On lattice, learning with errors, random linear codes, and cryptography. In: Proc. of the 37th Annual ACM SIGACT Symp. on Theory of Computing – STOC 2005, ACM, 2005, 84 - 93. doi:10.1145/1060590.1060603
- [2] Lyubashevsky V, Peikert C, Regev O. On ideal lattices and learning with errors over rings. In: Gilbert H, ed. Advances in Cryptology – CRYPTO 2010, vol. 6110, Springer Berlin Heidelberg, 2010, 1 - 23. doi: 10.1007/978-3-642-13190-5_1
- [3] Langlois A, Stehle D. Worst-case to average-case reductions for module lattices. Designs, Codes and Cryptography, 2015, 75(3): 565-599. doi: 10.1007/s10623-014-9938-4
- [4] Jia WJ, Zhang J, Xiang BW, Wang BC. Hardness of (M)LWE with semi-uniform seeds. Theoretical Computer Science, 2024, 994: 114481. doi: 10.1016/j.tcs.2024.114481
- [5] Brakerski Z, Langlois A, Peikert C, Regev O, Stehle D. Classical hardness of learning with errors. In: Proc. of the 45th annual ACM symp. on Theory of Computing – STOC 2013, ACM, 2005, 575 - 584. doi: 10.1145/2488608.2488680
- [6] Goldwasser S, Kalai YT, Peikert C, Vaikuntanathan V: Robustness of the learning with errors assumption. In: Innovations in Computer Science – ICS 2010, Tsinghua University Press, 2010, pp. 230–240. url: <http://hdl.handle.net/1721.1/73191>
- [7] Micciancio D. On the hardness of learning with errors with Binary Secrets. Theory of Computing, 2018, 14(13): 1–17. doi: 10.4086/toc.2018.v014a013

- [8] Brakerski Z, Dottling N: Hardness of LWE on general entropic distributions. In: Canteaut A, Ishai Y, ed. *Advances in Cryptology – EUROCRYPT 2020*, vol. 12106, Springer Berlin Heidelberg, 2020, 551–575. doi: 10.1007/978-3-030-45724-2_19
- [9] Boudgoust K, Jeudy C, Roux-Langlois A, Wen WQ. On the hardness of module learning with errors with short distributions. *Journal of Cryptology*, 2023, 36(1). doi: 10.1007/s00145-022-09441-3
- [10] Kim D, Lee D, Seo J, Song Y. Toward practical lattice-based proof of knowledge from hint-MLWE. In: Handschuh H, Lysyanskaya A, ed. *Advances in Cryptology – CRYPTO 2023*, vol. 14085, Springer Berlin Heidelberg, 2023, 549 - 580. doi: 10.1007/978-3-031-38554-4_18
- [11] Dottling N, Kolonelos D, Lai R, Lin CW, Malavolta G, Rahimi A. Efficient laconic cryptography from learning with errors. In: Hazay C, Stam M, ed. *Advances in Cryptology – EUROCRYPT 2023*, vol. 14006, Springer Berlin Heidelberg, 2023, 417 - 446. doi: 10.1007/978-3-031-30620-4_14
- [12] Agrawal S, Libert B, Stehle D. Fully secure functional encryption for inner products, from standard assumptions. In: Robshaw M, Katz J, ed. *Advances in Cryptology – CRYPTO 2016*, vol. 9816, Springer Berlin Heidelberg, 2023, 333 - 362. doi: 10.1007/978-3-662-53015-3_12
- [13] Mera JMB, Karmakar A, Marc T, Soleimani A. Efficient lattice-based inner-product functional encryption. In: Hanaoka G, Shikata J, Watanabe Y, ed. *Public-Key Cryptography – PKC 2022*, vol. 13178, Springer Berlin Heidelberg, 2022, 163 - 193. doi: 10.1007/978-3-030-97131-1_6
- [14] Alperin-Sheriff J, Peikert C. Circular and KDM security for identity-based encryption. In: Fischlin M, Buchmann J, Manulis M, ed. *Public-Key Cryptography – PKC 2012*, vol. 7293, Springer Berlin Heidelberg, 2012, 334 - 352. doi: 10.1007/978-3-642-30057-8_20
- [15] O’Neill A, Peikert C, Waters B. Bi-deniable public-key encryption. In: Rogaway P, ed. *Advances in Cryptology – CRYPTO 2011*, vol. 6841, Springer Berlin Heidelberg, 2011, 525 - 542. doi: 10.1007/978-3-642-22792-9_30
- [16] Boneh D, Lewi K, Montgomery H, Raghunathan A. Key homomorphic PRFs and their applications. In: Canetti R, Garay JA, ed. *Advances in Cryptology – CRYPTO 2013*, vol. 8042, Springer Berlin Heidelberg, 2013, 410–428. doi: 10.1007/978-3-642-40041-4_23
- [17] Jia WJ, Zhang J, Wang BC. Hardness of module-LWE with semiuniform seeds from module-NTRU. *IET Information Security*, 2023: 2969432. doi: 10.1049/2023/2969432
- [18] Jia WJ, Wang BC. Hardness of (Semiuniform) MLWE with short distributions using the Renyi divergence. *IET Information Security*, 2023: 2104380. doi: 10.1049/2023/2104380
- [19] NIST: Post-Quantum Cryptography - Round 1 Submissions. <https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-1-submissions>.
- [20] Bos J, Ducas L, Kiltz E, Lepoint T, Lyubashevsky V, Mchanck JM, Schwabe P, Seiler G, Stehle D. CRYSTALS – Kyber: a cca-secure module lattice-based KEM. In: *IEEE European Symp. on Security and Privacy – Euro S&P 2018*, IEEE, 353 - 367. doi: 10.1109/EuroSP.2018.00032
- [21] Lindner R, Peikert C. Better key sizes (and attacks) for LWE-based encryption. In: Kiayias A, ed. *Topics in Cryptology – CT-RSA 2011*, vol. 6558, Springer Berlin Heidelberg, 2011, 319 - 339. doi: 10.1007/978-3-642-19074-2_21
- [22] NIST: Post-Quantum Cryptography - Round 2 Submissions. <https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-2-submissions>
- [23] NIST: Post-Quantum Cryptography - Round 3 Submissions. <https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions>
- [24] NIST: Post-Quantum Cryptography - Selected Algorithms 2022. <https://csrc.nist.gov/Projects/post-quantum-cryptography/selected-algorithms-2022>
- [25] Jin ZZ, Zhao YL. Optimal key consensus in presence of noise. *Cryptology ePrint Archive*, ePrint 2017/1058.
- [26] Jin ZZ, Zhao YL. Generic and practical key establishment from lattices. In: Deng R, Gauthier-Umana V, Ochoa M, Yung M, ed. *Applied Cryptography and Network Security – ACNS 2019*, vol. 11464, Springer Berlin Heidelberg, 2019, 302 - 322. doi: 10.1007/978-3-030-21568-2_15

- [27] Zhang J, Yu Y, Fan SQ, Zhang ZF, Yang K. Tweaking the asymmetry of asymmetric-key cryptography on lattices: KEMs and signatures of smaller sizes. In: Kiayias A, Kohlweiss M, Wallden P, Zikas V, ed. *Public-Key Cryptography – PKC 2020*, vol. 12111, Springer Berlin Heidelberg, 2020, 37 – 65. doi: 10.1007/978-3-030-45388-6_2
- [28] Chinese Association for Cryptologic Research: Public key algorithms selected to the second round competition of national cryptographic algorithm competitions. http://sfjs.cacernet.org.cn/site/term/list_77_1.html.
- [29] Chinese Association for Cryptologic Research: Announcement of the selection results of the national cryptographic algorithm competition. <https://www.cacernet.org.cn/site/content/854.html>.
- [30] Micciancio D. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 2007, 16: 365 - 411. doi: 10.1007/s00037-007-0234-9
- [31] Bruna J, Regev O, Song MJ, Tang Y. Continuous LWE. In: *Proc. of the 53th Annual ACM SIGACT Symp. on Theory of Computing – STOC 2021*, ACM, 2021, 694 - 707. doi: 10.1145/3406325.3451000
- [32] Gupte A, Vafa N, Vaikuntanathan V. Continuous LWE is as hard as LWE & applications to learning Gaussian mixtures. In: *IEEE 63th Annual symp. on Foundations of Computer Science – FOCS 2022*, Denver, 2022, 1162 - 1173. doi: 10.1109/FOCS54457.2022.00112
- [33] Peikert C. An efficient and parallel Gaussian sampler for lattices. In: Rabin T, ed. *Advances in Cryptology – CRYPTO 2010*, vol. 6223, Springer Berlin Heidelberg, 2010, 80 - 97. doi: 10.1007/978-3-642-14623-7_5
- [34] Gentry C, Peikert C, Vaikuntanathan V. Trapdoors for hard lattices and new cryptographic constructions. In: *Proc. of the 40th annual ACM symp. on Theory of Computing – STOC 2008*, ACM, 2008, 197 – 206. doi: 10.1145/1374376.1374407
- [35] Micciancio D, Regev O. Worst-case to average-case reductions based on Gaussian measures. *SIAM Journal on Computing*, 2007, 37 (1): 267 - 302. doi: 10.1137/S0097539705447360
- [36] Applebaum B, Cash D, Peikert C, Sahai A. Fast cryptographic primitives and circular secure encryption based on hard learning problems. In: Halevi S, ed. *Advances in Cryptology – CRYPTO 2009*, vol. 5677, Springer Berlin Heidelberg, 2009, 595 – 618. doi: 10.1007/978-3-642-03356-8_35

附中文参考文献:

- [28] 中国密码学会: 全国密码算法设计竞赛进入第二轮公钥算法. http://sfjs.cacernet.org.cn/site/term/list_77_1.html.
- [29] 中国密码学会: 关于全国密码算法设计竞赛算法评选结果的公示. <https://www.cacernet.org.cn/site/content/854.html>.