ABSTRACT

Title of dissertation: Improved Robustness and Versatility of Lattice-Based Cryptography

Huijing Gong
Doctor of Philosophy, 2021

Dissertation directed by: Dana Dachman-Soled
Department of Electrical and Computer Engineering

Current public key cryptosystems that are based on the hardness of integer factorization and discrete logarithm are insecure in the presence of large-scale quantum computers. Much effort has been devoted to replacing the quantum-insecure cryptosystems with newly developed "post-quantum" cryptosystem candidates, conjectured to be secure against quantum attack. Lattice-based cryptography has been widely recognized as a prominent candidate for practical post-quantum security.

This dissertation improves the robustness and versatility of lattice-based cryptography through the following three contributions:

1. Chapter 3 introduces a constant-round protocol for unauthenticated group key exchange (i.e., with security against a passive eavesdropper). Group key-exchange protocols allow a set of N parties to agree on a shared, secret key by communicating over a public network. Our protocol is based on the hardness of a lattice problem, which hence yields (plausible) post-quantum security.

2. In Chapter 4, we propose a framework for cryptanalysis of lattice-based schemes when certain types of information about the secret are leaked. Our framework generalizes the primal lattice reduction attack. The generalization allows for integrating the leaked information progressively before running a final lattice reduction step. Our framework can estimate the amount of security loss caused by the leaked information, and perform lattice reduction attacks with leaked information when computationally feasible.

3. Chapter 5 introduces an approach towards a ring analogue of the Leftover Hash Lemma (LHL). The LHL is a mathematical tool often used in the analysis of various lattice-based cryptosystems, as well as their leakage-resilient counterparts. However, it does not hold in the ring setting, which is typical for efficient cryptosystems. Lyubashevsky et al. (Eurocrypt '13) proved a "regularity lemma," which is used in the ring setting instead of the LHL; however, this applies only for centered, spherical Gaussian inputs, while the LHL applies when the input is drawn from any high min-entropy distribution. Our approach generalizes the "regularity lemma" of Lyubashevsky et al. to certain conditional distributions. A number of Ring-Learning with Errors based cryptosystems can achieve certain leakage resilience properties using our results.

Improved Robustness and Versatility of Lattice-Based Cryptography

by

Huijing Gong

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2021

Advisory Committee:
Professor Dana Dachman-Soled, Chair/Advisor
Dr. Gorjan Alagic
Professor William Gasarch
Professor Michelle Mazurek
Professor Lawrence C. Washington

Dedication

I dedicate this dissertation to my whole beloved family.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my doctoral advisor, Professor Dana Dachman-Soled, for her continuous support and encouragement over the years. Her deep expertise and extensive knowledge of the field provided me with invaluable guidance. I have enjoyed numerous helpful and thought-provoking conversations with her. Her insightful feedback pushed me to sharpen my thinking and brought my work to a higher level. In addition, her diligence and perseverance in research have set a strong inspiration for the rest of my career. Most importantly, she has always been patient with and accessible to me and has shown confidence in my ability, even when I didn't have much confidence in myself. This dissertation would not have been possible without her support and nurturing.

I am very grateful to all of my committee members, Dana Dachman-Soled, Gorjan Alagic, William Gasarch, Michelle Mazurek, and Lawrence Washington, for taking the time to review my dissertation and participate in my defense.

I would like to extend my gratitude to my amazing groupmates, Mukul Kulkarni and Aria Shahverdi. I had the great pleasure of working with them and received an enormous amount of helpful advice and support from them. I feel most grateful and fortunate for the moment that I asked to join Dana's group, Dana, Aria, and Mukul extended their helpful and warm hands to me – a lonely, desperate explorer – and pulled me from the marsh. I gratefully acknowledge the assistance of Dana, Mukul, Aria and Hunter Michael Kippen in giving me plenty of helpful advice on

how to improve my presentations and slides.

I also want to thank my fellow graduate students and all my collaborators. Special thanks goes to Aishwarya Thiruvengadam and Daniel Apon for the helpful discussions and suggestions for my research. Thanks should also go to Léo Ducas and Mélissa Rossi for getting the opportunity to work on the LWE with side information project. I gained a tremendous amount from this collaboration and learned to look at some research topics from a different perspective. I would like to thank Yilei Chen, Daniel Masny, and Thuy-Duong Vuong for the interesting conversations and collaborations throughout my internship at Visa Research. In particular, I want to thank Yilei Chen for his mentorship, his career advice, and his patient, insightful introduction to quantum hardness reduction. In addition, I'm grateful to friends and researchers in the crypto community including but not limited to Wen Wang, Zhenfei Zhang, and Wei Dai who generously provided me with a wealth of information and advice on finding a job.

I would like to express my sincere gratitude to my undergraduate advisor Professor Kenneth Ribet. He opened the door to the wonderful world of cryptology to me. I am also very much blessed and grateful to be taught by the awesome teachers and professors: Huang Yuling, Liu Hanlu and Professor Paul Yun. Without their excellent teaching, guidance, and support, I would not be where I am now.

I'm also grateful to precious decade-long sisterhood of friends, Chen Jiexiu, Dong Xiaoyu, Liu Ruixue, for discussing life and the state of the world and for sharing "I'm doomed!" moments since high school. Luckily, Jiexiu and I both have been pursuing doctoral degrees in the past few years, so we could also share our

Ph.D. struggles. This sharing has provided me with great mental support. I was also fortunate to live with my wonderful roommates Qiwen Zheng, and Xinyi Wang. I very much enjoyed "hanging out" for food with them. During the "dark" period of my Ph.D., whenever I went back to our apartment and opened the door, I felt warm and less stressed.

I want to sincerely thank my parents, whose love and support are with me in whatever I pursue. When, as a child, I was particularly fond of *Wind Watcher*, a drama about cryptologists, my parents encouraged me to study math well if I wanted to be a cryptographer. I wish to thank my loving and supportive husband, Yiming. The best thing that happened during my graduate program was that I met Yiming. I want to thank Yiming for the love and companionship. I cherish the time we spent together as graduate students, during which, we were sometimes happy together, sometimes struggling together, sometimes pushing each other to study hard, and sometimes lazy together. We supported and encouraged each other through this challenging but wonderful journey.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

As Shor's algorithm solves integer factorization and discrete logarithm in polynomial time [106], all public key cryptosystems that are based on hardness of integer factorization and discrete logarithm are insecure in the presence of large-scale quantum computers. Much effort has been made to replace standardized public key cryptosystems, which are quantum-insecure, with newly developed *post-quantum* cryptosystems, conjectured to be secure against quantum attack.

One promising candidate for practical, post-quantum cryptography are the cryptosystems based on the hard lattice problems – lattice-based cryptography. There are several lattice problems, which have been extensively studied for decades, that are believed to be hard, even against a quantum computer; for example, shortest vector problem (SVP), closest vector problem (CVP), etc. Cryptosystems have been built from such problems, beginning from a seminal work by Ajtai [3]. One relevant hard lattice problem to our work is the Learning with Errors (LWE) problem introduced by Regev [103]. The (Decisional) LWE problem is defined as the problem of distinguishing between the two distributions $(A, A\mathbf{s} + \mathbf{e})$ and $(A, \mathbf{u})$, where $\mathbf{s}$ is a secret vector, matrix A and vector $\mathbf{u}$ are uniform, and vector $\mathbf{e}$ has a small norm. LWE is proved to be as hard to solve as several worst-case standard lattice prob-

lems (e.g. [95,103]). To improve the efficiency of lattice-based cryptosystems, lattices with additional algebraic structure were introduced [84,109]. Specifically, a ring version of the LWE (Ring-LWE) problem is defined as the problem of distinguishing $(a, b = a \cdot s + e) \in R_q \times R_q$ from random pairs, where $s \in R_q$ is a secret, $a \in R_q$ is uniform and the error term $e \in R$ has a small norm, where $R_q := \mathbb{Z}_q[x]/x^n + 1$.

Lattice-based cryptography has been recognized for its versatility in realizing cryptographic applications. Hardness of lattice problems, especially LWE and Ring-LWE, have been relied upon as security assumptions for key exchange (e.g. [9]), public key encryption (e.g., [85]), digital signature (e.g., [56]), pseudorandom functions (e.g., [12]), attribute-based encryption (e.g., [21]), fully homomorphic encryption (e.g., [29]), non-interactive zero-knowledge [98], etc.

NIST has initiated a standardization process for quantum-resistant public-key cryptographic algorithms. One concern that arises is the potential security loss in the process of transitioning theoretical work with provable security into practice. Although the best-known algorithms for solving the LWE problem run in exponential time, faulty parameter instantiation, incorrect implementation, or attacks on implementation may lead to severe security risks, which is often not considered in the scenarios of the original provable security claim. Such scenarios are referred to as "imperfect" scenarios. Various efforts have been made to analyze security of lattice schemes under "imperfect" scenarios (see [6, 64, 97] for examples).

1.1 Our Contributions

This dissertation improves the robustness and versatility of lattice-based cryptography through analyzing the security impact of potential information leakage on lattice cryptosystems [44, 47] and designing a lattice-based crypto application [10].

1.1.1 Lattice Cryptographic Protocol Construction

Group Key Exchange Protocol. In Chapter 3, we propose a constant-round group key exchange protocol based on the hardness of the Ring-LWE problem. Group key exchange is a protocol that allows a set of N parties to agree on a shared, secret key by communicating over a public network. A number of solutions to this problem have been proposed over the years, mostly based on variants of Diffie-Hellman (two-party) key exchange. There has been relatively little work, however, looking at candidate post-quantum group key-exchange protocols. Our protocol is constructed by generalizing the Burmester-Desmedt protocol to the Ring-LWE setting, which requires addressing several technical challenges.

This work was originally published in PQC 2019 [10], in collaboration with Daniel Apon, Dana Dachman-Soled, and Jonathan Katz.

1.1.2 Security of Lattice-Based Cryptography under "Imperfect" Scenarios

Security of Learning with Errors with Side Information. In Chapter 4, we propose a framework for the cryptanalysis of lattice-based schemes for which certain types of side information about the secret and/or error are available.

While there are many prior works addressing the concern of security impact under "imperfect" scenarios for specific cryptosystems (See [6, 25] for side channel attacks examples), these prior works use either ad-hoc statistical methods to recover the secret key, requiring new techniques to be developed for each setting, or require substantial amounts of information leakage to efficiently recover the secret. As a general framework, our framework can estimate how much the leaked information reduces the security level, with no requirement for the amount of information leakage.

Our framework generalizes the primal lattice reduction attack and allows the progressive integration of the side information before running a final lattice reduction step. Our techniques for integrating side information include sparsifying the lattice, projecting onto and intersecting with hyperplanes, and/or altering the distribution of the secret vector. Our main contribution is to propose a toolbox and a methodology to integrate such information into lattice reduction attacks and to predict the cost of these lattice attacks with side information. In addition, we implement our framework on Sage 9.0 and provide three examples that exploit side information leaked through side-channel information, decryption failures, and constraints imposed by certain

schemes (LAC, Round5, NTRU).

This work was originally published in Crypto 2020 [44], in collaboration with Dana Dachman-Soled, Léo Ducas, and Mélissa Rossi.

A Generalized Regularity Lemma over Ideal Lattice. In Chapter 5, we focus on proving a ring analogue of the leftover hash lemma (LHL). LHL is used in the analysis of various integer lattice-based cryptosystems as well as their leakage-resilient counterparts; however, it does not hold in the ring setting, which is typical for efficient cryptosystems. Lyubashevsky et al. (Eurocrypt '13) proved a "regularity lemma," which can be used instead of the LHL, but applies only for centered, spherical Gaussian inputs, while LHL applies when the input is drawn from any high min-entropy distribution. We present an approach for generalizing the "regularity lemma" of Lyubashevsky et al. to certain conditional distributions.We assume the input was sampled from a discrete Gaussian distribution and consider the induced distribution, given side-channel leakage on the input. We present three instantiations of our approach, proving that the regularity lemma holds for three natural conditional distributions. Since applications of the regularity lemma in lattice-based cryptography are widespread, a number of Ring-LWE cryptosystems can achieve certain leakage resilience properties using our results.

This work was originally published in MathCrypt 2019 [47], in collaboration with Dana Dachman-Soled, Mukul Kulkarni, and Aria Shahverdi.

Chapter 2: Preliminary

2.1 Notation

Let $\mathbb{Z}$ be the ring of integers, and let $[N] = \{1, \ldots, N\}$ for a positive integer N. If S is a set, then $x_1, x_2, \ldots, x_\ell \leftarrow S$ denotes uniformly sampling each x_i from S; if χ is a probability distribution, then $x_1, x_2, \ldots, x_\ell \leftarrow \chi$ denotes independently sampling each x_i according to that distribution. Let $\chi(E)$ denote the probability that event E occurs under distribution χ. We let $\mathrm{Supp}(\chi) = \{x : \chi(x) \neq 0\}$. Given an event E, we let $\overline{E}$ denote its complement. Given a polynomial p_i, let $(p_i)_j$ denote the jth coefficient of p_i. We use $\log(X)$ to denote $\log_2(X)$, and $\exp(X)$ to denote e^X.

We let λ denote a computational security parameter, and ρ a statistical security parameter.

We denote vectors in boldface $\mathbf{x}$ and matrices using capital letters A. For vector $\mathbf{x}$ over $\mathbb{R}^n$ or $\mathbb{C}^n$, define the ℓ_2 norm as $\|\mathbf{x}\|_2 = \left(\sum_i |x_i|^2\right)^{1/2}$. We write this as $\|\mathbf{x}\|$ for simplicity.

2.2 Probability and Statistics

2.2.1 Rényi divergence

For two discrete probability distributions P and Q with $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$, their *Rényi divergence* is defined as

$$\mathrm{RD}_2(P\|Q) = \sum_{x \in \mathrm{Supp}(P)} \frac{P(x)^2}{Q(x)}.$$

We use the following results (see [79, 84, 111] for proofs):

Theorem 2.1. *For discrete distributions P and Q with $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$ and any f, we have*

$$\mathrm{RD}_2(f(P)||f(Q)) \leq \mathrm{RD}_2(P||Q).$$

Theorem 2.2. *For discrete distributions P and Q with $\mathrm{Supp}(P) \subseteq \mathrm{Supp}(Q)$, let $E \subseteq \mathrm{Supp}(Q)$ be an arbitrary event. We have*

$$Q(E) \geq P(E)^2/\mathrm{RD}_2(P||Q).$$

The second property implies, roughly, that as long as $\mathrm{RD}_2(P\|Q)$ is bounded by some polynomial, then any event E that occurs with negligible probability $Q(E)$ under distribution Q also occurs with negligible probability $P(E)$ under distribution P.

The following theorem bounds the Rényi divergence between the 1-dimensional

discrete Gaussian distribution centered at the origin and one centered at a point near the origin.

Theorem 2.3 ([17]). *Fix $m, q, \lambda \in \mathbb{Z}$, a bound $\beta_{\text{Rényi}}$, and σ with $\beta_{\text{Rényi}} < \sigma < q$. Let $e \in \mathbb{Z}$ be such that $|e| \leq \beta_{\text{Rényi}}$. Then*

$$\mathrm{RD}_2((e + D_{\mathbb{Z}_q,\sigma})^m || D^m_{\mathbb{Z}_q,\sigma}) \leq \exp(2\pi m(\beta_{\text{Rényi}}/\sigma)^2).$$

(Here, χ^m denotes m independent samples from distribution χ.)

The above theorem implies that if $\sigma = \Omega(\beta_{\text{Rényi}}\sqrt{m/\log\lambda})$ for some security parameter λ, then $\mathrm{RD}_2((e + D_{\mathbb{Z}_q,\sigma})^m || D^m_{\mathbb{Z}_q,\sigma}) = \mathrm{poly}(\lambda)$.

2.2.2 Statistics

Random variables, i.e., variables whose values depend on outcomes of a random phenomenon, are denoted in lowercase calligraphic letters, e.g., $\mathcal{a}, \mathcal{b}, \mathcal{e}$. Random vectors are denoted in uppercase calligraphic letters, e.g., $\mathcal{C}, \mathcal{X}, \mathcal{Z}$.

Before hints are integrated, we will assume that the secret and error vectors follow a multidimensional normal (Gaussian) distribution. Hints will typically correspond to learning a (noisy, modular or perfect) linear equation on the secret. We must then consider the altered distribution on the secret, conditioned on this information. Fortunately, this will also be a multidimensional normal distribution with an altered covariance and mean. In the following, we present the precise formulae for the covariance and mean of these conditional distributions.

Definition 2.4 (Multidimensional normal distribution)**.** *Let $d \in \mathbb{Z}$, for $\boldsymbol{\mu} \in \mathbb{Z}^d$ and $\boldsymbol{\Sigma}$ being a symmetric matrix of dimension $d \times d$, we denote by $D^d_{\boldsymbol{\Sigma},\boldsymbol{\mu}}$ the multidimensional normal distribution supported by $\boldsymbol{\mu} + \mathrm{Span}(\boldsymbol{\Sigma})$ by the following*

$$\mathbf{x} \mapsto \frac{1}{\sqrt{(2\pi)^{\mathrm{rank}(\boldsymbol{\Sigma})} \cdot \mathrm{rdet}(\boldsymbol{\Sigma})}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}) \cdot \boldsymbol{\Sigma}^{\sim} \cdot (\mathbf{x}-\boldsymbol{\mu})^T\right).$$

The following states how a normal distribution is altered under linear transformation.

Lemma 2.5. *Suppose $\mathcal{X}$ has a $D^d_{\boldsymbol{\Sigma},\boldsymbol{\mu}}$ distribution. Let $\mathbf{A}$ be a $n \times d$ matrix. Then $\mathcal{X}\mathbf{A}^T$ has a $D^n_{\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T,\boldsymbol{\mu}\mathbf{A}^T}$ distribution.*

Lemma 2.6 shows the altered distribution of a normal random variable conditioned on its noisy linear transformation value, following from [82, Equations (6) and (7)].

Lemma 2.6 (Conditional distribution $\mathcal{X}|\mathcal{X}\mathbf{A}^T + \mathcal{b}$ from [82])**.** *Suppose that $\mathcal{X} \in \mathbb{Z}^d$ has a $D^d_{\boldsymbol{\Sigma},\boldsymbol{\mu}}$ distribution, and $\mathcal{b} \in \mathbb{Z}^n$ has a $D^n_{\boldsymbol{\Sigma}_{\mathcal{b}},\mathbf{0}}$ distribution. Let us fix $\mathbf{A}$ as a $n \times d$ matrix and $\mathbf{z} \in \mathbb{Z}^n$. The conditional distribution of $\mathcal{X}\Big|\left(\mathcal{X}\mathbf{A}^T + \mathcal{b} = \mathbf{z}\right)$ is $D^d_{\boldsymbol{\Sigma}',\boldsymbol{\mu}'}$, where*

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + (\mathbf{z} - \boldsymbol{\mu}\mathbf{A}^T)(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\mathcal{b}})^{-1}\mathbf{A}\boldsymbol{\Sigma}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}\mathbf{A}^T(\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\mathcal{b}})^{-1}\mathbf{A}\boldsymbol{\Sigma}.$$

Corollary 2.7 (Conditional distribution $\mathcal{X}|\langle \mathcal{X}, \mathbf{v}\rangle + \mathcal{e}$)**.** *Suppose that $\mathcal{X} \in \mathbb{Z}^d$ has a $D^d_{\boldsymbol{\Sigma},\boldsymbol{\mu}}$ distribution and $\mathcal{e}$ has a $D^1_{\sigma^2_{\mathcal{e}},0}$ distribution. Let us fix $\mathbf{v} \in \mathbb{R}^d$ as a nonzero*

vector and $z \in \mathbb{Z}$. *We define the following scalars:*

$$y = \langle \mathcal{X}, \mathbf{v} \rangle + e,\ \mu_2 = \langle \mathbf{v}, \boldsymbol{\mu} \rangle \text{ and } \sigma_2 = \mathbf{v}\boldsymbol{\Sigma}\mathbf{v}^T + \sigma_e^2$$

If $\sigma_2 \neq 0$, *the conditional distribution of* $\mathcal{X} \Big| \big(y = z\big)$ *is* $D^d_{\boldsymbol{\Sigma}', \boldsymbol{\mu}'}$, *where*

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + \frac{(z - \mu_2)}{\sigma_2}\mathbf{v}\boldsymbol{\Sigma}, \qquad \boldsymbol{\Sigma}' = \boldsymbol{\Sigma} - \frac{\boldsymbol{\Sigma}\mathbf{v}^T\mathbf{v}\boldsymbol{\Sigma}}{\sigma_2}. \tag{2.1}$$

If $\sigma_2 = 0$, *the conditional distribution of* $\mathcal{X} \Big| \big(y = z\big)$ *is* $D^d_{\boldsymbol{\Sigma}, \boldsymbol{\mu}}$.

Remark 2.8. *We note that Corollary 2.7 is also useful to describe for* $\mathcal{X} | \langle \mathcal{X}, \mathbf{v} \rangle$ *by letting* $\sigma_e = 0$.

2.3 Lattice-based Cryptography

2.3.1 Lattice over $\mathbb{R}^m$

A *lattice*, denoted as Λ, is a discrete additive subgroup of $\mathbb{R}^m$, which is generated as the set of all linear integer combinations of n $(m \geq n)$ linearly independent basis vectors $\{\mathbf{b}_j\} \subset \mathbb{R}^m$, namely,

$$\Lambda := \left\{ \sum\nolimits_j z_j \mathbf{b}_j : z_j \in \mathbb{Z} \right\},$$

We say that m is the *dimension* of Λ and n is its rank. A lattice is *full rank* if $n = m$. A matrix $\mathbf{B}$ having the basis vectors as rows is called a *basis*. The *volume* of a lattice Λ is defined as $\mathrm{Vol}(\Lambda) := \sqrt{\det(\mathbf{B}\mathbf{B}^T)}$. The *dual lattice* of Λ in $\mathbb{R}^n$ is

defined as follows.

$$\Lambda^* := \{\mathbf{y} \in \mathrm{Span}(\mathbf{B}) \mid \forall \mathbf{x} \in \Lambda, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}.$$

Note that, $(\Lambda^*)^* = \Lambda$, and $\mathrm{Vol}(\Lambda^*) = 1/\mathrm{Vol}(\Lambda)$.

Lemma 2.9 ([87, Proposition 1.3.4]). *Let Λ be a lattice and let F be a subspace of $\mathbb{R}^n$. If $\Lambda \cap F$ is a lattice, then the dual of $\Lambda \cap F$ is the orthogonal projection onto F of the dual of Λ. In other words, each element of Λ^* is multiplied by the projection matrix $\mathbf{\Pi}_F$:*

$$(\Lambda \cap F)^* = \Lambda^* \cdot \mathbf{\Pi}_F.$$

Lemma 2.10 ([87, Proposition 1.2.9]). *Let Λ be a lattice in $\mathbb{R}^n$, let F be a subspace of $\mathbb{R}^n$ such that $\Lambda \cap F$ is a lattice and let $\mathbf{\Pi}_F^{\perp}$ be the orthogonal projection onto $F^{\perp}$. Then*

$$\mathrm{Vol}(\Lambda \cdot \mathbf{\Pi}_F^{\perp}) = \mathrm{Vol}(\Lambda)(\mathrm{Vol}(\Lambda \cap F)^{-1}).$$

Definition 2.11 (Primitive vectors). *A set of vectors $\mathbf{y}_1, \ldots, \mathbf{y}_k \in \Lambda$ is said to be primitive with respect to Λ if $\Lambda \cap \mathrm{Span}(\mathbf{y}_1, \ldots, \mathbf{y}_k)$ is equal to the lattice generated by $\mathbf{y}_1, \ldots, \mathbf{y}_k$. Equivalently, it is primitive if it can be extended to a basis of Λ. If $k = 1$, this is equivalent to $\mathbf{y}_1/i \notin \Lambda$ for any integer $i \geq 2$.*

2.3.2 Lattice over the space H

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denote the cycle, i.e. the additive group of reals modulo 1. We also denote by $\mathbb{T}_q$ its cyclic subgroup of order q, i.e., the subgroup given by

$\{0, 1/q, \ldots, (q-1)/q\}$.

Let H be a subspace, defined as $H \subseteq \mathbb{C}^{\mathbb{Z}_m^*}$, (for some integer $m \geq 2$),

$$H = \{\mathbf{x} \in \mathbb{C}^{\mathbb{Z}_m^*} : x_i = \overline{x_{m-i}}, \forall i \in \mathbb{Z}_m^*\}.$$

A *lattice* is a discrete additive subgroup of H. We exclusively consider the full-rank lattices, which are generated as the set of all linear integer combinations of some set of n linearly independent *basis* vectors $B = \{\mathbf{b}_j\} \subset H$:

$$\Lambda = \mathcal{L}(B) = \left\{ \sum_j z_j \mathbf{b}_j : z_j \in \mathbb{Z} \right\}.$$

The *determinant* of a lattice $\mathcal{L}(B)$ is defined as $|\det(B)|$, which is independent of the choice of basis B. The *minimum distance* $\lambda_1(\Lambda)$ of a lattice Λ (in the Euclidean norm) is the length of a shortest nonzero lattice vector.

The *dual lattice* of $\Lambda \subset H$ is defined as following, where $\langle \cdot, \cdot \rangle$ denotes the inner product.

$$\Lambda^\vee = \{\mathbf{y} \in H : \forall \mathbf{x} \in \Lambda, \langle \mathbf{x}, \overline{\mathbf{y}} \rangle = \sum_i x_i y_i \in \mathbb{Z}\}.$$

Note that, $(\Lambda^\vee)^\vee = \Lambda$, and $\det(\Lambda^\vee) = 1/\det(\Lambda)$.

2.3.2.1 Discretization

Discretization is an important procedure used in applications based on lattices, such as converting continuous Gaussian distribution (defined in Appendix 2.4.2) into a discrete Gaussian distribution (Definition 2.29). Given a lattice $\Lambda = \mathcal{L}(B)$

represented by some "good" basis $B = \{\mathbf{b}_i\}$, a point $\mathbf{x} \in H$, and a point $\mathbf{c} \in H$ representing a lattice coset $\Lambda+\mathbf{c}$, the discretization process outputs a point $\mathbf{y} \in \Lambda+\mathbf{c}$ such that the length of $\mathbf{y} - \mathbf{x}$ is not too large. This is denoted as $\mathbf{y} \leftarrow \lfloor \mathbf{x} \rceil_{\Lambda+\mathbf{c}}$. A discretization procedure is called *valid* if it is efficient; and depends only on the lattice coset $\Lambda + (\mathbf{c} - \mathbf{x})$, not on particular representative used to specify it. Note that for a valid discretization, $\lfloor \mathbf{z} + \mathbf{x} \rceil_{\Lambda+\mathbf{c}}$ and $\mathbf{z} + \lfloor \mathbf{x} \rceil_{\Lambda+\mathbf{c}}$ are identically distributed for any $\mathbf{z} \in \Lambda$. For more details and actual description of algorithms used for discretization we refer the interested reader to [85].

2.3.3 Algebraic Number Theory

For a positive integer m, the m^{th} *cyclotomic number field* is a field extension $K = \mathbb{Q}(\zeta_m)$ obtained by adjoining an element ζ_m of order m (i.e., a primitive m^{th} root of unity) to the rationals. The minimal polynomial of ζ_m is the m^{th} *cyclotomic polynomial*

$$\Phi_m(X) = \prod_{i \in \mathbb{Z}_m^*} (X - \omega_m^i) \in \mathbb{Z}[X],$$

where $\omega_m \in \mathbb{C}$ is any primitive m^{th} root of unity in $\mathbb{C}$.

For every $i \in \mathbb{Z}_m^*$, there is an embedding $\sigma_i : K \to \mathbb{C}$, defined as $\sigma_i(\zeta_m) = \omega_m^i$. Let $n = \varphi(m)$, the totient of m. The *trace* $\mathrm{Tr} : K \to \mathbb{Q}$ and *norm* $\mathrm{N} : K \to \mathbb{Q}$ can be defined as the sum and product, respectively, of the embeddings:

$$\mathrm{Tr}(x) = \sum_{i \in [n]} \sigma_i(x) \quad \text{and} \quad \mathrm{N}(x) = \prod_{i \in [n]} \sigma_i(x).$$

For any $x \in K$, the l_p *norm* of x is defined as $\|x\|_p = \|\sigma(x)\|_p = (\sum_{i\in[n]} |\sigma_i(x)|^p)^{1/p}$. We omit p when $p = 2$. Note that the appropriate notion of norm $\|\cdot\|$ is used throughout this paper depending on whether the argument is a vector over $\mathbb{C}^n$, or whether the argument is an element from K; whenever the context is clear.

2.3.3.1 Ring of Integers and Its Ideals

Let $R \subset K$ denote the set of all algebraic integers in a number field K. This set forms a ring (under the usual addition and multiplication operations in K), called the *ring of integers* of K. The ring of integers in K is written as $R = \mathbb{Z}[\zeta_m]$.

The (absolute) discriminant Δ_K of K measures the geometric sparsity of its ring of integers. The discriminant of the m^{th} cyclotomic number field K is

$$\Delta_K = \left(\frac{m}{\prod\limits_{\text{prime } p|m} p^{1/(p-1)}} \right)^n \leq n^n,$$

in which the product in the denominator runs over all the primes dividing m.

An (*integral*) *ideal* $\mathcal{I} \subseteq R$ is a non-trivial (i.e. $\mathcal{I} \neq \varnothing$ and $\mathcal{I} \neq \{0\}$) additive subgroup that is closed under multiplication by R, i,e., $r \cdot a \in \mathcal{I}$ for any $r \in R$ and $a \in \mathcal{I}$. The *norm* of an ideal $\mathcal{I} \subseteq R$ is the number of cosets of $\mathcal{I}$ as an addictive subgroup in R, defined as *index* of $\mathcal{I}$, i.e., $\mathrm{N}(\mathcal{I}) = |R/\mathcal{I}|$. Note that $\mathrm{N}(\mathcal{I}\mathcal{J}) = \mathrm{N}(\mathcal{I})\mathrm{N}(\mathcal{J})$.

A *fractional* ideal $\mathcal{I}$ in K is defined as a subset such that $\mathcal{I} \subseteq R$ is an integral ideal for some nonzero $d \in R$. Its norm is defined as $\mathrm{N}(\mathcal{I}) = \mathrm{N}(d\mathcal{I})/\mathrm{N}(d)$. An *ideal lattice* is a lattice $\sigma(\mathcal{I})$ embedded from a fractional ideal $\mathcal{I}$ by σ in H. The

determinant of an ideal lattice $\sigma(\mathcal{I})$ is $\det(\sigma(\mathcal{I})) = \mathrm{N}(\mathcal{I}) \cdot \sqrt{\Delta_K}$. For simplicity, however, most often when discussing about ideal lattice, we omit mention of σ since no confusion is likely to arise.

Lemma 2.12 ([85]). *For any fractional ideal $\mathcal{I}$ in a number field K of degree n,*

$$\sqrt{n} \cdot N^{1/n}(\mathcal{I}) \leq \lambda_1(\mathcal{I}) \leq \sqrt{n} \cdot N^{1/n}(\mathcal{I}) \cdot \sqrt{\Delta_K^{1/n}}.$$

For any *fractional* ideal $\mathcal{I}$ in K, its *dual* ideal is defined as

$$\mathcal{I}^\vee = \{a \in K : \mathrm{Tr}(a\mathcal{I}) \subset \mathbb{Z}\}.$$

Definition 2.13. *For $R = \mathbb{Z}[\zeta_m]$, define $g = \prod_p (1 - \zeta_p) \in R$, where p runs over all* odd *primes dividing m. Also, define $t = \frac{\hat{m}}{g} \in R$, where $\hat{m} = \frac{m}{2}$ if m is even, otherwise $\hat{m} = m$.*

The dual ideal $R^\vee$ of R is defined as $R^\vee = \langle t^{-1} \rangle$, satisfying $R \subseteq R^\vee \subseteq \hat{m}^{-1} R$. For any fractional ideal $\mathcal{I}$, its dual is $\mathcal{I}^\vee = \mathcal{I}^{-1} \cdot R^\vee$. The quotient $R_q^\vee$ is defined as $R_q^\vee = R^\vee / qR^\vee$.

Fact 2.14 ([85]). *Assume that q is a prime satisfying $q = 1 \mod m$, so that $\langle q \rangle$ splits completely into n distinct ideals of norm q. The prime ideal factors of $\langle q \rangle$ are $\mathfrak{q}_i = \langle q \rangle + \langle \zeta_m - \omega_m^i \rangle$, for $i \in \mathbb{Z}_m^*$. By Chinese Reminder Theorem, the natural ring homomorphism $R/\langle q \rangle \to \prod_{i \in \mathbb{Z}_m^*} (R/\mathfrak{q}_i) \cong (\mathbb{Z}_q^n)$ is an isomorphism.*

2.3.4 Ring Learning with Errors

Informally, the (decisional) version of the Ring Learning with Errors (Ring-LWE) problem is: for some secret ring element s, distinguish many random "noisy ring products" with s from elements drawn uniformly from the ring. More precisely, the Ring-LWE problem is parameterized by (R, q, χ, ℓ) where:

1. $R = \mathbb{Z}[X]/(f(X))$ is a ring, where $f(X)$ is an irreducible polynomial $f(X)$ in the indeterminate X. In this paper, we restrict to the case of $f(X) = X^n + 1$, where n is a power of 2.

2. q is a modulus defining the quotient ring $R_q := R/qR = \mathbb{Z}_q[X]/(f(X))$. We restrict to the case where q is prime with $q = 1 \bmod 2n$.

3. $\chi = (\chi_s, \chi_e)$ is a pair of noise distributions over R_q (with χ_s the *secret-key* distribution and χ_e the *error* distribution) that are concentrated on "short" elements, for an appropriate definition of "short."

4. ℓ is the number of samples provided to the adversary.

Formally, the Ring-LWE problem is to distinguish between ℓ samples independently drawn from one of two distributions. In the first case, the samples are generated by choosing $s \leftarrow \chi_s$ and then outputting

$$(a_i, b_i = s \cdot a_i + e_i) \in R_q \times R_q$$

for $i \in [\ell]$, where each a_i is uniform in R_q and each $e_i \leftarrow \chi_e$ is drawn from the

error distribution χ_e. In the second case, each sample (a_i, b_i) is uniformly and independently drawn from $R_q \times R_q$. We let $\mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_s,\chi_e,\ell}(\mathcal{B})$ denote the advantage of algorithm $\mathcal{B}$ in distinguishing these two cases, and define $\mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_s,\chi_e,\ell}(t)$ to be the maximum advantage of any algorithm running in time t. If $\chi = \chi_s = \chi_e$, we write $\mathsf{Adv}_{n,q,\chi,\ell}$ for simplicity.

The noise distribution. The noise distribution $\chi = \chi_s = \chi_e$ is usually a discrete Gaussian distribution on R_q. For power-of-2 cyclotomic rings of the form we consider here, it is possible to sample a polynomial from this distribution by drawing each coefficient of the polynomial independently from the 1-dimensional discrete Gaussian distribution over $\mathbb{Z}_q$ with parameter σ. This distribution, supported on $\{x \in \mathbb{Z}; -q/2 < x < q/2\}$, has density function

$$D_{\mathbb{Z}_q,\sigma}(x) = \frac{e^{\frac{-\pi x^2}{\sigma^2}}}{\sum_{x=-\infty}^{\infty} e^{\frac{-\pi x^2}{\sigma^2}}}.$$

2.3.4.1 Formal Definitions of Ring-LWE in LPR13 [85]

Lemma 2.15. *[85, Lemma 2.23] Let p and q be positive coprime integers, and $\lfloor \cdot \rceil$ be a valid discretization to (cosets of) $pR^\vee$. There exists an efficient transformation that on input $w \in R_p^\vee$ and a pair in $(a', b') \in R_q \times (K_\mathbb{R}/qR^\vee)$, outputs a pair $(a = pa' \bmod qR, b) \in R_q \times R_q^\vee$ with the following guarantees: if the input pair is uniformly distributed then so is the output pair; and if the input pair is distributed according to the RLWE distribution $A_{s,\psi}$ for some (unknown) $s \in R^\vee$ and distribution ψ over $K_\mathbb{R}$, then the output pair is distributed according to $A_{s,\chi}$, where $\chi = \lfloor p \cdot \psi \rceil_{w+pR^\vee}$.*

Lemma 2.16. *[85, Lemma 2.24] Let p and q be positive coprime integers, $\lfloor \cdot \rceil$ be a valid discretization to (cosets of) $pR^\vee$, and w be an arbitrary element in $R_p^\vee$. If $R\text{-}\mathsf{DLWE}_{q,\psi}$ is hard given l samples, then so is the variant of $R\text{-}\mathsf{DLWE}_{q,\psi}$ in which the secret is sampled from $\chi := \lfloor p \cdot \psi \rceil_{w+pR^\vee}$, given $l-1$ samples.*

2.4 The Remaining Mathematical Background

2.4.1 Linear Algebra

We use bold lower case letters to denote vectors, and bold upper case letters to denote matrices. We use row notations for vectors, and start indexing from 0. Let $\mathbf{I}_d$ denote the d-dimensional identity matrix. Let $\langle \cdot, \cdot \rangle$ denote the inner product of two vectors of the same size. Let us introduce the row span of a matrix (denoted Span$(\cdot)$) as the subspace generated by all $\mathbb{R}$-linear combinations of the rows of its input.

Definition 2.17 (Positive Semidefinite)**.** *A $n \times n$ symmetric real matrix $\mathbf{M}$ is positive semidefinite if scalar $\mathbf{x}\mathbf{M}\mathbf{x}^T \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$; if so we write $\mathbf{M} \geq 0$. Given two $n \times n$ real matrix $\mathbf{A}$ and $\mathbf{B}$, we note $\mathbf{A} \geq \mathbf{B}$ if $\mathbf{A} - \mathbf{B}$ is positive semidefinite.*

Definition 2.18. *A matrix $\mathbf{M}$ is a square root of $\mathbf{\Sigma}$, denoted $\sqrt{\mathbf{\Sigma}}$, if*

$$\mathbf{M}^T \cdot \mathbf{M} = \mathbf{\Sigma},$$

Our techniques involve keeping track of the covariance matrix $\mathbf{\Sigma}$ of the secret and error vectors as hints are progressively integrated. The covariance matrix may

become singular during this process and will not have an inverse. Therefore, in the following we introduce some degenerate notions for the inverse and the determinant of a square matrix. Essentially, we restrict these notions to the row span of their input. For $\mathbf{X} \in \mathbb{R}^{d\times k}$ (with any $d, k \in \mathbb{N}$), we will denote $\mathbf{\Pi}_{\mathbf{X}}$ the orthogonal projection matrix onto $\mathrm{Span}(\mathbf{X})$. More formally, let $\mathbf{Y}$ be a maximal set of independent row-vectors of $\mathbf{X}$; the orthogonal projection matrix is given by $\mathbf{\Pi}_{\mathbf{X}} = \mathbf{Y}^T \cdot (\mathbf{Y} \cdot \mathbf{Y}^T)^{-1} \cdot \mathbf{Y}$. Its complement (the projection orthogonally to $\mathrm{Span}(\mathbf{X})$) is denoted $\mathbf{\Pi}_{\mathbf{X}}^{\perp} := \mathbf{I}_d - \mathbf{\Pi}_{\mathbf{X}}$. We naturally extend the notation $\mathbf{\Pi}_F$ and $\mathbf{\Pi}_F^{\perp}$ to subspaces $F \subset \mathbb{R}^d$. By definition, the projection matrices satisfy $\mathbf{\Pi}_F^2 = \mathbf{\Pi}_F$, $\mathbf{\Pi}_F^T = \mathbf{\Pi}_F$ and $\mathbf{\Pi}_F \cdot \mathbf{\Pi}_F^{\perp} = \mathbf{\Pi}_F^{\perp} \cdot \mathbf{\Pi}_F = \mathbf{0}$.

Definition 2.19 (Restricted inverse and determinant)**.** *Let $\mathbf{\Sigma}$ be a symmetric matrix. We define a* restricted inverse *denoted $\mathbf{\Sigma}^{\sim}$ as*

$$\mathbf{\Sigma}^{\sim} := (\mathbf{\Sigma} + \mathbf{\Pi}_{\mathbf{\Sigma}}^{\perp})^{-1} - \mathbf{\Pi}_{\mathbf{\Sigma}}^{\perp}.$$

It satisfies $\mathrm{Span}(\mathbf{\Sigma}^{\sim}) = \mathrm{Span}(\mathbf{\Sigma})$ *and* $\mathbf{\Sigma} \cdot \mathbf{\Sigma}^{\sim} = \mathbf{\Pi}_{\mathbf{\Sigma}}$.

We also denote $\mathrm{rdet}(\mathbf{\Sigma})$ *as the* restricted determinant *defined as follows.*

$$\mathrm{rdet}(\mathbf{\Sigma}) := \det(\mathbf{\Sigma} + \mathbf{\Pi}_{\mathbf{\Sigma}}^{\perp}).$$

The idea behind Definition 2.19 is to provide an (artificial) invertibility property to the input $\mathbf{\Sigma}$ by adding the missing orthogonal part and to remove it afterwards. For example, if $\mathbf{\Sigma} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & 0 \end{bmatrix}$ where $\mathbf{A}$ is invertible,

$$\mathbf{\Sigma}^{\sim} = \left(\begin{bmatrix} \mathbf{A} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} - \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \operatorname{rdet} \mathbf{\Sigma} = \det(\mathbf{A}).$$

Fact 2.20 (Lattice volume under linear transformations)**.** *Let* Λ *be a lattice in* $\mathbb{R}^n$*, and* $\mathbf{M} \in \mathbb{R}^{n \times n}$ *a matrix such that* $\ker \mathbf{M} = \operatorname{Span}(\Lambda)^{\perp}$*. Then we have* $\operatorname{Vol}(\Lambda \cdot \mathbf{M}) = \operatorname{rdet}(\mathbf{M}) \operatorname{Vol}(\Lambda)$.

2.4.2 Regularity and Fourier Transforms

Let $\rho_{\mathbf{s},\mathbf{c}}$ denote an n-dimensional Gaussian function with parameter $\mathbf{s}$ and mean $\mathbf{c}$.

One and Multi-Dimensional Gaussians. For $s > 0$, $c \in \mathbb{R}$, $x \in \mathbb{R}$, define the Gaussian function $\rho^1_{s,c} : \mathbb{R} \to (0, 1]$ as

$$\rho^1_{s,c}(x) := e^{\frac{-\pi(x-c)^2}{s^2}}.$$

When $c = 0$, we write for simplicity,

$$\rho^1_s(x) := e^{\frac{-\pi(x)^2}{s^2}}.$$

By normalizing this function we obtain the *continuous* Gaussian probability distribution $\psi^1_{s,c}$ (resp. ψ^1_s) of parameter s, whose density is given by $s^{-1} \cdot \rho^1_{s,c}(x)$ (resp. $s^{-1} \cdot \rho^1_s(x)$).

We denote by $\rho_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}$ the distribution over $\mathbb{R}^n$ with the following pdf:

Let $\rho^1_{s,c}$ denote a one-dimensional Gaussian function as above with standard deviation s and mean c.We denote by $\rho_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}$ the distribution over $\mathbb{R}^n$ with the following pdf:

$$\rho_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}(x_1,\ldots,x_n) := \rho^1_{s_1,c_1}(x_1)\cdots\rho^1_{s_n,c_n}(x_n).$$

When $c = \mathbf{0}$, we again write for simplicity, $\rho_{(s_1,\ldots,s_n)}$. Moreover, when $s_1 = \cdots = s_n$ and the dimension is clear from context we write for simplicity $\rho_{s,(c_1,\ldots,c_n)}$ (resp. ρ_s). Normalizing as above, we obtain the corresponding *continuous* Gaussian probability distribution $\psi_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}$ (resp. $\psi_{(s_1,\ldots,s_n)}$, $\psi_{s,(c_1,\ldots,c_n)}$, ψ_s).

Definition 2.21 (**Fourier Transform**). *Given an integrable function $f : \mathbb{R}^n \to \mathbb{C}$, we denote by $\widehat{f} : \mathbb{R}^n \to \mathbb{C}$ the Fourier transform of f, defined as*

$$\widehat{f}(\mathbf{y}) := \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-2\pi i\langle \mathbf{x},\mathbf{y}\rangle}\, \mathrm{d}\mathbf{x}.$$

Theorem 2.22 (Poisson Summation Formula). *Let $\Lambda \subset \mathbb{R}^n$ be an arbitrary lattice of dimension n, and let $f : \mathbb{R}^n \to \mathbb{C}$ be an appropriate function* [1] *Then*

$$f(\Lambda) = \frac{1}{\det(\Lambda)}\widehat{f}(\Lambda^\vee),$$

where $\Lambda^\vee$ is the dual lattice of Λ and $\widehat{f}$ is a Fourier transform of f.

[1] Assume that (1). $\int_{\mathbb{R}^n} |f(x)|dx < \infty$. (2). Function $f(\Lambda + u)$ is continuous on $\mathbb{R}^n$. (3). The series $\widehat{f}(\Lambda^\vee)$ is absolutely convergent. (See [59] for details)

Definition 2.23. *For an n-dimensional lattice Λ, and positive real $\varepsilon > 0$, we define its smoothing parameter $\eta_\varepsilon(\Lambda)$ to be the smallest s such that $\rho_{1/s}(\Lambda^\vee \setminus \{\mathbf{0}\}) \leq \varepsilon$.*

Lemma 2.24. *[42,91] For any n-dimensional lattice Λ, we have $\frac{\sqrt{\ln(1/\varepsilon)}}{\sqrt{\pi}\lambda_1(\Lambda^\vee)} \leq \eta_\varepsilon(\Lambda) \leq \frac{\sqrt{n}}{\lambda_1(\Lambda^\vee)}$, for $\varepsilon \in [2^{-n}, 1]$.*

Claim 2.25 ([85]). *For any n-dimensional lattice Λ and $\varepsilon, s > 0$,*

$$\rho_{1/s}(\Lambda) \leq \max\left(1, \left(\frac{\eta_\varepsilon(\Lambda^\vee)}{s}\right)^n\right)(1+\varepsilon).$$

Lemma 2.26. *For any n-dimensional lattice Λ and $\varepsilon > 0$, $\mathbf{s} := (s_1, \ldots, s_n) \in \mathbb{R}^n_{>0}$, and $\mathbf{c} := (c_1, \ldots, c_n) \in \mathbb{R}^n$, if all of $s_1, \ldots, s_n < \eta_\varepsilon(\Lambda^\vee)$ then*

$$\rho_{(1/s_1,\ldots,1/s_n),(c_1,\ldots,c_n)}(\Lambda) \leq \left(\frac{\eta_\varepsilon(\Lambda^\vee)}{s_1} \cdots \frac{\eta_\varepsilon(\Lambda^\vee)}{s_n}\right)(1+\varepsilon).$$

Proof. Applying Poisson summation formula twice, using the fact that for all vectors $\mathbf{x} \in \mathbb{R}^n$, $\widehat{\rho}_{(1/s_1,\ldots,1/s_n),(c_1,\ldots,c_n)}(\mathbf{x}) \leq (s_1)^{-1} \cdots (s_n)^{-1} \cdot \rho_{(s_1,\ldots,s_n)}(\mathbf{x})$, and the fact that $\widehat{\rho}_{\eta_\varepsilon(\Lambda^\vee)} = \eta_\varepsilon(\Lambda^\vee)^n \cdot \rho_{1/\eta_\varepsilon(\Lambda^\vee)}$, we have:

$$\begin{aligned}
\rho_{(1/s_1,\ldots,1/s_n),(c_1,\ldots,c_n)}(\Lambda) &\leq \det(\Lambda)^{-1}(s_1)^{-1} \cdots (s_n)^{-1} \cdot \rho_{(s_1,\ldots,s_n)}(\Lambda^\vee) \\
&\leq \det(\Lambda)^{-1}(s_1)^{-1} \cdots (s_n)^{-1} \cdot \rho_{\eta_\varepsilon(\Lambda^\vee)}(\Lambda^\vee) \\
&= (s_1)^{-1} \cdots (s_n)^{-1} \cdot \eta_\varepsilon(\Lambda^\vee)^n \cdot \rho_{1/\eta_\varepsilon(\Lambda^\vee)}(\Lambda) \\
&\leq \left(\frac{\eta_\varepsilon(\Lambda^\vee)}{s_1} \cdots \frac{\eta_\varepsilon(\Lambda^\vee)}{s_n}\right)(1+\varepsilon).
\end{aligned}$$

where the last inequality follows from the definition of $\eta_\varepsilon(\Lambda^\vee)$.

□

Lemma 2.27. *[91, Lemma 3.6] For any lattice Λ, positive real $s > 0$ and a vector $\mathbf{c}$, $\rho_{s,\mathbf{c}}(\Lambda) \leq \rho_s(\Lambda)$.*

Definition 2.28. *Let Λ be an n-dimensional lattice and Ψ a probability distribution over $\mathbb{R}^n$. Define the discrete probability distribution of Ψ over Λ to be:*

$$D_{\Lambda,\Psi}(\mathbf{x}) = \frac{\Psi(\mathbf{x})}{\Psi(\Lambda)}, \forall \mathbf{x} \in \Lambda.$$

Definition 2.29. *Let Λ be an n-dimensional lattice, define the discrete Gaussian probability distribution over Λ with parameter $(s_1, \ldots, s_n)$ and center $(c_1, \ldots, c_n)$ as*

$$D_{\Lambda,(s_1,\ldots,s_n),(c_1,\ldots,c_n)}(\mathbf{x}) = \frac{\rho_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}(\mathbf{x})}{\rho_{(s_1,\ldots,s_n),(c_1,\ldots,c_n)}(\Lambda)}, \forall \mathbf{x} \in \Lambda.$$

Remark 2.30. *Whenever Ψ is Gaussian with parameter $(s_1, \ldots, s_n)$ and center $(c_1, \ldots, c_n)$ we denote it's discrete Gaussian probability by $D_{\Lambda,(s_1,\cdots,s_n),(c_1,\ldots,c_n)}$. If $s = s_1 = \cdots = s_n$ (resp. $c = c_1 = \cdots = c_n$) we write $D_{\Lambda,s,(c_1,\ldots,c_n)}$ (resp. $D_{\Lambda,(s_1,\ldots,s_n),c}$). If $c_1 = \cdots = c_n = 0$ we write $D_{\Lambda,(s_1,\cdots,s_n)}$.*

Lemma 2.31. *[91, Lemma 4.4] For any n'-dimensional lattice Λ, and reals $0 < \varepsilon < 1, s \geq \eta_\varepsilon(\Lambda)$, we have*

$$\Pr_{\mathbf{x} \sim D_{\Lambda,\psi_s}} \left(\|\mathbf{x}\| > s\sqrt{n'} \right) \leq \frac{1+\varepsilon}{1-\varepsilon} \cdot 2^{-n'}.$$

The following is a modified version of Lemma 3.8 from [103].

Lemma 2.32. *Let Λ be an n-dimensional lattice and Ψ a probability distribution over $\mathbb{R}^n$. If $|\widehat{\Psi}|(\Lambda^\vee \setminus \{\mathbf{0}\}) \leq \varepsilon$, then for any $\mathbf{c} \in \mathbb{R}^n$, $\Psi(\Lambda + \mathbf{c}) \in \det(\Lambda^\vee)(1 \pm \varepsilon)$, where $|\widehat{\Psi}|(\Lambda^\vee \setminus \{\mathbf{0}\})$ denotes the summation of the absolute value of the function at each point in $\Lambda^\vee \setminus \{\mathbf{0}\}$.*

Proof. First, since Ψ is a pdf, we have that $\widehat{\Psi}(\mathbf{0}) = 1$. We have:

$$\begin{aligned}
\Psi(\Lambda + \mathbf{c}) &= \det(\Lambda^\vee) \sum_{\mathbf{y} \in \Lambda^\vee} \widehat{\Psi}(\mathbf{y}) e^{2\pi i < \mathbf{c}, \mathbf{y} >} \\
&\in \det(\Lambda^\vee) \left(1 \pm \sum_{\mathbf{y} \in \Lambda^\vee \setminus \{\mathbf{0}\}} |\widehat{\Psi}(\mathbf{y}) e^{2\pi i < \mathbf{c}, \mathbf{y} >}| \right) \\
&\subseteq \det(\Lambda^\vee) \left(1 \pm \sum_{\mathbf{y} \in \Lambda^\vee \setminus \{\mathbf{0}\}} \widehat{\Psi}(\mathbf{y}) \right) \\
&\subseteq \det(\Lambda^\vee)(1 \pm \varepsilon),
\end{aligned}$$

where the equality follows from properties of the Fourier transform. □

The proof of the following lemma proceeds as the proof of Corollary 2.8 in [63].

Lemma 2.33. *Let Λ' be an n-dimensional lattice and Ψ a probability distribution over $\mathbb{R}^n$. Assume that for all $\mathbf{c} \in \mathbb{R}^n$ it is the case that*

$$\Psi(\Lambda' + \mathbf{c}) \in \left[\frac{1 - \varepsilon}{1 + \varepsilon}, \frac{1 + \varepsilon}{1 - \varepsilon} \right] \cdot \Psi(\Lambda'),$$

Let Λ be an n-dimensional lattice such that $\Lambda' \subseteq \Lambda$ then the distribution of $(D_{\Lambda,\Psi} \mod \Lambda')$ is within statistical distance of at most 4ε of uniform over $(\Lambda \mod \Lambda')$.

Definition 2.34. *For a matrix $A \in R_q^{k \times l}$ we define $\Lambda^\perp(A) = \{\mathbf{z} \in R^l : A\mathbf{z} =$*

$0 \bmod qR\}$, *which we identify with a lattice in* H^l. *Its dual lattice (which is again a lattice in* H^l*) is denoted by* $\Lambda^{\perp}(A)^{\vee}$.

Theorem 2.35. *[85] Let* R *be the ring of integers in the* m^{th} *cyclotomic number field* K *of degree* n, *and* $q \geq 2$ *an integer. For positive integers* $k \leq l \leq \mathrm{poly}(n)$, *let* $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$, *where* $I_k \in (R_q)^{k\times k}$ *is the identity matrix and* $\bar{A} \in (R_q)^{k\times(l-k)}$ *is uniformly random. Then for all* $s \geq 2n$,

$$\mathbb{E}_{\bar{A}}\left[\rho_{1/s}\left(\Lambda^{\perp}(A)^{\vee}\right)\right] \leq 1 + 2(s/n)^{-nl}q^{kn+2} + 2^{-\Omega(n)}.$$

In particular, if $s > 2n \cdot q^{k/l+2/(nl)}$ *then* $\mathbb{E}_{\bar{A}}\left[\rho_{1/s}\left(\Lambda^{\perp}(A)^{\vee}\right)\right] \leq 1 + 2^{-\Omega(n)}$, *and so by Markov's inequality,* $\eta_{2^{-\Omega(n)}}(\Lambda^{\perp}(A)) \leq s$ *except with probability at most* $2^{-\Omega(n)}$.

The following corollary was presented in [85].

Corollary 2.36. *Let* R, n, q, k *and* l *be as in Theorem 2.35. Assume that* $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$ *is chosen as in Theorem 2.35. Then, with probability* $1 - 2^{-\Omega(n)}$ *over the choice of* $\bar{A}$, *the distribution of* $A\mathbf{x} \in R_q^k$, *where each coordinate of* $\mathbf{x} \in R_q^l$ *is chosen from a discrete Gaussian distribution of parameter* $s > 2n \cdot q^{k/l+2/(nl)}$ *over* R, *satisfies that the probability of each of the* q^{nk} *possible outcomes is in the interval* $(1 \pm 2^{-\Omega(n)})q^{-nk}$ *(and in particular is within statistical distance* $2^{-\Omega(n)}$ *of the uniform distribution over* R_q^k*).*

We next state an additional corollary of the regularity theorem from [85].

Corollary 2.37. *Let* R, n, q, k *and* l *be as in Theorem 2.35. Assume that* $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$ *is chosen as in Theorem 2.35. Then, with probability* $1 - 2^{-\Omega(n)}$

over the choice of $\bar{A}$*, the shortest non-zero vector in* $\Lambda^{\perp}(A)^{\vee}$ *has length at least* $\frac{\sqrt{n/\pi}}{2n \cdot q^{k/l+2/(nl)}}$.

Chapter 3: Constant-Round Group Key Exchange

Protocols for (authenticated) key exchange are among the most fundamental and widely used cryptographic primitives. They allow parties communicating over an insecure public network to establish a common secret key, called a *session key*, permitting the subsequent use of symmetric-key cryptography for encryption and authentication of sensitive data. They can be used to instantiate so-called "secure channels" upon which higher-level cryptographic protocols often depend.

Most work on key exchange, beginning with the classical paper of Diffie and Hellman, has focused on two-party key exchange. However, many works have also explored extensions to the *group* setting [1, 2, 13, 14, 18, 19, 30–33, 35, 36, 41, 71, 73, 76, 78, 107, 110, 113] in which N parties wish to agree on a common session key that they can each then use for encrypted communication with the rest of the group.

The recent effort by NIST to evaluate and standardize one or more quantum-resistant public-key cryptosystems is entirely focused on digital signatures and two-party key encapsulation/key exchange,[1] and there has been an extensive amount of research over the past decade focused on designing such schemes. In contrast, we are aware of almost no[2] work on *group* key-exchange protocols with post-quantum

[1] Note that CPA-secure key encapsulation is equivalent to two-round key-exchange (with passive security).

[2] Exceptions include the work of Ding et al. [52], which lacks a proof of security; the work of

security beyond the observation that a post-quantum group key-exchange protocol can be constructed from any post-quantum two-party protocol by having a designated group manager run independent two-party protocols with the $N-1$ other parties, and then send a session key of its choice to the other parties encrypted/authenticated using each of the resulting keys. Such a solution is often considered unacceptable since it is highly asymmetric, requires additional coordination, is not contributory, and puts a heavy load on a single party who becomes a central point of failure.

3.1 Our High-Level Approach

In this work, we propose a constant-round group key-exchange protocol based on the hardness of the Ring-LWE problem [84], and hence with (plausible) post-quantum security. In this work, we focus on constructing an *unauthenticated* protocol—i.e., one secure against a passive eavesdropper—since known techniques such as the Katz-Yung compiler [75] can then be applied to obtain an *authenticated* protocol secure against an active attacker.

The starting point for our work is the two-round group key-exchange protocol by Burmester and Desmedt [35, 36, 76], which is based on the decisional Diffie-Hellman assumption. Assume a group $\mathbb{G}$ of prime order q and a generator $g \in \mathbb{G}$ are fixed and public. The Burmester-Desmedt protocol run by parties $P_0, \ldots, P_{N-1}$ then works as follows:

Boneh et al. [22] shows a framework for group key-exchange protocols with plausible post-quantum security but without a concrete instantiation.

1. In the first round, each party P_i chooses uniform $r_i \in \mathbb{Z}_q$ and broadcasts $z_i = g^{r_i}$ to all other parties.

2. In the second round, each party P_i broadcasts $X_i = (z_{i+1}/z_{i-1})^{r_i}$ (where the parties' indices are taken modulo N).

Each party P_i can then compute its session key sk_i as

$$\mathsf{sk}_i = (z_{i-1})^{Nr_i} \cdot X_i^{N-1} \cdot X_{i+1}^{N-2} \cdots X_{i+N-2}.$$

One can check that all the keys are equal to the same value $g^{r_0 r_1 + \cdots + r_{N-1} r_0}$.

In attempting to adapt their protocol to the Ring-LWE setting, we could fix a public ring R_q and a uniform element $a \in R_q$. Then:

1. In the first round, each party P_i chooses "small" secret value $s_i \in R_q$ and "small" noise term $e_i \in R_q$ (with the exact distribution being unimportant in the present discussion), and broadcasts $z_i = as_i + e_i$ to the other parties.

2. In the second round, each party P_i chooses a second "small" noise term $e'_i \in R_q$ and broadcasts $X_i = (z_{i+1} - z_{i-i}) \cdot s_i + e'_i$.

Each party can then compute a session key b_i as

$$b_i = N \cdot s_i \cdot z_{i-1} + (N-1) \cdot X_i + (N-2) \cdot X_{i+1} + \cdots + X_{i+N-2}.$$

The problem, of course, is that (due to the noise terms) these session keys computed by the parties will *not* be equal. They will, however, be "close" to each other if the

$\{s_i, e_i, e'_i\}$ are all sufficiently small, so we can add an additional reconciliation step to ensure that all parties agree on a common key k.

This gives a protocol that is correct, but proving security (even for a passive eavesdropper) is more difficult than in the case of the Burmester-Desmedt protocol. Here we informally outline the main difficulties and how we address them. First, we note that trying to prove security by direct analogy to the proof of security for the Burmester-Desmedt protocol (cf. [75]) fails; in the latter case, it is possible to use the fact that, for example,

$$(z_2/z_0)^{r_1} = z_1^{r_2 - r_0},$$

whereas in our setting the analogous relation does not hold. In general, the natural proof strategy here is to switch all the $\{z_i\}$ values to uniform elements of R_q, and similarly to switch the $\{X_i\}$ values to uniform subject to the constraint that their sum is approximately 0 (i.e., subject to the constraint that $\sum_i X_i \approx 0$). Unfortunately this cannot be done by simply invoking the Ring-LWE assumption $O(N)$ times; in particular, the first time we try to invoke the assumption, say on the pair $(z_1 = as_1 + e_1,\ X_1 = (z_2 - z_0) \cdot s_1 + e'_1)$, we need $z_2 - z_0$ to be uniform—which, in contrast to the analogous requirement in the Burmester-Desmedt protocol (for the value z_2/z_0), is not the case here. Thus, we must somehow break the circularity in the mutual dependence of the $\{z_i, X_i\}$ values.

Toward this end, let us look more carefully at the distribution of $\sum_i X_i$. We may write

$$\textstyle\sum_i X_i = \sum_i (e_{i+1}s_i - e_{i-1}s_i) + \sum_i e'_i.$$

Consider now changing the way X_0 is chosen: that is, instead of choosing $X_0 = (z_1 - z_{N-1})s_0 + e'_0$ as in the protocol, we instead set $X_0 = -\sum_{i=1}^{N-1} X_i + e'_0$ (where e'_0 is from the same distribution as before). Intuitively, as long as the standard deviation of e'_0 is large enough, these two distributions of X_0 should be "close" (as they both satisfy $\sum_i X_i \approx 0$). This, in particular, means that we need the distribution of e'_0 to be different from the distribution of the $\{e'_i\}_{i>0}$, as the standard deviation of the former needs to be larger than the latter.

We can indeed show that when we choose e'_0 from an appropriate distribution then the Rényi divergence between the two distributions of X_0, above, is bounded by a polynomial. With this switch in the distribution of X_0, we have broken the circularity and can now use the Ring-LWE assumption to switch the distribution of z_0 to uniform, followed by the remaining $\{z_i, X_i\}$ values.

Unfortunately, bounded Rényi divergence does not imply statistical closeness. However, polynomially bounded Rényi divergence *does* imply that any event occurring with negligible probability when X_0 is chosen according to the second distribution also occurs with negligible probability when X_0 is chosen according to the first distribution. For these reasons, we change our security goal from an "indistinguishability-based" one (namely, requiring that the real session key k is indistinguishable from uniform) to an "unpredictability-based" one (namely, requiring that it is infeasible for an attacker to compute the real session key k). In the end, though, once the parties agree on an unpredictable value k they can hash it to obtain the final session key $\mathsf{sk} = \mathcal{H}(k)$; this final value sk will be indistinguishable from uniform if $\mathcal{H}$ is modeled as a random oracle.

3.2 Security Model

A group key-exchange protocol allows a session key to be established among $N > 2$ parties. Following prior work [31–33, 75], we will use the term *group key exchange* (GKE) to denote a protocol secure against a *passive* (eavesdropping) adversary, and use the term *authenticated group key exchange* (GAKE) to denote a protocol secure against an *active* adversary who controls all communication channels. Fortunately, the work of Katz and Yung [75] presents a compiler that takes any GKE protocol and transforms it into a GAKE protocol. The underlying tool required for this transform is any secure signature scheme; if post-quantum security is needed, then any post-quantum signature scheme can be used. We thus focus our attention on achieving GKE in the remainder of this work.

In the security definition for group key exchange, the adversary observes a single transcript generated by an execution of the protocol. The adversary's goal is then to distinguish the real session key generated in that execution of the protocol from a key that is generated uniformly and independently of that transcript. Formally, given a GKE protocol Π we let $\mathsf{Execute}_{\Pi}(\lambda)$ denote an execution of the protocol (on security parameter λ), resulting in a transcript trans of all messages sent during the course of that execution, along with the session key sk computed by the parties. Protocol Π is secure if the following distribution ensembles are

computationally indistinguishable:

$$\{(\mathsf{trans}, \mathsf{sk}) \leftarrow \mathsf{Execute}_{\Pi}(1^{\lambda}) : (\mathsf{trans}, \mathsf{sk})\}_{1^{\lambda} \in \mathbb{N}},$$

$$\{(\mathsf{trans}, \mathsf{sk}) \leftarrow \mathsf{Execute}_{\Pi}(1^{\lambda}), \mathsf{sk}' \leftarrow \{0,1\}^{1^{\lambda}} : (\mathsf{trans}, \mathsf{sk}')\}_{1^{\lambda} \in \mathbb{N}}.$$

Our protocol Π will be analyzed in the random-oracle model. In this case, fixing some λ, we let $\mathsf{Adv}^{\mathsf{GKE}}_{\Pi}(\mathcal{A})$ denote the advantage of an adversary $\mathcal{A}$ in distinguishing between the distributions above, and define $\mathsf{Adv}^{\mathsf{GKE}}_{\Pi}(t, \mathsf{q})$ to be the maximum advantage of any adversary running in time t and making at most q queries to the random oracle.

3.3 Group Key Exchange Protocol

In this section, we present a group key exchange protocol Π for N parties $P_0, \ldots, P_{N-1}$. Our protocol relies on a key-reconciliation mechanism KeyRec (parameterized by a bound β_{Rec}) as a subroutine.

3.3.1 Generic Key Reconciliation

In this subsection, we define a generic, one round, two-party key reconciliation mechanism (tailored to the Ring-LWE setting) that allows two parties to derive a shared key if they begin holding "close" ring elements. Formally, a key-reconciliation mechanism KeyRec consists of two algorithms recMsg and recKey, parameterized by a bound β_{Rec} (that may depend on the security parameter). The first algorithm takes as input the security parameter 1^λ and a value $b \in R_q$, and outputs a reconciliation message rec and a key $k \in \{0,1\}^\lambda$. The second algorithm takes as input 1^λ, a value $b' \in R_q$, and rec, and outputs $k' \in \{0,1\}^\lambda$.

Correctness requires that whenever b, b' are "close," then $k' = k$. Specifically, for any b, b' for which each coefficient of $b - b'$ is bounded by β_{Rec}, if we run $(\mathsf{rec}, k) \leftarrow \mathsf{recMsg}(1^\lambda, b)$ followed by $k' := \mathsf{recKey}(1^\lambda, b', \mathsf{rec})$ then $k = k'$.

Security requires that if b is uniform and we derive $(\mathsf{rec}, k) \leftarrow \mathsf{recMsg}(1^\lambda, b)$, then k is computationally indistingiushable from uniform even for an attacker given rec. Formally, the following two distribution ensembles must be computationally indistinguishable:

$$\left\{b \leftarrow R_q; (\mathsf{rec}, k) \leftarrow \mathsf{recMsg}(1^\lambda, b) : (\mathsf{rec}, k)\right\}_{\lambda \in \mathbb{N}},$$

$$\left\{b \leftarrow R_q; (\mathsf{rec}, k) \leftarrow \mathsf{recMsg}(1^\lambda, b); k' \leftarrow \{0,1\}^\lambda : (\mathsf{rec}, k')\right\}_{\lambda \in \mathbb{N}},$$

For some fixed value of λ we denote by $\mathsf{Adv}_{\mathsf{KeyRec}}(\mathcal{B})$ the advantage of adversary $\mathcal{B}$ in distinguishing these distributions, and let $\mathsf{Adv}_{\mathsf{KeyRec}}(t)$ be the maximum advantage

of any such adversary running in time t.

Key-reconciliation mechanisms from the literature. The notion of key reconciliation was first introduced by Ding et al. [52], and was later used in several works on two-party key exchange [9, 96, 115]. In the key reconciliation mechanisms of Peikert [96], Zhang et al. [115] and Alkim et al. [9], the agreed-upon key $k = k'$ is close to each of the original values b, b' held by the parties. When instantiating our group key exchange (GKE) protocol with this type of key-reconciliation mechanism, our final GKE protocol is contributory. In other cases [8], the agreed-upon key is determined by the randomness used when running recMsg; instantiating our GKE protocol with this type of key-reconciliation mechanism yields a non-contributory protocol.

3.3.2 Construction

The overall structure of the protocol is as follows. The first two rounds allow the parties to agree on "close" keys $b_0 \approx \cdots \approx b_{N-1}$. Player $N-1$ then initiates the key-reconciliation mechanism to allow all parties to agree on the same key $k = k_0 = \cdots = k_{N-1} \in \{0,1\}^\lambda$. Since we are only able to prove that k is difficult to compute for an eavesdropping adversary (but may not be indistinguishable from random), we then have each party hash k (using a hash function $\mathcal{H}$) to obtain the final shared key sk.

Our protocol is parameterized by noise distributions $\chi_{\sigma_1}, \chi_{\sigma_2}$, and assumes public parameters $R_q = \mathbb{Z}_q[x]/(x^n+1)$ along with a uniform value $a \in R_q$. The protocol proceeds as follows:

Round 1: Each player P_i samples $s_i, e_i \leftarrow \chi_{\sigma_1}$ and broadcasts $z_i = as_i + e_i$.

Round 2: Player P_0 samples $e'_0 \leftarrow \chi_{\sigma_2}$ and each of the other players P_i samples $e'_i \leftarrow \chi_{\sigma_1}$. Each P_i broadcasts $X_i = (z_{i+1} - z_{i-1})s_i + e'_i$.

Round 3: Player P_{N-1} samples $e''_{N-1} \leftarrow \chi_{\sigma_1}$ and computes

$$b_{N-1} = z_{N-2}Ns_{N-1} + (N-1)\cdot X_{N-1} + (N-2)\cdot X_0 + \cdots + X_{N-3} + e''_{N-1}.$$

It then computes $(\mathsf{rec}, k_{N-1}) = \mathsf{recMsg}(b_{N-1})$ and broadcasts rec. Finally, it outputs the session key $\mathsf{sk}_{N-1} = \mathcal{H}(k_{N-1})$.

Key computation: Each player P_i (except P_{N-1}) computes

$$b_i = z_{i-1} N s_i + (N-1) \cdot X_i + (N-2) \cdot X_{i+1} + \cdots + X_{i+N-2}.$$

It then sets $k_i = \mathsf{recKey}(b_i, \mathsf{rec})$, and outputs the session key $\mathsf{sk}_i = \mathcal{H}(k_i)$.

The following shows a condition under which each party derives the same session key with all but negligible probability.

Theorem 3.1. *Fix ρ, and assume*

$$(N^2 + 2N) \cdot \sqrt{n}\, \rho^{3/2}\, \sigma_1^2 + (\frac{N^2}{2} + 1) \cdot \sigma_1 + (N-2) \cdot \sigma_2 \leq \beta_{\mathsf{Rec}}.$$

Then all parties output the same key except with probability at most $2^{-\rho+1}$.

Proof. We begin by introducing the following lemmas to analyze probabilities that each coordinate of $s_i, e_i, e'_i, e''_{N-1}, e'_0$ are "short" for all i, and conditioned on the first event, $s_i e_i$ is "short".

Lemma 3.2. *Given $s_i, e_i, e'_i, e''_{N-1}, e'_0$ for all i as defined in the group key exchange protocol, fix $c = \sqrt{\frac{2\rho}{\pi \log e}}$, and let bound_ρ denote the event that for all i and all coordinate indices j, $|(e'_0)_j| \leq c\sigma_2$ and $|(s_i)_j|$, $|(e_i)_j|$, $|(e''_{N-1})_j| \leq c\sigma_1$, and that for all $i > 0$ and all j it holds that $|(e'_i)_j| \leq c\sigma_1$, we have*

$$\Pr[\mathsf{bound}_\rho] \geq 1 - 2^{-\rho}.$$

Proof. Using the fact that $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt \leq e^{-x^2}$, we obtain

$$\begin{aligned}\Pr[|v| \geq c\sigma + 1; v \leftarrow D_{\mathbb{Z}_q,\sigma}] &\leq 2 \sum_{x=\lfloor c\sigma+1 \rceil}^{\infty} D_{\mathbb{Z}_q,\sigma}(x) \leq \frac{2}{\sigma} \int_{c\sigma}^{\infty} e^{-\frac{\pi x^2}{\sigma^2}} dx \\ &= \frac{2}{\sqrt{\pi}} \int_{\frac{\sqrt{\pi}}{\sigma}(c\sigma)}^{\infty} e^{-t^2} dt \leq e^{-c^2\pi}.\end{aligned}$$

Note that there are $3nN$ coordinates sampled from distribution $D_{\mathbb{Z}_q,\sigma_1}$, and n coordinates sampled from distribution $D_{\mathbb{Z}_q,\sigma_2}$ in total. Under the assumption that $3nN + n \leq e^{c^2\pi/2}$ (which holds for all reasonable settings for the parameters), we have:

$$\begin{aligned}\Pr[\mathsf{bound}_\rho] &= \left(1 - \Pr[|v| \geq c\sigma_1 + 1; v \leftarrow D_{\mathbb{Z}_q,\sigma_1}]\right)^{3nN} \\ &\qquad \cdot \left(1 - \Pr[|e_0'| \geq c\sigma_2 + 1; e_0' \leftarrow D_{\mathbb{Z}_q,\sigma_2}]\right)^{n} \\ &\geq 1 - (3nN + n)e^{-c^2\pi} \geq 1 - e^{-c^2\pi/2} \geq 1 - 2^{-\rho}.\end{aligned}$$

□

Lemma 3.3. *Given* bound_ρ *as defined in Lemma 3.2, let* $\mathsf{product}_{\mathsf{s_i},\mathsf{e_j}}$ *denote the event that, for all* v, $|(s_i e_j)_v| \leq \sqrt{n}\rho^{3/2}\sigma_1^2$,

$$\Pr[\mathsf{product}_{\mathsf{s_i},\mathsf{e_j}} \mid \mathsf{bound}_\rho] \geq 1 - 2n \cdot 2^{-2\rho}.$$

Proof. For $t \in \{0, \ldots, n-1\}$, Let $(s_i)_t$ denote the t^{th} coefficient of $s_i \in R_q$, namely, $s_i = \sum_{t=0}^{n-1} (s_i)_t X^i$. $(e_j)_t$ is defined analogously. Since we have $X^n + 1$ as modulo of R, it is easy to see that $(s_i e_j)_v = c_v X^v$, where $c_v = \sum_{u=0}^{n-1} (s_i)_u (e_j)^*_{v-u}$. If $v - u \geq 0$,

$(e_j)^*_{v-u} = (e_j)_{v-u}$. $(e_j)^*_{v-u} = -(e_j)_{v-u+n}$ otherwise. Thus, conditioned on $|(s_i)_t| \leq c\sigma_1$ and $|(e_j)_t| \leq c\sigma_1$ (for all i, j, t) where $c = \sqrt{\frac{2\rho}{\pi \log e}}$, by Hoeffding's Inequality [68], we derive

$$\Pr[|(s_i e_j)_v| \geq \delta \mid \mathsf{bound}_\rho] = \Pr\left[\left|\sum_{u=0}^{n-1} (s_i)_u (e_j)^*_{v-u}\right| \geq \delta\right] \leq 2\exp\left(\frac{-2\delta^2}{n(2c^2\sigma_1^2)^2}\right),$$

as each product $(s_i)_u (e_j)^*_{v-u}$ in the sum is an independent random variable with mean 0 in the range $[-c^2\sigma_1^2, c^2\sigma_1^2]$. By fixing $\delta = \sqrt{n}\rho^{3/2}\sigma_1^2$, we obtain

$$\Pr[|(s_i e_j)_v| \geq \sqrt{n}\rho^{3/2}\sigma_1^2 \mid \mathsf{bound}_\rho] \leq 2^{-2\rho+1}. \tag{3.1}$$

Finally, via a union bound, we thus have

$$\Pr[\mathsf{product}_{\mathsf{s_i},\mathsf{e_j}} | \mathsf{bound}_\rho] = \Pr[\forall v : |(s_i e_j)_v| \leq \sqrt{n}\rho^{3/2}\sigma_1^2] \geq 1 - 2n \cdot 2^{-2\rho}. \tag{3.2}$$

□

Now we begin analyzing the chance that not all parties agree on the same final key. The correctness of KeyRec guarantees that this group key exchange protocol has agreed session key among all parties. Formally, if for all i and j that the j^{th} coefficient of $|b_{N-1} - b_i| \leq \beta_{\mathsf{Rec}}$, then for all i, $k_i = k_{N-1}$.

For better illustration, we first write $X_0, \ldots, X_{N-1}$ in form of linear system as

follows. $\mathbf{X} = [X_0 \;\; X_1 \;\; X_2 \;\; \cdots \;\; X_{N-1}]^T$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & -1 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{bmatrix}}_{\mathbf{M}} \underbrace{\begin{bmatrix} as_0s_1 \\ as_1s_2 \\ as_2s_3 \\ as_3s_4 \\ \vdots \\ as_{N-2}s_{N-1} \\ as_{N-1}s_0 \end{bmatrix}}_{\mathbf{S}} + \underbrace{\begin{bmatrix} s_0e_1 - s_0e_{N-1} + e'_0 \\ s_1e_2 - s_1e_0 + e'_1 \\ s_2e_3 - s_2e_1 + e'_2 \\ s_3e_4 - s_3e_2 + e'_3 \\ \vdots \\ s_{N-2}e_{N-3} - s_{N-2}e_{N-3} + e'_{N-2} \\ s_{N-1}e_0 - s_{N-1}e_{N-2} + e'_{N-1} \end{bmatrix}}_{\mathbf{E}}. \tag{3.3}$$

We denote the matrices above by $\mathbf{M}, \mathbf{S}, \mathbf{E}$ from left to right and have the linear system as $\mathbf{X} = \mathbf{MS} + \mathbf{E}$. Let $\mathbf{B}_i = [i-1 \;\; i-2 \;\; \cdots \;\; 0 \;\; N-1 \;\; N-2 \;\; \cdots \;\; i]$ as a N-dimensional row vector. We can then write b_i as $\mathbf{B}_i \cdot \mathbf{X} + N(as_is_{i-1} + s_ie_{i-1}) = \mathbf{B}_i\mathbf{MS} + \mathbf{B}_i\mathbf{E} + N(as_is_{i-1} + s_ie_{i-1})$ for $i \neq N-1$ and write b_{N-1} as $\mathbf{B}_{N-1}\mathbf{MS} + \mathbf{B}_{N-1}\mathbf{E} + N(as_{N-1}s_{N-2} + s_{N-1}e_{N-2}) + e''_{N-1}$. It is straightforward to see that, entries of $\mathbf{MS}$ and Nas_is_{i-1} are eliminated through the process of computing $b_{N-1} - b_i$. Thus we obtain

$$b_{N-1} - b_i = (\mathbf{B}_{N-1} - \mathbf{B}_i)\,\mathbf{E} + N(s_{N-1}e_{N-2} - s_i e_{i-1}) + e''_{N-1}$$

$$= (N-i-1)\cdot\left(\sum_{\substack{j\in\mathbb{Z}\cap[0,i-1]\\ \text{and } j=N-1}} s_j e_{j+1} - s_j e_{j-1} + e'_j\right) + e''_{N-1}$$

$$+ (-i-1)\left(\sum_{j=i}^{N-2} s_j e_{j+1} - s_j e_{j-1} + e'_j\right) + N(s_{N-1}e_{N-2} - s_i e_{i-1})$$

Observe that for an arbitrary $i \in \{0, 1, ..., N-1\}$, and in any coordinate of the sum above, there are at most $(N^2 + 2N)$ terms in form of $s_u e_v$, at most $N^2/2$ terms in form of e'_w sampled from χ_{σ_1}, at most $N-2$ terms of e'_0 sampled from χ_{σ_2}, and one term of e''_{N-1}.

Let $\mathsf{product}_{\mathsf{ALL}}$ denote the event that for all the terms in form of $s_u e_v$ observed above, each coefficient of such term is bounded by $\sqrt{n}\rho^{3/2}\sigma_1^2$. Under that assumption that assuming $2n(N^2 + 2N) \leq 2^\rho$ (which holds for all reasonable settings of the parameters) and using a union bound, it is straightforward to see

$$\Pr[\overline{\mathsf{product}_{\mathsf{ALL}}}|\mathsf{bound}_\rho] \leq (N^2 + 2N)\cdot 2n2^{-2\rho} \leq 2^{-\rho}.$$

Let fail be the event that not all parties agree on the same final key. Given the constraint $(N^2 + 2N)\cdot\sqrt{n}\rho^{3/2}\sigma_1^2 + (\frac{N^2}{2} + 1)\sigma_1 + (N-2)\sigma_2 \leq \beta_{\mathsf{Rec}}$ satisfied, we

have

$$\Pr[\mathsf{fail}] = \Pr[\mathsf{fail}|\mathsf{bound}_\rho] \cdot \Pr[\mathsf{bound}_\rho] + \Pr[\mathsf{fail}|\overline{\mathsf{bound}_\rho}] \cdot \Pr[\overline{\mathsf{bound}_\rho}] \tag{3.4}$$

$$\leq \Pr[\overline{\mathsf{product}_{\mathsf{ALL}}}] \cdot 1 + 1 \cdot \Pr[\overline{\mathsf{bound}_\rho}] \leq 2 \cdot 2^{-\rho}, \tag{3.5}$$

which completes the proof.

□

3.4 Proof of Security

Here we prove security of our protocol Π. We remark that our proof considers only a classical attacker; in particular, we only allow the attacker classical access to $\mathcal{H}$. We leave proving the protocol can be proven secure even against attackers that are allowed to make quantum queries to $\mathcal{H}$ to future work.

Theorem 3.4. *Assume* $2N\sqrt{n}\,\lambda^{3/2}\,\sigma_1^2 + (N-1)\cdot\sigma_1 \leq \beta_{\mathsf{R\acute{e}nyi}}$ *and* $\beta_{\mathsf{R\acute{e}nyi}} < \sigma_2 < q$*, and model* $\mathcal{H}$ *as a random oracle. Then*

$$\mathsf{Adv}_{\Pi}^{\mathsf{GKE}}(t,\mathsf{q}) \leq 2^{-\lambda+1} + \sqrt{\left(N \cdot \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1) + \mathsf{Adv}_{\mathsf{KeyRec}}(t_2) + \frac{\mathsf{q}}{2^{1^\lambda}}\right) \cdot \frac{\exp\left(2\pi n\left(\beta_{\mathsf{R\acute{e}nyi}}/\sigma_2\right)^2\right)}{1-2^{-\lambda+1}}},$$

where $t_1 = t + \mathcal{O}(N \cdot t_{\mathsf{ring}}), t_2 = t + \mathcal{O}(N \cdot t_{\mathsf{ring}})$ *and* t_{ring} *is the time required to perform operations in* R_q*.*

Proof. Let Expt_0 refer to the experiment in which protocol Π is executed to obtain

output $(\mathsf{T}, \mathsf{sk})$, where $\mathsf{T} = (\{z_i\}, \{X_i\}, hint)$ is the transcript of the execution and sk is the final shared session key (more formally, the session key output by P_{N-1}). We also then provide the attacker $\mathcal{A}$ with $(\mathsf{T}, \mathsf{sk})$, and then allow $\mathcal{A}$ to interact with the random oracle used when executing Π. Our goal is to bound the advantage of an attacker in distinguishing between samples $(\mathsf{T}, \mathsf{sk})$ distributed according to Expt_0 and samples $(\mathsf{T}, \mathsf{sk}')$ in which T is distributed the same way but sk' is a uniform key (chosen independently of T). To do so, we show that the probability that $\mathcal{A}$ queries k_{N-1} to the random oracle (which we denote by the event Query) is small; since that is the only way an attacker can distinguish $\mathsf{sk} = \mathcal{H}(k_{N-1})$ from an independent, uniform value, that allows us to prove our desired result. In proving our result, we consider a sequence of experiments, and let $\Pr_i[\cdot]$ denote the probability of an event in Experiment i.

For completeness, we write out the distribution of $(\mathsf{T}, \mathsf{sk})$ in Expt_0:

$$\mathsf{Expt}_0 := \left\{ \begin{array}{l} a \leftarrow R_q;\ \forall i : s_i, e_i \leftarrow \chi_{\sigma_1}; z_i = as_i + e_i; \\ \\ e'_1, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}; e'_0 \leftarrow \chi_{\sigma_2}; \\ \\ \forall i : X_i = (z_{i+1} - z_{i-1})s_i + e'_i; \\ \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; \qquad\qquad\qquad\qquad\qquad\qquad : (\mathsf{T}, \mathsf{sk}) \\ \\ b_{N-1} = e''_{N-1} + z_{N-2}Ns_{N-1} + X_{N-1} \cdot (N-1) + \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; \\ \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); \\ \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint) \end{array} \right\}.$$

Since $\mathsf{Adv}^{\mathsf{GKE}}_{\Pi}(t, \mathsf{q}) \leq \mathrm{Pr}_0[\mathsf{Query}]$, we focus on bounding $\mathrm{Pr}_0[\mathsf{Query}]$ for the rest of the proof.

Experiment 1. In this experiment, X_0 is replaced by $X'_0 = -\sum_{i=1}^{N-1} X_i + e'_0$. The

corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows:

$$\mathsf{Expt}_1 := \left\{ \begin{array}{ll} a \leftarrow R_q;\ \forall i : s_i, e_i \leftarrow \chi_{\sigma_1}; z_i = a s_i + e_i; & \\ e'_1, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}; e'_0 \leftarrow \chi_{\sigma_2} & \\ X'_0 = -\sum_{i=1}^{N-1} X_i + e'_0; & \\ \forall i > 0 : X_i = (z_{i+1} - z_{i-1}) s_i + e'_i & : (\mathsf{T}, \mathsf{sk}) \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; & \\ b_{N-1} = e''_{N-1} + z_{N-2} N s_{N-1} + X_{N-1} \cdot (N-1) + & \\ \qquad X'_0 \cdot (N-2) + \cdots + X_{N-3}; & \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X'_0, \ldots, X_{N-1}, hint) & \end{array} \right\}.$$

The following claim, which is the crux of our proof, relates the probabilities of Query in Expt_0 and Expt_1.

Claim 3.5. *If* $2N\sqrt{n}\,\lambda^{3/2}\,\sigma_1^2 + (N-1)\cdot\sigma_1 \leq \beta_{\mathsf{R\acute{e}nyi}}$, *then*

$$\Pr_0[\mathsf{Query}] \leq \sqrt{\Pr_1[\mathsf{Query}] \cdot \frac{\exp(2\pi n(\beta_{\mathsf{R\acute{e}nyi}}/\sigma_2)^2)}{1 - 2^{-\lambda+1}}} + 2^{-\lambda+1}. \tag{3.6}$$

Proof. Note that we may define the random variables X_0, X'_0 in both experiments Expt_0 and Expt_1. Define the random variable Error (in either experiment) as

$$\mathsf{Error} = \sum_{i=0}^{N-1} (s_i e_{i+1} + s_i e_{i-1}) + \sum_{i=1}^{N-1} e'_i.$$

Defining

$$\mathsf{main} = as_1s_0 - as_{N-1}s_0 - \mathsf{Error},$$

it is straightforward to verify that

$$\begin{aligned} X_0 &= \mathsf{main} + \mathsf{Error} + e'_0 \\ X'_0 &= \mathsf{main} + e'_0, \end{aligned}$$

where e'_0 is sampled from χ_{σ_2}. Our aim is to apply Theorem 2.3 to show that the Rényi divergence between X_0 and X'_0 (and hence between Expt_0 and Expt_1) is small. To do so, we must first show that the absolute value of each coefficient of Error is bounded by $\beta_{\mathsf{Rényi}}$ with all but negligible probability.

Let $\mathsf{bound}_{\mathsf{Err}}$ be the event that for all j we have $|\mathsf{Error}_j| \leq \beta_{\mathsf{Rényi}}$. Note that

$$|\mathsf{Error}_j| = \left| \left(\sum_{i=0}^{N-1} (s_i e_{i+1} + s_i e_{i-1}) + \sum_{i=1}^{N-1} e'_i \right)_j \right|.$$

Fix $c = \sqrt{\frac{2\lambda}{\pi \log e}}$, and let bound be the event that for all i, j we have $|(e'_0)_j| \leq c\sigma_2$ and $|(s_i)_j|,\ |(e_i)_j|,\ |(e''_{N-1})_j| \leq c\sigma_1$, and that for all $i > 0$ and all j it holds that $|(e'_i)_j| \leq c\sigma_1$. Applying Lemmas 3.2 and 3.3 (with $\rho = \lambda$), we see that

$$\Pr[\mathsf{bound}] \geq 1 - 2^{-\lambda}$$

and

$$\Pr\left[|(s_i e_j)_v| \geq \sqrt{n}\lambda^{3/2}\sigma_1^2 \mid \mathsf{bound}\right] \leq 2^{-2\lambda+1}.$$

Via a union bound, we thus have

$$\Pr\left[\forall j : |\mathsf{Error}_j| \leq 2N\sqrt{n}\lambda^{3/2}\sigma_1^2 + (N-1)\sigma_1 \mid \mathsf{bound}\right] \geq 1 - 4N \cdot n \cdot 2^{-2\lambda}.$$

Under the assumption that $4Nn \leq 2^\lambda$ (which holds for all reasonable settings of the parameters) and using a similar argument as in the proof of Lemma 3.3, we conclude that

$$\Pr[\mathsf{bound}_{\mathsf{Err}}] \geq 1 - 2^{-\lambda+1}. \tag{3.7}$$

When $\mathsf{bound}_{\mathsf{Err}}$ occurs, Theorem 2.3 tells us that

$$\mathrm{RD}_2(\mathsf{Error} + \chi_{\sigma_2} || \chi_{\sigma_2}) \leq \exp(2\pi n(\beta_{\mathsf{Rényi}}/\sigma_2)^2). \tag{3.8}$$

Therefore,

$$\begin{aligned}
\Pr_0[\mathsf{Query}] &\leq \Pr_0[\mathsf{Query} \mid \mathsf{bound}_{\mathsf{Err}}] + \Pr_0[\overline{\mathsf{bound}_{\mathsf{Err}}}] \\
&\leq \Pr_0[\mathsf{Query} \mid \mathsf{bound}_{\mathsf{Err}}] + 2^{-\lambda+1} \\
&\leq \sqrt{\Pr_1[\mathsf{Query} \mid \mathsf{bound}_{\mathsf{Err}}] \cdot \exp(2\pi n(\beta_{\mathsf{Rényi}}/\sigma_2)^2)} + 2^{-\lambda+1} \\
&\leq \sqrt{\Pr_1[\mathsf{Query}] \cdot \frac{\exp(2\pi n(\beta_{\mathsf{Rényi}}/\sigma_2)^2)}{\Pr_1[\mathsf{bound}_{\mathsf{Err}}]}} + 2^{-\lambda+1} \\
&\leq \sqrt{\Pr_1[\mathsf{Query}] \cdot \frac{\exp(2\pi n(\beta_{\mathsf{Rényi}}/\sigma_2)^2)}{1 - 2^{-\lambda+1}}} + 2^{-\lambda+1}.
\end{aligned}$$

This completes the proof of the claim. □

Recall that Experiment 0 is the real world experiment. We have that $\mathsf{Adv}_\Pi^{\mathsf{GKE}}(t, \mathsf{q}) \leq$

$\Pr_0[\mathsf{Query}]$, where Query is the event that k_{N-1} is among the adversary $\mathcal{A}$'s random oracle queries and $\Pr_i[\mathsf{Query}]$ is the probability that event Query happens in Experiment i.

In Experiment 1, we switched from X_0 as sampled in the real world to $X_0' = -\sum_{i=1}^{N-1} X_i + e_0'$ and showed (see Equation 3.6) that

$$\Pr_0[\mathsf{Query}] \leq \sqrt{\Pr_1[\mathsf{Query}] \cdot \frac{\exp(2\pi n(\beta_{\mathsf{R\acute{e}nyi}}/\sigma_2)^2)}{1-2^{-\lambda+1}}} + 2^{-\lambda+1}.$$

Therefore, to prove the theorem, it remains to show that

$$\Pr_1[\mathsf{Query}] \leq \left(N \cdot \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1) + \mathsf{Adv}_{\mathsf{KeyRec}}(t_2) + \frac{\mathsf{q}}{2^{1^\lambda}}\right).$$

We do so by considering a sequence of experiments as follows:

Experiment 2. In this experiment, z_0 is replaced by a uniform element in R_q. The

corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows:

$$\mathsf{Expt}_2 := \left\{ \begin{array}{l} a, z_0 \leftarrow R_q;\ \forall i \geq 1 : s_i, e_i \leftarrow \chi_{\sigma_1}; z_i = a s_i + e_i; \\ \\ e'_1, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}; e'_0 \leftarrow \chi_{\sigma_2} \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, \forall i \geq 1 : X_i = (z_{i+1} - z_{i-1}) s_i + e'_i \quad : (\mathsf{T}, \mathsf{sk}) \\ \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; \\ \\ b_{N-1} = e''_{N-1} + z_{N-2} N s_{N-1} + X_{N-1} \cdot (N-1) + \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; \\ \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); \\ \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). \end{array} \right\}.$$

Claim 3.6. *For any algorithm $\mathcal{A}$ running in time t, we have*

$$|\Pr_2[\mathsf{Query}] - \Pr_1[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1), \tag{3.9}$$

where $t_1 = t + \mathcal{O}(N \cdot t_{\mathsf{ring}})$ and t_{ring} is the time required to perform operations in R_q.

Proof. We first consider an experiment Expt'_1 which is identical to Expt_1 except for (a, z_0) given as input. For algorithm $\mathcal{A}$ running in time t, let $\mathcal{B}$ be an algorithm running in time t_1 which takes as input (a, z_0), generates $(\mathsf{T}, \mathsf{sk})$ according to Expt'_1, runs $\mathcal{A}(\mathsf{T}, \mathsf{sk})$ as a subroutine and outputs whatever $\mathcal{A}$ outputs. t_1 is then equal to t plus a minor overhead for the simulation of the security experiment for $\mathcal{A}$.

It is straightforward to see that if (a, z_0) is sampled from $A_{n,q,\chi_{\sigma_1}}$, then Expt'_1

is identical to Expt_1, and if (a, z_0) is sampled from R_q^2, Expt_1' is identical to Expt_2.

Therefore the difference of algorithm $\mathcal{A}$'s success probability in Experiment 1 and Experiment 2 is bounded by probability that $\mathcal{B}$ running in time t_1 distinguishes $A_{n,q,\chi_{\sigma_1}}$ from R_q^2 given one sample. Since

$$\mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1) \geq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},2}(t_1) \geq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},1}(t_1),$$

for simplicity, we conclude that:

$$|\Pr_2[\mathsf{Query}] - \Pr_1[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1), \tag{3.10}$$

□

Recall that in the previous experiment, we switched z_0 to be uniformly distributed in R_q. In next two experiments, we switch z_1, X_1 to be elements uniformly distributed in R_q.

Experiment 3. In this experiment, z_0 is replaced by $z_2 - r_1$, and X_1 is replaced by $r_1 s_1 + e_1'$, where r_1 is uniform in R_q. The corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows:

Since r_1 is uniform, then $z_2 - r_1$ is also uniform. Thus, we conclude that Experiment 3 is identical to Experiment 2 up to variable substitution, namely

$$\Pr_3[\mathsf{Query}] = \Pr_2[\mathsf{Query}]. \tag{3.11}$$

$$\mathsf{Expt}_3 := \left\{ \begin{array}{l} a, r_1 \leftarrow R_q;\ \forall i \geq 1 : s_i, e_i \leftarrow \chi_{\sigma_1}; z_i = as_i + e_i; \\ \\ z_0 = z_2 - r_1; \\ \\ \forall i \geq 1 : e'_i \leftarrow \chi_{\sigma_1};\ e'_0 \leftarrow \chi_{\sigma_2}; \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0; X_1 = r_1 s_1 + e'_1; \qquad\qquad : (\mathsf{T}, \mathsf{sk}) \\ \\ \forall i \geq 2 : X_i = (z_{i+1} - z_{i-1}) s_i + e'_i; \\ \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; \\ \\ b_{N-1} = e''_{N-1} + z_{N-2} N s_{N-1} + X_{N-1} \cdot (N-1) + \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; \\ \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); \\ \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). \end{array} \right\}.$$

Experiment 4. In this experiment, z_1, X_1 are replaced by uniform elements in R_q. The corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows:

$$\mathsf{Expt}_4 := \left\{ \begin{array}{ll} a, r_1 \leftarrow R_q;\ \forall i \geq 2 : s_i, e_i \leftarrow \chi_{\sigma_1}; z_i = as_i + e_i; & \\ z_0 = z_2 - r_1, z_1 \leftarrow R_q; & \\ e'_2, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}; e'_0 \leftarrow \chi_{\sigma_2}; & \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1 \leftarrow R_q; & \\ \forall i \geq 2 : X_i = (z_{i+1} - z_{i-1})s_i + e'_i, & : (\mathsf{T}, \mathsf{sk}) \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; & \\ b_{N-1} = e''_{N-1} + z_{N-2}Ns_{N-1} + X_{N-1} \cdot (N-1) + & \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; & \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). & \end{array} \right\}.$$

Claim 3.7. *For any algorithm $\mathcal{A}$ running in time t, we have*

$$|\Pr_4[\mathsf{Query}] - \Pr_3[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1), \tag{3.12}$$

where $t_1 = t + \mathcal{O}(N \cdot t_{\mathsf{ring}})$ and t_{ring} is the time required to perform operations in R_q.

Proof. We first consider an experiment Expt'_3 which is identical to Expt_3 except for (a, z_1), (r_1, X_1) given as input. For algorithm $\mathcal{A}$ running in time t, let $\mathcal{B}$ be an algorithm running in time t_1 that takes as input (a, z_1), (r_1, X_1), generates $(\mathsf{T}, \mathsf{sk})$ according to Expt'_3. $\mathcal{B}$ then runs $\mathcal{A}(\mathsf{T}, \mathsf{sk})$ as a subroutine and outputs whatever $\mathcal{A}$

outputs. t_1 is then equal to t plus a minor overhead for the simulation of the security experiment for $\mathcal{A}$.

It is clear to see that if (a, z_1) and (r_1, X_1) are sampled from $A_{n,q,\chi_{\sigma_1}}$, then Expt_3' is identical to Expt_3. If (a, z_1) and (r_1, X_1) are sampled from $\mathcal{U}(R_q^2)$, Expt_3' is identical to Expt_4.

Therefore the difference of algorithm $\mathcal{A}$ successful probability in Experiment 3 and Experiment 4 is bounded by the advantage of adversary $\mathcal{B}$ running in time t_1 in distinguishing $A_{n,q,\chi_{\sigma_1}}$ from $\mathcal{U}(R_q^2)$ given two samples. Thus, we conclude

$$|\Pr_4[\mathsf{Query}] - \Pr_3[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1). \tag{3.13}$$

□

Experiment 5. In this experiment, z_0 is replaced by a uniform element in R_q. The corresponding distribution is denoted as Expt_5. We leave the formal definition of Expt_5 implicit for simplicity

It is easy to see that the corresponding distribution Expt_5 is identical to Expt_4 by substituting variable z_0 for $z_2 - r_1$. Thus,

$$\Pr_5[\mathsf{Query}] = \Pr_4[\mathsf{Query}]. \tag{3.14}$$

In the case that $N \geq 3$, we present the following sequence of experiments from Experiment 6 to Experiment $3N - 4$. For $i = 2, 3, \ldots, N - 2$, we define three experiments Experiment $3i$, Experiment $3i + 1$, Experiment $3i + 2$. It is ensured

that in the experiments prior to Experiment $3i$, we already switched z_j, X_j for all $0 \leq j \leq i-1$. In Experiment $3i$, Experiment $3i+1$ and Experiment $3i+2$, we replace z_i and X_i by random elements in R_q. Experiment $3i$, Experiment $3i+1$, Experiment $3i+2$ are formally defined as follows:

Experiment $3i$**.** The experiment proceeds exactly the same as Experiment $3i-1$, except for setting $z_{i-1} = z_{i+1} - r_i, X_i = r_i s_i + e_i'$, where r_1 is uniform in R_q. The corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows, denoted Expt_{3i}:

Experiment $3i+1$**.** In this experiment, z_i, X_i are replaced by uniform elements in R_q. The corresponding distribution of $(\mathsf{T}, \mathsf{sk})$ is thus as follows, denoted Expt_{3i+1}:

Experiment $3i+2$**.** In this experiment, z_{i-1} is replaced by a uniform element in R_q. The corresponding distribution is denoted as Expt_{3i+2}. We leave the formal definition of Expt_{3i+2} implicit for simplicity.

$$\mathsf{Expt}_{3i} := \left\{ \begin{array}{ll} a, r_i \leftarrow R_q;\ \forall j \geq i : s_j, e_j \leftarrow \chi_{\sigma_1}; z_j = a s_j + e_j; & \\ & \\ z_0, \ldots, z_{i-2} \leftarrow R_q, z_{i-1} = z_{i+1} - r_i; & \\ & \\ e'_i, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}, e'_0 \leftarrow \chi_{\sigma_2}; & \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1, \ldots, X_{i-1} \leftarrow R_q; & : (\mathsf{T}, \mathsf{sk}) \\ & \\ X_i = r_i s_i + e'_i;\ \forall j \geq i : X_{j+1} = (z_{j+2} - z_j) s_{j+1} + e'_{j+1} & \\ & \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; & \\ & \\ b_{N-1} = e''_{N-1} + z_{N-2} N s_{N-1} + X_{N-1} \cdot (N-1) + & \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; & \\ & \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); & \\ & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). & \end{array} \right\}.$$

$$\mathsf{Expt}_{3i+1} := \left\{ \begin{array}{ll} a, r_i \leftarrow R_q;\ \forall j \geq i+1 : s_j, e_j \leftarrow \chi_{\sigma_1}; z_j = as_j + e_j; & \\ z_0, \ldots, z_{i-2} \leftarrow R_q, z_{i-1} = z_{i+1} - r_i, z_i \leftarrow R_q, & \\ e'_1, \ldots, e'_{N-1} \leftarrow \chi_{\sigma_1}; e'_0 \leftarrow \chi_{\sigma_2} & \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1, \ldots, X_i \leftarrow R_q, & : (\mathsf{T}, \mathsf{sk}) \\ \forall j \geq i+1, X_j = (z_{j+1} - z_{j_1})s_j + e'_j; & \\ e''_{N-1} \leftarrow \chi_{\sigma_1}; & \\ b_{N-1} = e''_{N-1} + z_{N-2} N s_{N-1} + X_{N-1} \cdot (N-1) + & \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; & \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). & \end{array} \right\}.$$

Using similar arguments as proving (in)equalities (3.11), (3.12) and (3.14), we conclude that:

$$\Pr_{3i}[\mathsf{Query}] = \Pr_{3i-1}[\mathsf{Query}]; \tag{3.15}$$

$$|\Pr_{3i+1}[\mathsf{Query}] - \Pr_{3i}[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1); \tag{3.16}$$

$$\Pr_{3i+2}[\mathsf{Query}] = \Pr_{3i+1}[\mathsf{Query}]; \tag{3.17}$$

Note that in Experiment $3N-4$, the last experiment of the experiment sequence above, we already switched all the z_i, X_i up to z_{N-1}, X_{N-1}. We construct

the next two experiments to switch $z_{N-1}, X_{N-1}, b_{N-1}$.

Experiment $3N-3$**.** The experiment proceeds exactly the same as Experiment $3N-4$, except for setting $z_{N-2} = r_2, X_{N-1} = r_1 s_{N-1} + e'_{N-1}, z_0 = r_1 + r_2$, where r_1, r_2 are uniform in R_q. The corresponding distribution is thus as follows:

Since r_1, r_2 are uniform, $r_1 + r_2$ is then also uniform. Thus we conclude that Experiment $3N-3$ is identical to Experiment $3N-4$ up to variable substitution, namely,

$$\Pr_{3N-3}[\mathsf{Query}] = \Pr_{3N-4}[\mathsf{Query}]; \tag{3.18}$$

$$\mathsf{Expt}_{3N-3} := \left\{ \begin{array}{ll} a, r_1, r_2 \leftarrow R_q, s_{N-1}, e_{N-1} \leftarrow \chi_{\sigma_1}; z_0 = r_1 + r_2, & \\ z_1, \ldots, z_{N-3} \leftarrow R_q, z_{N-2} = r_2, & \\ z_{N-1} = a s_{N-1} + e_{N-1}; e'_0 \leftarrow \chi_{\sigma_2}; e'_{N-1} \leftarrow \chi_{\sigma_1}; & \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1, \ldots, X_{N-2} \leftarrow R_q, & \\ X_{N-1} = r_1 s_{N-1} + e'_{N-1}; e''_{N-1} \leftarrow \chi_{\sigma_1}; & : (\mathsf{T}, \mathsf{sk}) \\ b_{N-1} = e''_{N-1} + r_2 N s_{N-1} + X_{N-1} \cdot (N-1) + & \\ \qquad X_0 \cdot (N-2) + \cdots + X_{N-3}; & \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). & \end{array} \right\}.$$

Experiment $3N-2$**.** In this experiment, $z_{N-1}, X_{N-1}, b_{N-1}$ are replaced by uniform elements in R_q. The corresponding distribution is thus as follows: :

$$\mathsf{Expt}_{3N-2} := \left\{ \begin{array}{l} a \leftarrow R_q; \forall i : z_i \leftarrow R_q; \\ \\ e'_0 \leftarrow \chi_{\sigma_2}; r_1, r_2 \leftarrow R_q \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1, \ldots, X_{N-1} \leftarrow R_q \qquad : (\mathsf{T}, \mathsf{sk}) \\ \\ b_{N-1} \leftarrow R_q; \\ \\ (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}); \mathsf{sk} = \mathcal{H}(k_{N-1}); \\ \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint). \end{array} \right\}.$$

Claim 3.8. *For any algorithm $\mathcal{A}$ running in time t, we have*

$$|\Pr_{3N-2}[\mathsf{Query}] - \Pr_{3N-3}[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1), \tag{3.19}$$

where $t_1 = t + \mathcal{O}(N \cdot t_{\mathsf{ring}})$ and t_{ring} is the time required to perform operations in R_q.

Proof. Since r_2 is uniform in R_q and N is invertible over R_q, then $r_2 N$ is uniformly distributed in R_q. It is easy to see that $(s_{N-1},\ r_2 N s_{N-1} + e''_{N-1})$ forms an RLWE instance. We let $b_{\mathsf{RLWE}} = r_2 N s_{N-1} + e''_{N-1}$.

We consider an experiment Expt'_{3N-3} which is identical to Expt_{3N-3} except for (a, z_{N-1}), (r_1, X_{N-1}), and $(r_2 N, b_{\mathsf{RLWE}})$ given as input. Given an algorithm $\mathcal{A}$ running in time t, let $\mathcal{B}$ be an algorithm that takes as input (a, z_{N-1}), (r_1, X_{N-1}), and $(r_2 N, b_{\mathsf{RLWE}})$, generates $(\mathsf{T}, \mathsf{sk})$ according to Expt'_{3N-3}. $\mathcal{B}$ runs $\mathcal{A}(\mathsf{T}, \mathsf{sk})$ as a subroutine and outputs whatever $\mathcal{A}$ outputs. Running time t_1 of $\mathcal{B}$ then equals to

t plus a minor overhead for the simulation of the security experiment for $\mathcal{A}$.

It is straightforward to see that if (a, z_{N-1}), (r_1, X_1), and $(r_2 N, b_{\mathsf{RLWE}})$ are sampled from $A_{n,q,\chi_{\sigma_1}}$, then Expt'_{3N-3} is identical to Expt_{3N-3}. If (a, z_{N-1}), (r_1, X_{N-1}), and $(r_2 N, b_{\mathsf{RLWE}})$ are sampled from R_q^2, then Expt'_{3N-3} is identical to Expt_{3N-2}, since when b_{RLWE} is sampled uniformly at random, $b_{\mathsf{RLWE}} + X_{N-1} \cdot (N-1) + X_0 \cdot (N-2) + \cdots + X_{N-3}$ is also uniformly distributed over R_q.

Therefore the difference of algorithm $\mathcal{A}$'s success probability in Experiment 3N - 2 and Experiment 3N - 3 is bounded by the advantage of adversary $\mathcal{B}$ running in time t_1 in distinguishing Ring-LWE from R_q given three samples. Thus, we conclude that

$$|\Pr_{3N-2}[\mathsf{Query}] - \Pr_{3N-3}[\mathsf{Query}]| \leq \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1), \tag{3.20}$$

□

Experiment $3N-1$**.** In this experiment, k_{N-1} is replaced by random element in $\{0,1\}^{1^\lambda}$. The corresponding distribution is thus as follows:

$$\mathsf{Expt}_{\mathsf{final}} := \left\{ \begin{array}{ll} a \leftarrow R_q; z_0, \ldots, z_{N-1} \leftarrow R_q; e'_0 \leftarrow \chi_{\sigma_1}; & \\ X_0 = -\sum_{i=1}^{N-1} X_i + e'_0, X_1, \ldots, X_{N-1} \leftarrow R_q & \\ b_{N-1} \leftarrow R_q; (hint, k_{N-1}) = \mathsf{recMsg}(b_{N-1}) & : (\mathsf{T}, \mathsf{sk}) \\ k'_{N-1} \leftarrow \{0,1\}^\lambda; \mathsf{sk} = \mathcal{H}(k'_{N-1}); & \\ \mathsf{T} = (z_0, \ldots, z_{N-1}, X_0, \ldots, X_{N-1}, hint); & \end{array} \right\}.$$

Given transcript T, and b_{N-1} which is uniformly distributed, using a straight forward reduction, we obtain advantage of adversary $\mathcal{B}$ running in time t_2 in distinguishing k_{N-1} computed by $\mathsf{recMsg}(b_{N-1})$ from a uniform bit string k'_{N-1} with length λ is at least $|\Pr_{3N-1}[\mathsf{Query}] - \Pr_{3N-2}[\mathsf{Query}]|$, namely,

$$|\Pr_{3N-1}[\mathsf{Query}] - \Pr_{3N-2}[\mathsf{Query}]| \leq \mathsf{Adv}_{\mathsf{KeyRec}}(t_2). \tag{3.21}$$

Note that t_2 equals to the running time of adversary $\mathcal{A}$ attacking the protocol Π, plus a minor overhead for simulating experiment for $\mathcal{A}$.

Finally, since adversary attacking the GKE protocol Π makes at most q queries to the random oracle, $\Pr_{3N-1}[\mathsf{Query}] = \frac{\mathsf{q}}{2^{1^\lambda}} \in \mathsf{negl}(1^\lambda)$. Combining Equations (3.9) - (3.21), we have

$$\Pr_1[\mathsf{Query}] \leq N \cdot \mathsf{Adv}^{\mathsf{RLWE}}_{n,q,\chi_{\sigma_1},3}(t_1) + \mathsf{Adv}_{\mathsf{KeyRec}}(t_2) + \frac{\mathsf{q}}{2^{1^\lambda}}. \tag{3.22}$$

The theorem now follows immediately from Equations (3.6), and (3.22). □

Parameter constraints. Beyond the parameter settings required for hardness of the Ring-LWE problem, the parameters $N, n, \sigma_1, \sigma_2, \lambda, \rho$ of the protocol are also required to satisfy the following:

$$(N^2 + 2N) \cdot \sqrt{n}\rho^{3/2}\sigma_1^2 + (\frac{N^2}{2} + 1)\sigma_1 + (N-2)\sigma_2 \leq \beta_{\mathsf{Rec}} \quad \text{(correctness)} \tag{3.23}$$

$$2N\sqrt{n}\lambda^{3/2}\sigma_1^2 + (N-1)\sigma_1 \leq \beta_{\mathsf{Rényi}} \quad \text{(security)} \tag{3.24}$$

$$\sigma_2 = \Omega(\beta_{\mathsf{Rényi}}\sqrt{n/\log 1^\lambda}). \quad \text{(security)} \tag{3.25}$$

Thus, fixing the ring, the noise distributions, and the security parameters λ, ρ induces a bound on the maximum number of parties the protocol can support.

Chapter 4: LWE with Side Information: Attacks and Concrete Security Estimation

4.1 Overview

The ongoing standardization process and anticipated deployment of lattice-based cryptography raises an important question: How resilient are lattices to side-channel attacks or other forms of side information? While there are numerous works addressing this question for specific cryptosystems (See [6,25,34,66,101,102] for side

Figure 4.1: Primal attack without hints (prior art).

channel attacks targeting lattice-based NIST candidates), these works use rather ad-hoc methods to reconstruct the secret key, requiring new techniques and algorithms to be developed for each setting. For example, the work of [25] uses brute-force methods for a portion of the attack, while [23] exploits linear regression techniques. Moreover, ad-hoc methods do not allow (1) to take advantage of decades worth of research and (2) optimization of standard lattice attacks. Second, most of the side-channel attacks from prior work consider substantial amounts of information leakage and show that it leads to feasible recovery of the entire key, whereas one may be interested in more precise tradeoffs in terms of information leakage versus concrete security of the scheme. The above motivates the focus of this work: Can one integrate side information into a standard lattice attack and if so, by how much does the information reduce the cost of this attack? Given that side-channel resistance is the next step toward the technological readiness of lattice-based cryptography, and that we expect numerous works in this growing area, we believe that a general framework and prediction software are in order.

Contributions. First, we propose a framework that generalizes the so-called primal lattice reduction attack, and allows the progressive integration of "hints" (i.e. side information that takes one of several forms) before running the final lattice reduction step. This contribution is summarized in Figures 4.1 and 4.2 and developed in

Section 4.2.1.

Second, we implement a Sage 9.0 toolkit to actually mount such attacks with hints when computationally feasible, and to predict their performance on larger instances. Our predictions are validated by extensive experiments. Our tool and these experiments are described in Section A1.2. Our toolkit is open-source, available at: `https://github.com/lducas/leaky-LWE-Estimator`.

Third, we demonstrate the usefulness of our framework and tool via three example applications. Our main example (Section 4.3.1) revisits the side channel information obtained from the first side-channel attack of [25] against Frodo. In that article, it was concluded that a divide-and-conquer side-channel template attack would not lead to a meaningful attack using standard combinatorial search for reconstruction of the secret. Our technique allows to integrate this side-channel information into lattice attacks, and to predict the exact security drop. For example, the CCS2 parameter set very conservatively aims for 128-bits of post-quantum security (or 448 "bikz" as defined in Section 4.2.5); but after the leakage of [25] we predict that its security drops to 29 "bikz", i.e. that it can be broken with BKZ-29, a computation that should be more than feasible, but would require a dedicated re-implementation of our framework.

Interestingly, we note that our framework is not only useful in the side-channel scenario; we are for example also able to model decryption failures as hints fitting our framework. This allows us to reproduce some predictions from [49]. This is discussed in Section 4.3.2.

Perhaps more surprisingly, we also find a novel improvement to attack a few

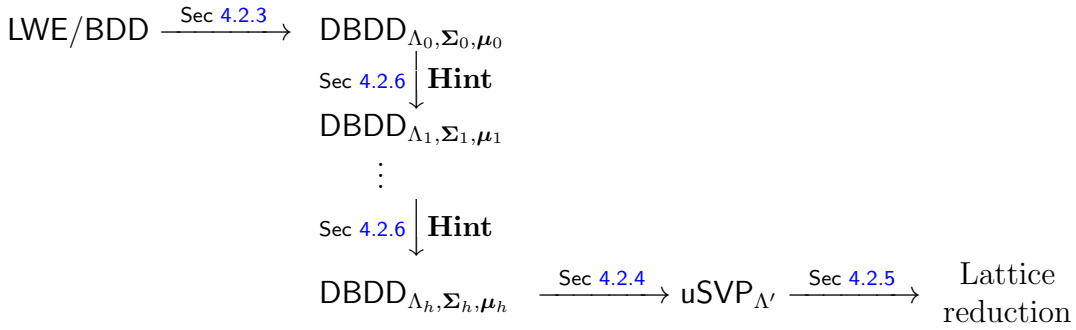


Figure 4.2: The primal attack with hints (our work).

schemes (LAC [83], Round5 [62], NTRU [116]) without any side-channel or oracle queries. Indeed, such schemes use ternary distribution for secrets, with a prescribed number of 1 and -1: this hint fits our framework, and leads to a (very) minor improvement, discussed in Section 4.3.3.

Lastly, our framework also encompasses and streamlines existing tweaks of the primal attack: the choice of ignoring certain LWE equations to optimize the volume-dimension trade-off, as well as the re-centering [94] and isotropization [39,69] accounting for potential a-priori distortions of the secret. It also implicitly solves the question of the optimal choice of the coefficient for Kannan's Embedding from the Bounded Distance Decoding problem (BDD) to the unique Shortest Vector Problem (uSVP) [72] (See Remark 4.9).

As a side contribution, we also propose in Section A1.1 a refined method to estimate the required blocksize to solve an LWE/BDD/uSVP instance. This refinement was motivated by the inaccuracy of the standard method from the literature [7,9] in experimentally reachable blocksizes, which was making the validation of our contribution difficult. While experimentally much more accurate, this new methodology certainly deserves further scrutiny.

4.2 Framework

4.2.1 Overview of Our Framework

Our work is based on a generalization of the Bounded Distance Decoding problem (BDD) to a Distorted version (DBDD), which allows to account for the potentially non-spherical covariance of the secret vector to be found.

Each hint will affect the lattice itself, the mean and/or the covariance parameter of the DBDD instance, making the problem easier (see Figure 4.2). At last, we make the distribution spherical again by applying a well-chosen linear transformation, reverting to a spherical BDD instance before running the attack. Thanks to the hints, this new instance will be easier than the initial one. Let us assume that $\mathbf{v}$, l, k and σ are parameters known by the attacker. Our framework can handle four types of hints on the secret $\mathbf{s}$ or on the lattice Λ.

- Perfect hints: $\langle \mathbf{s},\ \mathbf{v} \rangle = l$ *intersect the lattice with an hyperplane.*
- Modular hints : $\langle \mathbf{s},\ \mathbf{v} \rangle = l \bmod k$ *sparsify the lattice.*
- Approximate hints : $\langle \mathbf{s},\ \mathbf{v} \rangle = l + \epsilon_\sigma$ *decrease the covariance of the secret.*
- Short vector hints : $\mathbf{v} \in \Lambda$ *project orthogonally to* $\mathbf{v}$.

While the first three hints are clear wins for the performance of lattice attacks, the last one is a trade-off between the dimension and the volume of the lattice. This last type of hint is in fact meant to generalize the standard trick consisting of 'ignoring'

certain LWE equations; ignoring such an equation can be interpreted geometrically as such a projection orthogonally to a so-called q-vector.

All the transformations of the lattice above can be computed in polynomial time. However, computing with general distribution in large dimension is not possible; we restrict our study to the case of Gaussian distributions of arbitrary covariance, for which such computations are also poly-time.

Some of these transformations remain quite expensive, in particular because they involve rational numbers with very large denominators, and it remains rather impractical to run them on cryptographic-grade instances. Fortunately, up to a necessary hypothesis of primitivity of the vector $\mathbf{v}$ (with respect to either Λ or its dual depending on the type of hint), we can also predict the effect of each hint on the lattice parameters, and therefore run faster predictions of the attack cost.

From Leaks to Hints. At first, it may not be so clear that the types of hints above are so useful in realistic applications, in particular since they need to be linear on the secret. Of course our framework can handle rather trivial hints such as the perfect leak of a secret coefficient $\mathbf{s}_i = l$. Slightly less trivial is the case where only the low-order bits leaks, a hint of the form $\mathbf{s}_i = l \bmod 2$.

We note that most of the computations done during an LWE decryption are linear: leaking any intermediate register during a matrix vector product leads to a hint of the same form (possibly mod q). Similarly, the leak of a NTT coefficient of a secret in a Ring/Module variant can also be viewed as such.

Admittedly, such ideal leaks of a full register are not the typical scenario and

leaks are typically not linear on the content of the register. However, such non-linearities can be handled by approximate hints. For instance, let $\mathbf{s}_0$ be a secret coefficient (represented by a signed 16-bits integer), whose a priori distribution is supported by $\{-5, \ldots, 5\}$. Consider the case where we learn the Hamming weight of $\mathbf{s}_0$, say $H(\mathbf{s}_0) = 2$. Then, we can narrow down the possibilities to $\mathbf{s}_0 \in \{3, 5\}$. This leads to two hints:

- a modular hint: $\mathbf{s}_0 = 1 \bmod 2$,
- an approximate hint: $\mathbf{s}_0 = 4 + \epsilon_1$, where ϵ_1 has variance 1.

While closer to a realistic scenario, the above example remains rather simplified. A detailed example of how realistic leaks can be integrated as hint will be given in Section 4.3.1, based on the leakage data from [25].

4.2.2 Definition of Distorted Bounded Distance Decoding

We first recall the definition of the (search) LWE problem, in its short-secret variant which is the most relevant to practical LWE-based encryption.

Definition 4.1 (Search LWE problem with short secrets.)**.** *Let n, m and q be positive integers, and let χ be a distribution over $\mathbb{Z}$. The* search LWE problem (with short secrets) *for parameters (n, m, q, χ) is:*

***Given** the pair $\left(\mathbf{A} \in \mathbb{Z}_q^{m \times n}, \mathbf{b} = \mathbf{z}\mathbf{A}^T + \mathbf{e} \in \mathbb{Z}_q^m\right)$ where:*

1. *$\mathbf{A} \in \mathbb{Z}_q^{m \times n}$ is sampled uniformly at random,*

2. $\mathbf{z} \leftarrow \chi^n$, *and* $\mathbf{e} \leftarrow \chi^m$ *are sampled with independent and identically distributed coefficients following the distribution* χ.

Find $\mathbf{z}$.

The primal attack (See for example [7]) against (search)-LWE proceeds by viewing the LWE instance as an instance of a Bounded Distance Decoding (BDD) problem, converting it to a uSVP instance (via Kannan's embedding [72]), and finally applying a lattice reduction algorithm to solve the uSVP instance. The central tool of our framework is a generalization of BDD that accounts for potential distortion in the distribution of the secret noise vector that is to be recovered.

Remark 4.2 (Adaptation to the dual attack)**.** *In principle, our techniques could be adapted to the dual attack as well. We focus on only one for conciseness, and the primal attack appears more pertient and more convenient. Indeed, the dual attack is very rarely better than the primal one [5], and this despite making more simplifications in favor of the attacker. Furthermore, the dual attack has not been the object of experimental verification studies, unlike the primal one. At last, the cost of the dual attack is not necessarily indepedent of the underlying SVP-algorithm; some analysis for example exploit the fact that sieving outputs many short vectors rather than one.*

Definition 4.3 (Distorted Bounded Distance Decoding problem)**.** *Let* $\Lambda \subset \mathbb{R}^d$ *be a lattice,* $\mathbf{\Sigma} \in \mathbb{R}^{d \times d}$ *be a symmetric matrix and* $\boldsymbol{\mu} \in \mathrm{Span}(\Lambda) \subset \mathbb{R}^d$ *such that*

$$\mathrm{Span}(\mathbf{\Sigma}) \subsetneq \mathrm{Span}(\mathbf{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}) = \mathrm{Span}(\Lambda). \tag{4.1}$$

The Distorted Bounded Distance Decoding problem $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ *is the following problem:*

Given $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ *and a basis of* Λ*.*

Find *the unique vector* $\mathbf{x} \in \Lambda \cap E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

where $E(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *denotes the ellipsoid*

$$E(\boldsymbol{\mu}, \boldsymbol{\Sigma}) := \{\mathbf{x} \in \boldsymbol{\mu} + \mathrm{Span}(\boldsymbol{\Sigma}) | (\mathbf{x} - \boldsymbol{\mu}) \cdot \boldsymbol{\Sigma}^{\sim} \cdot (\mathbf{x} - \boldsymbol{\mu})^T \leq \mathrm{rank}(\boldsymbol{\Sigma})\}.$$

We will refer to the triple $\mathcal{I} = (\Lambda, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ *as the instance of the* $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ *problem.*

Intuitively, Definition 4.3 corresponds to knowing that the secret vector $\mathbf{x}$ to be recovered follows a distribution of variance $\boldsymbol{\Sigma}$ and average $\boldsymbol{\mu}$. The quantity $(\mathbf{x} - \boldsymbol{\mu}) \cdot \boldsymbol{\Sigma}^{\sim} \cdot (\mathbf{x} - \boldsymbol{\mu})^T$ can be interpreted as a non-canonical Euclidean squared distance $\|\mathbf{x} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Sigma}}$, and the expected value of such a distance for a Gaussian $\mathbf{x}$ of variance $\boldsymbol{\Sigma}$ and average $\boldsymbol{\mu}$ is $\mathrm{rank}(\boldsymbol{\Sigma})$. One can argue that, for such a Gaussian, there is a constant probability that $\|\mathbf{x} - \boldsymbol{\mu}\|^2_{\boldsymbol{\Sigma}}$ is slightly greater than $\mathrm{rank}(\boldsymbol{\Sigma})$. Since we are interested in the average behavior of our attack, we ignore this benign technical detail. In fact, we will typically interpret DBDD as the promise that the secret follows a Gaussian distribution of center $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

The ellipsoid can be seen as an affine transformation (that we call "distortion") of the centered hyperball of radius $\mathrm{rank}(\boldsymbol{\Sigma})$. Let us introduce a notation for the

hyperball; for any $d \in \mathbb{N}$

$$B_d := \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\|_2 \leq d\}. \tag{4.2}$$

One can thus write using Definition 2.18:

$$E(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = B_{\mathrm{rank}(\boldsymbol{\Sigma})} \cdot \sqrt{\boldsymbol{\Sigma}} + \boldsymbol{\mu}. \tag{4.3}$$

From the Span inclusion in Equation (4.1), one can deduce that the condition is equivalent to requiring $\boldsymbol{\mu} \notin \mathrm{Span}(\boldsymbol{\Sigma})$ and $\mathrm{rank}(\boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}) = \mathrm{rank}(\boldsymbol{\Sigma}) + 1 = \mathrm{rank}(\Lambda)$. This technical detail is necessary for embedding it properly into a uSVP instance (See later in Section 4.2.4).

Particular cases of Definition 4.3. Let us temporarily ignore the condition in Equation (4.1) to study some particular cases. As shown in Figure 4.3, when $\boldsymbol{\Sigma} = \mathbf{I}_d$, $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\mathbf{I}_d}$ is BDD instance. Indeed, the ellipsoid becomes a shifted hyperball $E(\boldsymbol{\mu}, \mathbf{I}_d) = \{\mathbf{x} \in \boldsymbol{\mu} + \mathbb{R}^{d \times d} \mid \|\mathbf{x} - \boldsymbol{\mu}\|_2 \leq d\} = B_d + \boldsymbol{\mu}$. If in addition $\boldsymbol{\mu} = 0$, $\mathsf{DBDD}_{\Lambda,\mathbf{0},\mathbf{I}_d}$ becomes a uSVP instance on Λ.

4.2.3 Embedding LWE into DBDD

In the typical primal attack framework (Figure 4.1), one directly views LWE as a BDD instance of the same dimension. For our purposes, however, it will be useful to apply Kannan's Embedding at this stage and therefore increase the dimension

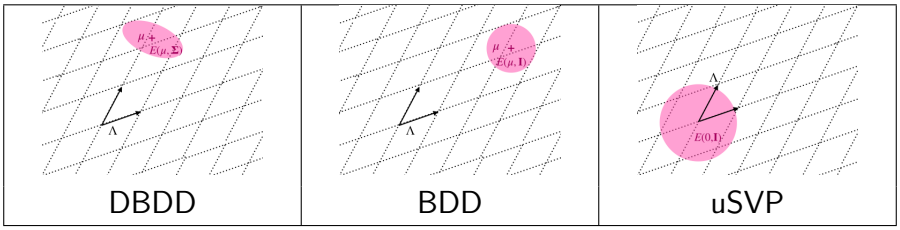


Figure 4.3: Graphical intuition of DBDD, BDD and uSVP in dimension two: the problem consists in finding a nonzero element of Λ in the colored zone. The identity hyperball is larger for uSVP to represent the fact that, during the reduction, the uSVP lattice has one dimension more than for BDD.

of the lattice by 1. While it could be delayed to the last stage of our attack, this extra fixed coefficient 1 will be particularly convenient when we integrate hints (see Remark 4.9 in Section 4.2.6). It should be noted that no information is lost through this transformation, since the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ allow us to encode the knowledge that the solution we are looking for has its last coefficient set to 1 and nothing else. In more details, the solution $\mathbf{s} := (\mathbf{e}, \mathbf{z})$ of an LWE instance is extended to

$$\bar{\mathbf{s}} := (\mathbf{e}, \mathbf{z}, 1) \tag{4.4}$$

which is a short vector in the lattice $\Lambda = \left\{(\mathbf{x}, \mathbf{y}, w) \,|\mathbf{x} + \mathbf{y}\mathbf{A}^{\mathrm{T}} - \mathbf{b}w = 0 \mod q\right\}$. A basis of this lattice is given by the row vectors of

$$\begin{bmatrix} q\mathbf{I}_m & 0 & 0 \\ \mathbf{A}^{\mathrm{T}} & -\mathbf{I}_n & 0 \\ \mathbf{b} & 0 & 1 \end{bmatrix}.$$

Denoting μ_χ and σ_χ^2 the average and variance of the LWE distribution χ (See Definition 4.1), we can convert this LWE instance to a $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ instance with

$\boldsymbol{\mu} = [\mu_\chi \cdots \mu_\chi \; 1]$ and $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_\chi^2 \mathbf{I}_{m+n} & 0 \\ 0 & 0 \end{bmatrix}$. The lattice Λ is of full rank in $\mathbb{R}^d$ where $d := m + n + 1$, and its volume is q^m. Note that the rank of $\boldsymbol{\Sigma}$ is only $d - 1$: the ellipsoid has one less dimension than the lattice. It then validates the requirement of Equation (4.1).

Remark 4.4. *Typically, Kannan's embedding from BDD to uSVP leaves the bottom right matrix coefficient as a free parameter, say c, to be chosen optimally. The optimal value is the one maximizing*

$$\frac{\|(\mathbf{z}; c)\|}{\det(\Lambda)^{1/d}} = \frac{(m+n)\sigma_\chi + c}{(c \cdot q^m)^{1/d}},$$

namely, $c = \sigma_\chi$ according to the arithmetic-geometric mean inequality. Some prior works [7,11] instead chose $c = 1$. While this is benign since σ_χ is typically not too far from 1, it remains a sub-optimal choice. Looking ahead, in our DBDD framework, this choice becomes irrelevant thanks to the isotropization *step introduced in the next section; we can therefore choose $c = 1$ without worwsening the attack.*

4.2.4 Converting DBDD to uSVP

In this Section, we explain how a DBDD instance $(\Lambda, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is converted into a uSVP one. Two modifications are necessary. First, we need to homogeneize the problem. Let us show that the ellipsoid in Definition 4.3 is contained in a larger centered ellipsoid (with one more dimension) as follows:

$$E(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \subset E(\mathbf{0}, \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}). \tag{4.5}$$

Using Equation (4.3), one can write

$$E(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = B_{\text{rank}(\boldsymbol{\Sigma})} \cdot \sqrt{\boldsymbol{\Sigma}} + \boldsymbol{\mu} \subset B_{\text{rank}(\boldsymbol{\Sigma})} \cdot \sqrt{\boldsymbol{\Sigma}} \pm \boldsymbol{\mu},$$

where $B_{\text{rank}(\boldsymbol{\Sigma})}$ is defined in Equation (4.2). And, with Equation (4.1), one can deduce $\text{rank}(\boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}) = \text{rank}(\boldsymbol{\Sigma}) + 1$, then:

$$B_{\text{rank}(\boldsymbol{\Sigma})} \cdot \sqrt{\boldsymbol{\Sigma}} \pm \boldsymbol{\mu} \subset B_{\text{rank}(\boldsymbol{\Sigma})+1} \cdot \begin{bmatrix} \sqrt{\boldsymbol{\Sigma}} \\ \boldsymbol{\mu} \end{bmatrix}.$$

We apply Definition 2.18 which confirms the inclusion of Equation (4.5):

$$E(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \subset B_{\text{rank}(\boldsymbol{\Sigma})+1} \cdot \begin{bmatrix} \sqrt{\boldsymbol{\Sigma}} \\ \boldsymbol{\mu} \end{bmatrix} = E(\mathbf{0}, \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}).$$

Thus, we can homogenize and transform the instance into a centered one with $\boldsymbol{\Sigma}' := \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu}$.

Secondly, to get an isotropic distribution (i.e. with all its eigenvalues being 1), one can just multiply every element of the lattice with the pseudoinverse of $\sqrt{\boldsymbol{\Sigma}'}$. We get a new covariance matrix $\boldsymbol{\Sigma}'' = \sqrt{\boldsymbol{\Sigma}'}^{\sim} \cdot \boldsymbol{\Sigma}' \cdot \sqrt{\boldsymbol{\Sigma}'}^{\sim T} = \boldsymbol{\Pi}_{\boldsymbol{\Sigma}'} \cdot {\boldsymbol{\Pi}_{\boldsymbol{\Sigma}'}}^T$. And since $\boldsymbol{\Pi}_{\boldsymbol{\Sigma}'} = \boldsymbol{\Pi}_{\boldsymbol{\Sigma}'}^T$ and $\boldsymbol{\Pi}_{\boldsymbol{\Sigma}'}^2 = \boldsymbol{\Pi}_{\boldsymbol{\Sigma}'}$ (see Section 2.4.1), $\boldsymbol{\Sigma}'' = \boldsymbol{\Pi}_{\boldsymbol{\Sigma}'} = \boldsymbol{\Pi}_{\Lambda}$, the last equality coming from Equation (4.1).

In summary, one must make by the two following changes:

$$\text{homogenize: } (\Lambda, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \mapsto (\Lambda, \mathbf{0}, \boldsymbol{\Sigma}' := \boldsymbol{\Sigma} + \boldsymbol{\mu}^T \cdot \boldsymbol{\mu})$$

$$\text{isotropize: } (\Lambda, \mathbf{0}, \boldsymbol{\Sigma}') \mapsto (\Lambda \cdot \mathbf{M}, \mathbf{0}, \boldsymbol{\Pi}_\Lambda)$$

where $\mathbf{M} := (\sqrt{\boldsymbol{\Sigma}'})^\sim$. From the solution $\mathbf{x}$ to the $\mathsf{uSVP}_{\Lambda \cdot \mathbf{M}}$ problem, one can derive $\mathbf{x}' = \mathbf{x}\mathbf{M}^\sim$ the solution to the $\mathsf{DBDD}_{\Lambda, \boldsymbol{\mu}, \boldsymbol{\Sigma}}$ problem.

Remark 4.5. *One may note that we could solve a DBDD instance without isotropization simply by including the ellipsoid in a larger ball, and directly apply lattice reduction before the second step. This leads, however, to less efficient attacks. One may also note that the first homogenization step "forgets" some information about the secret's distribution. This, however, is inherent to the conversion to a unique-SVP problem which is geometrically homogeneous, and is already present in the original primal attack.*

4.2.5 Security estimates of uSVP: bikz versus bits

The attack on a uSVP instance consists of applying BKZ-β on the uSVP lattice Λ for an appropriate block size parameter β. The cost of the attack grows with β, however, modeling this cost precisely is at the moment rather delicate, as the state of the art seems to still be in motion. Numerous NIST candidates choose to underestimate this cost, keeping a margin to accommodate future improvements, and there seems to be no clear consensus on which model to use (see [5] for a summary of existing cost models).

While this problem is orthogonal to our work, we still wish to be able to formulate quantitative security losses. We therefore express all concrete security estimates using the blocksize β as our measure of the level of security, and treat the latter as a measurement of the security level in a unit called the *bikz*. We thereby leave the question of the exact bikz-to-bit conversion estimate outside the scope of this paper, and recall that those conversion formulae are not necessarily linear, and may have small dependency in other parameters. For the sake of concreteness, we note that certain choose, for example, to claim 128 bits of security for 380 bikz, and in this range, most models suggest a security increase of one bit every 2 to 4 bikz.

Remark 4.6. *We also clarify that the estimates given in this paper only concern the pure lattice attack via the uSVP embedding discussed above. In particular, we note that some NIST candidates with ternary secrets [83] also consider the hybrid attack of [70], which we ignore in this work. We nevertheless think that the compatibility with our framework is plausible, with some effort.*

Predicting β from a uSVP instance The state-of-the-art predictions for solving uSVP instances using BKZ were given in [7, 9]. Namely, for Λ a lattice of dimension $\dim(\Lambda)$, it is predicted that BKZ-β can solve a uSVP_Λ instance with secret $\mathbf{s}$ when

$$\sqrt{\beta/\dim(\Lambda)} \cdot \|\mathbf{s}\| \leq \delta_\beta^{2\beta-\dim(\Lambda)-1} \cdot \mathrm{Vol}(\Lambda)^{1/\dim(\Lambda)} \tag{4.6}$$

where δ_β is the so called root-Hermite-Factor of BKZ-β. For $\beta \geq 50$, the Root-Hermite-Factor is predictable using the Gaussian Heuristic [38]:

$$\delta_\beta = \left((\pi\beta)^{\frac{1}{\beta}} \cdot \frac{\beta}{2\pi e} \right)^{1/(2\beta-2)}. \tag{4.7}$$

Note that the uSVP instances we generate are isotropic and centered so that the secret has covariance $\mathbf{\Sigma} = \mathbf{I}$ (or $\mathbf{\Sigma} = \mathbf{\Pi}_\Lambda$ if Λ is not of full rank) and $\mu = \mathbf{0}$. Thus, on average, we have $\|\mathbf{s}\|^2 = \mathrm{rank}(\mathbf{\Sigma}) = \dim(\Lambda)$. Therefore, β can be estimated as the minimum integer that satisfies

$$\sqrt{\beta} \leq \delta_\beta^{2\beta - \dim(\Lambda) - 1} \cdot \mathrm{Vol}(\Lambda)^{1/\dim(\Lambda)}. \tag{4.8}$$

While β must be an integer as a BKZ parameter, we nevertheless provide a continuous value, for a finer comparison of the difficulty of an instance. Below, we will call this method the "GSA-Intersect" method.

Remark 4.7. *To predict security, one does not need the basis of Λ, but only its dimension and its volume. Similarly, it is not necessary to explicitly compute the isotropization matrix $\mathbf{M}$ of Section 4.2.4, thanks to Fact 2.20:* $\mathrm{Vol}(\Lambda \cdot \mathbf{M}) = \mathrm{rdet}(\mathbf{M})\,\mathrm{Vol}(\Lambda) = \mathrm{rdet}(\mathbf{\Sigma}')^{-1/2}\,\mathrm{Vol}(\Lambda)$. *These two shortcuts will allow us to efficiently make predictions for cryptographically large instances, in our* lightweight *implementation of Section A1.2.*

Refined prediction for small blocksizes For experimental validation purposes of our work, we prefer to have accurate prediction even for small blocksizes; a regime

where those predictions are not accurate with the current state of the art. We therefore present a refined strategy using BKZ-simulation and a probabilistic model in Appendix A1.1.

4.2.6 Hints and their integration

In thie section, we define several categories of hints—**perfect hints**, **modular hints**, **approximate hints** (**conditioning** and ***a posteriori***), and **short vector hints**—and show that these types of hints can be integrated into a DBDD instance. Hints belonging to these categories typically have the form of a linear equation in $\mathbf{s}$ (and possibly additional variables). As emphasized in Section 4.1, these hints have lattice-friendly forms and their usefulness in realistic applications may not be obvious. We refer to Section 4.3 for detailed applications of these hints.

The technical challenge, therefore, is to characterize the effect of such hints on the DBDD instance—i.e. determine the resulting $(\Lambda', \boldsymbol{\mu}', \boldsymbol{\Sigma}')$ of the new DBDD instance, after the hint is incorporated.

Henceforth, let $\mathcal{I} = \mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ be a fixed instance constructed from an LWE instance with secret $\mathbf{s} = (\mathbf{z}, \mathbf{e})$. Each hint will introduce new constraints on $\mathbf{s}$ and will ultimately decrease the security level.

Non-Commutativity. It should be noted that many types of hints commute: Integrating them in any order will lead to the same DBDD instance. Potential exceptions are **non-smooth modular hints** (See later in Section 4.2.6.2) and ***a posteriori* approximate hints** (See later in Section 4.2.6.4): they do not always

commute with the other types of hints, and do not always commute between themselves, unless the vectors $\mathbf{v}$'s of those hints are all orthogonal to each other. The reason is: in these cases, the distribution in the direction of $\mathbf{v}$ is redefined which erases the prior information.

4.2.6.1 Perfect Hints

Definition 4.8 (Perfect hint)**.** *A perfect hint on the secret* $\mathbf{s}$ *is the knowledge of* $\mathbf{v} \in \mathbb{Z}^{d-1}$ *and* $l \in \mathbb{Z}$*, such that*

$$\langle \mathbf{s},\ \mathbf{v} \rangle = l.$$

A perfect hint is quite strong in terms of additional knowledge. It allows decreasing the dimension of the lattice by one and increases its volume. One could expect such hints to arise from the following scenarios:

- The full leak without noise of an original coefficient, or even an unreduced intermediate register since most of the computations are linear. For the second case, one may note that optimized implementations of NTT typically attempt to delay the first reduction modulo q, so leaking a register on one of the first few levels of the NTT would indeed lead to such a hint.

- A noisy leakage of the same registers, but with still a rather high guessing confidence. In that case it may be worth making the guess while decreasing the success probability of the attack.[1] This could happen in a cold-boot attack

[1] One may then re-amplify the success probability by retrying the attack making guesses at different locations.

scenario. This is also the case in the single trace attack on Frodo [25] that we will study as one of our examples in Section 4.3.1.

- More surprisingly, certain schemes, including some NIST candidates offer such a hint 'by design'. Indeed, LAC, Round5 and NTRU-HPS all choose ternary secret vectors with a prescribed number of 1's and -1's, which directly induce one or two such perfect hints. This will be detailed in Section 4.3.3.

Integrating a perfect hint into a DBDD instance Let $\mathbf{v} \in \mathbb{Z}^{d-1}$ and $l \in \mathbb{Z}$ be such that $\langle \mathbf{s}, \mathbf{v} \rangle = l$. Note that the hint can also be written as

$$\langle \bar{\mathbf{s}},\ \bar{\mathbf{v}} \rangle = 0,$$

where $\bar{\mathbf{s}}$ is the extended LWE secret as defined in Equation (4.4) and $\bar{\mathbf{v}} := (\mathbf{v}\,;\, -l)$.

Remark 4.9. *Here we understand the interest of using Kannan's embedding* before *integrating hints rather than after: it allows to also homogenize the hint, and therefore to make Λ' a proper lattice rather than a lattice coset (i.e. a shifted lattice).*

Including this hint is done by modifying the $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$ to $\mathsf{DBDD}_{\Lambda',\boldsymbol{\mu}',\boldsymbol{\Sigma}'}$, where:

$$\Lambda' = \Lambda \cap \left\{ \mathbf{x} \in \mathbb{Z}^d \mid \langle \mathbf{x}, \bar{\mathbf{v}} \rangle = 0 \right\}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} - \frac{(\bar{\mathbf{v}}\boldsymbol{\Sigma})^T \bar{\mathbf{v}}\boldsymbol{\Sigma}}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T} \tag{4.9}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} - \frac{\langle \bar{\mathbf{v}}, \boldsymbol{\mu} \rangle}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T} \bar{\mathbf{v}}\boldsymbol{\Sigma} \tag{4.10}$$

We now explain how to derive the new mean $\boldsymbol{\mu}'$ and the new covariance $\boldsymbol{\Sigma}'$. Let y be the random variable $\langle \bar{\mathbf{s}}, \bar{\mathbf{v}} \rangle$, where $\bar{\mathbf{s}}$ has mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Then $\boldsymbol{\mu}'$ is the mean of $\bar{\mathbf{s}}$ conditioned on $y = 0$, and $\boldsymbol{\Sigma}'$ is the covariance of $\bar{\mathbf{s}}$ conditioned on $y = 0$. Using Corollary 2.7, we obtain the corresponding conditional mean and covariance.

We note that lattice Λ' is an intersection of Λ and a hyperplane orthogonal to $\bar{\mathbf{v}}$. Given $\mathbf{B}$ as basis of Λ, by Lemma 2.9 a basis of Λ' can be computed as follows:

1. Let $\mathbf{D}$ be dual basis of $\mathbf{B}$. Compute $\mathbf{D}_\perp := \mathbf{D} \cdot \boldsymbol{\Pi}^\perp_{\bar{\mathbf{v}}}$.

2. Apply the LLL algorithm on $\mathbf{D}_\perp$ to eliminate linear dependencies. Then delete the first row of $\mathbf{D}_\perp$ (which is $\mathbf{0}$ because with the hyperplane intersection, the dimension of the lattice is decremented).

3. Output the dual of the resulting matrix.

While polynomial time, the above computation is quite heavy, especially as there is no convenient library offering a parallel version of LLL. Fortunately, for predicting attack costs, one only needs the dimension of the lattice Λ and its volume. These can easily be computed assuming $\bar{\mathbf{v}}$ is a primitive vector (see Definition 2.11) of the dual lattice: the dimension decreases by 1, and the volume increases by a factor $||\bar{\mathbf{v}}||$. This is proved by the following Lemma.

To predict the hardness of the lattice reduction on altered instances, we must compute the volume of the final transformed lattice. We devise a highly efficient way to do this, by observing that each time a hint is integrated, we can update the

volume of the transformed lattice, given only the volume of the previous lattice and information about the current hint (under mild restrictions on the form of the hint).

Lemma 4.10 (Volume of a lattice slice)**.** *Given a lattice* Λ *with volume* $\mathrm{Vol}(\Lambda)$*, and a primitive vector* $\mathbf{v}$ *with respect to* Λ^**. Let* $\mathbf{v}^\perp$ *denote subspace orthogonal to* $\mathbf{v}$*. Then* $\Lambda \cap \mathbf{v}^\perp$ *is a lattice with volume* $\mathrm{Vol}(\Lambda \cap \mathbf{v}^\perp) = \|\mathbf{v}\| \cdot \mathrm{Vol}(\Lambda)$*.*

Proof. Let use denote $\Lambda' = (\Lambda \cap \mathbf{v}^\perp) = \{\mathbf{x} \in \Lambda \mid \langle \mathbf{x}, \mathbf{v} \rangle = 0\}$. We now compute $\mathrm{Vol}(\Lambda')$ as follows

$$\begin{aligned} \mathrm{Vol}(\Lambda') = \frac{1}{\mathrm{Vol}(\Lambda'^*)} &= \frac{1}{\mathrm{Vol}(\Lambda^* \cdot \mathbf{\Pi}_{\mathbf{v}}^{\perp})} && (4.11) \\ &= \frac{\mathrm{Vol}\left(\Lambda^* \cap \mathrm{Span}(\mathbf{v})\right)}{\mathrm{Vol}(\Lambda^*)} && (4.12) \\ &= \mathrm{Vol}\left(\Lambda^* \cap \mathrm{Span}(\mathbf{v})\right) \mathrm{Vol}(\Lambda), \end{aligned}$$

where Equation (4.11) follows from Lemma 2.9, and Equation (4.12) follows from Lemma 2.10. By Definition 2.11, $\mathbf{v}$ generates the one-dimensional lattice $\Lambda^* \cap \mathrm{Span}(\mathbf{v})$, and $\mathrm{Vol}(\Lambda^* \cap \mathrm{Span}(\mathbf{v})) = \|\mathbf{v}\|$. Therefore we have $\mathrm{Vol}(\Lambda') = \|\mathbf{v}\| \cdot \mathrm{Vol}(\Lambda)$.

□

Intuitively, the primitivity condition is needed since then one can scale the leak to $\langle \mathbf{s}, f\mathbf{v} \rangle = fl$ for any non-zero factor $f \in \mathbb{R}$ and get an equivalent leak; however there is only one factor f that can ensure that $f\bar{\mathbf{v}} \in \Lambda^*$, and is primitive in it.

Remark 4.11. *Note that if* $\bar{\mathbf{v}}$ *is not in the span of* Λ*—as typically occurs if other non-orthogonal perfect hints have already been integrated—Lemma 4.10 should be*

applied to the orthogonal projection $\bar{\mathbf{v}}' = \bar{\mathbf{v}} \cdot \mathbf{\Pi}_\Lambda$ *of* $\bar{\mathbf{v}}$ *onto* Λ*. Indeed, the perfect hint* $\langle \bar{\mathbf{s}},\ \bar{\mathbf{v}}' \rangle = 0$ *replacing* $\bar{\mathbf{v}}$ *by* $\bar{\mathbf{v}}'$ *is equally valid.*

4.2.6.2 Modular Hints

Definition 4.12 (Modular hint)**.** *A modular hint on the secret* $\mathbf{s}$ *is the knowledge of* $\mathbf{v} \in \mathbb{Z}^{d-1}$*,* $k \in \mathbb{Z}$ *and* $l \in \mathbb{Z}$*, such that*

$$\langle \mathbf{s},\ \mathbf{v} \rangle = l \mod k.$$

We can expect such hints to arise from several scenarios:

- obtaining the value of an intermediate register during LWE decryption would likely correspond to giving such a modular equation modulo q. This is also the case if an NTT coefficient leaks in a Ring-LWE scheme. It can also occur "by design" if the LWE secret is chosen so that certain NTT coordinates are fixed to 0 modulo q, as is the case in some instances of Order LWE [20].

- obtaining the absolute value $a = |s|$ of a coefficient s implies $s = a \bmod 2a$, and such a hint could be obtained by a timing attack on an unprotected implementation of a table-based sampler, in the spirit of [34].

- obtaining the Hamming weight of the string $b_1 b_2 \dots b'_1 b'_2 \dots$ used to sample a centered binomial coefficient $s = \sum b_i - \sum b'_i$ (as done in NewHope and Kyber [100, 105]) reveals in particular $s \bmod 2$. Indeed, the latter string (or at least some parts of it) is more likely to be leaked than the Hamming weight

of s.

Integrating a modular hint into a DBDD instance. Let $\mathbf{v} \in \mathbb{Z}^{d-1}$; $k \in \mathbb{Z}$ and $l \in \mathbb{Z}$ be such that $\langle \mathbf{s}, \mathbf{v} \rangle = l \mod k$. Note that the hint can also be written as

$$\langle \bar{\mathbf{s}}, \bar{\mathbf{v}} \rangle = 0 \mod k \tag{4.13}$$

where $\bar{\mathbf{s}}$ is the extended LWE secret as defined in Equation 4.4 and $\bar{\mathbf{v}} := (\mathbf{v}\,;\, -l)$. We refer to Remark 4.9 for the legitimacy of such dimension increase.

Smooth case. Intuitively, such a hint should only sparsify the lattice, and leave the average and the variance unchanged. This is not entirely true, this is only (approximately) true when the variance is sufficiently large in the direction of $\mathbf{v}$ to ensure smoothness, i.e. when $k^2 \ll \mathbf{v}\boldsymbol{\Sigma}\mathbf{v}^T$; one can refer to [91, Lemma 3.3 and Lemma 4.2] for the quality of that approximation. In this smooth case, we therefore have:

$$\Lambda' = \Lambda \cap \left\{ \mathbf{x} \in \mathbb{Z}^d \mid \langle \mathbf{x}, \bar{\mathbf{v}} \rangle = 0 \mod k \right\} \tag{4.14}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} \tag{4.15}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} \tag{4.16}$$

On the other hand, if $k^2 \gg \mathbf{v}\boldsymbol{\Sigma}\mathbf{v}^T$, then the residual distribution will be highly concentrated on a single value, and one should therefore instead use a perfect $\langle \mathbf{s}, \mathbf{v} \rangle = l + ik$ for some i.

General case. In the general case, one can resort to a numerical computation of the average μ_c and the variance σ_c^2 of the one-dimensional centered discrete Gaussian of variance $\sigma^2 = \mathbf{v}\boldsymbol{\Sigma}\mathbf{v}^T$ over the coset $l + k\mathbb{Z}$, and apply the corrections:

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + \frac{\mu_c - \langle \bar{\mathbf{v}}, \boldsymbol{\mu} \rangle}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T} \bar{\mathbf{v}}\boldsymbol{\Sigma} \tag{4.17}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} + \left(\frac{\sigma_c^2}{(\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T)^2} - \frac{1}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T} \right) (\bar{\mathbf{v}}\boldsymbol{\Sigma})^T (\bar{\mathbf{v}}\boldsymbol{\Sigma}) \tag{4.18}$$

Intuitively, these formulae completely erase prior information on $\langle \mathbf{s}, \bar{\mathbf{v}} \rangle$, before it is replaced by the new average and variance in the adequate direction. Both can be derived[2] using Corollary 2.7.

As for perfect hints, the computation of Λ' can be done by working on the dual lattice. More specifically:

1. Let $\mathbf{D}$ be dual basis of $\mathbf{B}$.
2. Redefine $\bar{\mathbf{v}} \leftarrow \bar{\mathbf{v}} \cdot \boldsymbol{\Pi}_\Lambda$, noting that this does not affect the validity of the hint.
3. Append $\bar{\mathbf{v}}/k$ to $\mathbf{D}$ and obtain $\mathbf{D}'$
4. Apply the LLL algorithm on $\mathbf{D}'$ to eliminate linear dependencies. Then delete the first row of $\mathbf{D}'$ (which is $\mathbf{0}$ since we introduced a linear dependency).
5. Output the dual of the resulting matrix.

Also, as for perfect hints the parameters of the new lattice Λ' can be predicted: the

[2] We are thankful to Thibauld Feneuil for pointing out an incorrect equation in a previous version of this paper.

dimension is unchanged, and the volume increases by a factor k under a primitivity condition, which is proved by the following lemma.

Lemma 4.13 (Volume of a sparsified lattice)**.** *Let Λ be a lattice, $\mathbf{v} \in \Lambda^*$ be a primitive vector of Λ^*, and $k > 0$ be an integer. Let $\Lambda' = \{\mathbf{x} \in \Lambda \mid \langle \mathbf{x}, \mathbf{v} \rangle = 0 \bmod k\}$ be a sublattice of Λ. Then $\mathrm{Vol}(\Lambda') = k \cdot \mathrm{Vol}(\Lambda)$.*

Proof. Because $\bar{\mathbf{v}}$ is a dual vector of Λ, we have $\langle \bar{\mathbf{v}}, \Lambda \rangle \subset \mathbb{Z}$. Let ℓ be such that, $\langle \bar{\mathbf{v}}, \Lambda \rangle = \ell\mathbb{Z}$. Note that $\bar{\mathbf{v}}/\ell \in \Lambda^*$, therefore, by primitivity of $\bar{\mathbf{v}}$, we have $\ell = 1$. In particular, the group morphism $\phi : \mathbf{x} \in \Lambda \mapsto \langle \mathbf{x}, \bar{\mathbf{v}} \rangle \bmod k$ is surjective. Note that $\Lambda' = \ker \phi$, therefore we have $|\Lambda/\Lambda'| = |\mathbb{Z}_k| = k$. We conclude. □

4.2.6.3 Approximate Hints (conditioning)

Definition 4.14 (Approximate hint)**.** *An approximate hint on the secret $\mathbf{s}$ is the knowledge of $\mathbf{v} \in \mathbb{Z}^{d-1}$ and $l \in \mathbb{Z}$, such that*

$$\langle \mathbf{s},\ \mathbf{v} \rangle + e = l,$$

where e models noise following a distribution $N_1(0, \sigma_e^2)$, independent of $\mathbf{s}$.

One can expect such hints from:

- any noisy side channel information about a secret coefficient. This is the case of our study in Section 4.3.1.
- decryption failures. In Section 4.3.2, we show how this type of hint can represent the information gained by a decryption failure.

To include this knowledge in the DBDD instance, we must combine this knowledge with the prior knowledge on the solution $\mathbf{s}$ of the instance.

Integrating an approximate hint into a DBDD instance Let $\mathbf{v} \in \mathbb{Z}^{d-1}$ and $l \in \mathbb{Z}$ be such that $\langle \mathbf{s}, \mathbf{v} \rangle \approx l$. Note that the hint can also be written as

$$\langle \bar{\mathbf{s}},\ \bar{\mathbf{v}} \rangle + e = 0 \tag{4.19}$$

where $\bar{\mathbf{s}}$ is the extended LWE secret as defined in Equation (4.4), $\bar{\mathbf{v}} := (\mathbf{v}\,;\, -l)$, and e has $N_1(0, \sigma_e^2)$ distribution. The unique shortest non-zero solution of $\mathsf{DBDD}_{\Lambda,\boldsymbol{\mu},\boldsymbol{\Sigma}}$, is also the unique solution of the instance $\mathsf{DBDD}_{\Lambda',\boldsymbol{\mu}',\boldsymbol{\Sigma}'}$ where

$$\Lambda' = \Lambda \tag{4.20}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} - \frac{(\bar{\mathbf{v}}\boldsymbol{\Sigma})^T \bar{\mathbf{v}}\boldsymbol{\Sigma}}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T + \sigma_e^2} \tag{4.21}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} - \frac{\langle \bar{\mathbf{v}}, \boldsymbol{\mu} \rangle}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T + \sigma_e^2}\bar{\mathbf{v}}\boldsymbol{\Sigma} \tag{4.22}$$

We note that Equation (4.20) comes from

$$\Lambda' := \Lambda \cap \left\{ \mathbf{x} \in \mathbb{Z}^d \mid \langle \mathbf{x}, \bar{\mathbf{v}} \rangle + e = 0,\ \text{for all possible } e \sim N_1(0, \sigma_e^2) \right\} = \Lambda.$$

The new covariance and mean follow from Corollary 2.7.

Consistency with Perfect Hint Note that if $\sigma_e = 0$, we fall back to a perfect hint $\langle \mathbf{s}, \mathbf{v} \rangle = l$. The above computation of $\boldsymbol{\Sigma}'$ (4.21) (resp. $\boldsymbol{\mu}'$ (4.22)) is indeed equivalent

to Equation (4.9) (resp. Equation (4.10)) from Section 4.2.6.1. Note however, in our implementation, that to avoid singularities, we require the span of $\text{Span}(\boldsymbol{\Sigma}+\boldsymbol{\mu}^T\boldsymbol{\mu}) = \text{Span}(\Lambda)$ (See the requirement in Equation (4.1)): If $\sigma_e = 0$, one *must* instead use a Perfect hint.

Multi-dimensional approximate hints The formulae of [82] are even more general, and one could consider a multidimensional hint of the form $\mathbf{sV} + \mathbf{e} = \mathbf{l}$, where $\mathbf{V} \in \mathbb{R}^{n\times k}$ and $\mathbf{e}$ a gaussian noise of any covariance $\boldsymbol{\Sigma}_\mathbf{e}$. However, those general formulae require explicit matrix inversion which becomes impractical in large dimension. We therefore only implemented full-dimensional ($k = n$) hint integration in the *super-lightweight* version of our tool, which assumes all covariance matrices to be diagonal. These will be used for hints obtained from decryption failures in Section 4.3.2.

4.2.6.4 Approximate Hint (*a posteriori*)

In certain scenarios, one may more naturally obtain directly the a posteriori distribution of $\langle \mathbf{s}, \mathbf{v}\rangle$, rather than a hint $\langle \mathbf{s}, \mathbf{v}\rangle + e = l$ for some error e independent of $\mathbf{s}$. Such a scenario is typical in template attacks, as we exemplify via the single trace attack on Frodo from [25], which we study in Section 4.3.1.

Given the a posteriori distribution of $\langle \bar{\mathbf{s}}, \bar{\mathbf{v}}\rangle$, one can derive its mean μ_{ap} and variance σ^2_{ap} and apply the corrections to compute the new mean and covariance

exactly as in Equations (4.17) and (4.18).

$$\Lambda' = \Lambda \tag{4.23}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} + \frac{\mu_{\mathrm{ap}} - \langle \bar{\mathbf{v}}, \boldsymbol{\mu} \rangle}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T}\bar{\mathbf{v}}\boldsymbol{\Sigma} \tag{4.24}$$

$$\boldsymbol{\Sigma}' = \boldsymbol{\Sigma} + \left(\frac{\sigma_{\mathrm{ap}}^2}{(\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T)^2} - \frac{1}{\bar{\mathbf{v}}\boldsymbol{\Sigma}\bar{\mathbf{v}}^T}\right)(\bar{\mathbf{v}}\boldsymbol{\Sigma})^T(\bar{\mathbf{v}}\boldsymbol{\Sigma}) \tag{4.25}$$

4.2.6.5 Short vector hints

Definition 4.15 (Short vector hint)**.** *A short vector hint on the lattice* Λ *is the knowledge of a short vector* $\bar{\mathbf{v}}$ *such that*

$$\bar{\mathbf{v}} \in \Lambda.$$

Note that such hints are not related to the secret, and are not expected to be obtained by side-channel information, but rather by the very design of the scheme. In particular, the lattice Λ underlying LWE instance modulo q contains the so-called q-vectors, i.e. the vectors $(q, 0, 0, \ldots, 0)$ and its permutations. These vectors are in fact implicitly exploited in the literature on the cryptanalysis of LWE since at least [81]. Indeed, in some regimes, the best attacks are obtained by 'forgetting' certain LWE equations, which can be geometrically interpreted as a projection orthogonally to a q-vector. Note that, among all hints, the short vector hints should be the last to be integrated. In our context, we need to generalize this idea beyond q-vector because the q-vectors may simply disappear after the integration of a perfect or modular

hint. For example, after the integration of a perfect hint $\langle \mathbf{s}, (1, 1, \ldots, 1) \rangle = 0$, all the q-vectors are no longer in the lattice, but $(q, -q, 0, \ldots, 0)$ still is, and so are all its permutations.

Resolving the DBDD problem resulting from this projection will not directly lead to the original secret, as projection is not injective. However, as long as we keep $n+1$ dimensions out of the $n+m+1$ dimensions of the original LWE instance, we can still efficiently reconstruct the full LWE secret by solving a linear system over the rationals.

Integrating a short vector hint into a DBDD instance It is the case when the secret vector is short enough to be a solution after applying projection $\boldsymbol{\Pi}_{\bar{\mathbf{v}}}^{\perp}$ on $\mathsf{DBDD}_{\Lambda,\boldsymbol{\Sigma},\boldsymbol{\mu}}$

.

$$\Lambda' = \Lambda \cdot \boldsymbol{\Pi}_{\bar{\mathbf{v}}}^{\perp} \tag{4.26}$$

$$\boldsymbol{\Sigma}' = (\boldsymbol{\Pi}_{\bar{\mathbf{v}}}^{\perp})^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\Pi}_{\bar{\mathbf{v}}}^{\perp} \tag{4.27}$$

$$\boldsymbol{\mu}' = \boldsymbol{\mu} \cdot \boldsymbol{\Pi}_{\bar{\mathbf{v}}}^{\perp} \tag{4.28}$$

To compute a basis of Λ' one can simply apply the projection to all the vectors of its current basis, and then eliminate linear dependencies in the resulting basis using LLL.

Remark 4.16. *Once a short vector hint $\bar{\mathbf{v}} \in \Lambda$ has been integrated, Λ has been transformed into Λ'. And, if one has to perform another short vector hint integration $\bar{\mathbf{v}}_1 \in \Lambda$, $\bar{\mathbf{v}}_1$ should be projected onto Λ' with $\bar{\mathbf{v}} \cdot \boldsymbol{\Pi}_{\Lambda'} \in \Lambda'$. In our implementation*

however, this has been taken into account and one can simply apply the same transformation as above, replacing a single vector $\bar{\mathbf{v}}$ *by a matrix* $\mathbf{V}$*.*

The dimension of the lattice decreases by one (or by k, if one directly integrates a matrix of k vectors) and the volume of the lattice also decreases according to Fact 4.17.

Fact 4.17 (Volume of a projected lattice)**.** *Let* Λ *be a lattice,* $\mathbf{v} \in \Lambda$ *be a primitive vector of* Λ*. Let* $\Lambda' = \Lambda \cdot \mathbf{\Pi}_{\mathbf{v}}^{\perp}$ *be a sublattice of* Λ*. Then* $\mathrm{Vol}(\Lambda') = \mathrm{Vol}(\Lambda)/\|\mathbf{v}\|$*. More generally, if* $\mathbf{V}$ *is a primitive set of vectors of* Λ*, then* $\Lambda' = \Lambda \cdot \mathbf{\Pi}_{\mathbf{V}}^{\perp}$ *has volume* $\mathrm{Vol}(\Lambda') = \mathrm{Vol}(\Lambda)/\sqrt{\det(\mathbf{V}\mathbf{V}^T)}$*.*

One can also predict the decrease of the determinant of $\mathbf{\Sigma}$ via the identity:

$$\mathrm{rdet}(\mathbf{\Sigma}') = \mathrm{rdet}(\mathbf{\Sigma}) \cdot \frac{\|\bar{\mathbf{v}}\|^2}{\bar{\mathbf{v}}\mathbf{\Sigma}\bar{\mathbf{v}}^T}, \quad \text{or } \mathrm{rdet}(\mathbf{\Sigma}') = \mathrm{rdet}(\mathbf{\Sigma}) \cdot \frac{\det(\mathbf{V}\mathbf{V}^T)}{\det(\mathbf{V}\mathbf{\Sigma}\mathbf{V}^T)}. \tag{4.29}$$

Worthiness and choice of short vector hints Integrating such a hint induces a trade-off between the dimension and the volume, and therefore it is not always advantageous to integrate.

This raises the following potentially hard problem: given a set $\mathbf{W}$ of short vectors of Λ (viewed as a matrix), which subset $\mathbf{V} \subset \mathbf{W}$ of size k lead to the easiest DBDD instance? Because the hardness of the new problem grows with

$$\frac{\mathrm{rdet}(\mathbf{\Sigma}')}{\mathrm{Vol}(\Lambda')^2} = \frac{\mathrm{rdet}(\mathbf{\Sigma})}{\mathrm{Vol}(\Lambda)^2} \cdot \frac{\det(\mathbf{V}\mathbf{V}^T)^2}{\det(\mathbf{V}\mathbf{\Sigma}\mathbf{V}^T)} \tag{4.30}$$

In the case of an un-hinted DBDD instance directly obtained from the LWE

problem, for $\mathbf{V}$ being the set of (primitive) q-vectors, the problem is easier: all subsets of size k lead to instances with the same parameters.

But this is not true anymore as soon as $\mathbf{\Sigma}$ has been altered or if the set $\mathbf{W}$ is arbitrary. For example, setting $\mathbf{\Sigma} = \mathbf{I}$, one simply wishes to minimize $\det(\mathbf{V}\mathbf{V}^T)$; but for an arbitrary set $\mathbf{W}$ the problem of finding the optimal subset $\mathbf{V} \subset \mathbf{W}$ is NP-hard [77], and remains NP-hard up to exponential approximation factors.

A natural approach to try to get an approximate solution in polynomial time consists in making sequential greedy choices. This involves computing $|\mathbf{V}|\cdot|\mathbf{W}|$ many matrix-vector products over increasingly large rationals, and appeared painfully slow in practice for making prediction on cryptographically large instances. Fortunately, in the typical cases where the vectors of $\mathbf{W}$ are the q-vectors, this can be made somewhat practical (See Section 4.3.3 for example).

Remark 4.18. *When the basis of an LWE-lattice is given in its systematic form, the q-vectors are already explicitly given to lattice reduction algorithms, and these algorithms will implicitly make use of them when they are worthy, as if we had integrated them. The reason is that lattice reduction algorithm naturally work with projected sublattices, and if a q-vector is shorter than what the algorithm can produce, those q-vectors will remain untouched at the beginning of the basis; the reduction algorithm will effectively work on the lattice projected orthogonally to them. In other words, integrating q-vectors is important to understand and predict how lattice reduction algorithm will work, but, in certain cases they may be automatically detected and exploited by lattice reduction algorithms themselves.*

4.3 Applications examples

4.3.1 Hints from side channels

In [25], W. Bos et al. study the feasibility of a single-trace power analysis of the Frodo Key Encapsulation Mechanism (FrodoKEM) [92]. Specifically, in the first approach, they analyze the possibility of a divide-and-conquer attack targeting a multiplication in the key generation. This attack was claimed unsuccessful in [25] because the bruteforce phase after recovering a candidate for the private key was too expensive. Along with this unsuccessful result, a successful powerful extend-and-prune attack is provided in [25].

We emphasize that the purpose of this section is to exemplify our tool on a standard side-channel attack, and this is why we choose the former unsuccessful divide-and-conquer attack of [25]. The point of this section is to show that our framework can indeed lead to improvements in the algorithmic phase of a side-channel attack, once the leak has been fixed.

FrodoKEM. FrodoKEM is based on small-secret-LWE; we outline here some details necessary to understand the attack. Note that we use different letter notations from [92] for consistency. For parameters n and q, the private key is $(\mathbf{z} \in \mathbb{Z}_q^n, \mathbf{e} \in \mathbb{Z}_q^n)$ where the coefficients of $\mathbf{z}$ and $\mathbf{e}$, denoted $\mathbf{z}_i$ and $\mathbf{e}_i$, can take several values in a small set that we denote L. The public key is $\big(\mathbf{A} \in \mathbb{Z}_q^{n \times n}, \mathbf{b} = \mathbf{z}\mathbf{A} + \mathbf{e}\big)$. The goal of the attack is to recover $\mathbf{z}$ by making measurements during the multiplication

between $\mathbf{z}$ and $\mathbf{A}$ when computing $\mathbf{b}$ in the key generation. Note that there is no multiplication involving $\mathbf{e}$ and thus it is not targeted in this attack. Six sets of parameters are considered: CCS1, CCS2, CCS3 and CCS4 introduced in [24] and NIST1 and NIST2 introduced in [92]. For example, with NIST1 parameters, $n = 640,\ q = 2^{15}$ and $L = \{-11, \cdots, 11\}$.

$$n = 640,\ q = 2^{15} \text{ and } L = \{-11, \cdots, 11\}.$$

Side-channel simulation. The divide-and-conquer attack provided by [25] simulates side-channel information using ELMO, a power simulator for a Cortex M0 [89]. This tool outputs simulated power traces using an elaborate leakage model with Gaussian noise. Thus, it is parametrized by the standard deviation of the side-channel noise. For proofs of concept, the authors of [89] suggest to choose the standard deviation of the simulated noise as $\sigma_{\text{SimNoise}} := 0.0045$ for realistic leakage modeling. This standard deviation was also the one chosen in [25, Fig. 2b] and W. Bos et al. implemented a Matlab script that calls ELMO to simulate the side-channel information applied on Frodo. This precise side-channel simulator was provided to us by the authors of [25] and we were able to re-generate all their data with Matlab, again using $\sigma_{\text{SimNoise}} = 0.0045$.

Template attack. The divide-and-conquer side-channel attack proposed by W. Bos et al. belongs in the template attack family. Template attacks were introduced in [37]. In a nutshell, these attacks include a profiling phase and an online phase.

Let us detail the template attack for Frodo implemented in [25].

1. The profiling phase consists in using a copy of the device and recording a large number of traces using many different known secret values. From these measures, the attacker can derive the multidimensional distribution of several points of interest when the traces share the same secret coefficient. More precisely, in the case of FrodoKEM, for a given index $i \in [0, n-1]$, the points of interest will be the instants in the trace when $\mathbf{z}_i$ is multiplied by the coefficients of $\mathbf{A}$ (n interest points in total). Let us define

$$\mathbf{c}_i := (T[t_{i,0}], \ldots, T[t_{i,n-1}]) \quad \mathbf{c} \in \mathbb{R}^n, \tag{4.31}$$

where T denotes the trace measurement and $(t_{i,k})$ denotes the instants of the multiplication of $\mathbf{z}_i$ with the coefficients $\mathbf{A}_{i,k}$ for $(i,k) \in [0, n-1]$. The random variable vector associated to $\mathbf{c}_i$ is denoted by $\mathcal{C}_i$. For each $i \in [0, n-1]$ and $x \in L$, the goal of the profiling phase is to learn the center of the probability distribution

$$A_{i,x}(\mathbf{c}) := P\left[\mathcal{C}_i = \mathbf{c} \mid \mathbf{z}_i = x\right].$$

By hypothesis, for template attacks (see [37, Section 2.1]), $A_{i,x}$ is assumed to follow a multidimenstional normal distribution of standard deviation $\sigma_{\text{SimNoise}} \cdot \mathbf{I}_n$. Thus, the attacker recovers the center of $A_{i,x}$ for each $i \in [0, n-1]$ and $x \in L$ by averaging all the measured $\mathbf{c}_i$ that validate $\mathbf{z}_i = x$. The center of $A_{i,x}$ is denoted $\mathbf{t}_{i,x}$ and we call it a *template*. W. Bos et al. [25] actually assume that

$\mathbf{t}_{i,x}$ depends only on x and is independent from the index i. Thus, $\mathbf{t}_{i,x} = \mathbf{t}_x$. Essentially, this common assumption implies that the index $i \in [0, n-1]$ of the target coefficient does not influence the leakage. Consequently, the attacker only has to derive $\mathbf{t}_{0,x}$, for example.

2. In a second step, the attacker knows the templates $\mathbf{t}_x$ for all $x \in L$. She also knows the points of interest $t_{i,k}$ as defined above in Equation 4.31. She will construct a candidate $\tilde{\mathbf{z}}$ for the secret $\mathbf{z}$ by recovering the coefficients one by one. For each unknown secret coefficient $\mathbf{z}_i$, she takes the measurement $\mathbf{c}_i$ as defined in Equation 4.31. Using this measurement, she can derive an a posteriori probability distribution: With her fixed $i \in [0, n-1]$ and measured $\mathbf{c}_i \in \mathbb{R}$, she computes for all $x \in L$,

$$P\left[\mathbf{z}_i = x \mid \mathcal{C}_i = \mathbf{c}_i\right] = \frac{P\left[\mathbf{z}_i = x\right]}{P\left[\mathcal{C}_i = \mathbf{c}_i\right]} \cdot P\left[\mathcal{C}_i = \mathbf{c}_i \mid \mathbf{z}_i = x\right] \tag{4.32}$$

$$\propto P\left[\mathbf{z}_i = x\right] \cdot \exp\left(-\frac{\|\mathbf{c}_i - \mathbf{t}_x\|_2^2}{2\sigma_{\text{SimNoise}}^2}\right) \tag{4.33}$$

In [25], a score table, denoted $(S_i[x])_{x \in L}$ is derived from the a posteriori distribution as follows,

$$S_i[x] := \ln\left(P\left[\mathbf{z}_i = x \mid \mathcal{C}_i = \mathbf{c}_i\right]\right) \tag{4.34}$$

$$= \ln\left(P\left[\mathbf{z}_i = x\right]\right) - \frac{\|\mathbf{c}_i - \mathbf{t}_x\|_2^2}{2\sigma_{\text{SimNoise}}^2}. \tag{4.35}$$

Finally, the output candidate for $\mathbf{z}_i$ is $\tilde{\mathbf{z}}_i := \text{argmax}_{x \in L}(S_i[x])$.

$\mathbf{z}_i$	S							
	-11	-10	-9	-8	-7	-6	-5	-4
0	-4098	-3918	-4344	-2580	-3212	-3108	-3758	-3155
1	-3273	-3114	-3491	-1951	-2495	-2405	-2972	-2445
-1	-341	-335	-352	-465	-358	-369	-329	-362
-1	-306	-298	-319	-414	-314	-323	**-290**	-317

	...	-3	-2	-1	0	1	2	3
0	...	-3583	-3498	-3900	**-340**	-380	-367	-452
1	...	-2819	-2744	-3098	-365	**-325**	-328	-338
-1	...	-331	-334	**-328**	-3712	-3079	-3195	-2656
-1	...	-291	-293	-291	-3608	-2982	-3097	-2564

	...	4	5	6	7	8	9	10	11
0	...	-818	-975	-933	-1084	-368	-459	-453	-592
1	...	-546	-657	-627	-737	-333	-344	-342	-407
-1	...	-1696	-1461	-1521	-1329	-3231	-2648	-2685	-2201
-1	...	-1617	-1385	-1444	-1256	-3132	-2556	-2593	-2115

Table 4.1: Examples of scores associated to the secret values $\mathbf{s}_i \in \{0, \pm 1\}$, after the side-channel analysis of [25] for NIST1 parameters. The best score in each score table is highlighted. This best guess is correct for the first 3 score table, but incorrect for the last one.

One can use the presented attack as a "black-box" to generate the score tables using the script from [25]. As an example, using the NIST1 parameters, we show several measured scores $(S[-11], \cdots, S[11])$ corresponding to several secret coefficients in Table 4.1. The first line corresponds to a secret equal to 0, the second line to 1 and the third and fourth line to -1. The last line is an example of failed guessing because we see that the outputted candidate is not -1. We remark that the values having the opposite sign are assigned a very low score, we conjecture that it is because the sign is filling the register and then the Hamming weight of the register will be very far from the correct one.

With this template attack, one can recover $\tilde{\mathbf{z}} \approx \mathbf{z}$. However, W. Bos et al. [25]

could not conclude the attack with a key recovery even though much information leaked about the secret. Frustratingly, a bruteforce phase to derive $\mathbf{z}$ from $\tilde{\mathbf{z}}$ did not lead to any security threat as stated in [25, Section 3]. They actually pointed out an interesting open question of whether "novel lattice reduction algorithms [can] take into account side-channel information". Our work solves this open question by combining the knowledge obtained in the divide-and-conquer template attack of [25] with our framework.

From scores to hints. We first instantiate a DBDD instance with a chosen set of parameters. Then we assume that, for each secret coefficient $\mathbf{z}_i$, we are given the associated score table S_i, thanks to the template attack that has already been carried out. We go back to the a posteriori distribution in Equation 4.33 by applying the exp() function and renormalizing the score table. As an example, we show the probability distributions derived from Table 4.1, along with their variances and centers, in Table 4.2.

Finally, we use our framework to introduce n a posteriori *approximate hints* to our DBDD instance with the derived centers and variances for each score table. When the variance is exactly 0, we integrate perfect hints instead.

Results. One can reproduce this attack using the Sage 9.0 script exploiting_SCA_from_Bos_et_al.sage. The experimentally derived data containing the score tables is in the folder Scores_tables_SCA for which, as mentioned earlier, was

$\mathbf{z}_i$	A posteriori distribution								
	-11	-10	-9	-8	-7	-6	-5	-4	-3
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
-1	0	0	0	0	0	0	0.26	0	0.04
-1	0	0	0	0	0	0	0.56	0	0.21

	...	-2	-1	0	1	2	3	4	5
0	...	0	0	1	0	0	0	0	0
1	...	0	0	0	0.95	0.04	0	0	0
-1	...	0.00	0.70	0	0	0	0	0	0
-1	...	0.03	0.21	0	0	0	0	0	0

	...	6	7	8	9	10	11	center	variance
0	...	0	0	0	0	0	0	0	0
1	...	0	0	0.01	0	0	0	1.05	0.06
-1	...	0	0	0	0	0	0	-2.11	3.11
-1	...	0	0	0	0	0	0	-3.68	2.63

Table 4.2: Probability distributions derived from Table 4.1, along with variances and centers.

generated with a simulated noise variance of 0.0045. One can note that the obtained security fluctuates a bit from instance to instance, as it depends on the strength of the hints, which themselves depend on the randomness of the scheme. In the first two lines of Table 4.3, we show the new security with the inclusion of the approximate hints averaged on 50 tests per set of parameters.

	NIST1	NIST2	CCS1	CCS2	CCS3	CCS4
Attack without hints	487	708	239	448	492	584
Attack with hints	**330**	**423**	**128**	**123**	**219**	**230**
Attack with hints & guesses	**292**	**298**	**70**	**29**	**124**	**129**
Number of guesses g	100	250	200	300	250	250
Success probability	0.86	0.64	0.87	0.77	0.81	0.84

Table 4.3: Cost of the attacks in bikz without/with hints without/with guesses.

Guessing. To improve the attack further, one can note from Table 4.2 that certain key values have a very high probability of being correct, and assuming each of these values are correct, one can replace an approximate hint with a perfect one. For example, considering the second line of Table 4.2, the secret has a probability of 0.95 to be 1 and thus guessing it trades a perfect hint for a decrease of the success probability of the attack by 5%. This hybrid attack exploiting hints, guesses and lattice reduction, works as follows. Let g be a parameter.

1. Include all the approximate and perfect hints given by the score tables,
2. Order the coefficients of the secret $\mathbf{z}_i$ according to the maximum value of their a posteriori distribution table,
3. Include perfect hints for the g first coefficients and then solve and check the solution.

Increasing the number of guesses g leads to a trade-off between the cost of the attack

and its success probability. We have chosen here a success probability larger than 0.6, while reducing the attack cost by 38 to 145 bikz depending on the parameter set. Given that 1 bit of security corresponds roughly to 3 or 4 bikz, this is undoubtedly advantageous.

Remark 4.19. *The refinement presented above are very recent (lastly improved on June 2020). We are grateful to the authors of [25] of for helping us reconstructing distributions from the score table.*

We remark that, with these results, the attacks with guesses on the parameters CCS1 and CCS2 seem doable in practice while it was not the case with our original results. However, some improvements of the implementation remain to be done in order to actually mount the attack. The full-fledged implementation cannot handle in reasonable time the large matrices of the original DBDD *instance. We require another class of implementation which fully maintains all information about the instance, like the* `DBDD` *class, and assumes that the covariance matrix* $\mathbf{\Sigma}$ *is diagonal to simplify the computations, like the* `DBDD_predict_diag` *class. We hope to report on such an implementation in a future update of this report.*

Remark 4.20. *It should be noted that, given a single trace, one cannot naively retry the attack to boost its success probability. Indeed, the "second-best" guess may already have a much lower success probability than the first. Setting up such an hybrid attack mixing lattice reduction within our framework and key-ranking appears to be an interesting problem.*

4.3.2 Hints from decryption failures

Another kind of hint our framework can model are hints provided by decryption failures. For a single-bit LWE encryption scheme, a decryption failure occurs when the random short vector $\mathbf{w}$ used during encryption is such that $|\langle \mathbf{s}, \mathbf{w} \rangle| \geq t$ for some t, typically $t = q/4$.

In fact, we can even assume to know the "side" of the decryption failure, i.e. we can assume we know that $\langle \mathbf{s}, \mathbf{w} \rangle \geq t$. Indeed, this can be guessed with probability $1/2$ for the first failure, and it can be deduced for subsequent failures using the fact that those sides are strongly correlated (see Section 4.3 in [49] for example). For multi-bit encryption, using either ring-element or matrices for secrets, similar techniques allow to "locate" the failure, and therefore obtain information of this form.

We will here consider the case of the Chosen-Ciphertext-Attack (CCA) secure variant of such schemes, typically obtained by variants of the Fujikasi-Okamoto transform. In this case, the attacker does not control the short vector $\mathbf{w}$, as it is generated following the randomness of a hash function.

Following our framework, it would be tempting to simply construct the conditional distribution of $\langle \mathbf{s}, \mathbf{w} \rangle$ given that $|\langle \mathbf{s}, \mathbf{w} \rangle| \geq t$, and integrate this as an a posteriori hint with $\mathbf{v} = \mathbf{w}$. However, this modeling would actually lose a lot of information. Indeed, such hints are designed in the case where one first chooses $\mathbf{w}$ independently of $\mathbf{s}$, and then learns partial information on $\langle \mathbf{s}, \mathbf{w} \rangle$. The setting here is quite different: one instead samples $\mathbf{w}$ following a prescribed distribution, until

failure occurs. In other word, $\mathbf{w}$ is sampled on a prescribed distribution, and conditioned on $\langle \mathbf{s}, \mathbf{w} \rangle \geq t$. In particular it is not sampled independently of the secret $\mathbf{s}$, and it carries information on $\mathbf{s}$ in all directions.

For the sake of simplicity, let us assume that the norm of $\mathbf{s}$ is exactly $\ell = \sqrt{n}\sigma$; making such a guess is rather inconsequential given how concentrated the norm of a high dimensional Gaussian is. Let us assume that $\mathbf{w}$ also follows a Gaussian of covariance $\tau^2\mathbf{I}$, before imposing the condition. After conditioning, $\mathbf{w}$ decomposes as $\mathbf{w} = \alpha\mathbf{s}/\ell + \mathbf{w}'$, where $\mathbf{w}'$ is a Gaussian of covariance $\tau^2\mathbf{\Pi}_{\mathbf{s}}^{\perp}$, and α is independent of $\mathbf{w}'$ and follows a distribution that we denote $G_{\tau}^{\geq t/\ell}$, the unidimensional Gaussian of variance τ^2 conditioned on $\alpha \geq t/\ell$. One can check that the $\mathbb{E}_{X \leftarrow G_{\tau}^{\geq t/\ell}}[(t/\ell - X)^2] \leq \tau^2$ for any $t/\ell \geq 0$. This means that we can write $\mathbf{w} = t/\ell^2 \cdot \mathbf{s} + \mathbf{e}$ for some error $\mathbf{e}$ of (ill-centered) covariance $\mathbf{\Sigma}_{\mathbf{e}} \leq \tau\mathbf{I}$.

Rewriting the above equality, we finally obtain a full dimensional approximate hint of the form

$$\mathbf{s} = \frac{\ell^2}{t}\mathbf{w} + \mathbf{e}'$$

with an error $\mathbf{e}' = -\frac{\ell^2}{t}\mathbf{e}$ of (uncentered) covariance $\tau^2\ell^4/t^2 \cdot \mathbf{I}$.

We can now compare the results of our prediction to prior work that used several other methodologies such as [48, 49, 58, 67]. We choose to compare with [49] on FRODOKEM-976, for which the data can be reproduced[3], and for which $\mathbf{w}$ is indeed very close to Gaussian. We note that both methods use different simplications or heuristics, nevertheless they produce essentially similar predictions, as shown

[3] `https://github.com/KULeuven-COSIC/PQCRYPTO-decryption-failures/`

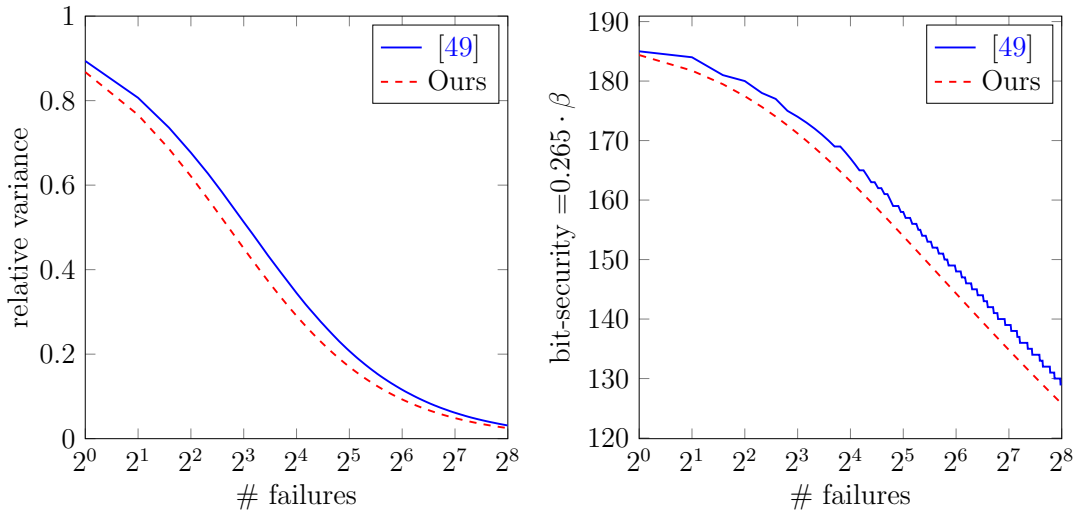


Figure 4.4: Security decrease as a function of the number of failure in FRODOKEM-976.

in Figure 4.4. The data using our framework has been acquired with the script exploiting_decryption_failures.sage.

Furthermore, one could try to refine the estimate of the average and variance of $\mathbf{e}$, which can improve in direction of $\mathbf{w}$. However, this would force us to deal with non-diagonal covariance matrices, which generically significantly slows down our script, and would require further optimizations to be doable in practice. The exploration of such improvements is left as future work.

4.3.3 Structural hints from Design

LAC is a Ring-LWE round two candidate of the NIST post-quantum competition [83]. The secrets are two polynomials $\mathbf{s}_0, \mathbf{s}_1$ (denoted $\mathbf{s}$ and $\mathbf{e}$ in the specifications) whose coefficients follow a distribution $\psi^{n,h}$, the uniform distribution over ternary vectors $\{-1, 0, 1\}^n$ with exactly $h/2$ ones and $h/2$ minus ones. Thus, two

	LAC-128	LAC-192	LAC-256
without hints	509.03	985.64	1104.83
with 2 hints	505.94	982.74	1101.61
	R5ND_{1}KEM_0d	R5ND_{3}KEM_0d	R5ND_{5}KEM_0d
without hints	494.39	658.67	877.71
with 1 hint	492.94	657.23	876.24
	ntruhps2048509	ntruhps2048677	ntruhps4096821
without hint	372.58	515.36	617.71
with 1 hint	371.23	513.95	616.39
with hint + 6 guesses	365.79	508.47	611.00

Table 4.4: New security estimates in bikz (GSA-Intersect method)

	ntruhps2048509	ntruhps2048677	ntruhps4096821
without hints	379.61	526.17	631.84
with 1 hint	378.22	524.74	630.49
with hint + 6 guesses	372.64	519.11	624.94
with hint + multi-target	367.58	512.68	618.24

Table 4.5: New security estimates in bikz (Probabilistic-Simulation method)

structural perfect hints can be derived:

$$\sum_{i=0}^{n-1} \mathbf{s}_0[i] = 0 \text{ and } \sum_{i=0}^{n-1} \mathbf{s}_1[i] = 0.$$

The same structure appears in the submissions Round5, and NTRU-HPS, but yields only one perfect hint on half of the secret as they also require the number of -1 coefficients to be balanced with the number of 1 coefficients of their ternary polynomial. In fact, exploiting this information was already mentionned in the cryptanalysis of the original NTRU scheme [43]. While it is clear that each such equation it decreases the dimension by 1, its effect on the volume of the lattice seems not to have been analyzed so far; according to Lemma 4.10, the volume is increased by a factor $\sqrt{n}$.

This new knowledge has been included in the security analysis and the results are stored in Table 4.4. One can check the experiments by running the scripts exploiting_design_LAC.sage, exploiting_design_round5.sage and exploiting_design_ntru.sage. For Round5, we arbitrarily chose for our testing the parameter set R5ND_$\{1,3,5\}$KEM_0d.

Remark 4.21. *Note, however, that integrating such hints removes some q-vectors from the lattice. For NTRU-HPS and Round5, there remain half of them, and this is sufficient to find the optimal volume-dimension trade-off.*[4] *For LAC, we note that while q-vectors are not in the lattice, a difference of 2 such vectors is still in it, for example the short vector hint $(q, -q, 0, 0, \ldots, 0) \in \Lambda$. We iteratively integrate $(q, -q, 0, 0, \ldots, 0)$, $(0, q, -q, 0, \ldots, 0)$, $(0, 0, q, -q, \ldots, 0)$, ... until such hints are not worthy anymore, i.e. until such hints do not decrease the cost of the attack anymore.*

[4] In a previous version of this paper, we treated NTRU-HPS and Round5 in the same way as LAC, and used $(q, -q)$-vectors rather than q-vectors, which lead to a somewhat suboptimal attack.

The case of NTRU. A first remark is that the NTRU problem is somewhat different from the BDD problem, in the sense that it is homogenoues already: there is no need to apply Kannan's embedding to make it into a short vector problem. This means in particular that the dimension of the input lattice is $2n$ and not $2n+1$. More specifically, the secret consists of two ternary elements of the cyclic convolution ring $f, g \in \mathbb{Z}[X]/(X^n - 1) =: R$, and the public key $h = f/g \bmod q$. One can directly construct the lattice $\Lambda = \{(x, y) \in R^2 \mid x - hy = 0 \bmod q\}$ and search for (f, g) as a short vector in that lattice.

Secondly, the lattice enjoys a rotational symmetry of order n; in particular there is not only a single short vector in that lattice, but n linearly independent such short vectors: $(f, g), (X \cdot f, X \cdot g), (X^2 \cdot f, X^2 g), \ldots (X^{n-1} \cdot f, X^{n-1} \cdot g)$.[5]

A third remark is that, even without hints, and using the same GSA-interesect method, our tool gives about 10 extra bikz of security to NTRU-HPS compared to the analysis given in the standardization document [116]. The largest part of this difference is to be accounted on the fact that [116] uses a lower-bound on the length of one half of the secret. Such a simplification avoid the need for an isotropization step, which would complicate an ad-hoc script, but is fully automatized by our tool.

One last remark is that [116] does also perform a dimension-reduction, but it is not equivalent to the one discussed above. More specifically, they suggest to reduce the problem modulo Φ_n where $\Phi_n = X^{n-1} + X^{n-2} + \cdots + X + 1$ denotes the n-th cyclotomic polynomial for n a prime. Using the coefficient embedding to define the

[5] We remark that such a symmetries can also appear in Cyclotomic Ring-LWE, but at the cost of increasing the lattice dimension from $2n + 1$ to $3n$.

geometry, this decreases dimension by 1, and leaves the volume unchanged; however such a reduction can significantly increase the length of the secret vectors to be found, when the leading coefficient of the secret polynomial (i.e. the last coefficient of the secret vector) is not 0. Fortunately, because of the rotational symmetries, there should some short vectors for which this reduction does not affect its length.[6] A posteriori, this dimension reduction techniques essentially boils down to making a guess $f_n = 0$, knowing that this guess are likely satified by one of the many short secrets; in our framework this is merely a perfect hint, and we predict, as in [116] that it decreases dimension by one without affecting the volume.

These remarks suggests several refinements. The first is that we can combine the integration of the hint $\sum f_i = 0$ and of a guess $f_n = 0$. In fact, we can follow the attack of May and Silverman [88], and integrate several such guesses so as to fully exploit symmetries. Roughly, given that the secret are essentially uniform and ternary, one can hope that one of the n short vectors will satisfy $\log_3(n) \approx 6$ equations of the form $f_i = 0$.

Yet, we can also wonder wether making such symmetry-breaking guesses is really advantageous, as it could be that lattice reduction already internally benefits from the presence of many short vectors. Under the GSA-interesect model, this does not seem to be the case, as this model would predict that all the short vectors are detected at the same time. However, the refined method of Section 4.2.5 can indeed the account for the accumulated probability over multiple targeted short vectors.[7]

[6] We are greatful to John Schanck for this clarification.

[7] To apply the probabilistic-simulation for such large parameters, we only account for the probability of detecting the vector at position $d-\beta$, so as to avoid numerical issues raised by the rest of this probability computation. However, the probability to be lifted back to the front once detected

Our results are compiled in Table 4.4 and Table 4.5, and the conclusion is, according to the probabilistic simulation method, that it is seems preferable to *not make any guesses*, and let lattice reduction naturally exploit the presence of many short vectors. However, due to the other approximations made in [116], our refined analysis does not invalidate the original security claims. We nevertheless think that this revised analysis clarify the phenomena at play during lattice attacks on NTRU.

Remark 4.22. *A similar structure is present in the candidate NTRU-Prime in its streamlined and LPR versions [15]. In the secret vector, the number of* ± 1*'s is fixed to an integer* w *without knowing the exact number of positive and negative ones. Thus, one can include a modular hint*

$$\sum_{i=0}^{n-1} \mathbf{s}_0[i] = w \mod 2.$$

The loss of security is however essentially negligible.

at position $d - \beta$ is very close to 1 for such a large β, as already argued in [7].

Chapter 5: Towards a Ring Analogue of the Leftover Hash Lemma

The leftover hash lemma (LHL) is used in the analysis of various lattice-based cryptosystems. Specifically, it is often useful to argue that for high-min entropy input $\mathbf{x} \in \mathbb{Z}_q^m$ and random matrix $A \leftarrow \mathbb{Z}_q^{n \times m}$, $A\mathbf{x}$ is uniform random, given A. The above fact is used in the proof of security for both the Regev and Dual-Regev encryption schemes. More sophisticated proof approaches that utilize the LHL along with the structure of the matrix A have been used to argue leakage resilience of these cryptosystems, such as in [4, 53].

Analogues of the statement above do not necessarily hold in the ring setting:

The ring setting. Consider the number field $K = \mathbb{Q}[x]/\Phi_m(x)$, where $\Phi_m(x)$ is the m-th cyclotomic polynomial of degree $\varphi(m)$. The ring of integers, $R \subset K$, is defined as $R = \mathbb{Z}[x]/\Phi_m(x)$. $R_q := \mathbb{Z}_q[x]/\Phi_m(x)$ denotes the set of polynomials obtained by taking an element of $\mathbb{Z}[x]/\Phi_m(x)$ and reducing each coefficient modulo q. In this paper, we further assume that m is a power of two, so $\Phi_m(x) = x^n + 1$ has degree $n = m/2$, and set q to be a prime such that $q \equiv 1 \mod m$. In this case $\Phi_m(x)$ completely splits into n factors in $\mathbb{Z}_q[x]$. This is the setting favored in practice since

[0]For example, techniques include decomposition of the matrix A into two random matrices of varying dimensions [4].

it allows for optimizations in the implementation, such as fast arithmetic over the ring R_q.

A Ring Analogue of the LHL. For rings R_q such as the above, a result analogous to the leftover hash lemma—proving that $a_{l+1} = \sum_{i\in[l]} a_i x_i$ is indistinguishable from random, given $a_2, \ldots, a_l$, as long as $x_1, \ldots, x_l$ has sufficiently high min-entropy—is impossible. For example, if the j-th NTT coordinate of each ring element in $\mathbf{x} = x_1, \ldots, x_l$ is leaked, then the j-th NTT coordinate of $a_{l+1} = \sum_{i\in[l]} a_i x_i$ is known[1], and so a_{l+1} is very far from uniform. Yet this is only a $1/n$ leakage rate![2]

Nevertheless, Lyubashevsky et al. [85] proved a "regularity lemma" showing that for matrix $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$, where $I_k \in (R_q)^{k\times k}$ is the identity matrix and $\bar{A} \in (R_q)^{k\times(l-k)}$ is uniformly random, and $\mathbf{x}$ chosen from a discrete Gaussian distribution (centered at 0) over R_q^l, the distribution over $A\mathbf{x}$ is (close to) *uniform random*. A similar result was proven by Micciancio [90], but requires super-constant dimension l, thus yielding non-compact cryptosystems. In contrast, the regularity lemma of [85] holds even for constant dimension l as small as 2. The fundamental technical question we consider in this work is:

> For which distributions $\mathcal{D}$ over $\mathbf{x} \in R_q^l$, is the distribution over $A\mathbf{x}$ (close to) *uniform random*, for R, q, A as above and constant l?

[1] Applying NTT to $a_i, x_i \in R_q$—resulting in n-dimensional vectors, $\widehat{a}_i, \widehat{x}_i \in \mathbb{Z}_q^n$—allows for *component-wise* multiplication/addition, so the j-th NTT coordinate of $a_i x_i, i \in [l]$ will be known and so the j-th NTT coordinate of a_{l+1} is known.

[2] We thank an anonymous reviewer for pointing out this counterexample to us.

5.1 Overview of Our Work

We prove a "regularity lemma" for three conditional distributions, which we describe next. Only the parameter s–the standard deviation of the discrete Gaussian for sampling each coordinate of $\mathbf{x}$–differs in each setting.

Conditional Distribution I. We assume a secret key $\mathbf{x} = (x_1, \ldots, x_l)$, where each $x_i \in R_q$. Moreover, each x_i itself is represented as an n-dimensional vector. So in total, $\mathbf{x}$ is an $l \cdot n$-dimensional vector. We consider the conditional distribution on $\mathbf{x}$ when the sum of $\mathbf{x}$ and $\mathbf{e}$ is revealed, where each coordinate of $\mathbf{e}$ is a Gaussian random variable with standard deviation at least s. This setting captures leakage on $\mathbf{x}$ by an adversary who uses a fast, but inaccurate device to obtain noisy measurements of *each* sampled coordinate of the secret key (e.g. through a power or timing channel). We prove that it is sufficient to set $s \geq \sqrt{2} \cdot 2n \cdot q^{k/l+2/(nl)}$. See Theorem 5.2 and Corollary 5.3.

Conditional Distribution II. We consider the conditional distribution over $\mathbf{x} = (x_1, \ldots, x_l)$ when we leak ℓ coordinates from each $x_i, i \in [l]$. and we set parameters such that the fraction of leaked coordinates–$\frac{\ell \cdot l}{n \cdot l}$–is constant. The ℓ leaked coordinates are arbitrary, but the same ℓ coordinates must be leaked from each $x_i, i \in [l]$.[3] *Low noise* is added to each leaked coordinate (only $2n$ standard deviation, as opposed to $\sqrt{2} \cdot 2n \cdot q^{k/l+2/(nl)}$ standard deviation as in Conditional Distribution I).

[3] Alternatively, we can view the leakage as ℓ completely arbitrary coordinates, with leakage rate of $\ell/(n \cdot l)$, which remains constant for constant l.

No information at all is leaked about the remaining coordinates. This setting corresponds to a side-channel attack launched during the sampling of $\mathbf{x}$, where the attacker has a slower, but more accurate device which allows it to obtain more accurate measurements for a *constant fraction* of the coordinates of the secret key, but *no* information for the remaining coordinates. [4] We prove that it is sufficient to set $s \geq 2n \cdot q^{\frac{kn+2}{l(n-\ell)}}$, where $\ell \cdot l$ is the number of leaked coordinates. See Theorem 5.4 and Corollary 5.8.

Conditional Distribution III. Here, we consider the conditional distribution on $\mathbf{x}$, when the magnitude of $\mathbf{x}$ with Gaussian channel error e is revealed (note that e is a scalar). We assume e is sampled from a univariate Gaussian with standard deviation s. A motivation for this type of leakage is that (discrete) Gaussian sampling of $\mathbf{x}$ is often implemented via rejection sampling in practice [28, 51]. E.g. a vector could be sampled from a "close" multi-dimensional binomial distribution and rejection sampling then used to obtain a sample from the correct distribution. The rejection condition depends on the weight of $\mathbf{x}$ under the target distribution, which in turn depends on the magnitude of $\mathbf{x}$, and so this information is vulnerable to leakage during computation. [5] We prove that it is sufficient to set $s \geq \sqrt{14/5 \cdot (n'/n) \cdot \ln n'} \cdot 2n \cdot q^{k/l+2/(nl)}$, where $n' = n \cdot l + 1$. See Theorem 5.12 and Corollary 5.13.

[4] Here we assume that the secret key is stored as a vector in the canonical embedding (in the other leakage scenarios, the result holds when the secret key is stored in using the polynomial representation or is stored as a vector in the canonical embedding).

[5] For example, a power analysis attack on the BLISS signature scheme [60] exploited the rejection sampling procedure to recover the magnitude (norm) of certain secret values, which then led to a full break of the scheme.

Applications to leakage resilience. Since applications of the LHL/Regularity Lemma in lattice-based cryptography are widespread, a number of Ring-LWE (RLWE) cryptosystems achieve certain leakage resilience properties using our results. Such cryptosystems include the ring analogues of Regev encryption [84], Dual-Regev encryption [85], and identity-based encryption (IBE) based on Dual-Regev encryption [63] (see ring version in [16]). Specifically, by substituting our "regularity lemma" for the original "regularity lemma" in the security proofs, those schemes still enjoy security guarantees even given certain leakage on the *randomness for encryption* (for Regev) the *secret key* (for Dual-Regev), and the *secret key corresponding to the challenge identity* (for IBE).

5.1.1 Our High-Level Approach

For a matrix $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$, where $I_k \in (R_q)^{k\times k}$ is the identity matrix and $\bar{A} \in (R_q)^{k\times(l-k)}$ is uniformly random, we define $\Lambda^{\perp}(A) = \{\mathbf{z} \in R^l : A\mathbf{z} = \mathbf{0} \bmod qR\}$. If $[\mathbf{x} \mod \Lambda^{\perp}(A)]$ is uniform random (over cosets of $\Lambda^{\perp}(A)$), then the distribution of $A\mathbf{x}$ is also uniform random over cosets of $(qR)^k$. The input/output distributions can then be discretized over the ring R. Therefore, the goal is to show that when $\mathbf{x}$ is sampled from continuous distribution $\mathcal{D}$, we have that $[\mathbf{x} \mod \Lambda^{\perp}(A)]$ is uniform random. Consider the case where the distribution $\mathcal{D}$ is exactly a Gaussian distribution with mean 0 and standard deviation s. In this case, if s is greater than or equal to the *smoothing parameter* of $\Lambda^{\perp}(A)$, this by definition ensures that the distribution $[\mathbf{x} \mod \Lambda^{\perp}(A)]$ is uniform random. Thus, [85] prove

their regularity lemma by showing that with high probability over choice of A, the smoothing parameter, $\eta_\varepsilon(\Lambda^\perp(A))$, is upperbounded by s.

Before presenting our approach to extending the above result, it is instructive to give a high-level recap of how to derive upper bounds on the smoothing parameter.

Let $\rho_s := e^{-\pi \frac{\langle \mathbf{x}, \mathbf{x} \rangle}{s^2}}$ and let ψ_s (the normalization of ρ_s) correspond to the probability density function (PDF) of the normalized n-dimensional Gaussian distribution with mean 0 and standard deviation s. In the following, for a function f we concisely represent $\sum_{\mathbf{v} \in \Lambda} f(\mathbf{v})$ by $f(\Lambda)$. To show that the distribution over $[\mathbf{x} \mod \Lambda]$ is (close to) uniform when $\mathbf{x}$ is sampled from a distribution with PDF ψ_s, one needs to show that for every coset $(\Lambda + \mathbf{c})$ of the lattice, $\psi_s(\Lambda + \mathbf{c}) \approx \frac{1}{\det(\Lambda)}$. Focusing on the zero coset, where $\mathbf{c} = \mathbf{0}$, we can prove this using the Poisson summation formula, which says that for any lattice Λ and integrable function ρ_s: $\psi_s(\Lambda) = \frac{1}{\det(\Lambda)} \cdot \widehat{\psi_s}(\Lambda^\vee)$, where for a function f, $\widehat{f}$ denotes the n-dimensional Fourier transform of f and $\Lambda^\vee$ is the dual lattice of Λ (see Section 2.3.2). It remains to show that $\widehat{\psi_s}(\Lambda^\vee)$ is close to 1 (i.e. is upperbounded by $1 + \varepsilon$).

The proof approach outlined above can be applied to (integrable) normalized PDF Ψ that are not Gaussians centered at 0: To show that the distribution over $[\mathbf{x} \mod \Lambda]$ is (close to) uniform when $\mathbf{x}$ is sampled from a distribution with PDF Ψ, it is sufficient to show that $\widehat{\Psi}(\Lambda^\vee)$ is upperbounded by $1 + \varepsilon$.

In this work, we consider PDF's, Ψ, that correspond to the PDF of $\mathbf{x}$, *from the point of view of the adversary*, given the leakage. The technical contribution of this work is to show that, for each conditional distribution, (with overwhelming probability over choice of $\bar{A}$) $\widehat{\Psi}(\Lambda^\perp(A)^\vee)$ is close to 1. Specifically, for each distribu-

tion, our approach requires: (1) Determining the PDF Ψ, (2) Computing (an upper bound for) the multi-dimensional Fourier transform of Ψ (denoted $\widehat{\Psi}$), (3) Proving that $\widehat{\Psi}((\Lambda^{\perp}(A))^{\vee})$ is upperbounded by $1+\varepsilon$ (or, equivalently that $\widehat{\Psi}((\Lambda^{\perp}(A))^{\vee} \setminus \{\mathbf{0}\})$ is upperbounded by ε).

5.1.2 Related Work

Leakage-resilient cryptography. There is a significant body of work on leakage-resilient cryptographic primitives, beginning with the work of Dziembowski and Pietrzak [57] on leakage-resilient stream-ciphers. Other constructions include [4, 26, 27, 54, 74, 74, 80, 80, 86, 93, 99]. With the exception of [4], most of these results construct new cryptosystems from the bottom up. In our work, we consider whether we can prove that an existing cryptosystem enjoys leakage resilience, without modification of the scheme.

Lattice-based & leakage-resilient cryptography. Goldwasser et al. [64] initiated the study of leakage resilience of lattice based cryptosystems. This was followed by series of works [4,53,55], all these papers however study leakage resilience of schemes based on standard LWE problem in both symmetric as well as public key setting.

Robustness of Ring-LWE To the best of our knowledge the ePrint version [45] of this work is the first effort to study the robustness of RLWE based cryptosystems under leakage. Subsequent to the publishing of ePrint [47], interest has sparked in analyzing the RLWE-based schemes and their leakage resilience. Albrecht et.al [6] investigatedcold boot attack on RLWE based KEM schemes and compared the num-

ber of operations required to mount the attack when secret is stored with different encodings. Recently, Bolboceanu et.al [20] studied the hardness of RLWE problem in cases where the secret is sampled from distributions other than uniform random distribution over the ring. In [46] it is shown that under specific structured leakage on the NTT encoding of secret key, it is possible to recover the entire secret key given multiple RLWE samples and they implement the attack to recover the secret in real world parameter settings.

Other variants of LHL Stehlé and Steinfeld [108] studied the leftover hash lemma in the ring setting for power of 2 cyclotomics and Rosca et.al [104] generalized their result to non-cyclotomic rings. However, both these results study the case where input is sampled from discrete Gaussian distribution.

5.2 Extending the Regularity Lemma

Our results are applicable when R is the ring of integers in the m^{th} cyclotomic number field K of degree n, $m = 2n$ is a power of 2 and prime q is s.t. $q \equiv 1 \mod m$. We denote by $I_k \in (R_q)^{k \times k}$ the identity matrix.

5.2.1 Conditional Distribution I

Recall that $\mathbf{x} = (x_1, \ldots, x_l)$, where each coordinate of each $x_i \in R_q$ is sampled from a discrete Gaussian with standard deviation s and each x_i is represented as a vector in either the polynomial or canonical basis.[6] We assume leakage of all

[6]Either representation works since for power-of-two cyclotomics, spherical Gaussians in the polynomial basis correspond to spherical Gaussians in the canonical basis.

coordinates, with Gaussian noise of standard deviation $v = \tau \cdot s$ added. It turns out that this conditional distribution is fairly simple to handle since if X and Y are independent Gaussian random variables, then the distribution of X conditioned on $X + Y$ is also a Gaussian that is *not* centered at 0. Fortunately, the regularity lemma of [85] straightforwardly extends to Gaussians that are not centered at 0. We discuss formal details next, however, we mainly view Conditional Distribution I as a warm-up to the more difficult Conditional Distributions II and III.

We begin by defining some notation, which will be useful in all of the Conditional Distributions when manipulating Gaussian-distributed random variables. We write probability density function of random variable X at value $\mathbf{x}$, sampled from n-dimensional Gaussian distribution with each component of variable pairwise independent, as

$$\psi_{\mathbf{s},\mathbf{u}}(X = \mathbf{x}) = \prod_{i \in [n]} \frac{1}{s_i} \exp\left(\frac{-\pi(x_i - u_i)^2}{s_i^2}\right),$$

with mean $\mathbf{u} = (u_1, \ldots, u_n)$ and standard deviation $\mathbf{s} = (s_1, \ldots, s_n)$. The probability density function of Y at value $\mathbf{y}$, sampled from n-dimensional Gaussian distribution with each component of variable pairwise independent, can be written as

$$\psi_{\mathbf{v},\boldsymbol{\mu}}(Y = \mathbf{y}) = \prod_{i \in [n]} \frac{1}{v_i} \exp\left(\frac{-\pi(y_i - \mu_i)^2}{v_i^2}\right),$$

with mean $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and standard deviation $\mathbf{v} = (v_1, \ldots, v_n)$.

We now consider the distribution of X, conditioned on knowledge of $X + Y$.

We proceed with the following straightforward lemma:

Lemma 5.1. *Given two independent random variables X and Y. Suppose that the distribution of X is a n-dimensional Gaussian distribution with mean $\mathbf{u}$ and standard deviation $\mathbf{s}$, each component of X pairwise independent, and the distribution of Y is a n-dimensional Gaussian distribution with mean $\boldsymbol{\mu}$ and standard deviation $\mathbf{v}$, each component of Y pairwise independent. Then the distribution of X conditioned on $X+Y$ is also a n-dimensional Gaussian distribution, where each component of X is pairwise-independent with mean $\mathbf{c} := (c_1, \ldots, c_n)$ where $c_i := \frac{\frac{u_i}{s_i^2} - \frac{\mu_i}{v_i^2} + \frac{z_i}{v_i^2}}{\left(\frac{1}{s_i^2} + \frac{1}{v_i^2}\right)}$ and standard deviation $\sigma := (\sigma_1, \ldots, \sigma_n)$, where $\sigma_i := \sqrt{\frac{1}{\frac{1}{s_i^2} + \frac{1}{v_i^2}}}$.*

Proof. We have $F_{Z|A}(Z = b)$ generically represent the probability density function of random variable Z at value b, conditioned on event A.

We can then derive the density function of X given the value $\mathbf{z} = (z_1, \ldots, z_n)$ of $X + Y$ by computing

$$
\begin{aligned}
F_{X|X+Y=\mathbf{z}}(X = \mathbf{x}) &= \frac{\psi_{\mathbf{s},\mathbf{u}}(X = \mathbf{x})\psi_{\mathbf{v},\boldsymbol{\mu}}(Y = \mathbf{y}),}{\int_{R^n} \psi_{\mathbf{s},\mathbf{u}}(X = \mathbf{x})\psi_{\mathbf{v},\boldsymbol{\mu}}(Y = \mathbf{y})\,\mathrm{d}\mathbf{x}} \\
&= \frac{\prod_{i\in[n]} \frac{1}{s_i v_i} e^{-\frac{\pi(x_i-u_i)^2}{v^2}} e^{-\frac{\pi(z_i-x_i-\mu)^2}{{v_i}^2}}}{\prod_{i\in[n]} \int_{-\infty}^{\infty} \frac{1}{s_i v_i} e^{-\frac{\pi(x_i-u_i)^2}{v^2}} e^{-\frac{\pi(z_i-x_i-\mu)^2}{{v_i}^2}}\,\mathrm{d}x} \\
&= \prod_{i\in[n]} \sqrt{\frac{1}{s_i^2} + \frac{1}{v_i^2}} \exp\left(-\pi\left(\frac{1}{s_i^2} + \frac{1}{v_i^2}\right)\left(x_i - \frac{\frac{u_i}{s_i^2} - \frac{\mu_i}{v_i^2} + \frac{z_i}{v_i^2}}{\frac{1}{s_i^2} + \frac{1}{v_i^2}}\right)^2\right)
\end{aligned}
$$

Hence $F_{X|X+Y=\mathbf{z}}(X = \mathbf{x})$ is also in the form of probability density function of X on value x sampled n-dimensional Gaussian distribution, where each component x_i is generated independently with mean $\frac{\frac{u_i}{s_i^2} - \frac{\mu_i}{v_i^2} + \frac{z_i}{v_i^2}}{\left(\frac{1}{s_i^2} + \frac{1}{v_i^2}\right)}$, and variance parameter $\frac{1}{\frac{1}{s_i^2} + \frac{1}{v_i^2}}$.

□

Specifically, Lemma 5.1 shows that, conditioned on leakage, each coordinate x_i of the secret key is sampled from a multivariate Gaussian distribution $\rho_{\sigma,\mathbf{c}^i}$ with mean $\mathbf{c}^i := (c_1^i, \ldots, c_n^i)$, where $c_j^i := \frac{z_j}{\tau^2+1}$ and $\sigma = s\sqrt{\frac{\tau^2}{\tau^2+1}}$. The entire secret key is then sampled from $\rho_{\sigma,\mathbf{c}}$, where $\mathbf{c} = [\mathbf{c}^i]_{i\in l}$. We have the following theorem:

Theorem 5.2. *For positive integers $k \leq l \leq \mathrm{poly}(n)$, let $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$, where $\bar{A} \in (R_q)^{k\times(l-k)}$ is uniformly random. Then for all $\sigma \geq 2n \cdot q^{k/l+2/(nl)}$ and $\mathbf{c} \in \mathbb{R}^{n\cdot l}$ then*

$$\widehat{\rho_{\sigma,\mathbf{c}}}\left(\Lambda^{\perp}(A)^{\vee}\right) \leq 1 + 2^{-\Omega(n)},$$

except with probability at most $2^{-\Omega(n)}$ over choice of $\bar{A}$.

Proof. The theorem follows from Lemma 2.27 and the regularity lemma from [85].

□

The following corollary follows from Lemmas 2.32 and 2.33 and Theorem 5.2.

Corollary 5.3. *Let $R, n, q, k, l, \mathbf{c}, \sigma$ be as in Theorem 5.2. Assume that $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$ is chosen as in Theorem 5.2. Then, with probability $1 - 2^{-\Omega(n)}$ over the choice of $\bar{A}$, the distribution of $A\mathbf{x} \in R_q^k$, where $\mathbf{x} \in R^l$ is chosen from $D_{\Lambda,\sigma,\mathbf{c}}$, the discrete Gaussian probability distribution over R^l with parameter σ and center $\mathbf{c}$, satisfies that the probability of each of the q^{nk} possible outcomes is in the interval $(1 \pm 2^{-\Omega(n)})q^{-nk}$ (and in particular is within statistical distance $2^{-\Omega(n)}$ of the uniform distribution over R_q^k).*

In particular, this means that the standard deviation used to sample $\mathbf{x}$ should be increased from $2n \cdot q^{k/l+2/(nl)}$ (as in [85]) to $\sqrt{\frac{1+\tau^2}{\tau^2}} \cdot 2n \cdot q^{k/l+2/(nl)}$. Setting $\tau = 1$, we obtain the parameters described in the introduction.

5.2.2 Conditional Distribution II

Recall that $\mathbf{x} = (x_1, \ldots, x_l)$, where each $x_i \in R_q$ and each x_i is represented as a vector in the canonical embedding. We assume leakage of ℓ coordinates—with low noise added—of each x_i for $i \in [l]$ and restrict the coordinates leaked across each x_i to be the same. Let $\mathcal{S} \subseteq [n]$, where $|\mathcal{S}| = \ell$ denote the set of positions (from each x_i) that are leaked. Lemma 5.1 shows that, conditioned on leakage, each component x_i^j, $i \in [l], j \in \mathcal{S}$, (resp. $\notin \mathcal{S}$) is sampled from Gaussian distribution with mean $c_i^j := \frac{nz_i^j}{n+\frac{1}{s^2}}$ (resp. 0), and variance $\sigma_j^2 \geq 4n^2$ (resp. $\sigma_j^2 = s^2$).

Theorem 5.4. *For positive integers $k \leq l \leq \mathrm{poly}(n)$, let $A = [I_k|\bar{A}] \in (R_q)^{k \times l}$, where $\bar{A} \in (R_q)^{k \times (l-k)}$ is uniformly random. Let $\sigma := (\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n_{>0}$ and $\mathbf{c} := (c_1, \ldots, c_{ln}) \in \mathbb{R}^{ln}$ be vectors, where ℓ positions in σ are set to $2n$, and all others are set to s. Let k, l, ℓ be such that $l - k - l \cdot \ell/n > 0$ and $l - k - 1 \geq 1$, and let $s \geq 2n \cdot q^{\frac{kn+2}{l(n-\ell)}}$ then $\widehat{\rho_{\sigma^l, \mathbf{c}}}\left(\Lambda^{\perp}(A)^{\vee}\right) \leq 1 + 2^{-\Omega(n)}$ except with probability at most $2^{-\Omega(n)}$ over choice of $\bar{A}$.*

For proving Theorem 5.4, we begin with exposition on the forms of the Ideals $qR^{\vee} \subseteq \mathcal{J} \subseteq R^{\vee}$ in power-of-two cyclotomics as well as some lemmas.

To generate the set T of ideals $\mathcal{J}$ such that $qR^{\vee} \subseteq \mathcal{J} \subseteq R^{\vee}$ we take each ideal $\mathcal{I}$ s.t. $qR \subseteq \mathcal{I} \subseteq R$ and set $\mathcal{J} := q\mathcal{I}^{\vee}$. Recall from Fact 2.14 that $\langle q \rangle$ splits

completely into n distinct ideals of norm q, i.e. $qR = \Pi_{i\in[n]}\mathfrak{p}_i$. Therefore, the set of all ideals $\mathcal{I}$ such that $qR \subseteq \mathcal{I} \subseteq R$, is exactly the set $\mathcal{S} := \{\Pi_{i\in S}\mathfrak{p}_i \mid S \subseteq [n]\}$. Thus, the number of ideals $\mathcal{I}$ such that $qR \subseteq \mathcal{I} \subseteq R$ (and hence also the number of ideals $\mathcal{J} \in T$) is exactly 2^n. Moreover, note that for each ideal $\mathcal{J} \in T$,

$$|\mathcal{J}/qR^\vee| = |R/q\mathcal{J}^\vee| = N(q\mathcal{J}^\vee).$$

Thus, we see that for each $\mathcal{J} \in T$, $1 \leq |\mathcal{J}/qR^\vee| \leq q^n$.

Let T_1 denote the set of ideals $\mathcal{J} \in T$ such that $|\mathcal{J}/qR^\vee| < 2^n$. Let T_2 denote the set of ideals $\mathcal{J}$ such that $|\mathcal{J}/qR^\vee| \geq 2^n$. Furthermore, let T_2^1 be the set of $\mathcal{J} \in T_2$ such that $s \geq \eta_{2^{-2n}}((\frac{1}{q}\mathcal{J})^\vee)$ (where $\eta_{2^{-2n}}$ denotes the smoothing parameter and s is fixed as above). Let $T_2^2 := T_2 \setminus T_2^1$. Let $\sigma := (\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n_{>0}$ be a vector with ℓ positions are set to $2n$, while the other positions are set to value s.

Lemma 5.5. *For ideals $\mathcal{J} \in T_1$,*

$$\eta_{2^{-2n}}\left((\frac{\mathcal{J}}{q})^\vee\right) \leq 2n.$$

Proof.

$$\eta_{2^{-2n}}\left((\frac{\mathcal{J}}{q})^\vee\right) \le \frac{\sqrt{n}}{\lambda_1\left((\frac{\mathcal{J}}{q})^\vee\right)} \tag{5.1}$$

$$\le \left(N\left(\frac{\mathcal{J}}{q}\right)\right)^{-1/n} \tag{5.2}$$

$$\le (|\mathcal{J}/qR^\vee| \cdot n^n)^{1/n} \tag{5.3}$$

$$\le (2^n \cdot n^n)^{1/n} \tag{5.4}$$

$$= 2n,$$

where (5.1) follows from Lemma 2.24, (5.2) follows from Lemma 2.12, and (5.3) follows from the fact that $\left(N\left(\frac{\mathcal{J}}{q}\right)\right)^{-1} = |\mathcal{J}/qR| = |R^\vee/R| \cdot |\mathcal{J}/qR^\vee| = \Delta_K|\mathcal{J}/qR|$ (for example, see [40, page. 63]), and (5.4) follows from the definition of T_1. □

Lemma 5.6. *For ideals $\mathcal{J} \in T_2^1$*

$$|\mathcal{J}/qR^\vee|^{-(l-k)}\left(\rho_{1/\sigma_1,\ldots,1/\sigma_n}\left(\frac{1}{q}\mathcal{J}\right)^l\right) \le 2^{-n(l-k)},$$

where $\rho_{1/\sigma_1,\ldots,1/\sigma_n}$ is an n-dimensional Gaussian function with coordinate-wise standard deviation $1/\sigma_i, i \in [n]$ and center 0 (see beginning of Appendix 2.4.2).

Proof. Recall that $\sigma := (\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^n_{>0}$ is defined as a vector such that ℓ positions are set to $2n$, while the other positions are set to s. Define $z_1, \ldots, z_n$ in the following way: For $i \in [n]$, if $\sigma_i = s$ then $z_i = \sigma_i$. Otherwise, $z_i = \eta_{2^{-2n}}\left((\frac{1}{q}\mathcal{J})^\vee\right)$. Applying Poisson summation twice we arrive at:

$$\rho_{1/\sigma_1,\ldots,1/\sigma_n}\left(\frac{1}{q}\mathcal{J}\right) = 1/\det(\frac{1}{q}\mathcal{J})\cdot(1/\sigma_1\cdots 1/\sigma_n)\rho_{\sigma_1,\ldots,\sigma_n}\left((\frac{1}{q}\mathcal{J})^\vee\right) \tag{5.5}$$

$$\leq 1/\det(\frac{1}{q}\mathcal{J})\cdot(1/\sigma_1\cdots 1/\sigma_n)\rho_{z_1,\ldots,z_n}\left((\frac{1}{q}\mathcal{J})^\vee\right) \tag{5.6}$$

$$= \left(\frac{\eta_{2^{-2n}}((\frac{1}{q}\mathcal{J})^\vee)}{2n}\right)^\ell \cdot \rho_{1/z_1,\ldots,1/z_n}\left(\frac{1}{q}\mathcal{J}\right) \tag{5.7}$$

$$\leq (1+2^{-2n})\cdot\left(\frac{\eta_{2^{-2n}}((\frac{1}{q}\mathcal{J})^\vee)}{2n}\right)^\ell, \tag{5.8}$$

where (5.6) follows from definitions of ρ and z_i. To derive (5.7), let us first introduce the following claim.

Claim 5.7. *For any lattice $L^\vee$,*

$$\rho_{s_1,\ldots,s_n}(L) = s_1\cdot s_2\cdot\ldots\cdot s_n\cdot\frac{1}{\det(L)}\cdot\rho_{1/s_1,\ldots,1/s_n}(L^\vee)$$

Proof. It can be easily verified by combining Poisson Summation formula and the fact that $\hat{\rho}_{s_1,\ldots,s_n} = s_1\cdots s_n\rho_{1/s_1,\ldots,1/s_n}$. □

By replacing s_i with $1/z_i$ for all i and replacing L with $\frac{1}{q}\mathcal{J}$, we have

$$1/\det(\frac{1}{q}\mathcal{J})\cdot\rho_{z_1,\ldots,z_n}\left((\frac{1}{q}\mathcal{J})^\vee\right) = z_1\cdots z_n\cdot\rho_{1/z_1,\ldots,1/z_n}\left(\frac{1}{q}\mathcal{J}\right).$$

By plugging into (5.6), we have

$$\left(\frac{z_1}{\sigma_1}\cdots\frac{z_n}{\sigma_n}\right)\cdot\rho_{1/z_1,\ldots,1/z_n}\left(\frac{1}{q}\mathcal{J}\right)$$

By definition of z_i, $\frac{z_i}{\sigma_i} = 1$ when $\sigma_i = s$ and $\frac{z_i}{\sigma_i} = \frac{\eta_{2^{-2n}}((\frac{1}{q}\mathcal{J})^\vee)}{2n}$, when $\sigma_i = 2n$. Since there are ℓ positions in σ when $\sigma_i = 2n$, we obtain (5.7). Finally (5.8) follows by definition of smoothing parameter $\eta_{2^{-2n}}((\frac{1}{q}\mathcal{J})^\vee)$.

Now, using the fact that $\eta_{2^{-2n}} \leq (\Delta_K |\mathcal{J}/qR^\vee|)^{1/n}$, the fact that $\Delta_K = n^n$ and the fact that $|\mathcal{J}/qR^\vee| \geq 2^n$, and the set of parameters, we have that

$$|\mathcal{J}/qR^\vee|^{-(l-k)} \left(\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J} \right)^l \right) \leq |\mathcal{J}/qR^\vee|^{-(l-k-l\cdot\ell/n)} (1 + 2^{-2n})^l \cdot 2^{-\ell\cdot l}$$
$$\leq 2^{-n(l-k)}$$

which completes the proof of the lemma. □

We now conclude the proof of Theorem 5.4.

Proof of Theorem 5.4. Since by Lemma 2.27 we have that for any $(n\cdot l)$-dimensional vectors, $\mathbf{c}$, $\mathbf{x}$ and any n-dimensional vector $\sigma = (\sigma_1, \ldots, \sigma_n)$:

$$\widehat{\rho_{\sigma^l,\mathbf{c}}}(\mathbf{x}) \leq \widehat{\rho_{\sigma^l}}(\mathbf{x}) = \rho_{(1/\sigma_1,\ldots,1/\sigma_n)^l}(\mathbf{x}),$$

then following the proof of [85] step-by-step, it is sufficient to show that

$$\sum_{\mathcal{J}\in T} |\mathcal{J}/qR^\vee|^{-(l-k)} \cdot \left(\rho_{(1/\sigma_1,\ldots,1/\sigma_n)} \left(\frac{1}{q}\mathcal{J} \right)^l - 1 \right) \leq 2^{-\Omega(n)}.$$

We will show that

$$\sum_{\mathcal{J}\in T_2^1} |\mathcal{J}/qR^\vee|^{-(l-k)} \left(\rho_{(1/\sigma_1,\ldots,1/\sigma_n)} \left(\frac{1}{q}\mathcal{J}\right)^l - 1 \right) \leq 2^{-\Omega(n)}, \tag{5.9}$$

and that

$$\sum_{\mathcal{J}\in(T_1\cup T_2^2)} |\mathcal{J}/qR^\vee|^{-(l-k)} \left(\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J}\right)^l - 1 \right) \leq 2^{-\Omega(n)} \tag{5.10}$$

To show (5.10), note that by Lemma 5.5, for ideals $\mathcal{J} \in T_1$ (we have that $\eta_{2^{-2n}}((\frac{\mathcal{J}}{q})^\vee) \leq 2n$. This means that for each $i \in [n]$, $\sigma_i \geq \eta_{2^{-2n}}$, which implies that $\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J}\right)^l \leq (1+2^{-2n})^l$.

On the other hand, by definition of T_2^2, for ideals $\mathcal{J} \in T_2^2$, we have that $\sigma_i < \eta_{2^{-2n}}$, for each $i \in [n]$. Thus, by Lemma 2.26 we have that $\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J}\right) \leq \left(\frac{\eta_{2^{-2n}}((\frac{\mathcal{J}}{q})^\vee)}{\sigma_1} \cdots \frac{\eta_{2^{-2n}}((\frac{\mathcal{J}}{q})^\vee)}{\sigma_n} \right) \cdot (1+2^{-2n})$. Since $\eta_{2^{-2n}}((\frac{\mathcal{J}}{q})^\vee)^n \leq |\mathcal{J}/qR^\vee|\Delta_K$, and plugging in the proper values for $\sigma_1,\ldots,\sigma_n$, we have that $\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J}\right)^l \leq (|\mathcal{J}/qR^\vee|\Delta_K s^{-n+\ell} \cdot (2n)^{-\ell})^l \cdot (1+2^{-2n})^l$. Combining the above, we get that for $\mathcal{J} \in T_1 \cup T_2^2$,

$$\rho_{1/\sigma_1,\ldots,1/\sigma_n} \left(\frac{1}{q}\mathcal{J}\right)^l \leq \max(1, (|\mathcal{J}/qR^\vee|\Delta_K s^{-n+\ell} \cdot (2n)^{-\ell})^l) \cdot (1+2^{-2n})^l.$$

Similarly to [85], using the lower bound of s from Theorem 5.4, we bound

$$\begin{aligned}
&\sum_{\mathcal{J}\in(T_1\cup T_2^2)} |\mathcal{J}/qR^\vee|^{-(l-k)} \left(\rho_{1/\sigma_1,\ldots,1/\sigma_n}\left(\frac{1}{q}\mathcal{J}\right)^l - 1\right) \\
&\leq \sum_{\mathcal{J}\in(T_1\cup T_2^2)} |\mathcal{J}/qR^\vee|^{-(l-k)} \cdot \max(1, (|\mathcal{J}/qR^\vee|\Delta_K s^{-n+\ell}\cdot(2n)^{-\ell})^l)\cdot(1+\varepsilon)^l \\
&\leq \sum_{\mathcal{J}\in T} |\mathcal{J}/qR^\vee|^{-(l-k)} \cdot \max(1, (|\mathcal{J}/qR^\vee|\Delta_K s^{-n+\ell}\cdot(2n)^{-\ell})^l)\cdot(1+\varepsilon)^l \\
&\leq 2^{-\Omega(n)} + 2(s/n)^{-nl} q^{kn+2}\left(\frac{s}{2n}\right)^{l\cdot\ell} \in 2^{-\Omega(n)}.
\end{aligned}$$

Moreover, by Lemma 5.6 and the fact that $|T_2^1| \leq |T| = 2^n$, we can bound

$$\sum_{\mathcal{J}\in T_2^1} |\mathcal{J}/qR^\vee|^{-(l-k)} \left(\rho_{1/\sigma_1,\ldots,1/\sigma_n}\left(\frac{1}{q}\mathcal{J}\right)^l - 1\right) \leq 2^n \cdot 2^{-n(l-k)} \in 2^{-\Omega(n)},$$

where the last line follows from the setting of parameters in Theorem 5.4.

This completes the proof. □

The following corollary follows from Lemmas 2.32 and 2.33 and Theorem 5.4.

Corollary 5.8. *Let k, l, ℓ, σ and $\mathbf{c}$ be as in Theorem 5.4. Assume that $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$ is chosen as in Theorem 5.4. Then, with probability $1 - 2^{-\Omega(n)}$ over the choice of $\bar{A}$, the distribution of $A\mathbf{x} \in R_q^k$, where $\mathbf{x} \in R^l$ is chosen from $D_{R^l,\sigma^l,\mathbf{c}}$, the discrete Gaussian probability distribution over R^l with parameter σ^l and center $\mathbf{c}$, satisfies that the probability of each of the q^{nk} possible outcomes is in the interval $(1\pm 2^{-\Omega(n)})q^{-nk}$ (and in particular is within statistical distance $2^{-\Omega(n)}$ of the uniform distribution over R_q^k).*

In particular, this means that the standard deviation used to sample $\mathbf{x}$ should be increased from $2n \cdot q^{k/l+2/(nl)}$ (as in [85]) to $2n \cdot q^{\frac{kn+2}{l(n-\ell)}}$.

5.2.3 Conditional Distribution III

We slightly change the dimensions so that $\mathbf{x}$ is represented by a vector of dimension $n' := l \cdot n + 1$. When n is a power of two, a spherical Gaussian in the coefficient representation is also a spherical Gaussian in the canonical embedding representation [84]. So we can assume that $\mathbf{x}$ is generated using the coefficient representation, where each coordinate is sampled independently from a discrete Gaussian, $D_{Z,s'}$. During sampling of $\mathbf{x}$, an additional coordinate is sampled and stored together with the remainder of the secret.

Recall that a generic PDF of one dimensional Gaussian distribution is defined as:

$$\psi_{s,u}(x) = \frac{1}{s} \exp\left(\frac{-\pi(x-u)^2}{s^2}\right),$$

where u is mean, and s is standard deviation of the distribution. We write probability density function of secret key X at value $\mathbf{x} = (x_1, \ldots, x_{n'})$, of which each coordinate is independently sampled from a Gaussian distribution with center at 0 and standard deviation s, as

$$\psi_s(X = \mathbf{x}) = \prod_{i \in [n']} \frac{1}{s} \exp\left(\frac{-\pi x_i^2}{s^2}\right) = \frac{1}{s^{n'}} \exp\left(\frac{-\pi r^2}{s^2}\right) = \psi_s(\|X\| = r),$$

where r is the magnitude of $\mathbf{x}$. It also can be viewed as probability density function

of secret key for its magnitude $\|X\| = r$, denoted as $\psi_s(\|X\| = r)$. The error is sampled from a 1-dimensional Gaussian distribution with center at 0. We write probability density function of error E at value y is

$$\psi_v(E = y) = \frac{1}{v} \exp\left(\frac{-\pi y^2}{v^2}\right).$$

Let $F_{Z|A}(f(Z) = b)$ generically represent the probability density function of random variable Z at value b of $f(Z)$, conditioned on event A.

We now derive the density function of secret key X given the value z of $|\|X\| + E|$. The weight placed on a value $\mathbf{x} = (x_1, \ldots, x_{n'})$ by the conditional distribution depends *only* on the magnitude of $\mathbf{x}$ (i.e. $r = \|\mathbf{x}\|$) and can be computed as:

$$\begin{aligned}
&F_{X\left|_{|\|X\|+E|=z}\right.}(\|X\| = r) = \frac{F_{X,E}(\|X\| = r, \|X\| + E| = z)}{F_{X,E}(\|X\| + E| = z)} \\
=&\frac{\psi_s(\|X\| = r)\psi_v(E = z - r) + \psi_s(\|X\| = r)\psi_v(E = -z - r)}{F_{X,E}(\|X\| + E = z) + F_{X,E}(\|X\| + E = -z)} \\
=&\frac{\psi_s(\|X\| = r)\psi_v(E = z - r) + \psi_s(\|X\| = r)\psi_v(E = -z - r)}{\int_{R^{n'}} \psi_s(\|X\| = \|\mathbf{x}\|)\psi_v(E = z - \|\mathbf{x}\|) + \psi_s(\|X\| = \|\mathbf{x}\|)\psi_v(E = -z - \|\mathbf{x}\|)\,\mathrm{d}\mathbf{x}} \\
=&\frac{e^{-\left(\frac{\pi}{s^2}+\frac{\pi}{v^2}\right)\left(r-\frac{zs^2}{v^2+s^2}\right)^2} + e^{-\left(\frac{\pi}{s^2}+\frac{\pi}{v^2}\right)\left(r+\frac{zs^2}{v^2+s^2}\right)^2}}{nV_n \int_{-\infty}^{\infty} e^{-\left(\frac{\pi}{s^2}+\frac{\pi}{v^2}\right)\left(r-\frac{zs^2}{v^2+s^2}\right)^2} r^{n-1} dr} \\
=&\frac{e^{-\left(\frac{\pi}{s^2}+\frac{\pi}{v^2}\right)\left(r-\frac{zs^2}{v^2+s^2}\right)^2} + e^{-\left(\frac{\pi}{s^2}+\frac{\pi}{v^2}\right)\left(r+\frac{zs^2}{v^2+s^2}\right)^2}}{N},
\end{aligned} \tag{5.11}$$

where N is the normalization factor.

$F_{X\left|_{|\|X\|+E|=z}\right.}(\|X\| = r)$ is the sum of two Gaussian functions centered at $\frac{zs^2}{v^2+s^2}$ and $-\frac{zs^2}{v^2+s^2}$ respectively with the same standard deviation σ. Suppose $v = s$, we have $\sigma = \frac{s}{\sqrt{2}}$.

Lemma 5.9. *Suppose $v = s$, we bound the center $\frac{zs^2}{v^2+s^2}$ from Equation 5.11 by $\Pr\left(\frac{zs^2}{v^2+s^2} \geq s\sqrt{n'}\right) \in 2^{-\Omega(n)}$, where the probability is taken over choice of $\mathbf{x}$ and e.*

We first present the following lemma, and then use it to prove Lemma 5.9.

Lemma 5.10. *Given a random variable Y chosen from a Gaussian distribution $G_E(y, v) = \frac{1}{v}\exp\left(\frac{-\pi y^2}{v^2}\right)$, Y is upper bounded by $v\sqrt{n'}$ except for negligible probability, written as $\Pr\left(Y \geq v\sqrt{n'}\right) \in 2^{-\Omega(n)}$.*

Proof. $\Pr(Y \geq y) = \Pr(X \geq x)$, where $X = \frac{\sqrt{2\pi}Y}{v}$ is a standard normal, $x = \frac{\sqrt{2\pi}y}{v}$. By using Chernoff bound and calculating exponential moment of standard normal distribution, we have, for any $\lambda > 0$.

$$\Pr(X \geq x) \leq \frac{\mathbb{E}\left[e^{\lambda X}\right]}{e^{\lambda x}} = \frac{e^{\lambda^2/2}}{e^{\lambda x}},$$

Set $\lambda = x$ and $y = v\sqrt{n'}$, then $\Pr\left(Y \geq v\sqrt{n'}\right) \leq e^{-x^2/2} = e^{-\pi n'}$. The lemma follows. □

Proof of Lemma 5.9. Using union bound, we have

$$\begin{aligned}
&\Pr\left(\frac{zs^2}{v^2+s^2} \geq s\sqrt{n'}\right) = \Pr\left(\frac{z}{2} \geq s\sqrt{n'}\right) \\
\leq &\Pr\left(R + E \geq 2s\sqrt{n'}\right) + \Pr\left(-R - E \geq 2s\sqrt{n'}\right) \\
\leq &\Pr\left(R \geq s \cdot \sqrt{n'}\right) + \Pr\left(E \geq v\sqrt{n'}\right) + \Pr\left(E \geq v\sqrt{n'}\right)
\end{aligned}$$

Note that since $s > n$, and using the fact that $\lambda_1((R^l \times \mathbb{Z})^\vee) \geq \lambda_1(R^\vee) \geq \sqrt{n}N^{\frac{1}{n}}(R^\vee) = \sqrt{n} \cdot (\Delta_k^{-1})^{\frac{1}{n}} \geq \sqrt{n}\left(\frac{1}{n^n}\right)^{\frac{1}{n}} = \frac{1}{\sqrt{n}}$ (See Lemma 2.12), by Lemma 2.24,

we ensure $s > \eta_{2^{-n}}(R^l \times \mathbb{Z})$. Then by Lemma 2.31 and Lemma 5.10, we deduce that $\Pr\left(\frac{zs^2}{v^2+s^2} \geq s\sqrt{n'}\right) \in 2^{-\Omega(n)}$. □

Let $\Psi_{\sigma,c}(\mathbf{x}) := F_{X\left|_{|\|X\|+E|=z}\right.}(\|X\| = \|\mathbf{x}\|)$ be the normalization of the function $f(\mathbf{x}) := e^{-\frac{\pi(\|\mathbf{x}\|-c)^2}{\sigma^2}} + e^{-\frac{\pi(\|\mathbf{x}\|+c)^2}{\sigma^2}}$. By Lemma 5.9, we have that with all but negligible probability, $c := \frac{zs^2}{v^2+s^2} \leq \sqrt{2} \cdot \sigma\sqrt{n'}$.

For the proof, we will require certain properties of the Fourier transform of $\Psi_{\sigma,c}$, when c is bounded as above. We state those properties in the following theorem, which is proved in Appendix A2.1.

Theorem 5.11. *Let $n' := l \cdot 2^a + 1$, where l, a are positive integers and $a > 2$, and $c \leq \sqrt{2} \cdot \sigma \cdot \sqrt{n'}$. Let $\Psi_{\sigma,c}$ denote the normalized pdf corresponding to the non-normalized function $f(\mathbf{x}) := e^{-\frac{\pi(\|\mathbf{x}\|-c)^2}{\sigma^2}} + e^{-\frac{\pi(\|\mathbf{x}\|+c)^2}{\sigma^2}}$, where $\mathbf{x}$ is a vector over n' dimensions. and let $\widehat{\Psi_{\sigma,c}}(\mathbf{y})$ denote the n'-dimensional Fourier transform of $\Psi_{\sigma,c}$. Then $|\widehat{\Psi_{\sigma,c}}(\mathbf{y})| \leq {n'}^{n'} \cdot e^{-\pi\|\mathbf{y}\|^2\sigma^2}$ for $\|\mathbf{y}\| > 1/\sigma$.*

We next present the main theorem of this section.

Theorem 5.12. *For positive integers $k \leq l \leq \text{poly}(n)$, let $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$, where $\bar{A} \in (R_q)^{k\times(l-k)}$ is uniformly random. Let $c \leq \sqrt{2} \cdot \sqrt{n'} \cdot \sigma$ and let $\sigma \geq \sqrt{\frac{7}{5} \cdot \frac{n'}{n} \ln n'} \cdot 2n \cdot q^{k/l+2/(nl)}$. Define $\Lambda^\perp(A)^+$ as a direct product of $\Lambda^\perp(A)$ and $\mathbb{Z}$, written as $\Lambda^\perp(A)^+ := \Lambda^\perp(A) \times \mathbb{Z}$. Then $\Psi_{\sigma,c}\left(\Lambda^\perp(A)^+\right) \leq \frac{1}{det(\Lambda^\perp(A)^+)}(1 + 2^{-\Omega(n)})$ except with probability at most $2^{-\Omega(n)}$.*

Proof. Note that $\Lambda^\perp(A)$ is a lattice of even dimension $l \cdot n$ (where n is a power of two), but Theorem 5.11 holds only for n' equal to $l \cdot 2^a + 1$. Therefore, we define

$n' := l \cdot n + 1$, and we have the n'-dimensional lattice $\Lambda^{\perp}(A)^{+} := \Lambda^{\perp}(A) \times \mathbb{Z}$. We have the following properties of $\Lambda^{\perp}(A)^{+}$, which can be verified by inspection:

(a) $(\Lambda^{\perp}(A)^{+})^{\vee} := \Lambda^{\perp}(A)^{\vee} \times \mathbb{Z}$;

(b) the shortest non-zero vector in $(\Lambda^{\perp}(A)^{+})^{\vee}$ is at least $\min(\lambda_1(\Lambda^{\perp}(A)^{\vee}), 1)$, where $\lambda_1(\Lambda^{\perp}(A)^{\vee})$ denotes the shortest non-zero vector in $\Lambda^{\perp}(A)^{\vee}$;

By Poisson summation formula, it is sufficient to show that with probability $1 - 2^{-\Omega(n)}$ over choice of A, $|\widehat{\Psi_{\sigma,c}}|(\Lambda^{\perp}(A)^{+})^{\vee}) \leq 1 + 2^{-\Omega(n)}$, where $\widehat{\Psi_{\sigma,c}}$ denotes the Fourier transform of $\Psi_{\sigma,c}$ over n' dimensions and the notation $|\widehat{\Psi_{\sigma,c}}|$ means the summation of the absolute value of the function over the lattice $\Lambda^{\perp}(A)^{+})^{\vee}$.

We first note that, over n' dimensions, $\widehat{\Psi_{\sigma,c}}(\mathbf{0}) = 1$. This follows due to the fact that by definition of Fourier transform, $\widehat{\Psi_{\sigma,c}}(\mathbf{0}) := \int_{\mathbb{R}^{n'}} \Psi_{\sigma,c}(\mathbf{x})\,\mathrm{d}\mathbf{x}$. Since $\Psi_{\sigma,c}$ is a normalized PDF, it must be the case that $\int_{\mathbb{R}^{n'}} \Psi_{\sigma,c}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 1$.

Thus, it remains to show that $\left|\widehat{\Psi_{\sigma,c}}\right| \left((\Lambda^{\perp}(A)^{+})^{\vee} \setminus \{\mathbf{0}\}\right) \leq 2^{-\Omega(n)}$.

Towards showing this, we first let $\beta = 2n \cdot q^{k/l+2/(nl)}$ for simplicity, and then use Theorem 5.11 to show that, when $\kappa = |\mathbf{y}| \geq \frac{\sqrt{n/\pi}}{\beta}$,

$$|\widehat{\Psi_{\sigma,c}}(\mathbf{y})| \leq n'^{n'} \cdot e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)} \leq n'^{n'} \cdot e^{-5(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \cdot e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \leq e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7},$$

where the last line follows since $\sigma := \sqrt{\frac{7n'}{5n} \ln n'} \cdot 2n \cdot q^{k/l+2/(nl)} = \sqrt{\left(\frac{7n'}{5n}\right) \ln n'} \cdot \beta$ is chosen so that when $\kappa \geq \frac{\sqrt{n/\pi}}{\beta}$, $e^{5(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \geq n'^{n'} = e^{n' \ln n'}$.

Let $Q := \sum_{\mathbf{y} \in (\Lambda^{\perp}(A)^{+})^{\vee} \setminus \{\mathbf{0}\}} e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7}$. Combining the above inequalities which hold when $\kappa \geq \frac{\sqrt{n/\pi}}{\beta}$, together with (b) and Corollary 2.37, which states

that with probability $1 - 2^{-\Omega(n)}$ over choice of A, the shortest non-zero vector in $\Lambda^{\perp}(A)^{\vee}$ has length $\kappa \geq \frac{\sqrt{n/\pi}}{\beta}$, we conclude that an upper bound on Q yields an upper bound on the desired quantity, $\left|\widehat{\Psi_{\sigma,c}}\right| \left((\Lambda^{\perp}(A)^{+})^{\vee} \setminus \{\mathbf{0}\}\right)$.

Additionally note that when $\kappa \geq \frac{\sqrt{n/\pi}}{\beta}$, then

$$e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7} = e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \cdot e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \leq e^{-1/5 \cdot n' \ln n'} \cdot e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7}, \tag{5.12}$$

where the inequality follows since (by above) $e^{5(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \geq n'^{n'} = e^{n' \ln n'}$. so $e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \leq n'^{-1/5 \cdot n'} = e^{-1/5 \cdot n' \ln n'}$. Moreover, recall that two applications of Poisson summation give:

$$\sum_{\mathbf{y} \in (\Lambda^{\perp}(A)^{+})^{\vee}} e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \leq 2^{n'} \cdot \sum_{\mathbf{y} \in (\Lambda^{\perp}(A)^{+})^{\vee}} e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \tag{5.13}$$

Combining the above, we have that

$$\begin{aligned} Q &\leq \sum_{\mathbf{y} \in (\Lambda^{\perp}(A)^{+})^{\vee}} e^{-1/5 \cdot n' \ln n'} \cdot e^{-(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \\ &\leq e^{-1/5 \cdot n' \ln n'} \cdot 2^{n'} \cdot \sum_{\mathbf{y} \in (\Lambda^{\perp}(A)^{+})^{\vee}} e^{-2(\sigma^2 \cdot \pi \cdot \kappa^2)/7} \\ &= e^{-1/5 \cdot n' \ln n'} \cdot 2^{n'}(1 + Q), \end{aligned}$$

where the first inequality follows from (5.12) and the definition of Q, the second inequality from (5.13), and the final equality from the definition of Q.

Thus we have that $(1 - e^{-1/5 \cdot n' \ln n'} \cdot 2^{n'})Q \leq e^{-1/5 \cdot n' \ln n'} \cdot 2^{n'}$ which implies that $Q \leq 2 \cdot e^{-1/5 n' \ln n'} \cdot 2^{n'} \leq 2^{-n'+1} \leq 2^{-\Omega(n)}$, assuming n' is at least 2^{10}. □

Corollary 5.13. *Let k, l, σ and $\mathbf{c}$ be as in Theorem 5.12. Assume that $A = [I_k|\bar{A}] \in (R_q)^{k\times l}$ is chosen as in Theorem 5.12. Then, with probability $1 - 2^{-\Omega(n)}$ over the choice of $\bar{A}$, the distribution of $A\mathbf{x} \in R_q^k$, where $(\mathbf{x}, x_{n'}) \in R^l \times Z$ is chosen from $D_{R^l\times Z,\Psi_{\sigma,c}}$ satisfies that the probability of each of the q^{nk} possible outcomes is in the interval $(1 \pm 2^{-\Omega(n)})q^{-nk}$ (and in particular is within statistical distance $2^{-\Omega(n)}$ of the uniform distribution over R_q^k).*

Proof. $\Psi_{\sigma,c}\left(\Lambda^{\perp}(A)^{+} + (\mathbf{b},\ b')\right) \in \det((\Lambda^{\perp}(A)^{+})^{\vee})(1 \pm 2^{-\Omega(n)})$, which means that if we choose a n'-dimensional vector from distribution $D_{R^l\times Z,\Psi_{\sigma,c}}$, written as $\mathbf{x}' = (\mathbf{x}, x_{n'})$, and let $(\mathbf{b}, b_{n'}) = \mathbf{x}' \mod (\Lambda^{\perp}(A)^{+})$, then the resulting distribution is within statistical distance $2^{-\Omega(n)}$ to uniform distribution over $(R^l \times Z)$ modulo $(\Lambda^{\perp}(A)^{+})$. Due to the structure of $\Lambda^{\perp}(A)^{+}$, this also implies that the marginal distribution over $\mathbf{b}$ is uniform over (R^l) modulo $(\Lambda^{\perp}(A))$. Moreover, we can easily see that for $\mathbf{x}' = (\mathbf{x}, x_{n'})$, if $\mathbf{x}' \mod (\Lambda^{\perp}(A)^{+}) = (\mathbf{b},\ b_{n'})$, then $A\mathbf{x} = A\mathbf{b}$. Finally, since when $\mathbf{b}$ is uniform random over R^l modulo $\Lambda^{\perp}(A)$, we have that $A\mathbf{b}$ is uniform random over R_q^k, the corollary follows. □

Given the corollary, the analysis of Conditional Distribution III is complete. In particular, this means that the standard deviation used to sample $\mathbf{x}$ should be increased from $2n \cdot q^{k/l+2/(nl)}$ (as in [85]) to $\sqrt{14/5 \cdot n'/n \cdot \ln n'} \cdot 2n \cdot q^{k/l+2/(nl)}$.

A1: Appendix of Learning with Errors with Side Information

A1.1 Refined prediction via BKZ-simulation and a probabilistic model

The work of [7] warns about a regime where those predictions are not accurate, due to a so-called *second-intersection* between the predicted lengths of the Gram-Schmidt vectors and the successive projections of the secret. This phenomenon only appears for small blocksizes β, which is not relevant for cryptographically hard instances. However, we would still like to be able to make reliable predictions for small blocksizes as well, so as to test the validity of our predictions with and without hints.

Other sources of inaccuracy of this model are the so-called head and tails phenomenon [11, 114], as well as the fact that one can be lucky: the projected length of the secret can vary, making it plausible that the secret will be found with a slightly smaller blocksize. For example, in [7] more than 50% of the attacks were already successful by running BKZ with blocksize $\beta_{\text{pred}} - 5$.

Furthermore, the predictions of [7] work under the assumption that as soon as the projected secret vector has been detected at position $d - \beta$, it will be "pulled-

back" to the front by the run of LLL that is typically executed between BKZ tours. For large block-sizes β this event is indeed very likely as argued and experimentally verified in [7], but may not occur in small or intermediate dimension. In fact, the issue of double-intersection is precisely related to this assumption.

For experimental validation purposes of our work, we prefer to have accurate prediction even for small blocksizes. We therefore devise a refined strategy. First, we resort to the so called BKZ-simulator [38] to predict more accurately the length ℓ_i of the Gram-Schmidt vectors. Secondly, we do not assume that the projected secret $\pi_i(\mathbf{s})$ (projected orthogonally to the $i-1$ first vectors of the reduced basis, as in [7]) has exactly length $\sqrt{n-i}$, but simply treat it as a spherical Gaussian. We can therefore compute the probability that it is detected at position i by considering the CDF of χ^2_{n-i}, the chi-square distribution with $n-i$ degrees of freedom.

At last, we do not only account for the detectability of the secret vector at position $i = n - \beta$, but also check whether it is likely that the vector will be pulled to the front (not by the interleaved LLL, by BKZ itself, which is more powerful). That is, we consider the probability that:

$$E_i : \|\pi_i(\mathbf{s})\| \leq \ell_i \text{ simultaneously for all } i \in \{d-\beta, d-2\beta+1, d-3\beta+2, \dots\}.$$

Those events are not perfectly independent, which makes computing the probability of the conjunction of those more painful.[1] For simplicity, we only account for dependence between consecutive events E_i and E_{i+1} and therefore avoid having

[1]The expert reader may note that, for $\mathbf{s}$ uniformly distributed over a sphere, such conjunction correspond to a cylinder interesection, as used for pruning in enumeration [61].

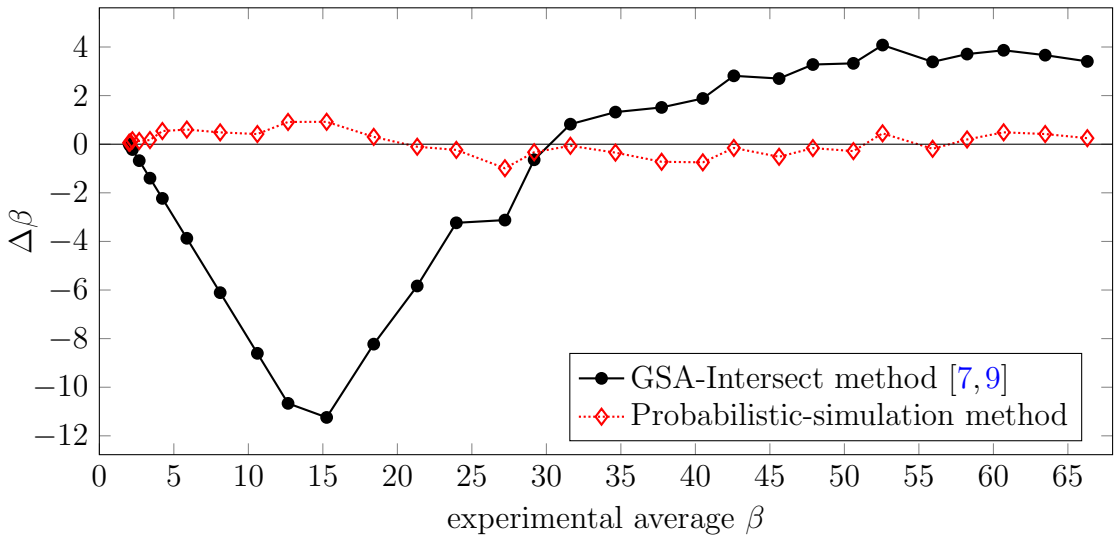


Figure A1.1: The difference $\Delta\beta = \text{real} - \text{predicted}$, as a function of the average experimental beta β. The experiment consists in running a single tour of BKZ-β for $\beta = 2, 3, 4, \ldots$ until the secret short vector is found. This was averaged over 256 many LWE instances per data-point, for parameters $q = 3301$, $\sigma = 20$ and $n = m \in \{30, 32, 34, \ldots, 88\}$.

to resort to numerical computation of nested integrals. We iteratively compute the success probability for each tour of BKZ-β for increasing β, and from there deduce the average successful β.

As depicted in Figure A1.1, this methodology (coined Probabilistic-simulation) leads to much more satisfactory estimates compared to the model from the literature [7, 9]. In particular, for low blocksize the literature widely underestimates the required blocksize, which is due to only considering detectability at position $d - \beta$. For large blocksize, it somewhat overestimates it, which could be attributed to the fact that it does not account for luck. On the contrary, our new methodology seems quite precise in all regimes, making errors of at most 1 bikz. This new methodology certainly deserves further study and refinement, which we leave to future work.

A1.2 Implementation

A1.2.1 Our Sage implementation

We propose three implementations of our framework, all following the same python/sage 9.0 API.[2] More specifically, the API and some common functions are defined in DBDD_generic.sage, as a class `DBDD_Generic`. Three derived classes are then given:

1. The class `DBDD` (provided in DBDD.sage) is the *full-fledged* implementation: i.e. it fully maintains all information about a DBDD instance as one integrates hints: the lattice Λ, the covariance matrix $\boldsymbol{\Sigma}$ and the average $\boldsymbol{\mu}$. While polynomial time, maintaining the lattice information can be quite slow, especially since consecutive intersections with hyperplanes can lead to manipulations on rationals with large denominators. It also allows to finalize the attack, running the homogenization, isotropization and lattice reduction, based on the fplll [50] library available through sage.

   We note that if one were to repeatedly use perfect or modular hints, a lot of effort would be spent on uselessly alternating between the primal and the dual lattice. Instead, we implement a caching mechanism for the primal and dual basis, and only update them when necessary.

2. The class `DBDD_predict` (provided in DBDD_predict.sage) is the *lightweight*

[2] While we would have preferred a full python implementation, we are making a heavy use of linear algebra over the rationals for which we could find no convenient python library.

implementation: it only fully maintains the covariance information, and the parameters of the lattice (dimension, volume). It must therefore work under assumptions about the primitivity of the vector $\mathbf{v}$; in particular, it cannot detect hints that are redundant. If one must resort to this faster variant on large instances, it is advised to consider potential (even partial) redundancy between the given hints, and to run a comparison with the previous on small instances with similarly generated hints.

3. The class `DBDD_predict_diag` (provided in DBDD_predict_diag.sage) is the *super-lightweight* implementation. It maintains the same information as the above, but requires the covariance matrix to remain diagonal at all times. In particular, one can only integrate hints for which the directional vector $\mathbf{v}$ is colinear with a canonical vector.

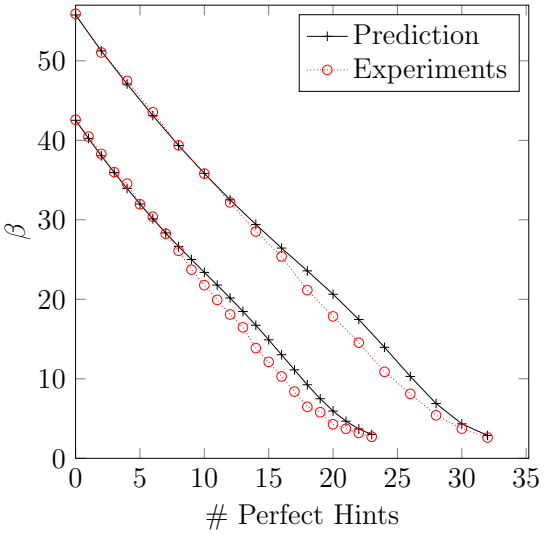
A1.2.2 Tests and validation

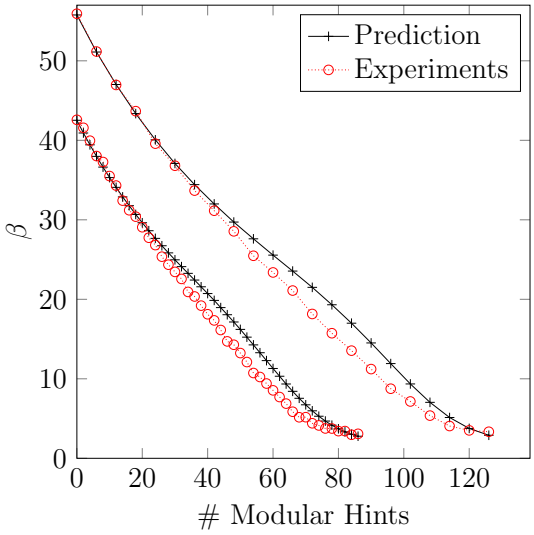
We implement two tests to verify the correctness of our scripts, and more generally the validity of our predictions.

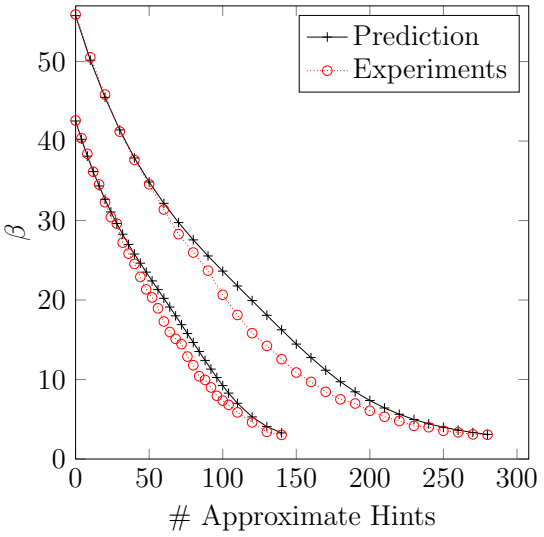
Consistency checks. Our first test (check_consistency.sage) simply verifies that all three classes always agree perfectly. More specifically we run all three versions on a given instances, integrating the same random hint in all of them, and compare their hardness prediction. We first test using the full-fledged version that the primitivity condition does hold, and discard the hint if not, as we know that predictions cannot be correct on such hints. This verification passes.

Prediction verifications. We now verify experimentally the prediction made by our tool for various types of hints, by comparing those predictions to actual attack experiments (see compare_usvp_models.sage for the prediction without hints and prediction_verifications.sage for the prediction with hints). This is done for a given set of LWE parameters, and increasing the number of hints. The details of the experiments and the results are given in Figure A1.2.

While our predictions seem overall accurate, we still note a minor discrepancy of up to 2 or 3 bikz in the low blocksize regime. This exceeds the error made by prediction on the attack without any hint, which was below 1 bikz, even in the same low blocksize regime. We suspected that this discrepancy is due to residual q-vectors, or small combinations of them, that are hard to predict for randomly generated hints, but would still benefit by lattice reduction. We tested that hypothesis by running similar experiments, but leaving certain coordinates untouched by hints, so to still explicitly know some q-vectors for short-vector hint integration, if they are "worthy". This didn't to improve the accuracy of our prediction, which infirms our suspected explanation. We are at the moment unable to explain this innacuracy. We nevertheless find our predictions satisfactory, considering that even without hints, previous predictions [7] were much less accurate (see Figure A1.1).







LWE Parameters		
$n = m = 70,$	$q = 3301,$	$\sigma = 20$
$n = m = 80,$	$q = 3301,$	$\sigma = 20$

Hint Type	Parameter
Perfect hints	–
Modular hints	mod 11
Approximate hints	$\sigma_\epsilon^2 = 3$

The hint vectors $\mathbf{v}$ were chosen as random ternary vectors of weight 5.

Figure A1.2: Experimental verification of the security decay predictions for each type of hints. Each data point was averaged over 256 samples.

A2: Appendix of Towards a Ring Analogue of the Leftover Hash Lemma

A2.1 Proof of Theorem 5.11

In this section, we prove the following theorem, which provides an upper bound on the Fourier transform of a pdf for the analysis of Conditional Distribution III in Section 5.2.3.

Theorem 5.11. *Let $n' := l \cdot 2^a + 1$, where l, a are positive integers and $a > 2$, and $c \leq \sigma \cdot \sqrt{2} \cdot \sqrt{n'}$. Let $\Psi_{\sigma,c}$ denote the normalized pdf corresponding to the non-normalized function $f(\mathbf{x}) := e^{-\frac{\pi(\|\mathbf{x}\|-c)^2}{\sigma^2}} + e^{-\frac{\pi(\|\mathbf{x}\|+c)^2}{\sigma^2}}$, where $\mathbf{x}$ is a vector over n' dimensions. and let $\widehat{\Psi_{\sigma,c}}(\mathbf{y})$ denote the n'-dimensional Fourier transform of $\Psi_{\sigma,c}$. Then $|\widehat{\Psi_{\sigma,c}}(\mathbf{y})| \leq n'^{n'} \cdot e^{-\pi\|\mathbf{y}\|^2\sigma^2}$ for $\|\mathbf{y}\| > 1/\sigma$.*

The following lemma computes a lower bound of the normalization factor of the pdf in Theorem 5.11. Once we prove the lemma, we proceed to the proof of Theorem 5.11.

Lemma A2.1. *Let $n' \in \mathbb{N}$ be odd, $\mathbf{x} \in \mathbb{R}^{n'}$, $c \in \mathbb{R}$. Then*

$$\int_{R^{n'}} e^{-\frac{\pi(\|\mathbf{x}\|-c)^2}{\sigma^2}} + e^{-\frac{\pi(\|\mathbf{x}\|+c)^2}{\sigma^2}} \, \mathrm{d}\mathbf{x} \geq \sigma^{n'}.$$

Proof. Let $f(\mathbf{x}) := e^{-\frac{\pi(\|\mathbf{x}\|-c)^2}{\sigma^2}} + e^{-\frac{\pi(\|\mathbf{x}\|+c)^2}{\sigma^2}}$. Let $r = \|\mathbf{x}\|$. Since f is a radial function, we slightly abuse notation and denote by $f(r) := e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}}$. Now, we have that

$$\int_{R^{n'}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = n' V_{n'} \int_0^\infty r^{n'-1} f(r)\,\mathrm{d}r, \tag{A2.1}$$

where $V_{n'}$ denotes the volume of n'-dimensional ball $V_{n'} = \frac{\pi^{n'/2}}{\Gamma(1+n'/2)}$. Since f is an even function and n' is odd, so $r^{n'-1}$ is an even function, we have that $r^{n'-1} f(r)$ is even and so

$$\int_0^\infty r^{n'-1} f(r)\,\mathrm{d}r = 1/2 \int_{-\infty}^\infty r^{n'-1} f(r)\,\mathrm{d}r. \tag{A2.2}$$

Let $a = \pi/\sigma^2$. Since n' is odd, we now have that

$$\begin{aligned}
&\int_{-\infty}^\infty e^{-a(r-c)^2} r^{n'-1}\,\mathrm{d}r \\
&= \int_{-\infty}^\infty e^{-at^2}(t+c)^{n'-1}\,\mathrm{d}t = \int_{-\infty}^\infty e^{-at^2} \sum_{j=0}^{n'-1} \binom{n'-1}{j} c^j t^{n'-1-j}\,\mathrm{d}t \\
&= \sum_{j=0}^{n'-1} \binom{n'-1}{j} c^j \int_{-\infty}^\infty e^{-at^2} t^{n'-1-j}\,\mathrm{d}t \\
&= \sum_{j=0}^{n'-1} \binom{n'-1}{j} c^j \frac{1}{2}(-1)^j \left((-1)^{n'+1} + (-1)^j\right) a^{\frac{1}{2}(-n'+j)} \Gamma\left(\frac{n'-j}{2}\right) \\
&= \sum_{j=0}^{\frac{n'-1}{2}} \binom{n'-1}{2j} c^{2j} a^{\frac{1}{2}(-n'+2j)} \Gamma\left(\frac{n'-2j}{2}\right) \\
&\geq a^{-\frac{1}{2}n'} \Gamma\left(\frac{n'}{2}\right)
\end{aligned}$$

Combining the above with (A2.1) and (A2.2) and substituting for a, we get that $\int_{R^{n'}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \geq \sigma^{n'}$, which completes the proof of the lemma. □

Proof of Theorem 5.11. Let N be the normalization of $f(\mathbf{x})$ over n' dimensions. We have from Lemma A2.1 that $N \geq \sigma^{n'}$ Thus, it remains to show that for $n' := l \cdot 2^a + 1$ and $c \leq \sigma \cdot \sqrt{2} \cdot \sqrt{n'}$, $\widehat{f}(\mathbf{y}) \leq \sigma^{n'} \cdot n'^{5/4} \cdot e^{-\pi \|\mathbf{y}\|^2 \sigma^2}$.

Let $r := \|\mathbf{x}\|$, we slightly abuse notation and view f as a function of r, $f(r) := e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}}$. Since $\Psi_{\sigma,c}$ is a radial function, so is its Fourier transform, thus, we again slightly abuse notation and view $F := \widehat{f}$ as a function of $\kappa := \|\mathbf{y}\|$. We may now use the formula for the radial Fourier transform of an n'-dimensional, radial function f to find F [65]:

$$F(\kappa) = \kappa^{\frac{-(n'-2)}{2}} (2\pi) \int_0^\infty r^{\frac{n'-2}{2}} f(r) J_{\frac{n'-2}{2}}(2\pi\kappa r) r \, \mathrm{d}r, \tag{A2.3}$$

where $J_{\frac{n'-2}{2}}$ denotes the Bessel function of the first kind of order $\frac{n'-2}{2}$. The Bessel function of first kind of order ν is defined as [112, Page 40]:

$$J_\nu(z) := \sum_{j=0}^{\infty} \frac{(-1)^j (\frac{1}{2}z)^{\nu+2j}}{\Gamma(\nu + j + 1) j!}. \tag{A2.4}$$

For half-integer order $\nu := n + \frac{1}{2}$, there is a closed-form representation of J_ν. Specifically, it can be expressed as [112, Page 298]:

$$J_{n+\frac{1}{2}}(z) := R_{n,\frac{1}{2}}(z) \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} \sin z - R_{n-1,\frac{3}{2}}(z) \left(\frac{2}{\pi z}\right)^{\frac{1}{2}} \cos z. \tag{A2.5}$$

where $R_{n,\frac{1}{2}}(z)$ and $R_{n-1,\frac{3}{2}}(z)$ are Lommel polynomials defined as [112, Page

296]:

$$R_{n,\nu}(z) = \sum_{j=0}^{[n/2]} \frac{(-1)^j (n-j)! \Gamma(\nu+n-j)}{j!(n-2j)!\Gamma(\nu+j)} \left(\frac{z}{2}\right)^{2j-n}, \tag{A2.6}$$

where the $[x]$ means the largest integer not exceeding x.

We now have:

$$\begin{aligned}
|F(\kappa)| &= \left| \kappa^{\frac{-(n'-2)}{2}} (2\pi) \int_0^\infty r^{\frac{n'-2}{2}} f(r) J_{\frac{n'-2}{2}}(2\pi\kappa r) r \, \mathrm{d}r \right| \\
&= \left| \kappa^{\frac{-(n'-2)}{2}} (2\pi) \left(\int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \sin(2\pi\kappa r) r \, \mathrm{d}r - \right.\right. \\
&\left.\left. \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-5}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \cos(2\pi\kappa r) r \, \mathrm{d}r \right) \right| \\
&\leq \kappa^{\frac{-(n'-2)}{2}} (2\pi) \left(\left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \sin(2\pi\kappa r) r \, \mathrm{d}r \right| + \right. \\
&\left. \left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-5}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \cos(2\pi\kappa r) r \, \mathrm{d}r \right| \right),
\end{aligned} \tag{A2.7}$$

where the first equality follows from (A2.3), the second equality follows from (A2.5), (A2.6) and the settings of $c_j := \frac{(-1)^j (\frac{n'-3}{2}-j)!\Gamma(\frac{1}{2}+\frac{n'-3}{2}-j)}{j!(\frac{n'-3}{2}-2j)!\Gamma(\frac{1}{2}+j)}$ and $c'_j := \frac{(-1)^j (\frac{n'-5}{2}-j)!\Gamma(\frac{1}{2}+\frac{n'-3}{2}-j)}{j!(\frac{n'-5}{2}-2j)!\Gamma(\frac{1}{2}+1+j)}$.

In order to bound (A2.7), we will individually upper bound

$$\mathrm{I}: \left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \sin(2\pi\kappa r) r \, \mathrm{d}r \right|$$

and

$$\mathrm{II}: \left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-5}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \cos(2\pi\kappa r) r \, \mathrm{d}r \right|.$$

Recalling that $f(r) = e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}}$, we have that

$$
\begin{aligned}
\mathrm{II} &= \left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2} \right)^{2j-\frac{n'-5}{2}} \right) \left(\frac{2}{2\pi^2\kappa r} \right)^{\frac{1}{2}} \cos(2\pi\kappa r) r \, \mathrm{d}r \right| \\
&= 1/2 \left| \int_{-\infty}^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2} \right)^{2j-\frac{n'-5}{2}} \right) \left(\frac{2}{2\pi^2\kappa r} \right)^{\frac{1}{2}} \left(\frac{e^{i2\pi\kappa r} + e^{-i2\pi\kappa r}}{2} \right) r \, \mathrm{d}r \right| \\
&= 1/2 \left(\frac{1}{4\pi^2\kappa} \right)^{\frac{1}{2}} \left| \int_{-\infty}^\infty r^{\frac{n'-1}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-5}{4}]} c'_j \left(\frac{2\pi\kappa r}{2} \right)^{2j-\frac{n'}{2}+\frac{5}{2}} \right) (e^{i2\pi\kappa r} + e^{-i2\pi\kappa r}) \, \mathrm{d}r \right| \\
&\leq 1/2 \left(\frac{1}{4\pi^2\kappa} \right)^{\frac{1}{2}} \sum_{j=0}^{[\frac{n'-5}{4}]} |c'_j| \left(\pi\kappa \right)^{2j-\frac{n'}{2}+\frac{5}{2}} \left| \int_{-\infty}^\infty r^{2j+2} \left(e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}} \right) (e^{i2\pi\kappa r} + e^{-i2\pi\kappa r}) \, \mathrm{d}r \right|, \quad \text{(A2.8)}
\end{aligned}
$$

where the second equality follows since $f(r)$ is an even function, $\cos(2\pi\kappa r)$ is an even function and for $n' = l \cdot 2^a + 1$, all powers of r in the integrand are even, which means that the entire integrand is an even function.

To compute an upper bound on

$$
\left| \int_{-\infty}^\infty r^{2j+2} \left(e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}} \right) (e^{i2\pi\kappa r} + e^{-i2\pi\kappa r}) \, \mathrm{d}r \right| \quad \text{(A2.9)}
$$

as above, we integrate each term separately. Since the analysis is essentially the same for each term, we focus on upper bounding the term $A := \left| \int_{-\infty}^\infty e^{-\frac{\pi(r-c)^2}{\sigma^2}} e^{i2\pi\kappa r} \, \mathrm{d}r \right| = \left| e^{-\pi\kappa^2\sigma^2 + 2\pi i\kappa c} \int_{-\infty}^\infty e^{-\pi\sigma^{-2}(r-(c+i\kappa\sigma^2))^2} \, \mathrm{d}r \right|$:

$$
\begin{aligned}
A &= \left|e^{-\pi\kappa^2\sigma^2+2\pi i\kappa c}\right| \cdot \left|\int_{-\infty}^{\infty} r^{2j+2} e^{-\pi\sigma^{-2}(r-(c+i\kappa\sigma^2))^2}\,\mathrm{d}r\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty}\Big(\frac{\sigma}{\sqrt{\pi}}r' + (c+i\kappa\sigma^2)\Big)^{2j+2} e^{-r'^2}\frac{\sigma}{\sqrt{\pi}}\,\mathrm{d}r'\right| \\
&= e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty}\sigma^{2j+2}\Big(\frac{1}{\sqrt{\pi}}r' + (\frac{c}{\sigma}+i\kappa\sigma)\Big)^{2j+2} e^{-r'^2}\frac{\sigma}{\sqrt{\pi}}\,\mathrm{d}r'\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty}\sigma^{2j+2}\Big(\frac{1}{\sqrt{\pi}}r' + (\frac{c}{\sigma}+\kappa\sigma)\Big)^{2j+2} e^{-r'^2}\frac{\sigma}{\sqrt{\pi}}\,\mathrm{d}r'\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\Big(\frac{\sigma}{\sqrt{\pi}}\Big)^{2j+3}\Big(\frac{c}{\sigma}+\kappa\sigma\Big)^{2j+2}\binom{2j+2}{j+1}\int_{-\infty}^{\infty} r'^{2j+2}e^{-r'^2}\,\mathrm{d}r \\
&\leq e^{-\pi\kappa^2\sigma^2}\Big(\frac{\sigma}{\sqrt{\pi}}\Big)^{2j+3}\Big(\frac{c}{\sigma}+\kappa\sigma\Big)^{2j+2}\binom{2j+2}{j+1}\frac{1}{2}(1+(-1)^{2j})\Gamma\Big(\frac{3}{2}+j\Big) \\
&\leq e^{-\pi\kappa^2\sigma^2}\Big(\frac{\sigma}{\sqrt{\pi}}\Big)^{2j+3}\Big(\frac{c}{\sigma}+\kappa\sigma\Big)^{2j+2}\binom{2j+2}{j+1}\Gamma\Big(\frac{3}{2}+j\Big)
\end{aligned}
$$

Thus, we have that

$$
(A2.9) \leq \Big(\frac{\sigma}{\sqrt{\pi}}\Big)^{2j+3} e^{-\pi\kappa^2\sigma^2}\Gamma(\frac{3}{2}+j)\binom{2j+2}{j+1}\left[4(\frac{c}{\sigma}+\kappa\sigma)^{2j+2}\right]
$$

Plugging the above back into (A2.8), and recalling that $|c'_j| = \frac{(\frac{n'-5}{2}-j)!\Gamma(\frac{1}{2}+\frac{n'-3}{2}-j)}{j!(\frac{n'-5}{2}-2j)!\Gamma(\frac{1}{2}+1+j)}$, we have that

$$
\begin{aligned}
\mathrm{II} &\leq 1/2\Big(\frac{1}{4\pi^2\kappa}\Big)^{\frac{1}{2}}\sum_{j=0}^{[\frac{n'-5}{4}]}|c'_j|\Big(\pi\kappa\Big)^{2j-\frac{n'}{2}+\frac{5}{2}}\Big(\frac{\sigma}{\sqrt{\pi}}\Big)^{2j+3}e^{-\pi\kappa^2\sigma^2}\Gamma(\frac{3}{2}+j)\Big(\frac{2j+2}{j+1}\Big)^2\Big(\frac{c}{\sigma}\Big)^{2j+2}\Big(\kappa\sigma\Big)^{2j+2} \\
&\leq 1/2\Big(\frac{1}{2\pi}\Big)e^{-\pi\kappa^2\sigma^2}\sum_{j=0}^{[\frac{n'-5}{4}]}(\pi)^{j-\frac{n'}{2}+1}\Big(\frac{\frac{n'-5}{2}-j}{j}\Big)\Big(\frac{2j+2}{j+1}\Big)^2\Gamma\Big(\frac{n'}{2}-1-j\Big)\,\sigma^{2j+3}\,c^{2j+2}\,\Big(\kappa\Big)^{4j-\frac{n'}{2}+4} \\
&\leq 1/2\Big(\frac{1}{2\pi}\Big)e^{-\pi\kappa^2\sigma^2}\,\Big(n'\cdot 2^{\frac{n'}{2}}\cdot n'^{\frac{n'}{2}}\Big)\sum_{j=0}^{[\frac{n'-5}{4}]}\sigma^{2j+3}\,c^{2j+2}\,\Big(\kappa\Big)^{4j-\frac{n'}{2}+4}
\end{aligned}
$$

Where the last inequality follows since $\binom{n}{i} \leq 2^n$ and $n! \leq n^n$. We now turn to

upper-bounding I. Recalling that $f(r) = e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}}$, we have that

$$\begin{aligned}
\mathrm{I} &= \left| \int_0^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \sin(2\pi\kappa r) r \, \mathrm{d}r \right| \\
&= 1/2 \left| \int_{-\infty}^\infty r^{\frac{n'-2}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) \left(\frac{2}{2\pi^2\kappa r}\right)^{\frac{1}{2}} \left(\frac{e^{i2\pi\kappa r} - e^{-i2\pi\kappa r}}{2i}\right) r \, \mathrm{d}r \right| \\
&\leq 1/2 \cdot \left(\frac{1}{4\pi^2\kappa}\right)^{\frac{1}{2}} \left| \int_{-\infty}^\infty r^{\frac{n'-1}{2}} f(r) \left(\sum_{j=0}^{[\frac{n'-3}{4}]} c_j \left(\frac{2\pi\kappa r}{2}\right)^{2j-\frac{n'-3}{2}} \right) (e^{i2\pi\kappa r} - e^{-i2\pi\kappa r}) \, \mathrm{d}r \right| \\
&\leq 1/2 \cdot \left(\frac{1}{4\pi^2\kappa}\right)^{\frac{1}{2}} \sum_{j=0}^{[\frac{n'-3}{4}]} |c_j| \left(\pi\kappa\right)^{2j-\frac{n'}{2}+\frac{3}{2}} \left| \int_{-\infty}^\infty r^{2j+1} \left(e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}} \right) (e^{i2\pi\kappa r} - e^{-i2\pi\kappa r}) \, \mathrm{d}r \right|,
\end{aligned} \tag{A2.10}$$

where the second equality follows since $f(r)$ is an even function, $\sin(2\pi\kappa r)$ is an odd function and for $n' = l \cdot 2^a + 1$, all powers of r in the integrand are odd, which means that the entire integrand is an even function.

To compute an upper bound on

$$\int_{-\infty}^\infty r^{2j+1} \left(e^{-\frac{\pi(r-c)^2}{\sigma^2}} + e^{-\frac{\pi(r+c)^2}{\sigma^2}} \right) (e^{i2\pi\kappa r} - e^{-i2\pi\kappa r}) \, \mathrm{d}r \tag{A2.11}$$

as above, we integrate each term separately. Since the analysis is essentially the same for each term, we focus on the term $B := \left| \int_{-\infty}^\infty e^{-\frac{\pi(r-c)^2}{\sigma^2}} e^{i2\pi\kappa r} \, \mathrm{d}r \right| = \left| e^{-\pi\kappa^2\sigma^2 + i2\pi\kappa c} \int_{-\infty}^\infty e^{-\pi\sigma^{-2}(r-(c+i\kappa\sigma^2))^2} \, \mathrm{d}r \right|$:

$$
\begin{aligned}
B &= \left|e^{-\pi\kappa^2\sigma^2+i2\pi\kappa c}\right| \cdot \left|\int_{-\infty}^{\infty} r^{2j+1} e^{-\pi\sigma^{-2}(r-(c+i\kappa\sigma^2))^2}\,\mathrm{d}r\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty} r^{2j+1} e^{-\pi\sigma^{-2}(r-(c+i\kappa\sigma^2))^2}\,\mathrm{d}r\right| \\
&= e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty} \left(\frac{\sigma}{\sqrt{\pi}}r' + (c+i\kappa\sigma^2)\right)^{2j+1} e^{-r'^2}\frac{\sigma}{\sqrt{\pi}}\,\mathrm{d}r'\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\left|\int_{-\infty}^{\infty} \left(\frac{\sigma}{\sqrt{\pi}}r' + (c+\kappa\sigma^2)\right)^{2j+1} e^{-r'^2}\frac{\sigma}{\sqrt{\pi}}\,\mathrm{d}r'\right| \\
&\leq e^{-\pi\kappa^2\sigma^2}\left(\frac{\sigma}{\sqrt{\pi}}\right)^{2j+2}\left(\frac{c}{\sigma}+\kappa\sigma\right)^{2j+1}\binom{2j+1}{j+1}\int_{-\infty}^{\infty} r'^{2j} e^{-r'^2}\,\mathrm{d}r \\
&\leq e^{-\pi\kappa^2\sigma^2}\left(\frac{\sigma}{\sqrt{\pi}}\right)^{2j+2}\left(\frac{c}{\sigma}+\kappa\sigma\right)^{2j+1}\binom{2j+1}{j+1}\frac{1}{2}(1+(-1)^{2j})\Gamma\left(\frac{1}{2}+j\right) \\
&\leq e^{-\pi\kappa^2\sigma^2}\left(\frac{\sigma}{\sqrt{\pi}}\right)^{2j+2}\left(\frac{c}{\sigma}+\kappa\sigma\right)^{2j+1}\binom{2j+1}{j+1}\Gamma\left(\frac{1}{2}+j\right)
\end{aligned}
$$

Thus, we have that

$$
(A2.11) \leq \left(\frac{\sigma}{\sqrt{\pi}}\right)^{2j+2} e^{-\pi\kappa^2\sigma^2}\Gamma(\frac{1}{2}+j)\binom{2j+1}{j+1}\left[4(\frac{c}{\sigma}+\kappa\sigma)^{2j+1}\right]
$$

Plugging the above back into (A2.10), and recalling that $|c_j| = \frac{(\frac{n'-3}{2}-j)!\Gamma(\frac{1}{2}+\frac{n'-3}{2}-j)}{j!(\frac{n'-3}{2}-2j)!\Gamma(\frac{1}{2}+j)}$, we have that

$$
\begin{aligned}
\mathrm{I} &\leq 1/2\left(\frac{1}{4\pi^2\kappa}\right)^{\frac{1}{2}}\sum_{j=0}^{[\frac{n'-3}{4}]}|c_j|\left(\pi\kappa\right)^{2j-\frac{n'}{2}+\frac{3}{2}}\left(\frac{\sigma}{\sqrt{\pi}}\right)^{2j+2} e^{-\pi\kappa^2\sigma^2}\Gamma(\tfrac{1}{2}+j)\binom{2j+1}{j+1}^2\left(\frac{c}{\sigma}\right)^{2j+1}\left(\kappa\sigma\right)^{2j+1} \\
&\leq 1/2\left(\frac{1}{2\pi}\right)e^{-\pi\kappa^2\sigma^2}\sum_{j=0}^{[\frac{n'-3}{4}]}(\pi)^{j-\frac{n'-1}{2}}\binom{\frac{n'-3}{2}-j}{j}\binom{2j+1}{j+1}^2\Gamma\left(\frac{n'}{2}-1-j\right)\sigma^{2j+2}\,c^{2j+1}\left(\kappa\right)^{4j-\frac{n'}{2}+3} \\
&\leq 1/2\left(\frac{1}{2\pi}\right)e^{-\pi\kappa^2\sigma^2}\left(n'\cdot 2^{\frac{n'}{2}}\cdot n'^{\frac{n'}{2}}\right)\sum_{j=0}^{[\frac{n'-3}{4}]}\sigma^{2j+2}\,c^{2j+1}\left(\kappa\right)^{4j-\frac{n'}{2}+3}
\end{aligned}
$$

Where the last inequality follows since $\binom{n}{i} \leq 2^n$ and $n! \leq n^n$. Finally, plugging

into (A2.7), and recalling that $c \leq \sigma \cdot \sqrt{2} \cdot \sqrt{n'}$ and $\kappa > \frac{1}{\sigma}$, we obtain:

$$|F(\kappa)| \leq 1/2\ e^{-\pi\kappa^2\sigma^2}\ \left(n' \cdot 2^{\frac{n'}{2}} \cdot {n'}^{\frac{n'}{2}}\right)\left(\sum_{j=0}^{[\frac{n'-5}{4}]}\ \sigma^{2j+3}\ c^{2j+2}\ \kappa^{4j-n'+5} + \sum_{j=0}^{[\frac{n'-3}{4}]}\ \sigma^{2j+2}\ c^{2j+1}\ \kappa^{4j-n'+4}\right)$$

$$\leq \sigma^{n'} \cdot {n'}^{n'} \cdot e^{-\pi\kappa^2\sigma^2}$$

□

Bibliography

[1] Michel Abdalla, Emmanuel Bresson, Olivier Chevassut, and David Pointcheval. Password-based group key exchange in a constant number of rounds. In Moti Yung, Yevgeniy Dodis, Aggelos Kiayias, and Tal Malkin, editors, *PKC 2006: 9th International Conference on Theory and Practice of Public Key Cryptography*, volume 3958 of *Lecture Notes in Computer Science*, pages 427–442, New York, NY, USA, April 24–26, 2006. Springer, Heidelberg, Germany.

[2] Michel Abdalla and David Pointcheval. A scalable password-based group key exchange protocol in the standard model. In Xuejia Lai and Kefei Chen, editors, *Advances in Cryptology – ASIACRYPT 2006*, volume 4284 of *Lecture Notes in Computer Science*, pages 332–347, Shanghai, China, December 3–7, 2006. Springer, Heidelberg, Germany.

[3] Miklós Ajtai. Generating hard instances of lattice problems. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 99–108. ACM, 1996.

[4] Adi Akavia, Shafi Goldwasser, and Vinod Vaikuntanathan. Simultaneous hardcore bits and cryptography against memory attacks. In Omer Reingold, editor, *TCC 2009: 6th Theory of Cryptography Conference*, volume 5444 of *Lecture Notes in Computer Science*, pages 474–495. Springer, Heidelberg, Germany, March 15–17, 2009.

[5] Martin R. Albrecht, Benjamin R. Curtis, Amit Deo, Alex Davidson, Rachel Player, Eamonn W Postlethwaite, Fernando Virdia, and Thomas Wunderer. Estimate all the LWE, NTRU schemes! In *International Conference on Security and Cryptography for Networks*, pages 351–367. Springer, 2018.

[6] Martin R. Albrecht, Amit Deo, and Kenneth G. Paterson. Cold boot attacks on ring and module LWE keys under the NTT. In *IACR Transactions on Cryptographic Hardware and Embedded Systems*, volume 2018, pages 173–213, Aug. 2018.

[7] Martin R. Albrecht, Florian Göpfert, Fernando Virdia, and Thomas Wunderer. Revisiting the expected cost of solving uSVP and applications to LWE. In

Tsuyoshi Takagi and Thomas Peyrin, editors, *Advances in Cryptology – ASIACRYPT 2017, Part I*, volume 10624 of *Lecture Notes in Computer Science*, pages 297–322, Hong Kong, China, December 3–7, 2017. Springer, Heidelberg, Germany.

[8] Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. NewHope without reconciliation. Cryptology ePrint Archive, Report 2016/1157, 2016. `http://eprint.iacr.org/2016/1157`.

[9] Erdem Alkim, Léo Ducas, Thomas Pöppelmann, and Peter Schwabe. Post-quantum key exchange - A new hope. In Thorsten Holz and Stefan Savage, editors, *USENIX Security 2016: 25th USENIX Security Symposium*, pages 327–343, Austin, TX, USA, August 10–12, 2016. USENIX Association.

[10] Daniel Apon, Dana Dachman-Soled, Huijing Gong, and Jonathan Katz. Constant-round group key exchange from the ring-lwe assumption. In Jintai Ding and Rainer Steinwandt, editors, *Post-Quantum Cryptography*, pages 189–205, Cham, 2019. Springer International Publishing.

[11] Shi Bai, Damien Stehlé, and Weiqiang Wen. Measuring, simulating and exploiting the head concavity phenomenon in BKZ. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 369–404. Springer, 2018.

[12] Abhishek Banerjee, Chris Peikert, and Alon Rosen. Pseudorandom functions and lattices. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 719–737. Springer, 2012.

[13] Klaus Becker and Uta Wille. Communication complexity of group key distribution. In *Proceedings of the 5th ACM Conference on Computer and Communications Security*, CCS '98, pages 1–6, New York, NY, USA, 1998.

[14] Mihir Bellare and Phillip Rogaway. Provably secure session key distribution: The three party case. In *27th Annual ACM Symposium on Theory of Computing*, pages 57–66, Las Vegas, NV, USA, May 29 – June 1, 1995. ACM Press.

[15] Daniel J. Bernstein, Chitchanok Chuengsatiansup, Tanja Lange, and Christine van Vredendaal. NTRU Prime. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.

[16] Pauline Bert, Pierre-Alain Fouque, Adeline Roux-Langlois, and Mohamed Sabt. Practical implementation of ring-SIS/LWE based signature and IBE. In Tanja Lange and Rainer Steinwandt, editors, *Post-Quantum Cryptography - 9th International Conference, PQCrypto 2018*, pages 271–291, Fort Lauderdale, Florida, United States, April 9–11 2018. Springer, Heidelberg, Germany.

[17] Andrej Bogdanov, Siyao Guo, Daniel Masny, Silas Richelson, and Alon Rosen. On the hardness of learning with rounding over small modulus. In *Theory of Cryptography Conference*, pages 209–224. Springer, 2016.

[18] Jens-Matthias Bohli, María Isabel González Vasco, and Rainer Steinwandt. Password-authenticated constant-round group key establishment with a common reference string. Cryptology ePrint Archive, Report 2006/214, 2006. `http://eprint.iacr.org/2006/214`.

[19] Jens-Matthias Bohli, María Isabel González Vasco, and Rainer Steinwandt. Secure group key establishment revisited. *International Journal of Information Security*, 6(4):243–254, Jul 2007.

[20] Madalina Bolboceanu, Zvika Brakerski, Renen Perlman, and Devika Sharma. Order-LWE and the hardness of ring-LWE with entropic secrets. In Steven D. Galbraith and Shiho Moriai, editors, *Advances in Cryptology – ASIACRYPT 2019, Part II*, volume 11922 of *Lecture Notes in Computer Science*, pages 91–120, Kobe, Japan, December 8–12, 2019. Springer, Heidelberg, Germany.

[21] Dan Boneh, Craig Gentry, Sergey Gorbunov, Shai Halevi, Valeria Nikolaenko, Gil Segev, Vinod Vaikuntanathan, and Dhinakaran Vinayagamurthy. Fully key-homomorphic encryption, arithmetic circuit abe and compact garbled circuits. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 533–556. Springer, 2014.

[22] Dan Boneh, Darren Glass, Daniel Krashen, Kristin Lauter, Shahed Sharif, Alice Silverberg, Mehdi Tibouchi, and Mark Zhandry. Multiparty non-interactive key exchange and more from isogenies on elliptic curves. *Journal of Mathematical Cryptology*, 14(1):5–14, 2020.

[23] Jonathan Bootle, Claire Delaplace, Thomas Espitau, Pierre-Alain Fouque, and Mehdi Tibouchi. LWE without modular reduction and improved side-channel attacks against BLISS. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 494–524. Springer, 2018.

[24] Joppe W. Bos, Craig Costello, Léo Ducas, Ilya Mironov, Michael Naehrig, Valeria Nikolaenko, Ananth Raghunathan, and Douglas Stebila. Frodo: Take off the ring! Practical, quantum-secure key exchange from LWE. In Edgar R. Weippl, Stefan Katzenbeisser, Christopher Kruegel, Andrew C. Myers, and Shai Halevi, editors, *ACM CCS 2016: 23rd Conference on Computer and Communications Security*, pages 1006–1018, Vienna, Austria, October 24–28, 2016. ACM Press.

[25] Joppe W. Bos, Simon Friedberger, Marco Martinoli, Elisabeth Oswald, and Martijn Stam. Assessing the feasibility of single trace power analysis of frodo. In *SAC*, 2018.

[26] Elette Boyle, Gil Segev, and Daniel Wichs. Fully leakage-resilient signatures. *Journal of Cryptology*, 26(3):513–558, July 2013.

[27] Zvika Brakerski, Yael Tauman Kalai, Jonathan Katz, and Vinod Vaikuntanathan. Overcoming the hole in the bucket: Public-key cryptography resilient to continual memory leakage. In *51st Annual Symposium on Foundations of Computer Science*, pages 501–510, Las Vegas, NV, USA, October 23–26, 2010. IEEE Computer Society Press.

[28] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *45th Annual ACM Symposium on Theory of Computing*, pages 575–584, Palo Alto, CA, USA, June 1–4, 2013. ACM Press.

[29] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) lwe. *SIAM Journal on Computing*, 43(2):831–871, 2014.

[30] Emmanuel Bresson and Dario Catalano. Constant round authenticated group key agreement via distributed computation. In Feng Bao, Robert Deng, and Jianying Zhou, editors, *PKC 2004: 7th International Workshop on Theory and Practice in Public Key Cryptography*, volume 2947 of *Lecture Notes in Computer Science*, pages 115–129, Singapore, March 1–4, 2004. Springer, Heidelberg, Germany.

[31] Emmanuel Bresson, Olivier Chevassut, and David Pointcheval. Provably authenticated group Diffie-Hellman key exchange – the dynamic case. In Colin Boyd, editor, *Advances in Cryptology – ASIACRYPT 2001*, volume 2248 of *Lecture Notes in Computer Science*, pages 290–309, Gold Coast, Australia, December 9–13, 2001. Springer, Heidelberg, Germany.

[32] Emmanuel Bresson, Olivier Chevassut, and David Pointcheval. Dynamic group Diffie-Hellman key exchange under standard assumptions. In Lars R. Knudsen, editor, *Advances in Cryptology – EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 321–336, Amsterdam, The Netherlands, April 28 – May 2, 2002. Springer, Heidelberg, Germany.

[33] Emmanuel Bresson, Olivier Chevassut, David Pointcheval, and Jean-Jacques Quisquater. Provably authenticated group Diffie-Hellman key exchange. In Michael K. Reiter and Pierangela Samarati, editors, *ACM CCS 2001: 8th Conference on Computer and Communications Security*, pages 255–264, Philadelphia, PA, USA, November 5–8, 2001. ACM Press.

[34] Leon Groot Bruinderink, Andreas Hülsing, Tanja Lange, and Yuval Yarom. Flush, gauss, and reload–a cache attack on the bliss lattice-based signature scheme. In *International Conference on Cryptographic Hardware and Embedded Systems*, pages 323–345. Springer, 2016.

[35] Mike Burmester and Yvo Desmedt. A secure and efficient conference key distribution system (extended abstract). In Alfredo De Santis, editor, *Advances in Cryptology – EUROCRYPT'94*, volume 950 of *Lecture Notes in Computer Science*, pages 275–286, Perugia, Italy, May 9–12, 1995. Springer, Heidelberg, Germany.

[36] Mike Burmester and Yvo Desmedt. A secure and scalable group key exchange system. *Information Processing Letters*, 94(3):137–143, May 2005.

[37] Suresh Chari, Josyula R. Rao, and Pankaj Rohatgi. Template attacks. In *Revised Papers from the 4th International Workshop on Cryptographic Hardware and Embedded Systems*, CHES '02, page 13–28, Berlin, Heidelberg, 2002. Springer-Verlag.

[38] Yuanmi Chen and Phong Q. Nguyen. BKZ 2.0: Better lattice security estimates. In Dong Hoon Lee and Xiaoyun Wang, editors, *Advances in Cryptology – ASIACRYPT 2011*, volume 7073 of *Lecture Notes in Computer Science*, pages 1–20, Seoul, South Korea, December 4–8, 2011. Springer, Heidelberg, Germany.

[39] Jung Hee Cheon, Duhyeong Kim, Joohee Lee, and Yongsoo Song. Lizard: Cut off the tail! A practical post-quantum public-key encryption from LWE and LWR. In *International Conference on Security and Cryptography for Networks*, pages 160–177. Springer, 2018.

[40] Dong Pyo Chi, Jeong Woon Choi, Jeong San Kim, and Taewan Kim. Lattice based cryptography for beginners. Cryptology ePrint Archive, Report 2015/938, 2015. `https://eprint.iacr.org/2015/938`.

[41] Kyu Young Choi, Jung Yeon Hwang, and Dong Hoon Lee. Efficient ID-based group key agreement with bilinear maps. In Feng Bao, Robert Deng, and Jianying Zhou, editors, *PKC 2004: 7th International Workshop on Theory and Practice in Public Key Cryptography*, volume 2947 of *Lecture Notes in Computer Science*, pages 130–144, Singapore, March 1–4, 2004. Springer, Heidelberg, Germany.

[42] Kai-Min Chung, Daniel Dadush, Feng-Hao Liu, and Chris Peikert. On the lattice smoothing parameter problem. In *Computational Complexity (CCC), 2013 IEEE Conference on*, pages 230–241. IEEE, 2013.

[43] Don Coppersmith and Adi Shamir. Lattice attacks on ntru. In *International Conference on the Theory and Applications of Cryptographic Techniques*, pages 52–61. Springer, 1997.

[44] Dana Dachman-Soled, Léo Ducas, Huijing Gong, and Mélissa Rossi. Lwe with side information: Attacks and concrete security estimation. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology – CRYPTO 2020*, pages 329–358, Cham, 2020. Springer International Publishing.

[45] Dana Dachman-Soled, Huijing Gong, Mukul Kulkarni, and Aria Shahverdi. On the leakage resilience of ring-lwe based public key encryption. Cryptology ePrint Archive, Report 2017/1127, 2017. https://eprint.iacr.org/2017/1127.

[46] Dana Dachman-Soled, Huijing Gong, Mukul Kulkarni, and Aria Shahverdi. (in)security of ring-lwe under partial key exposure. *Journal of Mathematical Cryptology*, 15(1):72–86, 2021.

[47] Dana Dachman-Soled, Huijing Gong, Mukul Kulkarni, and Aria Shahverdi. Towards a ring analogue of the leftover hash lemma. *Journal of Mathematical Cryptology*, 15(1):87–110, 2021.

[48] Jan-Pieter D'Anvers, Mélissa Rossi, and Fernando Virdia. (One) failure is not an option: Bootstrapping the search for failures in lattice-based encryption schemes. Cryptology ePrint Archive, Report 2019/1399, 2019.

[49] Jan-Pieter D'Anvers, Frederik Vercauteren, and Ingrid Verbauwhede. On the impact of decryption failures on the security of LWE/LWR based schemes. *IACR Cryptology ePrint Archive*, 2018:1089, 2018.

[50] The FPLLL development team. fplll, a lattice reduction library. Available at https://github.com/fplll/fplll, 2016.

[51] Luc Devroye. Sample-based non-uniform random variate generation. In *Proceedings of the 18th conference on Winter simulation*, pages 260–265. ACM, 1986.

[52] Jintai Ding, Xiang Xie, and Xiaodong Lin. A simple provably secure key exchange scheme based on the learning with errors problem. Cryptology ePrint Archive, Report 2012/688, 2012. http://eprint.iacr.org/2012/688.

[53] Yevgeniy Dodis, Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Public-key encryption schemes with auxiliary inputs. In Daniele Micciancio, editor, *TCC 2010: 7th Theory of Cryptography Conference*, volume 5978 of *Lecture Notes in Computer Science*, pages 361–381, Zurich, Switzerland, February 9–11, 2010. Springer, Heidelberg, Germany.

[54] Yevgeniy Dodis, Kristiyan Haralambiev, Adriana López-Alt, and Daniel Wichs. Cryptography against continuous memory attacks. In *51st Annual Symposium on Foundations of Computer Science*, pages 511–520, Las Vegas, NV, USA, October 23–26, 2010. IEEE Computer Society Press.

[55] Yevgeniy Dodis, Yael Tauman Kalai, and Shachar Lovett. On cryptography with auxiliary input. In Michael Mitzenmacher, editor, *41st Annual ACM Symposium on Theory of Computing*, pages 621–630, Bethesda, MD, USA, May 31 – June 2, 2009. ACM Press.

[56] Léo Ducas, Alain Durmus, Tancrède Lepoint, and Vadim Lyubashevsky. Lattice signatures and bimodal gaussians. In *Annual Cryptology Conference*, pages 40–56. Springer, 2013.

[57] Stefan Dziembowski and Krzysztof Pietrzak. Leakage-resilient cryptography. In *49th Annual Symposium on Foundations of Computer Science*, pages 293–302, Philadelphia, PA, USA, October 25–28, 2008. IEEE Computer Society Press.

[58] Jan-Pieter D'Anvers, Qian Guo, Thomas Johansson, Alexander Nilsson, Frederik Vercauteren, and Ingrid Verbauwhede. Decryption failure attacks on IND-CCA secure lattice-based schemes. In *IACR International Workshop on Public Key Cryptography*, pages 565–598. Springer, 2019.

[59] Wolfgang Ebeling. Lattices and codes. In *Lattices and Codes*, pages 1–32. Springer, 2013.

[60] Thomas Espitau, Pierre-Alain Fouque, Benoit Gerard, and Mehdi Tibouchi. Side-channel attacks on BLISS lattice-based signatures – exploiting branch tracing against strongSwan and electromagnetic emanations in microcontrollers. Cryptology ePrint Archive, Report 2017/505, 2017. `http://eprint.iacr.org/2017/505`.

[61] Nicolas Gama, Phong Q Nguyen, and Oded Regev. Lattice enumeration using extreme pruning. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 257–278. Springer, 2010.

[62] Oscar Garcia-Morchon, Zhenfei Zhang, Sauvik Bhattacharya, Ronald Rietman, Ludo Tolhuizen, Jose-Luis Torre-Arce, Hayo Baan, Markku-Juhani O. Saarinen, Scott Fluhrer, Thijs Laarhoven, and Rachel Player. Round5. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.

[63] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In Richard E. Ladner and Cynthia Dwork, editors, *40th Annual ACM Symposium on Theory of Computing*, pages 197–206, Victoria, BC, Canada, May 17–20, 2008. ACM Press.

[64] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In Andrew Chi-Chih Yao, editor, *ICS 2010: 1st Innovations in Computer Science*, pages 230–240, Tsinghua University, Beijing, China, January 5–7, 2010. Tsinghua University Press.

[65] Loukas Grafakos and Gerald Teschl. On fourier transforms of radial functions and distributions. *Journal of Fourier Analysis and Applications*, 19(1):167–179, Feb 2013.

[66] Leon Groot Bruinderink and Peter Pessl. Differential fault attacks on deterministic lattice signatures. In *IACR Transactions on Cryptographic Hardware and Embedded Systems*, volume 2018, pages 21–43, Aug. 2018.

[67] Qian Guo, Thomas Johansson, and Alexander Nilsson. A generic attack on lattice-based schemes using decryption errors. Cryptology ePrint Archive, Report 2019/043, 2019.

[68] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.

[69] Jeff Hoffstein, Nick Howgrave-Graham, Jill Pipher, and William Whyte. Practical lattice-based cryptography: NTRUEncrypt and NTRUSign. In *The LLL Algorithm*, pages 349–390. Springer, 2009.

[70] Nick Howgrave-Graham. A hybrid lattice-reduction and meet-in-the-middle attack against NTRU. In Alfred Menezes, editor, *Advances in Cryptology – CRYPTO 2007*, volume 4622 of *Lecture Notes in Computer Science*, pages 150–169, Santa Barbara, CA, USA, August 19–23, 2007. Springer, Heidelberg, Germany.

[71] I. Ingemarsson, D. Tang, and C. Wong. A conference key distribution system. *IEEE Trans. Inf. Theor.*, 28(5):714–720, September 1982.

[72] Ravi Kannan. Minkowski's convex body theorem and integer programming. In *Mathematics of operations research*, volume 12, pages 415–440. INFORMS, 1987.

[73] Jonathan Katz and Ji Sun Shin. Modeling insider attacks on group key-exchange protocols. In *Proceedings of the 12th ACM Conference on Computer and Communications Security*, CCS '05, pages 180–189, New York, NY, USA, 2005. ACM.

[74] Jonathan Katz and Vinod Vaikuntanathan. Signature schemes with bounded leakage resilience. In Mitsuru Matsui, editor, *Advances in Cryptology – ASIACRYPT 2009*, volume 5912 of *Lecture Notes in Computer Science*, pages 703–720, Tokyo, Japan, December 6–10, 2009. Springer, Heidelberg, Germany.

[75] Jonathan Katz and Moti Yung. Scalable protocols for authenticated group key exchange. In Dan Boneh, editor, *Advances in Cryptology – CRYPTO 2003*, volume 2729 of *Lecture Notes in Computer Science*, pages 110–125, Santa Barbara, CA, USA, August 17–21, 2003. Springer, Heidelberg, Germany.

[76] Jonathan Katz and Moti Yung. Scalable protocols for authenticated group key exchange. *Journal of Cryptology*, 20(1):85–113, January 2007.

[77] Leonid Khachiyan. On the complexity of approximating extremal determinants in matrices. volume 11, pages 138–153. Elsevier, 1995.

[78] Yongdae Kim, Adrian Perrig, and Gene Tsudik. Simple and fault-tolerant key agreement for dynamic collaborative groups. In *Proceedings of the 7th ACM Conference on Computer and Communications Security*, CCS '00, pages 235–244, New York, NY, USA, 2000.

[79] Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In Phong Q. Nguyen and Elisabeth Oswald, editors, *Advances in Cryptology – EUROCRYPT 2014*, volume 8441 of *Lecture Notes in Computer Science*, pages 239–256, Copenhagen, Denmark, May 11–15, 2014. Springer, Heidelberg, Germany.

[80] Allison B. Lewko, Mark Lewko, and Brent Waters. How to leak on key updates. In Lance Fortnow and Salil P. Vadhan, editors, *43rd Annual ACM Symposium on Theory of Computing*, pages 725–734, San Jose, CA, USA, June 6–8, 2011. ACM Press.

[81] Richard Lindner and Chris Peikert. Better key sizes (and attacks) for LWE-based encryption. In Aggelos Kiayias, editor, *Topics in Cryptology – CT-RSA 2011*, pages 319–339. Springer Berlin Heidelberg, 2011.

[82] Li-Ping Liu. Linear transformation of multivariate normal distribution: Marginal, joint and posterior, Accessed on September 2019. `http://www.cs.columbia.edu/~liulp/pdf/linear_normal_dist.pdf`.

[83] Xianhui Lu, Yamin Liu, Dingding Jia, Haiyang Xue, Jingnan He, Zhenfei Zhang, Zhe Liu, Hao Yang, Bao Li, and Kunpeng Wang. LAC. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.

[84] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In Henri Gilbert, editor, *Advances in Cryptology – EUROCRYPT 2010*, volume 6110 of *Lecture Notes in Computer Science*, pages 1–23, French Riviera, May 30 – June 3, 2010. Springer, Heidelberg, Germany.

[85] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. A toolkit for ring-LWE cryptography. In Thomas Johansson and Phong Q. Nguyen, editors, *Advances in Cryptology – EUROCRYPT 2013*, volume 7881 of *Lecture Notes in Computer Science*, pages 35–54, Athens, Greece, May 26–30, 2013. Springer, Heidelberg, Germany.

[86] Tal Malkin, Isamu Teranishi, Yevgeniy Vahlis, and Moti Yung. Signatures resilient to continual leakage on memory and computation. In Yuval Ishai, editor, *TCC 2011: 8th Theory of Cryptography Conference*, volume 6597 of *Lecture Notes in Computer Science*, pages 89–106, Providence, RI, USA, March 28–30, 2011. Springer, Heidelberg, Germany.

[87] Jacques Martinet. *Perfect lattices in Euclidean spaces*, volume 327. Springer Science & Business Media, 2013.

[88] Alexander May and Joseph H Silverman. Dimension reduction methods for convolution modular lattices. In *International Cryptography and Lattices Conference*, pages 110–125. Springer, 2001.

[89] David McCann, Elisabeth Oswald, and Carolyn Whitnall. Towards practical tools for side channel aware software engineering: 'grey box' modelling for instruction leakages. In *26th USENIX Security Symposium (USENIX Security 17)*, pages 199–216, Vancouver, BC, August 2017. USENIX Association.

[90] Daniele Micciancio. Generalized compact knapsacks, cyclic lattices, and efficient one-way functions. *Computational Complexity*, 16(4):365–411, 2007.

[91] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM Journal on Computing*, 37(1):267–302, 2007.

[92] Michael Naehrig, Erdem Alkim, Joppe Bos, Léo Ducas, Karen Easterbrook, Brian LaMacchia, Patrick Longa, Ilya Mironov, Valeria Nikolaenko, Christopher Peikert, Ananth Raghunathan, and Douglas Stebila. FrodoKEM. Technical report, National Institute of Standards and Technology, 2017. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-1-submissions`.

[93] Moni Naor and Gil Segev. Public-key cryptosystems resilient to key leakage. *SIAM J. Comput.*, 41(4):772–814, 2012.

[94] Phong Nguyen. Giophanthus and *LWR-based submissions, 2019. Comment on the NIST PQC forum, `https://groups.google.com/a/list.nist.gov/d/msg/pqc-forum/nZBIBvYmmUI/J0pug16CBgAJ`.

[95] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 333–342, 2009.

[96] Chris Peikert. Lattice cryptography for the internet. Cryptology ePrint Archive, Report 2014/070, 2014. `http://eprint.iacr.org/2014/070`.

[97] Chris Peikert. How (not) to instantiate ring-lwe. In *International Conference on Security and Cryptography for Networks*, pages 411–430. Springer, 2016.

[98] Chris Peikert and Sina Shiehian. Noninteractive zero knowledge for np from (plain) learning with errors. In *Annual International Cryptology Conference*, pages 89–114. Springer, 2019.

[99] Krzysztof Pietrzak. A leakage-resilient mode of operation. In Antoine Joux, editor, *Advances in Cryptology – EUROCRYPT 2009*, volume 5479 of *Lecture Notes in Computer Science*, pages 462–482, Cologne, Germany, April 26–30, 2009. Springer, Heidelberg, Germany.

[100] Thomas Poppelmann, Erdem Alkim, Roberto Avanzi, Joppe Bos, Léo Ducas, Antonio de la Piedra, Peter Schwabe, Douglas Stebila, Martin R. Albrecht, Emmanuela Orsini, Valery Osheter, Kenneth G. Paterson, Guy Peer, and Nigel P. Smart. NewHope. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.

[101] Prasanna Ravi, Mahabir Prasad Jhanwar, James Howe, Anupam Chattopadhyay, and Shivam Bhasin. Side-channel assisted existential forgery attack on Dilithium - A NIST PQC candidate. Cryptology ePrint Archive, Report 2018/821, 2018.

[102] Prasanna Ravi, Mahabir Prasad Jhanwar, James Howe, Anupam Chattopadhyay, and Shivam Bhasin. Exploiting determinism in lattice-based signatures: Practical fault attacks on pqm4 implementations of nist candidates. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, Asia CCS '19, page 427–440, New York, NY, USA, 2019. Association for Computing Machinery.

[103] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):34, 2009.

[104] Miruna Rosca, Damien Stehlé, and Alexandre Wallet. On the ring-LWE and polynomial-LWE problems. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology – EUROCRYPT 2018, Part I*, volume 10820 of *Lecture Notes in Computer Science*, pages 146–173, Tel Aviv, Israel, April 29 – May 3, 2018. Springer, Heidelberg, Germany.

[105] Peter Schwabe, Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Gregor Seiler, and Damien Stehlé. CRYSTALS-KYBER. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.

[106] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[107] D. G. Steer and L. Strawczynski. A secure audio teleconference system. In *MILCOM 88, 21st Century Military Communications - What's Possible?'. Conference record. Military Communications Conference*, Oct 1988.

[108] Damien Stehlé and Ron Steinfeld. Making NTRU as secure as worst-case problems over ideal lattices. In Kenneth G. Paterson, editor, *Advances in Cryptology – EUROCRYPT 2011*, volume 6632 of *Lecture Notes in Computer Science*, pages 27–47, Tallinn, Estonia, May 15–19, 2011. Springer, Heidelberg, Germany.

[109] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 617–635. Springer, 2009.

[110] M. Steiner, G. Tsudik, and M. Waidner. Key agreement in dynamic peer groups. *IEEE Transactions on Parallel and Distributed Systems*, 11(8):769–780, Aug 2000.

[111] Tim Van Erven and Peter Harremos. Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014.

[112] G.N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library. Cambridge University Press, 1995.

[113] Qianhong Wu, Yi Mu, Willy Susilo, Bo Qin, and Josep Domingo-Ferrer. Asymmetric group key agreement. In Antoine Joux, editor, *Advances in Cryptology – EUROCRYPT 2009*, volume 5479 of *Lecture Notes in Computer Science*, pages 153–170, Cologne, Germany, April 26–30, 2009. Springer, Heidelberg, Germany.

[114] Yang Yu and Léo Ducas. Second order statistical behavior of LLL and BKZ. In *International Conference on Selected Areas in Cryptography*, pages 3–22. Springer, 2017.

[115] Jiang Zhang, Zhenfeng Zhang, Jintai Ding, Michael Snook, and Özgür Dagdelen. Authenticated key exchange from ideal lattices. In Elisabeth Oswald and Marc Fischlin, editors, *Advances in Cryptology – EUROCRYPT 2015, Part II*, volume 9057 of *Lecture Notes in Computer Science*, pages 719–751, Sofia, Bulgaria, April 26–30, 2015. Springer, Heidelberg, Germany.

[116] Zhenfei Zhang, Cong Chen, Jeffrey Hoffstein, William Whyte, John M. Schanck, Andreas Hulsing, Joost Rijneveld, Peter Schwabe, and Oussama Danba. NTRUEncrypt. Technical report, National Institute of Standards and Technology, 2019. available at `https://csrc.nist.gov/projects/post-quantum-cryptography/round-2-submissions`.