Numéro National de Thèse : 2021LYSEN071

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON
opérée par
l'École Normale Supérieure de Lyon

École Doctorale N°512
École Doctorale en Informatique et Mathématiques de Lyon

Discipline : Informatique

Soutenue publiquement le 15/11/2021, par :
Thanh-Huyen NGUYEN

Cryptographic aspects of orthogonal lattices

Aspects cryptographiques des réseaux orthogonaux

Devant le jury composé de :

Roux-Langlois Adeline, Chargée de recherche, CNRS	Rapporteure
Tillich Jean-Pierre, Directeur de recherche, INRIA	Rapporteur
Zémor Gilles, Professeur des universités, Université de Bordeaux	Examinateur
Stehlé Damien, Professeur des universités, ENS de Lyon	Directeur de thèse
Kirshanova Elena, Professeure, I.Kant Baltic Federal University	Co-encadrante

Résumé

La cryptographie à base de réseaux euclidiens vise à faire reposer la sécurité des primitives cryptographiques sur la difficulté conjecturée de problèmes algorithmiques bien identifiés et bien étudiés impliquant les réseaux euclidiens. Cette approche conduit à des primitives plus efficaces, à une sécurité accrue (les problèmes de réseaux les plus courants sont conjecturés quantiquement difficiles) et à des fonctionnalités cryptographiques améliorées (chiffrement entièrement homomorphe, chiffrement fonctionnel, obscurcissement de programme, etc). Une famille de réseaux moins courante mais récurrente sont les réseaux dits orthogonaux où la matrice est souvent échantillonnée à partir d'une distribution gaussienne. Dans cette thèse, nous étudions certains aspects cryptographiques des réseaux orthogonaux. Lorsque les réseaux sont devenus un élément majeur de la conception de primitives cryptographiques, les réseaux orthogonaux ont été utilisés dans diverses constructions telles que les fonctions multilinéaires cryptographiques, les schémas de traçage des traîtres et le chiffrement fonctionnel pour le produit scalaire.

Tout d'abord, nous considérons les minima successifs et le paramètre de lissage de réseaux orthogonaux aléatoires. La motivation principale (et notre résultat) est une généralisation du lemme des restes du hachage (LHL pour *Left Over Hash Lemma*) aux réseaux et aux distributions gaussiennes discrètes. Nos résultats améliorent la borne supérieure probabiliste sur le paramètre de lissage et donnent une borne supérieure probabiliste sur le plus grand minimum du réseau orthogonal.

Ensuite, nous étudions le chiffrement de diffusion avec révocation anonyme, dans lequel les chiffrés ne révèlent aucune information portant sur les utilisateurs qui ont été révoqués. Les réseaux orthogonaux sont impliqués dans les preuves de sécurité de ces protocoles. Nous décrivons une transformation générique du chiffrement fonctionnel linéaire vers des systèmes de diffusion supportant le traçage et la révocation, avec comme nouveauté l'obtention d'une propriété d'anonymat.

Enfin, un problème fondamental lié aux réseaux est l'apprentissage avec erreurs (LWE pour *Learning With Errors*) car il est une base polyvalente pour les constructions cryptographiques. Nous introduisons une nouvelle variante de LWE, dite sur les entiers car elle ne fait pas intervenir de réduction modulaire. Nous montrons que ce nouveau problème est au moins aussi difficile que des problèmes standard portant sur les réseaux euclidiens.

Abstract

Lattice-based cryptography aims at harnessing the security of cryptographic primitives in the conjectured hardness of well-identified and well-studied algorithmic problems involving Euclidean lattices. This approach leads to more efficient primitives, increased security (the most common lattice problems are conjectured quantum-hard), and improved cryptographic functionalities (fully homomorphic encryption, functional encryption, program obfuscation, etc). A less common but still recurring family of lattices are the so-called orthogonal lattices where the matrix is often sampled from a Gaussian distribution. In this thesis, we study the cryptographic aspects of orthogonal lattices. When lattices have turned into a major build block in designing cryptographic primitives, orthogonal lattices have been used in various constructions such as cryptographic multilinear maps, traitor-tracing schemes, and inner product functional encryption.

Firstly, we consider the successive minima and the smoothing parameter of random orthogonal lattices. The main motivation (and our result) is a generalization of the leftover hash lemma (LHL) to lattices and discrete Gaussian distributions. Our results improve the probabilistic upper bound on the smoothing parameter and give a probabilistic upper bound on the last minimum of the orthogonal lattice.

Next, we investigate broadcast encryption with anonymous revocation, in which ciphertexts do not reveal any information about which users have been revoked. The orthogonal lattices are involved in the security proofs of these protocols. We describe a generic transformation of linear functional encryption toward the broadcast systems supporting the trace and revoke with the novelty of achieving anonymity.

Finally, a fundamental problem related to lattices is the learning with errors (LWE) problem which is an amazingly versatile basis for cryptographic constructions. We introduce a new variant of LWE, called on the integers because it does not involve any modular reduction. We show that the new problem is at least as hard the standard problems over lattices.

Contents

List of Symbols

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$	.....	sets of natural, integer, rational, real numbers
$\mathbb{Z}_q$	.....	the ring of integers modulo q
$\mathbb{Z}^n$	.....	the set of integer vectors of dimension n
$\mathbb{Z}_q^n$	.....	the set of integer vectors modulo q of dimension n
$\mathbb{Q}^n, \mathbb{R}^n$	.....	vector-spaces of dimension n
$[n]$	.....	the set $\{1, \ldots, n\}$
$[a, b]$	.....	the set $\{a, \ldots, b\}$
$\mathbf{x}$	.....	column vector
$\mathbf{x}^\mathbf{t}$	.....	row vector
$\mathbf{e}_i$	.....	canonical unit vectors of $\mathbb{Z}^m$
$\mathbf{0}_{m\times n}$	.....	zero matrix of dimensions $m \times n$
$\mathbf{0}_m$	.....	zero-vector of dimension m
$\mathbf{1}_m$	.....	vector with all m entries equal to 1
$E^\perp$	.....	the orthogonal of a vector space E
$\|\mathbf{x}\|$	.....	Euclidean norm of vector $\mathbf{x}$
$\|\mathbf{x}\|_\infty$	.....	ℓ_∞-norm of $\mathbf{x}$: $\max_i \|x_i\|$
I_n	.....	$n \times n$ identity matrix
$X^\mathbf{t}$	.....	the transpose of matrix X
X^{-1}	.....	the inverse of matrix X
$\ln x$	.....	the natural logarithm with base e
$\log x$	.....	the logarithm with base 2
$x \leftarrow D$	.....	x is sampled from a probability distribution D
$x \leftarrow U(S)$	.....	x is uniformly sampled from a finite measure set S
$\mathrm{Span}(W)$	.....	the real span of column vectors of W

We use the Landau notations $\mathcal{O}(\cdot), \Theta(\cdot), \Omega(\cdot), \omega(\cdot), o(\cdot)$.

Résumé long en français

Contexte

La cryptographie basée sur les réseaux euclidiens a été introduite par Ajtai [Ajt96]. Elle fait reposer la sécurité des primitives cryptographiques sur la difficulté conjecturée de problèmes algorithmiques bien identifiés et bien étudiés impliquant des réseaux euclidiens [Pei15]. Elle a été beaucoup développée depuis ces dernières années, car les problèmes à base de réseaux euclidiens sont très prometteurs comme source de sécurité cryptographique, pour quand un ordinateur quantique sera construit. En effet, tous les protocoles utilisés aujourd'hui sur Internet (RSA, ECDSA,...) sont basés sur des problèmes faciles à résoudre si l'on dispose d'un ordinateur quantique. Il y a eu un certain nombre de propositions de schémas cryptographiques dont la sécurité repose de manière prouvée sur des problèmes difficiles portant sur les réseaux euclidiens de grande dimension. Ces problèmes devraient être exponentiellement difficiles à résoudre (la dimension du réseau), même avec des ordinateurs quantiques. Par exemple, *le problème du vecteur le plus court* (SVP pour *Shortest Vector Problem*, voir Figure 1) dans ℓ_2 est NP-difficile pour des réductions probabilistes [Ajt98]. À part être présumée résistante aux ordinateurs quantiques, la cryptographie basée sur les réseaux conduit à des primitives plus efficaces, à une sécurité accrue et à des fonctionnalités cryptographiques améliorées (chiffrement entièrement homomorphe, chiffrement fonctionnel, obscurcissement du programme, etc).

Lemme des restes du hachage sur les réseaux. De nombreux schémas basés sur les réseaux nécessitent un échantillonnage à partir d'une distribution Gaussienne discrète. Il a été largement utilisé dans tous les aspects de la cryptographie basée sur les réseaux. De plus, D. Micciancio et O. Regev montrent dans [MR04] que les distributions Gaussiennes partagent de nombreuses propriétés intéressantes avec leurs contreparties continues, et démontrent leur utilité pour la cryptographie basée sur les réseaux.

Le *lemme des restes du hachage* (LHL pour *Leftover Hash Lemma*) est un outil très puissant. Son application la plus simple est la suivante : échantillonnons $a_1, \ldots, a_m$ uniformément dans $\mathbb{Z}_q$ pour un entier $q > 1$; échantillonnons $z_1, \ldots, z_m$ petits entiers gaussiens. Alors, conditionnée aux a_i, la valeur $\sum_i a_i z_i \bmod q$ "ressemble" à un élément uniforme de $\mathbb{Z}_q$. Cette observation, due à [GPV08, Lemme 4.2], est une variation du lemme des restes du hachage [ILL89], et sa preuve est basée sur le paramètre de lissage du réseau correspondant au noyau de la fonction $\mathbf{z} \in \mathbb{Z}^m \mapsto \langle \mathbf{z}, \mathbf{a} \rangle \bmod q$. Un tel résultat nous permet par exemple d'argumenter sur l'impossibilité de distinguer les clés publiques des éléments uniformes : dans [Reg05], Regev utilise le LHL sur $\mathbb{Z}_q$, pour montrer que $\sum s_i a_i$ donne une clé publique uniforme, où les $s_i \in \mathbb{Z}_q$ sont la clé secrète.

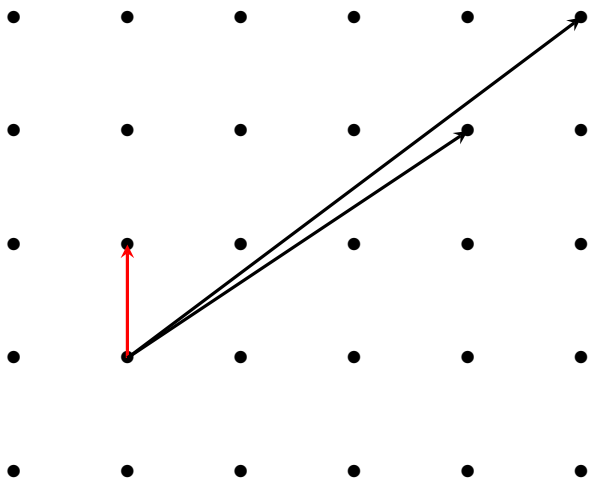

Figure 1 : Un exemple du problème du vecteur le plus court en 2 dimensions : donnons une base de réseau (les vecteurs noirs) comme entrée, la sortie est un vecteur le plus court (le vecteur rouge).

Plus généralement, le LHL conduit à des extracteurs d'aléa simples et efficaces, et peut être utilisé dans de nombreuses applications nécessitant un bon aléa. Il trouve donc de nombreuses applications en cryptographie : dérivation de clés, générateurs de nombres aléatoires, etc [Reg05, GPV08].

La variante classique du LHL relie une distribution uniforme fixe sur un support *fini* à une autre distribution qui provient d'une construction cryptographique spécifique. Cependant, pour certaines primitives à base de réseaux euclidiens, nous ne pouvons pas utiliser directement le LHL. En effet, pour les constructions à base de réseaux, l'application du LHL peut être limitée pour deux raisons. La raison principale est que nous considérons des distributions dont le support est un réseau euclidien, qui est un domaine *infini*. De plus, un choix populaire de distribution à considérer est une distribution Gaussienne discrète, au lieu d'une distribution uniforme (qui n'existe pas sur le domaine infini). Par conséquent, il est important d'étendre le LHL à un tel cadre.

Une autre application simple du LHL est la construction d'un échantillonneur gaussien discret. Plus précisément, nous considérons l'échantillonneur suivant. Dans une phase hors ligne, pour $m > n$, nous échantillonnons l'ensemble de vecteurs courts $\mathbf{x}_1, \ldots, \mathbf{x}_m$ à partir d'un réseau L. Par la suite, dans la phase en ligne, l'échantillon génère $\mathbf{z} \in \mathbb{Z}^m$ selon une Gaussienne discrète et renvoie simplement $\sum_{i=1}^{m} z_i \mathbf{x}_i$. Dans [AGHS13], les auteurs analysent la distribution de $X^{\mathsf{t}}\mathbf{z}$ où X est une matrice Gaussienne aléatoire dans $\mathbb{Z}^{m \times n}$ et $\mathbf{z}$ est un vecteur gaussien aléatoire dans $\mathbb{Z}^m$. Leurs résultat principal est le suivant : si X satisfait une certaine contrainte et si l'écart type de la distribution Gaussienne de $\mathbf{z}$ est suffisamment grand, alors la distribution de $X^{\mathsf{t}}\mathbf{z}$ est statistiquement proche de la distribution Gaussienne discrète avec une covariance appropriée (voir Figure 2). Par la suite, dans [AR16], D. Aggarwal et O. Regev améliorent le résultat principal de [AGHS13] pour certains jeux de paramètres. Dans les deux résultats, la borne de l'écart type de la distribution Gaussienne de $\mathbf{z}$ provient du *paramètre de lissage* de *réseau orthogonal* $\Lambda^{\perp}(X)$ qui est l'ensemble de tous les vecteurs $\mathbf{v} \in \mathbb{Z}^m$ qui appartien-

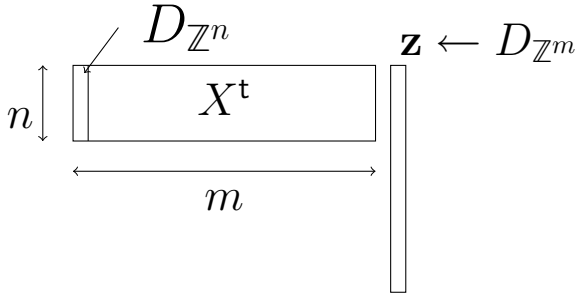

Figure 2 : Lemme des restes du hachage sur les réseaux [AGHS13] : pour les paramètres appropriés, la distribution $X^{\mathsf{t}}\mathbf{z}$ conditionnée par X est proche d'une Gaussienne sur $\mathbb{Z}^n$.

nent au noyau (gauche) de X. Le paramètre de lissage a été introduit dans [MR04]. Intuitivement, cela dit que lorsque l'écart type est suffisamment grand, alors les propriétés statistiques de la distribution Gaussienne discrète sur les réseaux sont très proches de celles de la distribution Gaussienne continue.

Ici, nous notons que le réseau $\Lambda^{\perp}(X)$ a pour dimension $m - \dim(\Lambda(X))$, et une base peut être calculée en temps polynomial à partir de X. De façon intéressante, les liens entre dualité et orthogonalité permettent de prouver que le volume de $\Lambda^{\perp}(X)$ est égal au volume du réseau $\mathrm{span}(\Lambda(X)) \cap \mathbb{Z}^m$ avec probabilité proche de 1 vis-à-vis du choix de X [Ngu99]. Ainsi, si un réseau dans $\mathbb{Z}^m$ consiste à faible dimension, son réseau orthogonal consiste à grande dimension avec un volume au plus égal : les minima successifs du réseau orthogonal sont susceptibles d'être beaucoup plus courts que ceux du réseau d'origine. En cryptographie, les réseaux orthogonaux sont d'abord apparus comme un outil cryptanalytique [NS97, NS99, DGHV10]. Des années plus tard, lorsque les réseaux sont devenus un élément majeur de la conception de primitives cryptographiques, les réseaux orthogonaux ont été utilisés dans diverses constructions telles que de fonctions multilinéaires cryptographiques [AGHS13], de schémas de traçage de traîtres [LPSS17] et de chiffrement fonctionnel pour le produit scalaire [ALS16].

Systèmes de traçage et de révocation. Un type de protocole qui utilise les réseaux orthogonaux dans les preuves de sécurité est le *chiffrement de diffusion* qui est une primitive cryptographique fondamentale qui donne la possibilité d'envoyer un message sécurisé à n'importe quel ensemble cible choisi parmi les utilisateurs enregistrés. La variation la plus intéressante du *chiffrement de diffusion* s'appelle *système de traçage-et-révocation* [BW06] qui est un système de chiffrement multi-destinataires où un distributeur de contenu peut trouver des *utilisateurs malveillants* et peut révoquer leur capacité de déchiffrement (voir Figure 3). De manière informelle, un tel schéma de chiffrement à clé publique, permet à un expéditeur de chiffrer des données sous une clé publique pk et chaque utilisateur légitime peut utiliser sa clé secrète sk_i pour déchiffrer les données. Un système de traçage des traîtres garantit que si une coalition d'utilisateurs (légitimes) mettent en commun leurs clés secrètes pour construire une *boîte décodeur* qui peut déchiffrer le texte chiffré, alors il existe un algorithme de traçage efficace pour trouver au moins un utilisateur coupable à condition que l'algorithme ait accès au décodeur. Par la suite, le distributeur de contenu peut utiliser la fonctionnalité de révocation pour interdire aux utilisateurs coupables d'accéder aux données à l'avenir. Un système de

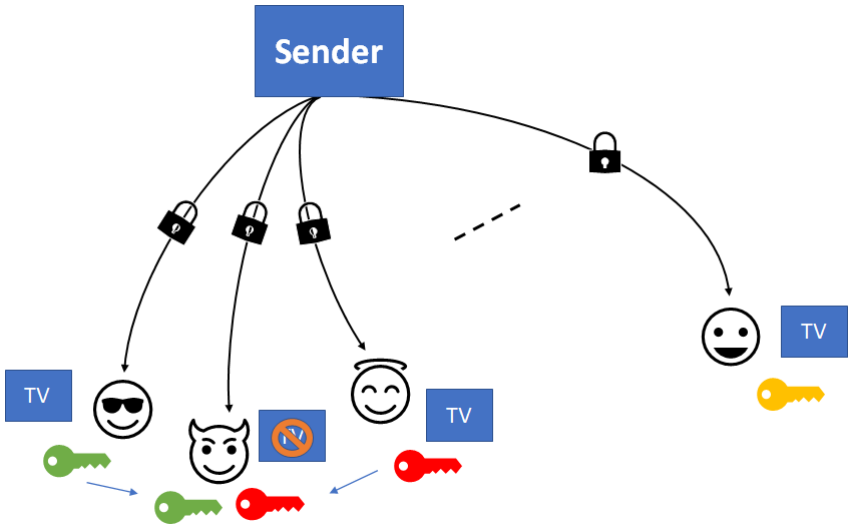

Figure 3 : Illustration d'un chiffrement de diffusion dans l'application de Pay-TV : le système a deux utilisateurs malveillants qui ont la clé rouge et la clé verte. Ils partagent leurs clés secrètes avec l'entité non légitime. Le système peut les identifier, de sorte qu'ils ne puissent plus accéder au nouveau contenu.

révocation garantit que si une coalition d'utilisateurs illégitimes mettent en commun leurs clés secrètes, ils ne peuvent toujours pas déchiffrer le texte chiffré.

Nous décrivons maintenant le protocole qui est basé sur [ABP+17] pour voir comment les réseaux orthogonaux aident pour cela. Dans un schéma de révocation, un utilisateur id avec la clé secrète $\mathsf{sk}_{\mathsf{id}}$ peut déchiffrer un texte chiffré $\mathsf{ct}_{\mathcal{R}}$ si l'utilisateur n'est pas révoqué. Agrawal *et al.* dans [ABP+17] ont construit un schéma de révocation avec traçabilité publique pour un nombre non borné d'utilisateur à partir du chiffrement fonctionnel pour le produit scalaire. Dans un chiffrement fonctionnel pour le produit scalaire [ABDCP15, ALS16], étant donné une clé secrète pour $\mathsf{sk}_{\mathbf{x}}$ le vecteur de clé $\mathbf{x}$ et un texte chiffré $\mathsf{ct}_{\mathbf{y}}$ pour le vecteur chiffré $\mathbf{y}$, le destinataire peut calculer le produit scalaire des deux vecteurs impliqués, c'est-à-dire $\langle \mathbf{x}, \mathbf{y} \rangle$. Intuitivement, dans la construction de révocation d'Agrawal *et al.*, chaque id est associé à un vecteur aléatoire $\mathbf{x}_{\mathsf{id}}$ et en conséquence un ensemble $\mathcal{R}$ est associé à l'espace vectoriel parcouru par $(\mathbf{x}_{\mathsf{id}})_{\mathsf{id} \in \mathcal{R}}$. Pour créer un texte chiffré avec l'ensemble révoqué $\mathcal{R}$, Agrawal *et al.* choisissent un vecteur $\mathbf{v}_{\mathcal{R}}$ othrogonal à $(\mathbf{x}_{\mathsf{id}})_{\mathsf{id} \in \mathcal{R}}$ et définissent $\mathbf{y}_{\mathcal{R}} = m \cdot \mathbf{v}_{\mathcal{R}}$ comme vecteur de texte chiffré. Notez que si $\mathsf{id} \in \mathcal{R}$, alors $\mathbf{x}_{\mathsf{id}}$ sera orthogonal à $\mathbf{v}_{\mathcal{R}}$ et ensuite à $\mathbf{y}_{\mathcal{R}}$ (c'est-à-dire $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = 0$). Dans le cas où $\mathsf{id} \notin \mathcal{R}$, alors avec une forte probabilité $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle \neq 0$ agit comme un facteur d'aveuglement multiplicatif pour le message en clair m : on a $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$. Maintenant, pendant le déchiffrement, le chiffrement fonctionnel pour le produit scalaire calcule $\mathsf{Res} = \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ comme valeur intermédiaire. Le destinataire, qui est doté de $\mathbf{v}_{\mathcal{R}}$ à côté du texte chiffré, recalcule le facteur d'aveuglement $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ pour récupérer le message m.

Une question naturelle consiste à savoir si l'on peut concevoir un protocole où l'utilisateur révoqué ne comprendra pas s'il a été révoqué. De plus, étant donné un texte chiffré, aucun utilisateur légitime n'obtiendra d'informations sur les utilisateurs qui ont été révoqués du déchiffrement du texte chiffré. Ce problème est appelé *chiffre-*

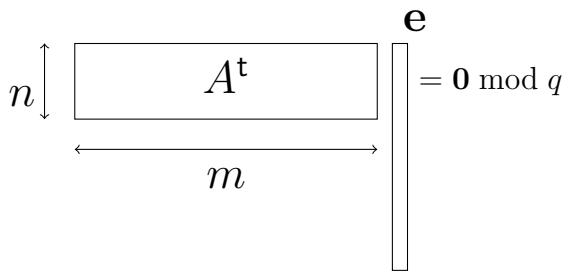

Figure 4 : Solution entière courte : donnons une matrice $A \leftarrow U(\mathbb{Z}_q^{m \times n})$ comme entrée, l'objectif est de calculer un tel vecteur court $\mathbf{e}$.

ment traçage et révocation avec anonymat.

Solution entière courte (Short Integer Solution). Un problème moyen-cas proposé par Ajtai consiste à trouver une solution entière non nulle courte $\mathbf{e} \in \mathbb{Z}^m$ au système linéaire homogène $A^{\mathsf{t}}\mathbf{e} = \mathbf{0} \bmod q$ pour $A \in \mathbb{Z}_q^{m \times n}$ uniforme. Cela équivaut à trouver un vecteur court non nul dans $\Lambda_q^{\perp}(A)$ (voir Figure 4). Ce problème est appelé *solution enièstre courte* (SIS pour *Short Integer Solution*). Il est la base de fonctions de hachage et résistantes aux collisions, de schémas d'identification, de signatures numériques, etc [GPV08, Lyu13, LS14, PR06, Pei15, LLLS13].

Apprentissage avec erreurs (Learning with errors). Un problème fondamental lié aux réseaux est le problème *d'apprentissage avec erreurs* (LWE pour *Learning With Errors*). Le problème LWE est introduit par Regev [Reg05]. Il consiste à récupérer un vecteur uniforme $\mathbf{s} \in \mathbb{Z}_q^n$, étant donné de nombreux échantillons de la forme $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)$, avec $\mathbf{a}$ uniforme dans $\mathbb{Z}_q^n$ et un bruit e échantillonné selon une distribution gaussienne sur $\mathbb{Z}_q$. Dans [Reg05], Regev a montré que pour des paramètres appropriés, ce problème est aussi difficile que des problèmes de réseau dans le pire des cas, et équivalent en temps polynomial à sa version de décision, qui demande de distinguer la distribution de l'échantillon LWE comme ci-dessus de la distribution uniforme sur $(\mathbb{Z}_q^n \times \mathbb{Z}_q)$.

Nous présentons maintenant le chiffrement à clé publique utilisant LWE de [Reg05]. Alice choisit une clé secrète $\mathbf{s}$ et une matrice aléatoire $A \in \mathbb{Z}_q^{m \times n}$. La clé publique d'Alice est

$$(A, \mathbf{b} = A\mathbf{s} + \mathbf{e}).$$

Elle est indiscernable d'une paire uniforme, sous l'hypothèse de difficulté de la version décisionnelle de LWE. Afin de chiffrer un message à Alice, Bob choisit un petit vecteur aléatoire $\mathbf{r}$, et calcule le texte chiffré

$$\mathbf{u} = A^{\mathsf{t}}\mathbf{r}, u' = \mathbf{b}^{\mathsf{t}}\mathbf{r} + \mu \cdot \lceil \frac{q}{2} \rceil$$

où μ est le bit de message. Dit autrement, il envoie soit $\mathbf{b}^{\mathsf{t}}\mathbf{r}$ pour $\mu = 0$, soit quelque chose de très éloigné de $\mathbf{b}^{\mathsf{t}}\mathbf{r}$, pour $\mu = 1$. Pour déchiffrer le message, Alice soustrait les informations en fonction de sa clé secrète $\mathbf{s}^{\mathsf{t}}\mathbf{u}$ de u', qui est soit approximativement 0 soit approximativement $\frac{q}{2}$, selon la valeur du bit de message μ.

De nombreuses variations du problème LWE ont été introduites, principalement dans le but d'améliorer l'efficacité de la cryptographie basée sur les réseaux. Par exemple, de nombreux articles ont été consacrés à l'analyse de LWE lorsque la clé secrète $\mathbf{s}$, ou l'erreur $\mathbf{e}$ ou le vecteur $\mathbf{a}$ suivent une distribution qui est différente de celle considérée par [Reg05], comme dans [MP13, BLP+13, HM17]. Des extensions de LWE sur des anneaux plus généraux ont également été largement étudiées, en commençant par les introductions du problème Polynomial-LWE [SSTX09] et du problème Ring-LWE [LPR10, RSW18].

Nos contributions

Dans cette thèse, nous nous concentrons sur les aspects cryptographiques des réseaux orthogonaux. Tout d'abord, nous considérons les minima successifs et le paramètre de lissage des réseaux orthogonaux aléatoires. Nous étudions ensuite le chiffrement de diffusion avec révocation anonyme. Enfin, nous introduisons une nouvelle variante du problème LWE sur les entiers, sans aucune réduction modulaire.

Nous notons que les résultats discutés ci-dessous ont été principalement tirés d'un article en préparation et des articles publiés suivants.

1. [KNSW20] E. Kirshanova, H. Nguyen, D. Stehlé, and A. Wallet. On the smoothing parameter and last minimum of random orthogonal lattices. Designs, Codes and Cryptography 2020.

2. [BMN+21] O. Blazy, S. Mukherjee, H. Nguyen, H. Phan and D. Stehlé. An Anonymous Trace-and-Revoke Broadcast Encryption Scheme. Accepté à Australasian Conference on Information Security and Privacy 2021.

Le paramètre de lissage des réseaux orthogonaux aléatoires

Le paramètre de lissage du réseau $\Lambda^{\perp}(X)$ généré par une gaussienne discrète $X \in \mathbb{Z}^{m \times n}$ a déjà été considéré par Agrawal *et al.* dans [AGHS13]. La motivation principale (et notre résultat) est une généralisation du lemme des restes du hachage (LHL) aux réseaux et aux distributions gaussiennes discrètes. Notre premier résultat est une borne supérieure probabiliste améliorée sur le paramètre de lissage, par rapport aux travaux antérieurs [AGHS13] et [AR16]. Notre deuxième résultat améliore la borne supérieure probabiliste sur le $(m-n)-$ième minimum du réseau orthogonal $\Lambda^{\perp}(X)$.

Chiffrement de traçage et révocation avec anonymat

Nous étudions le chiffrement de diffusion avec révocation anonyme. Notre contribution est triple. Tout d'abord, nous développons une transformation générique du chiffrement fonctionnel linéaire vers des systèmes de suivi et de révocation. Il s'inspire de la transformation d'Agrawal *et al.* [ABP+17] avec la nouveauté d'atteindre l'anonymat. Notre deuxième contribution est d'instancier le chiffrement fonctionnel linéaire sous-jacent à partir d'hypothèses standards. Nous proposons une construction basée sur le problème Diffie-Hellman décisionnel (DDH pour *Decisional Diffie–Hellman*) qui améliore considérablement les performances par rapport à la construction basée sur DDH d'Agrawal

et al. Dans le ces de LWE, nous avons essayé d'instancier notre construction en nous appuyant sur le schéma de Wang *et al.* [WFL19] mais nous avons finalement trouvé une attaque contre ce schéma. Notre troisième contribution est d'étendre le chiffrement à 1 bit de la transformation générique au chiffrement à n bits. En introduisant le chiffrement fonctionnel par multiplication matricielle, qui effectue un nombre fixe d'appels parallèles sur des chiffrements fonctionnels avec le même caractère aléatoire, nous pouvons prouver la sécurité du schéma final avec une réduction serrée qui ne dépend pas de n, contrairement à ce que l'on obtient en utilisant l'argument hybride.

Nouveau problème LWE sur les entiers

Nous introduisons une nouvelle variante du problème LWE sur les entiers, sans aucune réduction modulaire.

Concrètement, nous considérons deux problèmes : integer-SIS, et integer-LWE. Pour integer-SIS, la différence avec le problème SIS est que les entrées de la matrice A sont gaussiennes sur les entiers. Par la suite, à partir de l'observation que le problème integer-SIS est syntaxiquement équivalent à trouver un vecteur court non nul dans le réseau orthogonal, et inspiré par la dualité avec le problème de décodage à distance bornée (BDD pour *Bounded Distance Decoding*), nous définissons le problème de recherche, integer-LWE comme un BDD sur le réseau dual : il consiste à récupérer $\mathbf{k}$ à partir de $(X, \pi_{\ker(X)}(\mathbf{k} + \mathbf{e}))$, où X et $\mathbf{k}$ sont gaussiens sur $\mathbb{Z}^{m\times n}$ et $\mathbb{Z}^m$, respectivement, et $\mathbf{e} \in \mathbb{R}^m$ est gaussien avec un écart type nettement inférieur à 1. Nous proposons une réduction en temps polynomial quantique du problème integer-SIS au problème integer-LWE.

Organisation du manuscrit

Au chapitre 2, nous rappelons toutes les définitions importantes qui sont utilisées au long de cette thèse. Celles-ci incluent les définitions de base liées aux réseaux. Au chapitre 3, nous exposons notre résultat sur le paramètre de lissage du réseau orthogonal. Le chapitre 4 concerne un schéma de chiffrement de diffusion de traçage avec révocation anonyme. Au chapitre 5, nous introduisons le nouveau problème integer-LWE. Enfin, nous concluons cette thèse et présentons quelques pistes de travaux futurs dans le chapitre 6.

Chapter 1

Introduction

1.1 Background

Lattice-based cryptography was introduced by Ajtai in 1996 [Ajt96]. It aims at harnessing the security of cryptographic primitives in the conjectured hardness of well-identified and well-studied algorithmic problems involving Euclidean lattices [Pei15]. It has been developed in recent years, as lattice-based problems are very promising as a source of cryptographic security, in the case when a quantum computer will be built. Indeed, all the protocols used today on the Internet (RSA, ECDSA,...) are based on problems that are easy to solve if one has a quantum computer. There have been a number of proposals of cryptographic schemes with security provably relying on hard problems over high-dimensional Euclidean lattices. Indeed, these problems are expected to be exponentially hard to solve (in the dimension of the lattice), even with quantum computers. For example, the shortest vector problem (SVP) (see Figure 1.1) in ℓ_2 is NP-hard for probabilistic reductions [Ajt98]. In addition to being supposedly resistant to quantum computers, this approach leads to more efficient primitives, increased security, and improved cryptographic functionalities (fully homomorphic encryption, functional encryption, program obfuscation, etc).

Leftover Hash Lemma over lattices. Many lattice-based schemes require sampling from discrete Gaussian distributions. They have been used extensively in all aspects of lattice-based cryptography. Moreover, D. Micciancio and O. Regev show in [MR04] that Gaussian distributions share many nice properties with their continuous counterparts, and demonstrate their usefulness for lattice-based cryptography.

The *leftover hash lemma* (LHL) is a very powerful tool. Its simplest application is the following: sample $a_1, \ldots, a_m$ uniformly in $\mathbb{Z}_q$ for some integer $q > 1$; sample $z_1, \ldots, z_m$ small Gaussian integers. Then, conditioned on the a_i's, the value $\sum_i a_i z_i \bmod q$ "looks like" a uniform element of $\mathbb{Z}_q$. This observation, due to [GPV08, Lemma 4.2], is a variant of the leftover hash lemma [ILL89], and its proof crucially relies on the smoothing parameter of the lattice corresponding to the kernel of the map $\mathbf{z} \in \mathbb{Z}^m \mapsto \langle \mathbf{z}, \mathbf{a} \rangle \bmod q$. Such a result enables us to argue about the indistinguishability of public keys from uniform elements. In [Reg05], Regev uses LHL over $\mathbb{Z}_q$, to show that $\sum s_i a_i$ gives a uniform public key, where $s_i \in \mathbb{Z}_q$ are the secret key. More general, LHL leads to simple and efficient randomness extractors, and can be used in many applications requiring good randomness. It therefore finds numerous applications in cryptography: key derivation, random number generators, etc [Reg05, GPV08].

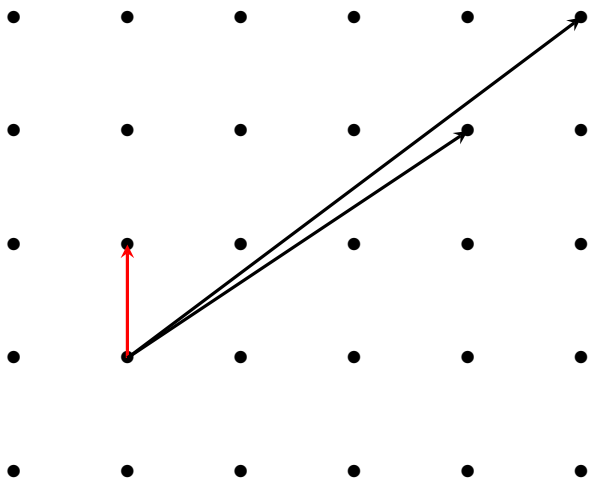

Figure 1.1: An example of shortest vector problem in 2 dimensions: given a lattice basis (the black vectors) as input, the output is a shortest vector in lattice (the red vector).

The classical statement of the LHL relates a fixed uniform distribution over a *finite* support to another distribution that comes from a specific cryptographic construction. However, for some lattice-based primitives, we cannot use the LHL directly. Indeed, for lattice-based constructions, the application of the LHL is limited for two reasons. The main reason is that we care about distributions whose support is a Euclidean lattice, which is an *infinite* domain. Moreover, a popular choice of a distribution to consider is a discrete Gaussian distribution, instead of a uniform one (which does not exist over an infinite domain). Hence, it is needed to extend LHL to such a setting.

Another application of LHL is an extremely simple discrete Gaussian sampler. Specifically, we consider the following sampler. In an offline phase, for $m > n$, we sample a set of short vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ from a lattice L. Then, in the online phase, the sample generates $\mathbf{z} \in \mathbb{Z}^m$ according to a discrete Gaussian and simply outputs $\sum_{i=1}^{m} z_i \mathbf{x}_i$. In [AGHS13], the authors analyze the distribution of $X^{\mathsf{t}}\mathbf{z}$ where X is random Gaussian matrix in $\mathbb{Z}^{m \times n}$ and $\mathbf{z}$ is random Gaussian vector in $\mathbb{Z}^m$. Their main result is the following: if X satisfies a certain constraint and if the standard deviation of the distribution of $\mathbf{z}$ is large enough, then the distribution of $X^{\mathsf{t}}\mathbf{z}$ is statistically close to a discrete Gaussian distribution with appropriate covariance (see Figure 1.2). Then in [AR16], D. Aggarwal and O. Regev improve over the main result of [AGHS13] for some parameter sets.

In both of the results, the bound on the standard deviation of the distribution of $\mathbf{z}$ comes from the so-called *smoothing parameter* of *orthogonal lattice* $\Lambda^{\perp}(X)$ which is a set of all vectors $\mathbf{v} \in \mathbb{Z}^m$ that belong to the (left) kernel of X. The smoothing parameter was introduced in [MR04]. Intuitively, it says that when the standard deviation is large enough, then the statistical properties of the discrete Gaussian distribution over lattices are very close to those of the continuous Gaussian distribution.

Here, we note that the lattice $\Lambda^{\perp}(X)$ has dimension $m - \dim(\Lambda(X))$, and its basis can be computed in polynomial time from X. Interestingly, the links between duality and orthogonality enable us to prove that the volume of $\Lambda^{\perp}(X)$ is equal to the volume of the lattice $\mathrm{span}(\Lambda(X)) \cap \mathbb{Z}^m$ with probability close to 1 over the choice of X [Ngu99].

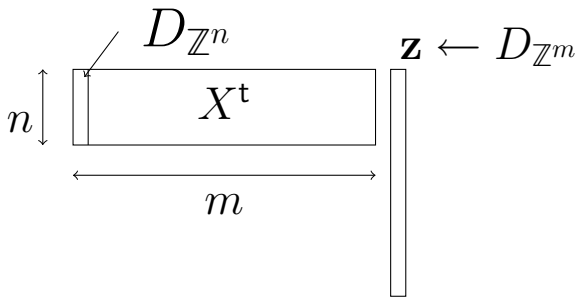

Figure 1.2: The leftover hash lemma over lattices from [AGHS13]: with appropriate parameters, the distribution of $X^{\mathsf{t}}\mathbf{z}$ conditioned on X is close to Gaussian distribution over $\mathbb{Z}^n$.

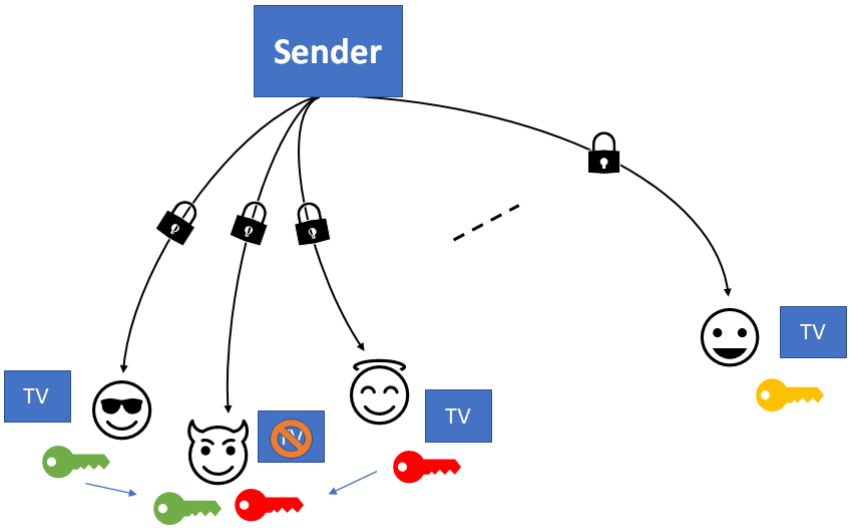

Figure 1.3: Illustration of a broadcast encryption in the Pay-TV application: the system has two malicious users who have the red key and the green key. They share their secret keys with the non-legitimate entity. The system can identify them, so that they can no longer access new content.

Thus, if a lattice in $\mathbb{Z}^m$ is low-dimensional, its orthogonal lattice is high-dimensional with volume at most equal: the successive minima of the orthogonal lattice are likely to be much shorter than the ones of the original lattice.

In cryptography, orthogonal lattices first appeared as a cryptanalytic tool in attacking several cryptographic constructions [NS97, NS99, DGHV10]. Years later, when lattices have turned into a major building block in designing cryptographic primitives, orthogonal lattices have been used in various constructions such as cryptographic multilinear maps [AGHS13], traitor-tracing schemes [LPSS17] and inner product functional encryption [ALS16].

Trace-and-revoke systems. A protocol that uses the orthogonal lattices in the security proofs is the so-called *broadcast encryption*. It is a fundamental cryptographic primitive that gives the ability to send a secure message to any chosen target set among registered users. An interesting variant of broadcast encryption is called *trace-and-revoke*

system [BW06]. It is a multi-recipient encryption system where a content distributor can find *malicious users* and can revoke their decryption capability (see Figure 1.3). Such a public-key encryption scheme, informally speaking, allows a sender to encrypt data under a public key pk and each legitimate user can use their secret key sk_i to decrypt the data. A traitor tracing system guarantees that if a coalition of (legitimate) users pool their secret keys to construct a *decoder box* that can decrypt the ciphertext, then there is an efficient trace algorithm to find at least one guilty user provided the algorithm is given access to the decoder. Then the content distributor can use the revocation functionality to prohibit the guilty users from accessing the data in the future. A revocation system ensures that if a coalition of illegitimate users pools their secret keys, they still cannot decrypt the ciphertext.

We now describe the protocol which is based on [ABP+17] to see how the orthogonal lattices work for it. In a revocation scheme, a user id with the secret key $\mathsf{sk}_{\mathsf{id}}$ can decrypt a ciphertext $\mathsf{ct}_{\mathcal{R}}$ if the user is not revoked. Agrawal *et al.* in [ABP+17] constructed a revocation scheme with public traceability for an unbounded number of users from inner product functional encryption. In an inner product functional encryption [ABDCP15, ALS16], given a secret key $\mathsf{sk}_{\mathbf{x}}$ for a key vector $\mathbf{x}$ and a ciphertext $\mathsf{ct}_{\mathbf{y}}$ for a vector $\mathbf{y}$, the decryptor can compute the inner product of the two vectors involved, i.e., $\langle \mathbf{x}, \mathbf{y} \rangle$. Intuitively, in the revocation construction by Agrawal *et al.*, each id is associated with a random vector $\mathbf{x}_{\mathsf{id}}$ and correspondingly a set $\mathcal{R}$ is associated with the vector space spanned by $(\mathbf{x}_{\mathsf{id}})_{\mathsf{id} \in \mathcal{R}}$. To create a ciphertext with the revoked set $\mathcal{R}$, Agrawal *et al.* choose a vector $\mathbf{v}_{\mathcal{R}}$ othrogonal to $(\mathbf{x}_{\mathsf{id}})_{\mathsf{id} \in \mathcal{R}}$ and define $\mathbf{y}_{\mathcal{R}} = m \cdot \mathbf{v}_{\mathcal{R}}$ as the ciphertext vector. Observe that if $\mathsf{id} \in \mathcal{R}$, then $\mathbf{x}_{\mathsf{id}}$ will be orthogonal to $\mathbf{v}_{\mathcal{R}}$ and subsequently to $\mathbf{y}_{\mathcal{R}}$ (i.e., $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = 0$). In case $\mathsf{id} \notin \mathcal{R}$, with high probability $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle \neq 0$ which acts as a multiplicative blinding factor for the plaintext message m as $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$. Now, during decryption, the inner product functional encryption computes $\mathsf{Res} = \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ as an intermediate step. The decryptor, who is provided with $\mathbf{v}_{\mathcal{R}}$ alongside the ciphertext, recomputes the blinding factor $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ to recover the message m.

A natural question is whether one can devise a protocol where the revoked user will not make out if it has been revoked. Moreover, given a ciphertext, no legitimate user will get any information about the users who have been revoked from decrypting the ciphertext. This problem is called *trace and revoke scheme with anonymity.*

Short integer solution. An average-case problem proposed by Ajtai [Ajt96] consists in finding a short non-zero integer solution $\mathbf{e} \in \mathbb{Z}^m$ to the homogeneous linear system $A^{\mathsf{t}}\mathbf{e} = \mathbf{0} \bmod q$ for a uniformly random $A \in \mathbb{Z}_q^{m \times n}$ This is equivalent to finding a short nonzero vector in $\Lambda_q^{\perp}(A)$ (see Figure 1.4). This problem is called *short integer solution* (SIS). It is a foundation for collision-resistant hash functions, identification schemes, digital signatures, etc [GPV08, Lyu13, LS14, PR06, Pei15, LLLS13].

Learning with errors. Another fundamental problem related to lattices is the *learning with errors* (LWE) problem. The LWE problem is introduced by Regev [Reg05]. It is the problem of recovering uniformly random vector $\mathbf{s} \in \mathbb{Z}_q^n$ given many samples of the form $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e)$, with $\mathbf{a}$ uniform in $\mathbb{Z}_q^n$, noise e sampled according to a Gaussian distribution over $\mathbb{Z}_q$. In [Reg05], Regev showed that for appropriate parameters, this problem is as hard as worst-case lattice problems and polynomial-time equivalent to its

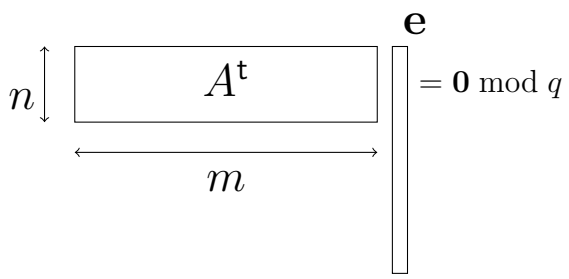

Figure 1.4: The short integer solution: given matrix $A \leftarrow U(\mathbb{Z}_q^{m \times n})$ as input, the goal is to compute such a short vector $\mathbf{e}$.

decision version, which asks to distinguish the distribution of a LWE sample as above from the uniform distribution over $(\mathbb{Z}_q^n \times \mathbb{Z}_q)$.

We now present a public key encryption using LWE. Alice chooses a secret key $\mathbf{s}$ and a random matrix $A \in \mathbb{Z}_q^{m \times n}$. Alice's public key is

$$(A, \mathbf{b} = A\mathbf{s} + \mathbf{e}).$$

It is indistinguishable from a uniform pair under the decision LWE assumption. In order to encrypt a message to Alice, Bob chooses a random small vector $\mathbf{r}$, and computes the ciphertext

$$\mathbf{u} = A^t\mathbf{r}, u' = \mathbf{b}^t\mathbf{r} + \mu \cdot \lceil \frac{q}{2} \rceil$$

where μ is the message bit. Essentially, he sends either $\mathbf{b}^t\mathbf{r}$ for $\mu = 0$, or something very far from $\mathbf{b}^t\mathbf{r}$, for $\mu = 1$. To decrypt the message, Alice subtracts information depending on her secret key $\mathbf{s}^t\mathbf{u}$ from u', which is either approximately 0 or approximately $\frac{q}{2}$, depending on the value of the message bit μ.

Many variants of the LWE problem have been introduced, mostly with the goal of improving the efficiency of lattice-based cryptography. For example, many papers have been devoted to the analysis of LWE when the secret key $\mathbf{s}$, or the error $\mathbf{e}$ follow a distribution which is different from that considered by [Reg05], as in [MP13, BLP+13, HM17]. Extensions of LWE over more general rings have also been broadly studied, starting with the introductions of the Polynomial-LWE problem [SSTX09] and of the Ring-LWE problem [LPR10, RSW18].

1.2 Our contributions

In this thesis, we focus on the cryptographic aspects of orthogonal lattices. First, we consider the successive minima and the smoothing parameter of random orthogonal lattices. We then investigate broadcast encryption with anonymous revocation. Lastly, we introduce a new variant of the LWE problem over the integers without any modular reduction.

We note that the results discussed below have been mainly taken from an article in preparation and from the following published articles.

1. [KNSW20] E. Kirshanova, H. Nguyen, D. Stehlé, and A. Wallet. On the smoothing parameter and last minimum of random orthogonal lattices. Designs, Codes and Cryptography 2020.

2. [BMN+21] O. Blazy, S. Mukherjee, H. Nguyen, H. Phan and D. Stehlé. An Anonymous Trace-and-Revoke Broadcast Encryption Scheme. Accepted to Australasian Conference on Information Security and Privacy 2021.

1.2.1 The smoothing parameter of random orthogonal lattices

The smoothing parameter of $\Lambda^{\perp}(X)$ generated by a discrete Gaussian $X \in \mathbb{Z}^{m \times n}$ was already considered by Agrawal *et al.* in [AGHS13]. Their main motivation (and result) was a generalization of the leftover hash lemma (LHL) to lattices and discrete Gaussian distributions. Our first result is an improved probabilistic upper bound on the smoothing parameter, compared to prior works in [AGHS13] and [AR16]. Our second result improves the probabilistic upper bound on the $(m-n)-$th minimum of the orthogonal lattice $\Lambda^{\perp}(X)$.

1.2.2 Trace and Revoke scheme with anonymity

We investigate broadcast encryption with anonymous revocation. Our contribution is threefold. First, we develop a generic transformation of linear functional encryption toward trace-and-revoke systems. It is inspired by the transformation by Agrawal *et al.* [ABP+17] with the novelty of achieving anonymity. Our second contribution is to instantiate the underlying linear functional encryption from standard assumptions. We propose a DDH-based construction which significantly improves the performance compared to the decisional Diffie–Hellman (DDH)-based construction of Agrawal *et al.* In the case of LWE, we tried to instantiate our construction by relying on the scheme from Wang *et al.* [WFL19] but we finally found an attack on this scheme. Our third contribution is to extend the $1-$bit encryption from the generic transformation to $n-$bit encryption. By introducing matrix multiplication functional encryption, which essentially performs a fixed number of parallel calls on functional encryptions with the same randomness, we can prove the security of the final scheme with a tight reduction that does not depend on n, in contrast to what we obtain by employing the hybrid argument.

1.2.3 New LWE problem over the integers

We introduce a new variant of the LWE problem over the integers, without any modular reduction.

Concretely, we consider two problems: integer-SIS, and integer-LWE. For integer-SIS, the difference from the standard SIS problem is that the entries of A are Gaussian over the integers. Then from the observation that the integer-SIS problem is syntactically equivalent to finding a nonzero short vector in the orthogonal lattice, and inspired by the duality with the bounded distance decoding problem (BDD), we define the search integer-LWE problem as a BDD problem on the dual lattice: it consists in recovering $\mathbf{k}$ from $(X, \pi_{\ker(X)}(\mathbf{k}+\mathbf{e}))$, where X and $\mathbf{k}$ are Gaussian over $\mathbb{Z}^{m \times n}$ and $\mathbb{Z}^m$, respectively,

$\mathbf{e} \in \mathbb{R}^m$ is Gaussian with standard deviation significantly below 1 and $\pi_{\ker(X)}(\mathbf{k}+\mathbf{e})$ is the orthogonal projection of $(\mathbf{k}+\mathbf{e})$ onto $\ker(X)$. We propose a quantum polynomial-time reduction from the integer-SIS problem to the search integer-LWE problem.

1.3 Organization of this thesis

In Chapter 2, we recall all necessary definitions that are used throughout this thesis. These include the basic definitions related to lattices. In Chapter 3, we expose our result on the smoothing parameter of the orthogonal lattice. Chapter 4 is about an anonymous trace-and-revoke broadcast encryption scheme. In Chapter 5, we introduce the new LWE problem over integers. Finally, we conclude this thesis and present some directions for future works in Chapter 6.

Chapter 2

Preliminaries

In this chapter, we first give some basic definitions and results on lattices that will be used throughout the thesis. These include the basic notations and properties of lattices, discrete Gaussian distributions over lattices, and the smoothing parameter.

2.1 Lattices

A lattice is a discrete additive subgroup of $\mathbb{R}^m$, for some integer $m \geqslant 1$. A set of linearly independent vectors $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_d\} \subset \mathbb{R}^m$ that generates a lattice via integer linear combinations is called a basis, and we write the lattice generated by B as

$$L(B) := \Big\{ B\mathbf{z} = \sum_{i \in [d]} z_i \mathbf{b}_i : \mathbf{z} \in \mathbb{Z}^d \Big\}.$$

The rank of this lattice is d and its embedding dimension is m. We define the determinant of L as $\det(L) := \sqrt{\det(B^{\mathsf{t}}B)}$. For a rank-$d$ matrix $B \in \mathbb{R}^{m \times d}$, there exist orthogonal matrices U, V and a diagonal matrix $\Sigma = \mathrm{Diag}(\sigma_1, \ldots, \sigma_d) \in \mathbb{R}^{m \times d}$ such that $B = U\Sigma V^{\mathsf{t}}$ and $\sigma_1 \geqslant \cdots \geqslant \sigma_d > 0$. From this decomposition, we see that $\det(L(B)) = \prod_{i \in [1,d]} |\sigma_i|$.

When $d = m$, we say that the lattice has full rank. Understanding geometric properties of high-dimensional lattices is a central topic in various areas of mathematics and computer science [NV09]. Among the most important invariants of a lattice are its so-called *successive minima* $\lambda_1(\Lambda), \ldots, \lambda_d(\Lambda)$. More precisely, for $i \in [d]$, the i-th successive minimum $\lambda_i(L)$ is defined as

$$\lambda_i(L) := \inf\{r : \dim(\mathrm{Span}(L \cap \mathcal{B}(r))) \geqslant i\},$$

where $\mathcal{B}(r)$ denotes the closed zero-centered Euclidean ball of radius r. We use the notation $\lambda_i^{\infty}(L)$ when we consider the infinity norm.

Any lattice $L \subseteq \mathbb{R}^m$ has a dual lattice $L^{\star}$. It consists of all the vectors in $\mathrm{Span}(L)$ that are orthogonal to L modulo 1, namely:

$$L^{\star} := \{\mathbf{y} \in \mathrm{Span}(L) : \forall\, \mathbf{x} \in L, \langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{Z}\}.$$

Note that $L^{\star\star} = L$.

Several families of lattices are considered in this work.

Definition 2.1. Let $m > n \geqslant 1$ and $q \geqslant 2$ be integers. Let $X \in \mathbb{Z}^{n \times m}$.

1. The **orthogonal lattice** $\Lambda^{\perp}(X)$ is the integral lattice whose vectors are orthogonal to the rows of X, i.e.,

$$\Lambda^{\perp}(X) := \{\mathbf{v} \in \mathbb{Z}^m : X\mathbf{v} = \mathbf{0}\}.$$

2. The lattice $\Lambda_q(X) \subseteq \mathbb{Z}^m$ is the full-rank lattice spanned by the rows of X and the vectors $q\mathbf{e}_i$, i.e.,

$$\Lambda_q(X) := \{X^{\mathsf{t}}\mathbf{z} + q\mathbf{y} \ : \ \mathbf{z} \in \mathbb{Z}^n, \mathbf{y} \in \mathbb{Z}^m\}.$$

3. The lattice $\Lambda_q^{\perp}(X) \subseteq \mathbb{R}^m$ is the dual of $\Lambda_q(X)$ scaled up by a factor of q, i.e.,

$$\Lambda_q^{\perp}(X) := \{\mathbf{v} \in \mathbb{R}^m : \forall\, \mathbf{u} \in \Lambda_q(X), \langle \mathbf{v}, \mathbf{u} \rangle \in q\mathbb{Z}\}.$$

We note that if X is of full row rank (over the integers), then $\Lambda^{\perp}(X)$ has rank $m-n$.

Definition 2.2 (Orthogonal projection)**.** Let L be a lattice and $E \subseteq \mathbb{R}^m$ be a vector subspace. The orthogonal projection of L onto E is:

$$\pi(L, E) = \{\mathbf{v}_1 \in E : \exists\ \mathbf{v}_2 \in E^{\perp}, \mathbf{v}_1 + \mathbf{v}_2 \in L\}.$$

Note that $\pi(L, E)$ is a finitely generated additive subgroup in $\mathbb{R}^m$, but not necessarily a lattice. The next lemma is standard (see, e.g., [CSV13, Lemma 3.4]).

Lemma 2.3. Let $E \subseteq \mathbb{R}^m$ be a vector space. For any lattice $L \in \mathbb{R}^m$ such that $\pi(L^{\star}, E)$ is a lattice, we have

$$L \cap E = \Big(\pi(L^{\star}, E)\Big)^{\star}.$$

Lemma 2.4. Let $X \in \mathbb{Z}^{n \times m}$ and $\Lambda^{\perp}(X)^{\star}$ be the dual lattice of $\Lambda^{\perp}(X)$. We have

$$\Lambda^{\perp}(X)^{\star} = (\mathbb{Z}^m + X^{\mathsf{t}}\mathbb{R}^n) \cap \ker(X).$$

Proof. Any $\mathbf{r} \in \pi(\mathbb{Z}^m, \ker(X))$ can be written as $\mathbf{r} = \mathbf{k} - X^{\mathsf{t}}(XX^{\mathsf{t}})^{-1}X\mathbf{k}$, for some $\mathbf{k} \in \mathbb{Z}^m$. It follows that $\pi(\mathbb{Z}^m, \ker(X)) \subseteq \frac{1}{\det(XX^{\mathsf{t}})} \cdot \mathbb{Z}^m$, hence $\pi(\mathbb{Z}^m, \ker(X))$ is a lattice. Now, we apply Lemma 2.3 with $L = \mathbb{Z}^m$ and $E = \ker(X)$ to obtain

$$\Lambda^{\perp}(X)^{\star} = (\mathbb{Z}^m \cap \ker(X))^{\star} = \pi(\mathbb{Z}^m, \ker(X)) = (\mathbb{Z}^m + X^{\mathsf{t}}\mathbb{R}^n) \cap \ker(X).$$

In the last equation, we use the fact that $(\ker(X))^{\perp} = X^{\mathsf{t}}\mathbb{R}^n$. □

2.2 Lattice Gaussian distributions and the smoothing parameter

For a rank-n matrix $S \in \mathbb{R}^{m \times n}$ and vector $\mathbf{c} \in \mathbb{R}^n$, the Gaussian function on $\mathbb{R}^n$ centered at $\mathbf{c}$ with covariance matrix $S^{\mathsf{t}}S$ is defined as:

$$\forall\, \mathbf{x} \in \mathbb{R}^n,\ \rho_{S,\mathbf{c}}(\mathbf{x}) = \exp(-\pi(\mathbf{x} - \mathbf{c})^{\mathsf{t}}(S^{\mathsf{t}}S)^{-1}(\mathbf{x} - \mathbf{c})).$$

Given a rank-d lattice $L \subset \mathbb{R}^m$, the discrete Gaussian distribution with support L, covariance parameter S and shift $\mathbf{c}$ is defined as:

$$\forall\, \mathbf{x} \in L, \mathcal{D}_{L,S,\mathbf{c}}(\mathbf{x}) = \frac{\rho_{S,\mathbf{c}}(\mathbf{x})}{\rho_{S,\mathbf{c}}(L)},$$

where $\rho_{S,\mathbf{c}}(L) = \sum_{\mathbf{x}\in L} \rho_{S,\mathbf{c}}(\mathbf{x})$. When $S = sI_n$ for some real $s > 0$, we write $\rho_{s,\mathbf{c}}$, resp. $D_{L,s,\mathbf{c}}$, the associated (spherical) function, resp. distribution, and we omit the subscript $\mathbf{c}$ when it is $\mathbf{0}$.

We will make use of the following tail bound for discrete Gaussians [Ban93]. This precise formulation is borrowed from [DRN14, Lemma 2.13].

Lemma 2.5. For any rank-d lattice L, $s > 0$ and $t \geqslant 1$, we have

$$\Pr_{\mathbf{v}\leftarrow D_{L,s}}\left[\|\mathbf{v}\| > s \cdot t\sqrt{\frac{d}{2\pi}}\right] \leqslant \exp\left(-\frac{d}{2}(t-1)^2\right).$$

We will use the following consequence of the Poisson summation formula:

$$\rho_S(\Lambda) = \det(\Lambda^\star) \cdot \sqrt{\det(S^{\mathsf{t}}S)} \cdot \rho_{S(S^{\mathsf{t}}S)^{-1}}(\Lambda^\star),$$

for any rank-d lattice Λ and any matrix $S \in \mathbb{R}^{m\times n}$ of rank n. The definition of the smoothing parameter is motivated by the Poisson summation formula. Given $\varepsilon > 0$ and a lattice L, the *smoothing parameter* $\eta_\varepsilon(L)$ is defined as the smallest real $s > 0$ such that $\rho_{1/s}(L^\star \setminus \{\mathbf{0}\}) \leqslant \varepsilon$. Introduced by Micciancio and Regev in 2004, the smoothing parameter has been used as a central tool in reductions between lattice problems [MR04] and in lattice-based cryptography. Also, the notion of smoothing parameter can be found in communication theory under the name 'flatness factor' [Bel11].

Chapter 3

On the smoothing parameter and last minimum of random orthogonal lattices

This chapter is based on a joint work with E. Kirshanova, D. Stehlé, and A. Wallet. It was published in Designs, Codes and Cryptography [KNSW20].

3.1 Introduction

In this work, we consider the successive minima and the smoothing parameter of random *orthogonal* lattices. For $X \in \mathbb{Z}^{n\times m}$ with $m > n$, the orthogonal lattice $\Lambda^\perp(X)$ is a set of all vectors $\mathbf{v} \in \mathbb{Z}^m$ that belong to the (right) kernel of X.

Given $X \in \mathbb{Z}^{n\times m}$, one can find a basis of $\Lambda^\perp(X)$ by a Hermite Normal Form computation (see, e.g., [HMM98]). Concretely: let $U \in \mathbb{Z}^{m\times m}$ be a unimodular transformation that brings X into HNF, i.e., $X \cdot U = X^{\mathtt{HNF}}$; if X is of full row-rank n, the last $m-n$ columns of $X^{\mathtt{HNF}}$ are zero vectors; viewing U as a block matrix $U = [U_1|U_2]$ for $U_2 \in \mathbb{Z}^{m\times(m-n)}$ of rank $m-n$, one can show that the columns of U_2 form a basis of $\Lambda^\perp(X)$. Similarly, Nguyen and Stern [NS97] show how to obtain a short basis of $\Lambda^\perp(X)$ by LLL-reducing [LLL82] the lattice spanned by the columns of $[cX^{\mathtt{t}}|I_m]^{\mathtt{t}}$ with some sufficiently large scalar c.

We study the $(m-n)$-th minimum and the smoothing parameter of the orthogonal rank-$(m-n)$ lattice $\Lambda^\perp(X)$, where each entry of X is independently and identically distributed according to an integer Gaussian distribution. In particular, we obtain probabilistic upper bounds on $\eta_\varepsilon(\Lambda^\perp(X))$ and $\lambda_{m-n}(\Lambda^\perp(X))$.

Prior results. With a motivation stemming from a cryptographic multilinear map construction [GGH13], Agrawal, Gentry, Halevi and Sahai in [AGHS13] considered the following variation: instead of starting from the finite set $\mathbb{Z}/q\mathbb{Z}$, they consider a matrix $X \in \mathbb{Z}^{n\times m}$ with entries sampled from an integer Gaussian distribution and focus on the closeness between the distribution of the vector $X\mathbf{z}$ and a discrete Gaussian distribution, for an appropriately chosen Gaussian multiplier $\mathbf{z}$ and conditioned over X. The main novelty was to replace the finite support $\mathbb{Z}/q\mathbb{Z}$ by the infinite support $\mathbb{Z}$. This question can be answered by considering the smoothing parameter of the lattice $\Lambda^\perp(X)$. Let us denote by $D_{\mathbb{Z}^n,S}$ the n-dimensional zero-centered discrete Gaussian distribution over $\mathbb{Z}^n$ with parameter a full column-rank matrix S with n columns (the probability of

a vector $\mathbf{k} \in \mathbb{Z}^n$ is proportional to $\exp(-\pi\|\mathbf{k}^{\mathsf{t}}(S^{\mathsf{t}}S)^{-1}\mathbf{k}\|)$ and by $D_{\mathbb{Z}^n,s}$ the case when $S = sI_n$. The following probabilistic bound is proved in [AGHS13]:

$$\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \mathcal{O}(mn\sqrt{\ln(m/\varepsilon)}),$$

where $X \leftarrow (D_{\mathbb{Z}^n,s})^m$, $s > \eta_\varepsilon(\mathbb{Z}^n)$ and $m = \Omega(n\ln(ns))$[1][2]. This statement holds with probability $\geqslant 1 - 2^{-\Omega(n)}$ over the choice of X. They obtain a LHL over lattices as a direct consequence of this result: for parameters satisfying the same conditions and if the above smoothing parameter bound holds, the statistical distance[3] between the distributions $X\mathbf{z}$ (conditioned on X) and $D_{\mathbb{Z}^m,s'X^{\mathsf{t}}}$ is at most ε, when $\mathbf{z}$ is sampled from $D_{\mathbb{Z}^m,s'}$ for $s' \geqslant \eta_\varepsilon(\Lambda^\perp(X))$.

Our objective here is to obtain a sufficient condition on s' which is as mild as possible. Note that improving the upper bound on $\eta_\varepsilon(\Lambda^\perp(X))$ directly leads to a milder condition on s'.

Following [AGHS13], Aggarwal and Regev [AR16] gave another bound on the smoothing parameter of $\Lambda^\perp(X)$. Namely, they showed that

$$\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \mathcal{O}\left(ns \cdot \sqrt{\ln(ns)\cdot\ln(m)\cdot\ln(m/\varepsilon)}\right),$$

with probability $\geqslant 1 - 2^{-\Omega(n)}$ over the choice of $X \leftarrow (D_{\mathbb{Z}^n,s})^m$. The bound holds for $s > \eta_\varepsilon(\mathbb{Z}^n)$ and $m = \Omega(n\ln(ns))$. The bound of [AR16] is lower for large m and small s, while the result of [AGHS13] is preferable for large s and small m.

Our results. Our first result is an improved upper bound on $\eta_\varepsilon(\Lambda^\perp(X))$.

Theorem 3.1. Let n be an integer growing to infinity, $\varepsilon > 0$, $s \geqslant \Omega(\sqrt{n})$ and $m = \Omega(n\ln s)$. Then, we have

$$\Pr_{X\leftarrow(D_{\mathbb{Z}^n,s})^m}\left[\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \mathcal{O}\left(\sqrt{(n+\ln m)\cdot\ln(m/\varepsilon)}\right)\right] \geqslant 1 - 2^{-\Omega(n)}.$$

Moreover, for any $\varepsilon \leqslant e^{-(m-n)}$,

$$\Pr_{X\leftarrow(D_{\mathbb{Z}^n,s})^m}\left[\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \mathcal{O}\left(\sqrt{\ln(1/\varepsilon)}\right)\right] \geqslant 1 - 2^{-\Omega(n)}.$$

Note that the second probabilistic upper bound is lower than the first, but is not applicable for every $\varepsilon > 0$. It holds for $\varepsilon \leqslant e^{-(m-n)}$, and, for larger values of ε, only the first bound applies. The reason why there are two possibilities for the upper bound stems from the two uncomparable relations between the smoothing parameter and the first minimum of the dual lattice $\Lambda^\perp(X)^\star$ for the Euclidean and infinity norms (see Lemmas 3.5 and 3.6).

A proof for this theorem can be found in Section 3.3. Our result improves over both bounds of [AGHS13] and [AR16] when $s = \Omega(\sqrt{n})$ and m is sufficiently large. In

[1]Note that an equivalent description of the distribution for X would be $X \leftarrow (D_{\mathbb{Z},s})^{n\times m}$. Our choice follows prior works.

[2]We recall that $\eta_\varepsilon(\mathbb{Z}^n) = \mathcal{O}(\sqrt{\ln(n/\varepsilon)})$ (see Section 3.2).

[3]The statistical distance between two distributions X and Y is half their ℓ_1-distance, i.e., $\Delta(X,Y) := \frac{1}{2}\|X - Y\|_1 = \frac{1}{2}\sum_{\omega\in\Omega}|X(\omega) - Y(\omega)|$.

particular, the minimum of our two upper bounds is smaller by an $\Omega(\sqrt{n})$ factor for the above ranges of m and s, is independent of s and depends at most logarithmically in m. However, our result requires $s = \Omega(\sqrt{n})$ (which is a consequence of our proof technique, in particular, Lemma 3.16), so for small values of s, the prior results of [AGHS13] and [AR16] remain the best known upper bounds on $\eta_\varepsilon(\Lambda^\perp(X))$. As an immediate corollary to Theorem 3.1, we obtain a tighter version the leftover hash lemma over lattices (see Corollary 3.12). We summarise our results and compare them with previous works in Table 3.1.

For applications, it is useful to keep in mind the following parameter set with respect to n: $\varepsilon = 2^{-\Theta(n)}$, $s = n^{\Theta(1)}$ and $m = \Theta(n \ln n)$. For these parameters, Theorem 3.1 yields $\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \widetilde{\mathcal{O}}(n)$ with probability $\geqslant 1 - 2^{-\Omega(n)}$. For the same parameters, the probabilistic bounds from [AGHS13] and [AR16] are $\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \widetilde{\mathcal{O}}(n^3)$ and $\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \widetilde{\mathcal{O}}(n^{3/2}s)$, respectively.

	Agrawal et al. [AGHS13]	Aggarwal-Regev [AR16]	[KNSW20]
s	$\Omega(\eta_\varepsilon(\mathbb{Z}^n))$	$\Omega(\eta_\varepsilon(\mathbb{Z}^n))$	$\Omega(\sqrt{n})$
m	$\Omega(n \ln(ns))$	$\Omega(n \ln(ns))$	$\Omega(n \ln s)$
$\eta_\varepsilon(\Lambda^\perp(X))$	$\mathcal{O}\left(mn\sqrt{\ln \frac{m}{\varepsilon}}\right)$	$\mathcal{O}\left(ns\sqrt{\ln(ns)\ln(m)\ln \frac{m}{\varepsilon}}\right)$	$\mathcal{O}\left(\sqrt{(n+\ln m)\cdot \ln \frac{m}{\varepsilon}}\right)$ or $\mathcal{O}\left(\sqrt{\ln \frac{1}{\varepsilon}}\right)$

Table 3.1: Probabilistic upper bounds on $\eta_\varepsilon(\Lambda^\perp(X))$ for $X \leftarrow (D_{\mathbb{Z}^n,s})^m$ together with requirements on s and m needed for the bounds to hold. The two cases in the last table entry depend on the range of ε, see Theorem 3.1.

.

Our second main result is an upper bound on the $(m-n)$-th minimum of the orthogonal lattice $\Lambda^\perp(X)$. Note that we could use our result of Theorem 3.1 to obtain an upper bound on $\lambda_{m-n}(\Lambda^\perp(X))$ via the relation $\lambda_{m-n}(\Lambda) \leqslant \sqrt{m-n} \cdot \eta_\varepsilon(\Lambda)$, which holds for any rank-$(m-n)$ lattice Λ and any $\varepsilon \in (0, 1/2)$) (see Lemma 3.7). Below we state the result, which gives a better bound for a large set of parameters. In particular, for many ranges of s and m, it improves over the bound $\lambda_{m-n}(\Lambda^\perp(X)) \leqslant \mathcal{O}(ns\sqrt{\ln(m)\ln(ns)})$ from [AR16], and the bound $\lambda_{m-n}(\Lambda^\perp(X)) \leqslant \mathcal{O}(mn)$ from [AGHS13].

Theorem 3.2. Let n be an integer growing to infinity, $s \geqslant \Omega(\sqrt{n})$ and m satisfying $m = \Omega(n \ln s)$ and $m \leqslant 2^{n/2}$. Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m}\left[\lambda_{m-n}(\Lambda^\perp(X)) \leqslant \mathcal{O}(n \ln s)\right] \geqslant 1 - 2^{-\Omega(n)}.$$

This theorem is proven in Section 3.4. An interesting fact to be noticed from this result is that, with overwhelming probability, the lattice $\Lambda^\perp(X)$ contains $(m-n)$ linearly independent vectors whose norms do not depend on m — a parameter which can be as large as $2^{n/2}$. On the other hand, our statement holds even when taking m as small as $\Theta(n \cdot \ln s)$. We summarise our results and compare them with previous works in Table 3.2.

	Agrawal et al. [AGHS13]	Aggarwal-Regev [AR16]	[KNSW20]
s	$\Omega(\eta_\varepsilon(\mathbb{Z}^n))$	$\Omega(\eta_\varepsilon(\mathbb{Z}^n))$	$\Omega(\sqrt{n})$
m	$\Omega(n\ln(ns))$	$\Omega(n\ln(ns))$	$\Omega(n\ln s)$ and $\leqslant 2^{n/2}$
$\lambda_{m-n}(\Lambda^\perp(X))$	$\mathcal{O}(mn)$	$\mathcal{O}\left(ns\sqrt{\ln(ns)\ln(m)}\right)$	$\mathcal{O}(n\ln s)$

Table 3.2: Probabilistic upper bounds on $\lambda_{m-n}(\Lambda^\perp(X))$ for $X \leftarrow (D_{\mathbb{Z}^n,s})^m$ together with requirements on s and m needed for the bounds to hold.

.

3.1.1 Techniques

The bound on the smoothing parameter is obtained via a chain of relations between successive minima of different lattices and the smoothing parameter of $\Lambda^\perp(X)$. Well-known transference relations between the smoothing parameter of a lattice and the first minimum of its dual lead us to study $\lambda_1(\Lambda^\perp(X)^\star)$. Namely, in order to obtain our result on $\eta_\varepsilon(\Lambda^\perp(X))$, we bound $\lambda_1(\Lambda^\perp(X)^\star)$ from below in both Euclidean and infinity norms.

To obtain these lower bounds, we consider the lattice $\Lambda_q(X) \subseteq \mathbb{Z}^m$ – the full-rank lattice spanned by the rows of X and $q\mathbb{Z}^m$ (in other words, we consider the so-called Construction A lattice of X, see [CS93, Chapter 5]). Following [AGHS13], our objective is to obtain a probabilistic lower bound on the norms of vectors from $\Lambda_q(X)\setminus X^\mathsf{t}\mathbb{Z}^n$. Note that this implies a probabilistic lower bound on $\lambda_{n+1}(\Lambda_q(X))$, as the vector space $X^\mathsf{t}\mathbb{Q}^n$ has dimension at most n.

At the heart of both proofs, ours and the one from [AGHS13], is a counting argument that allows to bound the norms of lattice vectors of the form $X^\mathsf{t}\mathbf{z} \bmod q \in \Lambda_q(X)\setminus X^\mathsf{t}\cdot\mathbb{Z}^n$. The counting argument is divided into several cases depending on the norm of $\mathbf{z}$. One source of improvement in our result is a more fine-grained division of these cases as well as a tighter treatment of interchange between different norms.

Once we have a lower bound on the norms of vectors from $\Lambda_q(X) \setminus X^\mathsf{t}\mathbb{Z}^n$, we relate the smallest norm of such vectors to $\lambda_1(\Lambda^\perp(X)^\star)$. This is where our approach differs most from [AGHS13]. The intuition behind the relation is the following: observe that for a sufficiently large q, the lattice $\frac{1}{q}\Lambda_q(X)$ can be thought of as an approximation to $\Lambda^\perp(X)^\star$, in the sense that any $\mathbf{u} \in \Lambda^\perp(X)^\star$ can be expressed as a vector in $\frac{1}{q}\Lambda_q(X)$ plus a small element in the row-span of X. Our lower bound on norms of vectors in $\Lambda_q(X) \setminus X^\mathsf{t}\cdot\mathbb{Z}^n$ and a Gaussian tail bound give a lower bound on $\lambda_1(\Lambda^\perp(X)^\star)$. This is in contrast to [AGHS13], which at this step invokes Banaszczyk's transference theorem [Ban93] in order to relate $\lambda_{n+1}(\Lambda_q(X))$ and $\lambda_{m-n}(\Lambda_q(X)^\star)$. Then, using the inclusion $\Lambda^\perp(X) \subseteq \frac{1}{q}\Lambda_q(X)^\star$, Agrawal et al. obtain their lower bound on $\lambda_1(\Lambda^\perp(X)^\star)$.

Summing up the above description, we:

1. obtain a (probabilistic) lower bound on $\|\mathbf{b}\|$ for $\mathbf{b} \in \Lambda_q(X) \setminus X^\mathsf{t}\mathbb{Z}^n$, improving the result of [AGHS13] by an $\Omega(n)$ factor (see Theorem 3.13);

2. relate the shortest norm of $\mathbf{b} \in \Lambda_q(X) \setminus X^\mathsf{t}\mathbb{Z}^n$ with $\lambda_1(\Lambda^\perp(X)^\star)$ (see Lemma 3.18);

3. apply known relations (Lemma 3.5 or Lemma 3.6, depending on the norm) between the first minimum of $\Lambda^\perp(X)^\star$ and $\eta_\varepsilon(\Lambda^\perp(X))$.

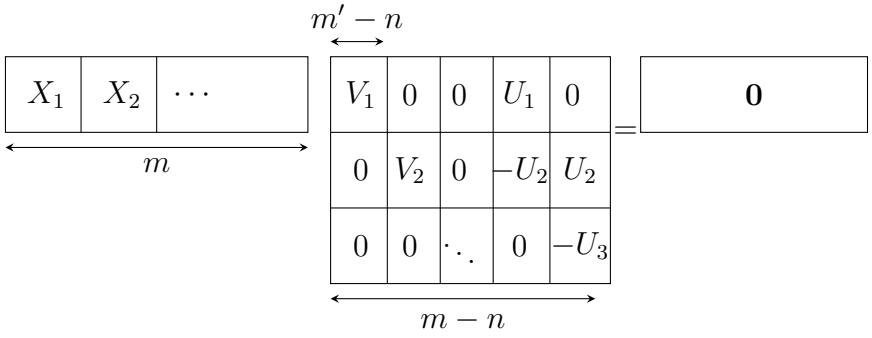

Figure 3.1: Technique used in bounding $\lambda_{m-n}(\Lambda^\perp(X))$. For each i, the matrix $V_i \in \mathbb{Z}^{m' \times m'-n}$ consists of $m'-n$ short linearly independent vectors orthogonal to X_i, the matrix consisting of the first m' columns of X. The vectors making up V_i are chosen so that they reach the first $(m'-n)$ successive minima of $\Lambda^\perp(X_i)$. Another $n\left(\frac{m}{m'}-1\right)$ short orthogonal to X vectors are obtained using matrices $U_i \in \mathbb{Z}^{m' \times n}$ that satisfy $X_i U_i = I_n$ and whose columns have small norms.

Our second result, stated in Theorem 3.2, gives an upper bound on the $(m-n)$-th minimum of $\Lambda^\perp(X)$. The main ingredient in the proof is an observation that we can subdivide, column-wise, a 'wide' matrix $X \in \mathbb{Z}^{n \times m}$ (here m is potentially much larger than n) into $\frac{m}{m'}$ smaller matrices $X_i \in \mathbb{Z}^{n \times m'}$, and obtain short vectors in each $\Lambda^\perp(X_i)$, which are also short vectors in $\Lambda^\perp(X)$. (For the sake of simplicity, we assume here that m' divides m.)

As a first step, we obtain an upper bound on $\lambda_{m'-n}(\Lambda^\perp(X_i))$ for all i. Such an upper bound on $\lambda_{m'-n}(\Lambda^\perp(X_i))$ is a corollary of our lower bound on $\lambda_{n+1}(\Lambda_q(X))$ and Banaszczyk's transference theorem [Ban93]. Thus, we obtain $\frac{m}{m'}(m'-n)$ relatively short vectors of dimension m'. Note that each such vector can be 'padded' with enough 0's in a way that the resulting m-dimensional vector belongs to $\Lambda^\perp(X)$. The latter is illustrated in Figure 3.1 as follows: the columns of each matrix V_i are linearly independent vectors reaching the minima of $\Lambda^\perp(X_i)$, the columns containing them in the center matrix in Figure 3.1 are short linearly independent vectors in $\Lambda^\perp(X)$.

The second step consists in obtaining $n\left(\frac{m}{m'}-1\right)$ additional short vectors (linearly independent with the previous ones), by applying a result due Aggarwal and Regev [AR16], which gives a probabilistic upper bound on the norm of the columns of a matrix $U \in \mathbb{Z}^{m \times n}$ such that $XU = I_n$. Hence, for each X_i, there exist $U_i \in \mathbb{Z}^{m' \times n}$ such that $X_i U_i = I_n$. Stacking the pairs $(U_i, -U_{i+1})$ as illustrated in Figure 3.1, we obtain the missing short vectors from $\Lambda^\perp(X)$.

3.1.2 Open problems

Our upper bound the last minimum of the lattice $\Lambda^\perp(X)$ improves the prior ones, but may not be the tightest possible. In fact, we suspect it is not sharp. The Gaussian matrix $X \leftarrow (D_{\mathbb{Z}^n,s})^m$ has rank n with overwhelming probability (see Lemma 3.9 below). Using the fact that $\det(\Lambda^\perp(X)) \leqslant \det(X^\mathsf{t}\mathbb{Z}^n)$ (see, e.g., [Ngu99, p. 30]) and Minkowski's

theorem, we have

$$\prod_{i\in[m-n]} \lambda_i(\Lambda^\perp(X)) \leqslant \sqrt{m-n}^{m-n} \cdot \det(X^\mathsf{t}\mathbb{Z}^n).$$

Then, by applying a Gaussian tail bound, all the columns of X^t have norms $\leqslant s\sqrt{m}$ with overwhelming probability. If we assume that the successive minima are essentially the same, we obtain from Hadamard's inequality that

$$\forall i \leqslant m-n, \quad \lambda_i(\Lambda^\perp(X)) \leqslant \sqrt{m-n} \cdot (s\sqrt{m})^{\frac{n}{m-n}}.$$

Consider now $m = \Theta(n \ln n)$ and $s \leqslant \mathrm{poly}(n)$, and assume the above inequality is essentially tight. Then it suggests that the minimum $\lambda_{m-n}(\Lambda^\perp(X))$ should be $\widetilde{\Theta}(\sqrt{n})$. However, Theorem 3.2 only states that $\lambda_{m-n}(\Lambda^\perp(X)) \leqslant \widetilde{\mathcal{O}}(n)$. This gap possibly stems from our counting arguments in Lemma 3.15 and Lemma 3.17. Indeed, we impose there that all points in an n-dimensional cube satisfy some property with a success probability that is exponentially close to 1. It could be the case that by weakening our constraints on the probabilities, e.g., by asking for a failure at most $n^{-\omega(1)}$, we could achieve sharper estimations for smaller parameters. However, it does not seem straightforward, because we also rely on union bounds over sets of exponential sizes.

On the other hand, for exponentially small ε, we also expect the smoothing parameter to be essentially the same as the minima. This would heuristically give $\eta_\varepsilon(\Lambda^\perp(X)) = \widetilde{\Theta}(\sqrt{n})$ when $m = \Theta(n \ln n)$ and $s \leqslant \mathrm{poly}(n)$. Theorem 3.1 provides a $\widetilde{\mathcal{O}}(\sqrt{n})$ bound for these parameters.

3.2 Preliminaries

For an integer $q > 2$, we use $[\mathbf{v}]_q$ to denote the modular reduction of all the entries of $\mathbf{v}$ into the interval $[-\frac{q}{2}, \frac{q}{2})$. The kernel of a matrix $X \in \mathbb{R}^{n\times m}$ seen as linear maps is denoted $\ker(X)$. For a vector subspace $V \subseteq \mathbb{R}^d$ and a vector $\mathbf{x} \in \mathbb{R}^d$, we let $\pi(\mathbf{x}, V)$ denote the orthogonal projection of $\mathbf{x}$ onto V. For two distributions D, D' over a common support Ω, their statistical distance is defined as $\Delta(D, D') = \frac{1}{2}\sum_{\omega\in\Omega} |D(\omega) - D'(\omega)|$. First, we need the following version of Hoeffding's inequality.

Lemma 3.3 (Hoeffding's inequality)**.** Let $X_1, \ldots, X_m$ be independent random variables such that $0 \leqslant X_i \leqslant 1$ for all i. Let $S_m = X_1 + \cdots + X_m$. Then for any $t > 0$, we have

$$\Pr\Big[|S_m - \mathbb{E}[S_m]| \geqslant t\Big] \leqslant 2\exp\Big(-2t^2/m\Big).$$

The following is a *transference theorem* as it allows to link the minima of a given lattice to those of its dual.

Theorem 3.4 ([Ban93])**.** For any rank-d lattice $L \subseteq \mathbb{R}^m$, and for all $i \in [d]$, we have

$$1 \leqslant \lambda_i(L) \cdot \lambda_{d-i+1}(L^\star) \leqslant d.$$

3.2.1 Bounds of the smoothing parameter

We recall below standard upper bounds on this parameter, involving lattice minima.

Lemma 3.5 ([Pei08, Lemma 3.5])**.** For any rank-d lattice L and $\varepsilon > 0$, we have

$$\eta_\varepsilon(L) \leqslant \frac{\sqrt{\ln(2d(1+\frac{1}{\varepsilon}))/\pi}}{\lambda_1^\infty(L^\star)}.$$

Lemma 3.6 ([PRS17, Lemma 2.6], [Ban93, Lemma 1.5])**.** For any rank-d lattice L and $\varepsilon \in (0, e^{-d}]$, we have

$$\eta_\varepsilon(L) \leqslant \frac{\sqrt{\ln(\frac{1}{\varepsilon})}}{\lambda_1(L^\star)}.$$

First, Lemma 3.5 and Lemma 3.6 differ in terms of the norm considered for the first minimum of the dual lattice. Second, these lemmas give different smoothing parameter bounds for different ε-regimes: depending on the smallness of ε, one of the lemmas may give a tighter statement than the other. In particular, in our results, we will obtain a probabilistic lower bound on $\lambda_1(L^\star)$ for a rank-d orthogonal lattice L that is larger than a lower bound on $\lambda_1^\infty(L^\star)$ by a factor quasi-linear in $\sqrt{d}$. It follows that for small ε (e.g., $\varepsilon = 2^{-o(d)}$), Lemma 3.6 is preferable to Lemma 3.5. When ε is large, Lemma 3.6 may not be applicable, whereas Lemma 3.5 still provides a bound.

The smoothing parameter of a lattice can alternatively be bounded using the last minimum of the (primal) lattice.

Lemma 3.7 ([APS18, Lemma 2.13] and [MR04, Lemma 3.3])**.** For any rank-d lattice L and $\varepsilon \in (0, 1/2)$, we have

$$\frac{\lambda_d(L)}{\sqrt{d}} \leqslant \eta_\varepsilon(L) \leqslant \lambda_d(L) \cdot \sqrt{\ln\left(2d\left(1+\frac{1}{\varepsilon}\right)\right)/\pi}.$$

3.2.2 Properties of smoothed Gaussians

The first lemma states that the Gaussian mass of a subset of a lattice L does not differ too much from the Gaussian mass of a small shift of it.

Lemma 3.8 ([AGHS13, Lemma 6])**.** Fix a rank-d lattice $L \subseteq \mathbb{R}^d$, $\varepsilon \in (0,1)$, $c > 2$, and $s \geqslant (1+c)\eta_\varepsilon(L)$. Then, for any subset $T \subseteq L$ and for any $\mathbf{v} \in L$, we have

$$D_{L,s}(T) - D_{L,s}(T-\mathbf{v}) \leqslant \frac{\mathrm{erf}(p(1+4/c)/2)}{\mathrm{erf}(2p)} \cdot \frac{1+\varepsilon}{1-\varepsilon},$$

where $p = \frac{\|\mathbf{v}\|\sqrt{\pi}}{s}$, and $\mathrm{erf}(\cdot)$ is the error function.

The following lemma implies that an integral lattice, generated by the columns sampled from a discrete Gaussian distribution over $\mathbb{Z}^n$, spans all of $\mathbb{Z}^n$ with overwhelming probability, if the standard deviation of this distribution is sufficiently large. It also provides information on the matrix that maps X to the canonical basis of $\mathbb{Z}^n$.

Lemma 3.9 (Adapted from [AR16, Lemma 4.2])**.** Let $n \geqslant 100$ and $\varepsilon \in (0, \frac{1}{1000})$. Further, let s, m be such that $s \geqslant 9\eta_\varepsilon(\mathbb{Z}^n)$, $m \geqslant 44n\ln(ns)$.Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m}\left[\exists U \in \mathbb{Z}^{m\times n} : XU = I_n \text{ and } \max_i \|\mathbf{u}_i\| \leqslant 2\sqrt{44n\ln(sn)}\right] \geqslant 1 - 2^{-n},$$

where the $\mathbf{u}_i$'s are the columns of U.

We now state the leftover hash lemma involving Gaussians over infinite domains, the topic of study of [AGHS13].

Lemma 3.10 ([AGHS13, Lemma 10])**.** Let $m > n \geqslant 1$ be integers and $\varepsilon \in (0, 1/3)$. Let $X \in \mathbb{Z}^{n\times m}$ such that the columns of X span all of $\mathbb{Z}^n$. If $s' \geqslant \eta_\varepsilon(\Lambda^\perp(X))$, then we have

$$\Delta(X \cdot D_{\mathbb{Z}^m,s'}, D_{\mathbb{Z}^m,s'X^\mathsf{t}}) \leqslant 2\varepsilon.$$

3.3 Smoothing parameter of the orthogonal lattice

This section is devoted to proving our first main result: a tighter upper bound on the smoothing parameter of $\Lambda^\perp(X)$, where the columns of $X \in \mathbb{Z}^{n\times m}$ are chosen from the discrete Gaussian $D_{\mathbb{Z}^n,s}$. In the rest of this chapter, we view all other parameters as functions of n. We stress that Theorem 3.11 below differs from Theorem 3.1 in that the asymptotic notations are made explicit by specifying the constants. In this section and the next, we keep these constants explicit. We do not claim that they are optimal in some sense: we provide them to help the reader follow the proofs.

Theorem 3.11. Let $n \geqslant 60$, $\varepsilon > 0$, $s \geqslant 20\sqrt{n}$, and $m \geqslant 1355n\ln s$. Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m}\left[\eta_\varepsilon(\Lambda^\perp(X)) \leqslant 77\sqrt{(n + \ln m)\cdot\ln(2m/\varepsilon)}\right] \geqslant 1 - 2^{-\Omega(n)}.$$

For any $\varepsilon \leqslant e^{-(m-n)}$, we also have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m}\left[\eta_\varepsilon(\Lambda^\perp(X)) \leqslant 96\sqrt{\ln(1/\varepsilon)}\right] \geqslant 1 - 2^{-\Omega(n)}.$$

We now give an informal description of the proof strategy. The first part of the proof is similar to the proof presented in [AGHS13]: we first embed the lattice $X^\mathsf{t}\mathbb{Z}^n$ into a full rank q-ary lattice $\Lambda_q(X)$ by "adding" all the vectors $q\mathbf{e}_i$ to $X^\mathsf{t}\mathbb{Z}^n$ (where the $\mathbf{e}_i$'s are the canonical basis vectors). If q is set sufficiently large, then the short vectors in $\Lambda_q(X)$ should come only from the rows of X (with overwhelming probability) and thus be common to the short vectors of $X^\mathsf{t}\mathbb{Z}^n$. Starting from this intuition, the authors of [AGHS13] provide a lower bound on the norms of the vectors *not* belonging to the row span of X. We improve their bound by an $\Omega(n)$ factor by using tighter arguments on several estimations during the proof. This lower bound also gives a lower bound on $\lambda_{n+1}(\Lambda_q(X))$ since $X^\mathsf{t}\mathbb{Q}^n$ spans an vector space of dimension at most n. We also observe that a lower bound on the infinity norms of vectors in $\Lambda_q(X) \setminus X^\mathsf{t}\mathbb{Z}^n$ can be derived from the proof, without relying on a loose norm equivalence.

The second part of the proof differs from the one of [AGHS13]: we observe that we can directly relate the $(n+1)$-th minimum of $\Lambda_q(X)$ to the first minimum of $\Lambda^\perp(X)^\star$.

This avoids relying twice on a transference argument, as in [AGHS13], which allows us to save another $\Omega(\sqrt{n})$ factor. The final result on the smoothing parameter is then a consequence of Lemmas 3.5 and 3.6.

As a direct corollary of Theorem 3.11, we obtain the following leftover hash lemma over lattices.

Corollary 3.12. Let $n \geqslant 100$, $\varepsilon \in \left(0, \frac{1}{1000}\right)$, $s \geqslant 20\sqrt{n}$ and $m > 1355n \ln s$. Let $s' \geqslant 77\sqrt{(n + \ln m) \cdot \ln(2m/\varepsilon)}$. Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m} \left[\Delta(X \cdot D_{\mathbb{Z}^m,s'}, D_{\mathbb{Z}^m,s'X^{\mathsf{t}}}) \leqslant 2\varepsilon\right] \geqslant 1 - 2^{-\Omega(n)}.$$

If moreover $\varepsilon \leqslant e^{-(m-n)}$, then the same result holds with $s' \geqslant 96\sqrt{\ln(1/\varepsilon)}$.

Proof. Using Lemma 3.9 with the parameters as in the statement, the columns of X span $\mathbb{Z}^n$ with probability $1 - 2^{-\Omega(n)}$. Now, from Theorem 3.11, these parameters also ensure that with probability at least $1 - 2^{-\Omega(n)}$, we have $\eta_\varepsilon(\Lambda^\perp(X)) \leqslant s'$. Finally, Lemma 3.10 states that when the columns of X span $\mathbb{Z}^n$ and s' is chosen such that $\eta_\varepsilon(\Lambda^\perp(X)) \leqslant s'$, we have $\Delta(X \cdot D_{\mathbb{Z}^m,s'}, D_{\mathbb{Z}^m,s'X^{\mathsf{t}}}) \leqslant 2\varepsilon$. □

3.3.1 Short vectors in the Construction A lattice of a Gaussian matrix

This section deals with short vectors in $\Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n$. More precisely, we prove the following theorems.

Theorem 3.13. Let $n \geqslant 60$, $q \geqslant 2$, $m \geqslant 335n \ln q$ be integers, and $s \geqslant 20\sqrt{n}$. Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m} \left[\exists \mathbf{b} \in \Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n : \|\mathbf{b}\| < \frac{q}{48}\right] \leqslant 2^{-\Omega(n)}.$$

As the vector space $X^{\mathsf{t}}\mathbb{Q}^n$ has dimension at most n, the above also gives a lower bound to $\lambda_{n+1}(\Lambda_q(X))$. We are also able to obtain a similar statement for the infinity norm. This result does not follow from just using the equivalence of norms and Theorem 3.13.

Theorem 3.14. Let $n \geqslant 7, q \geqslant 2$, $m \geqslant 20n \ln q$ be integers, and $s \geqslant 20\sqrt{n}$. Then, we have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m} \left[\exists \mathbf{b} \in \Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n : \|\mathbf{b}\|_\infty < \frac{q}{48\sqrt{n + \ln m}}\right] \leqslant 2^{-\Omega(n)}.$$

For the sake of readability, we split the proofs into several lemmas. The theorems follow from these lemmas and their proofs are given at the end of this subsection. We now give an overview of the proofs of the lemmas.

Recall that $\Lambda_q(X)$ is a lattice spanned by the rows of X and the vectors $q\mathbf{e}_i$.In particular, it contains the integer span of the rows of X, which is of dimension at most n. The purpose of the following lemmas is to prove that every vector in $\Lambda_q(X)$ that is not in the linear span of the rows of X, is of Euclidean norm $\Omega(q)$. In order to show this, we look at the vectors of the form $[X^{\mathsf{t}}\mathbf{z}]_q \in \Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n$ for $\mathbf{z} \in \mathbb{Z}^n$. This is

indeed enough as any vector in $\Lambda_q(X)$ can be written $X^{\mathsf{t}}\mathbf{z}+q\mathbf{y}$ for some $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ and $\mathbf{y} \in \mathbb{Z}^m$. To obtain a lower bound on the norms of such vectors, we divide the proof into two cases depending on the norm of $\mathbf{z}$.

In the first lemma, we prove for a "short" $\mathbf{z}$ that, with all but probability $2^{-\Omega(n)}$, the vector $[X^{\mathsf{t}}\mathbf{z}]_q$ belongs to the row-span of X. This part of our proof differs from the one of [AGHS13] as we bypass norm equivalence between Euclidean and infinity norms. The second lemma deals with the other ranges of $\mathbf{z}$: we obtain a lower bound on the entries of $[X^{\mathsf{t}}\mathbf{z}]_q$ by first proving a probabilistic lower bound on $[\langle \mathbf{x}\,, \mathbf{z}\rangle]_q$ taken over a Gaussian vector $\mathbf{x}$. For a "large" $\mathbf{z}$, the proof is identical to the proof of [AGHS13]. This is detailed in the proof of Lemma 3.16. Finally, we extend this argument from a vector $\mathbf{x}$ to a matrix X using Hoeffding's inequality. This part of the proof is also new.

Lemma 3.15. Let $m, n, q \geqslant 2$ be integers. Then we have

$$\Pr\left[\exists\, \mathbf{z} \in \mathbb{Z}^n \text{ with } \|\mathbf{z}\| < \frac{q}{4s\sqrt{n + \ln m}} \,:\, [X^{\mathsf{t}}\mathbf{z}]_q \in \Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n\right] \leqslant 2^{-n},$$

where the probability is taken over $X \leftarrow (D_{\mathbb{Z}^n,s})^m$.

Proof. Each row $\mathbf{x}_i$ is distributed as $D_{\mathbb{Z}^n,s}$. Let $t = \sqrt{2\pi(1 + (\ln m)/n)} + 1$. Lemma 2.5 gives that $\Pr_X[\|\mathbf{x}_i\| > st\sqrt{n/2\pi}] \leqslant 2^{-n}/m$. When this does not occur, we have, for any integer vector $\mathbf{z}$ with $\|\mathbf{z}\| \leqslant \frac{q}{4s\sqrt{n+\ln m}} \leqslant \frac{q\sqrt{2\pi}}{2st\sqrt{n}}$:

$$|\langle \mathbf{z}, \mathbf{x}_i\rangle| \leqslant \|\mathbf{z}\| \cdot \|\mathbf{x}_i\| < \frac{q}{2}.$$

The result follows by union bound over $i \in [m]$. □

We now consider longer $\mathbf{z}$'s. We show that the probability that their inner product with a Gaussian vector is quite smaller than q is bounded away from 1 by a constant.

Lemma 3.16. Let $m, n \geqslant 7$ and $q \geqslant 2$ be integers and $s \geqslant 20\sqrt{n}$. For any $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ such that $\|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}}$, we have

$$\Pr_{\mathbf{x}\leftarrow D_{\mathbb{Z}^n,s}}\left[|[\langle \mathbf{x}, \mathbf{z}\rangle]_q| < \frac{q}{48\sqrt{n + \ln m}}\right] \leqslant 0.95.$$

We first outline the main ideas of the proof. For a fixed $\mathbf{z}$, our concern is the vectors $\mathbf{x} \in \mathbb{Z}^n$ whose inner-products with $\mathbf{z}$ are "small" when reduced modulo q: they lead to vectors in the lattice $X^{\mathsf{t}}\mathbb{Z}^n$ that are shorter than what we would expect. Thus, we shall call them "Bad$_\mathbf{z}$" vectors. Then we show that a suitably chosen translation maps any "Bad$_\mathbf{z}$" vector $\mathbf{x}$ to a "Good$_\mathbf{z}$" vector $\mathbf{x}'$, such that the inner-product between $\mathbf{x}'$ and $\mathbf{z}$ is "large". This proof technique is borrowed from [AGHS13]; however, we refine it by splitting the ranges of $\|\mathbf{z}\|$ further and finding a better translation map for medium $\|\mathbf{z}\|$. In either case, the translation vectors turn out to be short enough to argue that the probabilities of sampling a "Bad$_\mathbf{z}$" $\mathbf{x}$ and a "Good$_\mathbf{z}$" $\mathbf{x}$ are relatively close. From there, we readily obtain an upper bound on the probability that $\mathbf{x}$ is "Bad$_\mathbf{z}$". Below we quantify the terms "large" and "short", "Bad$_\mathbf{z}$" and "Good$_\mathbf{z}$", and provide formal arguments.

Proof. Fix a $\mathbf{z}$ as in the statement and define the set of "$\mathrm{Bad}_{\mathbf{z}}$" vectors as

$$\mathrm{Bad}_{\mathbf{z}} := \left\{ \mathbf{x} \in \mathbb{Z}^n : |[\langle \mathbf{z}, \mathbf{x} \rangle]_q| < \frac{q}{48\sqrt{n + \ln m}} \right\}.$$

We also define $\mathrm{Good}_{\mathbf{z}} = \mathbb{Z}^n \setminus \mathrm{Bad}_{\mathbf{z}}$, i.e., vectors outside the set $\mathrm{Bad}_{\mathbf{z}}$ will be considered as "good".

Case 1: "Medium" $\mathbf{z}$, i.e., $\frac{q}{4s\sqrt{n+\ln m}} \leqslant \|\mathbf{z}\| < \frac{q}{2s}$.

We let $\mu = \lceil \frac{s}{6\sqrt{n}} \rceil$. If $\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}$, then we can obtain a $\mathrm{Good}_{\mathbf{z}}$ vector using the injective map

$$\begin{array}{ccc} \mathrm{Bad}_{\mathbf{z}} & \longrightarrow & \mathrm{Good}_{\mathbf{z}} \\ \mathbf{x} & \longmapsto & \mathbf{x} + \mu \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor, \end{array}$$

where the ceiling is taken coordinate-wise. Now, we show that the map indeed sends $\mathrm{Bad}_{\mathbf{z}}$ to $\mathrm{Good}_{\mathbf{z}}$. First, we note that $\langle \mathbf{z}, \lceil 2\mathbf{z}\sqrt{n}/\|\mathbf{z}\| \rceil \rangle \geqslant 0$, because a and $\lceil a \rceil$ have the same sign for any $a \in \mathbb{R}$. Also, by the choice of μ, we have

$$\begin{aligned} 0 \leqslant \mu \left\langle \mathbf{z}, \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \right\rangle &\leqslant \mu \sum_{i \in [n]} \left(|z_i| \cdot \frac{2|z_i|\sqrt{n}}{\|\mathbf{z}\|} + |z_i| \right) \\ &\leqslant 2\frac{s}{6\sqrt{n}}(2\|\mathbf{z}\|\sqrt{n} + \|\mathbf{z}\|\sqrt{n}) \\ &< \frac{q}{2}, \end{aligned}$$

where for the second inequality we use the fact that $s > 6\sqrt{n}$, and, for the last inequality, the fact that $\|\mathbf{z}\| < \frac{q}{2s}$. Combining this with the fact that $|[a+b]_q| \geqslant |[a]_q| - |[b]_q|$ for all $a, b \in \mathbb{R}$, we obtain for $\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}$ that

$$\left| \left[\left\langle \mathbf{z}, \mathbf{x} + \mu \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \right\rangle \right]_q \right| \geqslant \mu \left\langle \mathbf{z}, \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \right\rangle - |[\langle \mathbf{z}, \mathbf{x} \rangle]_q|.$$

Since $\|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}}$ and $|[\langle \mathbf{z}, \mathbf{x} \rangle]_q| \leqslant \frac{q}{48\sqrt{n+\ln m}}$, we have

$$\begin{aligned} \mu \left\langle \mathbf{z}, \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \right\rangle - |[\langle \mathbf{z}, \mathbf{x} \rangle]_q| &\geqslant \mu \sum_{i \in [n]} \left(|z_i| \cdot \frac{2|z_i|\sqrt{n}}{\|\mathbf{z}\|} - |z_i| \right) - \frac{q}{48\sqrt{n + \ln m}} \\ &\geqslant \frac{s}{6\sqrt{n}}(2\|\mathbf{z}\|\sqrt{n} - \|\mathbf{z}\|\sqrt{n}) - \frac{q}{48\sqrt{n + \ln m}} \\ &\geqslant \frac{q}{48\sqrt{n + \ln m}}. \end{aligned}$$

This implies that $\mathbf{x} + \mu \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \in \mathrm{Good}_{\mathbf{z}}$.

Now, we want to apply Lemma 3.8 with $\mathbf{v} = \mu \lceil 2\mathbf{z}\sqrt{n}/\|\mathbf{z}\| \rceil$. For this, we bound $\|\mathbf{v}\|$ from above. Using that $\lceil a \rceil^2 \leqslant (|a| + 1)^2$ for any $a \in \mathbb{R}$, we have

$$\left\| \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rfloor \right\|^2 \leqslant \sum_{i \in [n]} \left(\frac{2\sqrt{n}|z_i|}{\|\mathbf{z}\|} + 1 \right)^2 = \left\| \frac{2\sqrt{n}|\mathbf{z}|}{\|\mathbf{z}\|} + \mathbf{1}_n \right\|^2,$$

where $|\mathbf{z}| = (|z_1|, \ldots, |z_n|)$. This gives us

$$\|\mathbf{v}\| = \mu \left\| \left\lceil \frac{2\mathbf{z}\sqrt{n}}{\|\mathbf{z}\|} \right\rceil \right\| \leqslant \mu \left(\left\| \frac{2\sqrt{n}|\mathbf{z}|}{\|\mathbf{z}\|} \right\| + \|\mathbf{1}_n\| \right) \leqslant \frac{2s}{6\sqrt{n}} \cdot 3\sqrt{n} = s.$$

Now, we apply Lemma 3.8 with parameters $L = \mathbb{Z}^n$, $\varepsilon = 1/1000$, $c = 14$, $\mathbf{v} = \mu\lceil 2\mathbf{z}\sqrt{n}/\|\mathbf{z}\| \rceil$ and $p = \|\mathbf{v}\|\sqrt{\pi}/s \leqslant \sqrt{\pi}$. The assumption of Lemma 3.8 is indeed satisfied for these parameters. This gives that

$$\Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}] - \Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Good}_{\mathbf{z}}] \leqslant \frac{\mathrm{erf}(p(1+\frac{4}{c})/2)}{\mathrm{erf}(2p)} \cdot \frac{1+\varepsilon}{1-\varepsilon} \leqslant 0.9.$$

Since we always have that

$$\Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}] + \Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Good}_{\mathbf{z}}] = 1,$$

it holds that

$$\Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}] \leqslant \frac{1+0.9}{2}.$$

We conclude that

$$\Pr_{\mathbf{x}}\left[|[\langle \mathbf{x}, \mathbf{z} \rangle]_q| < \frac{q}{48\sqrt{n + \ln m}} \right] \leqslant 0.95.$$

Case 2: "Long" $\mathbf{z}$, i.e., $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ and $\|\mathbf{z}\| \geqslant \frac{q}{2s}$.

This part of the proof is the same as in [AGHS13, Lemma 11]. We reproduce it for the sake of completeness. Consider a "long" $\mathbf{z}$, i.e., $\|\mathbf{z}\| \geqslant \frac{q}{2s}$, which implies $\|\mathbf{z}\|_\infty \geqslant \frac{q}{2s\sqrt{n}}$. We modify the mapping defined in Case 1 from $\mathrm{Bad}_{\mathbf{z}}$ to $\mathrm{Good}_{\mathbf{z}}$ vectors by letting $\mu := \min\{\lceil s \rceil, \lfloor \frac{q}{2\|\mathbf{z}\|_\infty} \rfloor\}$ and defining:

$$\begin{aligned} \mathrm{Bad}_{\mathbf{z}} &\to \mathrm{Good}_{\mathbf{z}} \\ \mathbf{x} &\mapsto \mathbf{x} + \mu \mathbf{e}_{i_{\max}}, \end{aligned}$$

where $i_{\max}$ is the index of a largest entry in $\mathbf{z}$ (in absolute value).

We now prove that the map indeed sends $\mathrm{Bad}_{\mathbf{z}}$ to $\mathrm{Good}_{\mathbf{z}}$. We have $\mu\|\mathbf{z}\|_\infty \leqslant \frac{q}{2\|\mathbf{z}\|_\infty}\|\mathbf{z}\|_\infty \leqslant \frac{q}{2}$. Therefore, it holds that

$$|[\langle \mathbf{z}, \mathbf{x} + \mu \mathbf{e}_{i_{\max}} \rangle]_q| = |[\langle \mathbf{z}, \mathbf{x} \rangle \pm \mu\|\mathbf{z}\|_\infty]_q| \geqslant \mu\|\mathbf{z}\|_\infty - |[\langle \mathbf{z}, \mathbf{x} \rangle]_q|.$$

First, assume that $\mu = \lceil s \rceil$. Using the facts that $\|\mathbf{z}\|_\infty > \frac{q}{2s\sqrt{n}}$ and $|[\langle \mathbf{z}, \mathbf{x} \rangle]_q| < \frac{q}{48\sqrt{n+\ln m}}$ for $\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}$, we obtain

$$\mu\|\mathbf{z}\|_\infty - |[\langle \mathbf{z}, \mathbf{x} \rangle]_q| > s\frac{q}{2s\sqrt{n}} - \frac{q}{48\sqrt{n + \ln m}} > \frac{q}{48\sqrt{n + \ln m}}.$$

Now, assume that $\mu = \left\lfloor \frac{q}{2\|\mathbf{z}\|_\infty} \right\rfloor$. If $\|\mathbf{z}\|_\infty \leqslant \frac{q}{6}$, then we have $\mu = \left\lfloor \frac{q}{2\|\mathbf{z}\|_\infty} \right\rfloor \geqslant \frac{q}{2\|\mathbf{z}\|_\infty} - 1 \geqslant \frac{q}{2\|\mathbf{z}\|_\infty} - \frac{q}{6\|\mathbf{z}\|_\infty} > \frac{q}{6\|\mathbf{z}\|_\infty}$, otherwise if $\frac{q}{6} < \|\mathbf{z}\|_\infty \leqslant \frac{q}{2}$, then $\mu = \left\lfloor \frac{q}{2\|\mathbf{z}\|_\infty} \right\rfloor \geqslant 1 \geqslant \frac{q}{6\|\mathbf{z}\|_\infty}$. Hence $\mu \geqslant \frac{q}{6\|\mathbf{z}\|_\infty}$ in this case. For $\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}$, it implies that

$$\mu\|\mathbf{z}\|_\infty - |[\langle \mathbf{z}, \mathbf{x} \rangle]_q| > \frac{q}{6\|\mathbf{z}\|_\infty} \cdot \|\mathbf{z}\|_\infty - \frac{q}{48\sqrt{n + \ln m}} > \frac{q}{48\sqrt{n + \ln m}}.$$

In both cases, we have $|[\langle \mathbf{z}, \mathbf{x} + \mu \mathbf{e}_{i_{\max}} \rangle]_q| > \frac{q}{48\sqrt{n+\ln m}}$.

We apply Lemma 3.8 with parameters $L = \mathbb{Z}^n$, $\varepsilon = 1/1000, c = 35$, and $\mathbf{v} = \mu \mathbf{e}_{i_{\max}}$. The assumption of Lemma 3.8 is indeed satisfied for these parameters. Note that $\|\mathbf{v}\| = \mu < s + 1$, and $p := \frac{\|\mathbf{v}\|\sqrt{\pi}}{s} < \frac{s+1}{s}\sqrt{\pi} < \frac{20\sqrt{n}+1}{20\sqrt{n}}\sqrt{\pi} < 1.02\sqrt{\pi}$. Similarly to the previous case it follows that $\Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Bad}_{\mathbf{z}}] - \Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Good}_{\mathbf{z}}] \leqslant 0.9$. Hence, we obtain $\Pr_{\mathbf{x}}[\mathbf{x} \in \mathrm{Bad}] \leqslant 0.95$. □

Using Lemma 3.16 and Hoeffding's bound, we can now show that with overwhelming probability over the choice of X, there are more than $n + \ln m$ entries of $[X^{\mathsf{t}}\mathbf{z}]_q$ that have magnitude larger than $\frac{q}{48\sqrt{n+\ln m}}$ for any not too short $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$. This implies the following result.

Lemma 3.17. Let $n \geqslant 60$, $q \geqslant 2$, $m \geqslant 335 n \ln q$ be integers, and $s \geqslant 20\sqrt{n}$.Then, we have

$$\Pr\left[\forall\, \mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n \text{ with } \|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}} : \|[X^{\mathsf{t}}\mathbf{z}]_q\| \geqslant \frac{q}{48}\right] > 1 - 2 \cdot 2^{-0.001m},$$

where the probability is taken over $X \leftarrow (D_{\mathbb{Z}^n,s})^m$.

Proof. Fix $\mathbf{z}$ with $\|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}}$. For $i \in [m]$, consider independent binary random variables Y_i, defined over the choice of the columns $\mathbf{x}_i$ of X:

$$\begin{cases} Y_i = 1 \text{ if } |[\langle \mathbf{x}_i, \mathbf{z} \rangle]_q| \geqslant \frac{q}{48\sqrt{n+\ln m}}, \\ Y_i = 0 \text{ otherwise.} \end{cases}$$

From Lemma 3.16, it follows that $\Pr_X[Y_i = 1] \geqslant 0.05$. Therefore by linearity of expectation, we have $\mathbb{E}\left[\sum_i Y_i\right] \geqslant 0.05m$. Using Hoeffding's bound (Lemma 3.3) with $t = 0.05m - (n + \ln m)$, we obtain

$$\begin{aligned} \Pr\left[|\sum_i Y_i - \mathbb{E}[\sum_i Y_i]| \geqslant 0.05m - (n + \ln m)\right] \\ \leqslant 2\exp\left(-2\frac{(0.05m - (n+\ln m))^2}{m}\right). \end{aligned}$$

Hence, for $m \geqslant 200(n + \ln m)$ (which is implied by the condition $m \geqslant 335 n \ln q$), we have

$$\begin{aligned} \Pr_X\left[\sum_i Y_i < n + \ln m\right] &\leqslant \Pr_X\left[0.05m - \sum_i Y_i \geqslant 0.05m - (n + \ln m)\right] \\ &\leqslant 2\exp\left(-2\frac{(0.05m - (n+\ln m))^2}{m}\right) \\ &\leqslant 2\exp(-0.004m). \end{aligned}$$

The inequality above holds for any $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ with $\|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}}$. Using the lower bound on m, $m \geqslant 335 n \ln q$, we conclude that

$$\Pr\left[\exists\, \mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n \text{ with } \|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}} : \sum_i Y_i < n + \ln m\right]$$

$$< 2q^n \cdot e^{-0.004m}$$
$$< 2 \cdot e^{-0.001m}.$$

Since $\sum_i Y_i \geqslant n + \ln m$ implies that $\|[X^{\mathsf{t}}\mathbf{z}]_q\| \geqslant \frac{q}{48}$, the result follows. □

We are now in a position to prove our first main results.

Proof of Theorem 3.13. The choice of parameters allows us to use both Lemma 3.15 and Lemma 3.17. Their combination tells us that, with all but probability $2^{-\Omega(n)}$ over the choice of X, there does not exist any vector $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ for which $[X^{\mathsf{t}}\mathbf{z}]_q \notin X^{\mathsf{t}}\mathbb{Z}^n$ and $\|[X^{\mathsf{t}}\mathbf{z}]_q\| < \frac{q}{48}$. This gives the first result. □

Proof of Theorem 3.14. We show that if $\mathbf{v} \in \Lambda_q(X) \setminus X^{\mathsf{t}}\mathbb{Z}^n$, then $\|\mathbf{v}\|_\infty \geqslant \frac{q}{48\sqrt{n+\ln m}}$ with overwhelming probability. For $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ and $\|\mathbf{z}\| \geqslant \frac{q}{4s\sqrt{n+\ln m}}$, we have from Lemma 3.16 that $\Pr_{\mathbf{x}}[|[\langle \mathbf{x}, \mathbf{z}\rangle]_q| < \frac{q}{48\sqrt{n+\ln m}}] \leqslant 0.95$. It follows that

$$\Pr_X\left[\|[X^{\mathsf{t}}\mathbf{z}]_q\|_\infty < \frac{q}{48\sqrt{n+\ln m}}\right] \leqslant 0.95^m \leqslant e^{-0.05m}.$$

By the union bound, we obtain

$$\Pr_X\left[\exists\, \mathbf{z} : \|[X^{\mathsf{t}}\mathbf{z}]_q\|_\infty < \frac{q}{48\sqrt{n+\ln m}}\right] \leqslant q^n \cdot e^{-0.05m} = 2^{-\Omega(n)}.$$

Combining the above with Lemma 3.15, we conclude that with all but probability $2^{-\Omega(n)}$ over the choice of X, there does not exist any vector $\mathbf{z} \in \mathbb{Z}^n \cap [-\frac{q}{2}, \frac{q}{2})^n$ for which $[X^{\mathsf{t}}\mathbf{z}]_q \notin X^{\mathsf{t}}\mathbb{Z}^n$ and $\|[X^{\mathsf{t}}\mathbf{z}]_q\|_\infty < \frac{q}{48\sqrt{n+\ln m}}$. This completes the proof. □

3.3.2 Using the dual of $\Lambda^\perp(X)$

We want to find an upper bound on the smoothing parameter of $\Lambda^\perp(X)$. Using Lemma 3.5, such a bound comes from a lower bound on the minimum of the dual lattice of $\Lambda^\perp(X)$. We now relate the $(n+1)$-th minimum of the lattice $\Lambda_q(X)$ and the norms of the shortest vectors in $\Lambda^\perp(X)^\star$. For the proof below, it is useful to recall that $\Lambda^\perp(X)^\star = (\mathbb{Z}^m + X^{\mathsf{t}}\mathbb{R}^n) \cap \ker(X)$, as showed in Lemma 2.4.

Lemma 3.18. Let $n \geqslant 60$ and $s \geqslant 20\sqrt{n}$. Let q and m be integers satisfying $m \geqslant 335n \ln q$ and that $q \geqslant 96sn\sqrt{m}$ We have

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n, s})^m}\left[\lambda_1^\infty(\Lambda^\perp(X)^\star) \geqslant \frac{1}{96\sqrt{n+\ln m}}\right] \geqslant 1 - 2^{-\Omega(n)},$$
$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n, s})^m}\left[\lambda_1(\Lambda^\perp(X)^\star) \geqslant \frac{1}{96}\right] \geqslant 1 - 2^{-\Omega(n)}.$$

Proof. Let $\mathbf{u}$ be any vector in $\Lambda^\perp(X)^\star$. From Lemma 2.4, we can write $\mathbf{u} = \mathbf{k} + X^{\mathsf{t}}\mathbf{y}$, for some $\mathbf{k} \in \mathbb{Z}^m$ and $\mathbf{y} \in \mathbb{R}^n$. Let $\mathbf{z} \in \mathbb{R}^n$ with $\|\mathbf{z}\|_\infty < \frac{1}{q}$ such that $\mathbf{y} = \mathbf{y}' + \mathbf{z}$ and $\mathbf{y}' \in \frac{1}{q}\mathbb{Z}^n$. Thus, we can write $\mathbf{u} = \mathbf{v} + X^{\mathsf{t}}\mathbf{z}$, where $\mathbf{v} = \mathbf{k} + X^{\mathsf{t}}\mathbf{y}' \in \frac{1}{q}\Lambda_q(X)$.

Assume now that $\mathbf{u}$ is a non-zero vector of $\Lambda^\perp(X)^\star$. We show by contradiction that $\mathbf{v}$ cannot be in the row-span of X. Assume on the contrary that $\mathbf{v} \in X^\intercal \mathbb{Q}^n$. Then, on the one hand, this implies that $\mathbf{u} \in X^\intercal \mathbb{R}^n = (\ker X)^\perp$. On the other hand, we have $\mathbf{u} \in \ker X$ by definition of $\Lambda^\perp(X)^\star$. Then we must have $\mathbf{u} = \mathbf{0}_m$, which contradicts the choice of $\mathbf{u}$ as a non-zero vector of $\Lambda^\perp(X)^\star$. In particular, from Theorems 3.13 and 3.14,we see that $\|\mathbf{v}\|_\infty \geqslant \frac{1}{48\sqrt{n+\ln m}}$ and $\|\mathbf{v}\| \geqslant \frac{1}{48}$ with probability at least $1 - 2^{-\Omega(n)}$.

Now, we let $\mathbf{u}$ be such that $\|\mathbf{u}\|_\infty = \lambda_1^\infty(\Lambda^\perp(X)^\star)$ and we compare it to $\|\mathbf{v}\|_\infty$, for $\mathbf{v}$ defined as above. Applying Lemma 2.5 with $t = \sqrt{2\pi}$, we obtain that with probability greater than $1 - 2^{-n}$, the rows of $X^\intercal$ have Euclidean norms smaller than $s\sqrt{n}$. It follows that $\|X^\intercal \mathbf{z}\|_\infty \leqslant \max(\|\mathbf{x}_i\|) \cdot \|\mathbf{z}\| \leqslant \frac{sn}{q}$. By the triangular inequality, we have $\|\mathbf{v}\|_\infty \leqslant \|\mathbf{u}\|_\infty + \|X^\intercal \mathbf{z}\|_\infty$, from which we deduce that

$$\|\mathbf{u}\|_\infty \geqslant \|\mathbf{v}\|_\infty - \frac{sn}{q}$$

with all but probability at most 2^{-n}. We then deduce using Theorem 3.14 and the assumptions on m and q that

$$\|\mathbf{v}\|_\infty - \frac{sn}{q} \geqslant \frac{1}{48\sqrt{n+\ln m}} - \frac{sn}{q} \geqslant \frac{1}{96\sqrt{n+\ln m}},$$

also with probability greater than $1 - 2^{-\Omega(n)}$.

Let now $\mathbf{u}$ be such that $\|\mathbf{u}\| = \lambda_1(\Lambda^\perp(X)^\star)$, and $\mathbf{v}$ be as defined above. By norm equivalence, we have $\|X^\intercal \mathbf{z}\| \leqslant \sqrt{m}\|X^\intercal \mathbf{z}\|_\infty \leqslant \frac{sn\sqrt{m}}{q}$ except with probability at most 2^{-n}. As above, we deduce that $\|\mathbf{u}\| \geqslant \|\mathbf{v}\| - \frac{sn\sqrt{m}}{q}$. Using Theorem 3.13 and the second assumption on q, we obtain

$$\|\mathbf{v}\| - \frac{sn\sqrt{m}}{q} \geqslant \frac{1}{48} - \frac{sn\sqrt{m}}{q} \geqslant \frac{1}{96},$$

except with probability at most $2^{-\Omega(n)}$. □

Finally, we complete the proof of our first main result.

Proof (Theorem 3.11). Let $q = \lceil 96sn\sqrt{m} \rceil$. With this choice, it turns out that any m satisfying $m \geqslant 1355n \ln s$ also satisfies $m \geqslant 335n \ln(97sn\sqrt{m})$. By Lemmas 3.5 and 3.18,we obtain that, with all but probability $2^{-\Omega(n)}$,

$$\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \frac{\sqrt{\ln\left(2(m-n)(1+\frac{1}{\varepsilon})\right)/\pi}}{\lambda_1^\infty(\Lambda^\perp(X)^\star)} \leqslant 96\sqrt{(n+\ln m) \cdot \frac{\ln\left(2(m-n)(1+\frac{1}{\varepsilon})\right)}{\pi}}.$$

Alternatively, for any $\varepsilon \leqslant 2^{-(m-n)}$, we can use Lemmas 3.6 and 3.18 to obtain that (with all but probability $2^{-\Omega(n)}$)

$$\eta_\varepsilon(\Lambda^\perp(X)) \leqslant \frac{\sqrt{\ln(1/\varepsilon)}}{\lambda_1(\Lambda^\perp(X)^\star)} \leqslant 96\sqrt{\ln(1/\varepsilon)}.$$

This completes the proof. □

3.4 Last minimum of $\Lambda^{\perp}(X)$

In this section we present our second result: an upper bound on the $(m-n)$-th minimum of the orthogonal lattice $\Lambda^{\perp}(X)$.

The question of finding an upper bound on $\lambda_{m-n}(\Lambda^{\perp}(X))$ was addressed in [AGHS13] and later in [AR16], with the aim of obtaining an upper bound on the smoothing parameter of $\Lambda^{\perp}(X)$. In particular, Agrawal et al. in [AGHS13] first give a lower bound on $\lambda_{n+1}(\Lambda_q(X))$, then use Banaszczyk's theorem (Theorem 3.4) to obtain an upper bound on $\lambda_{m-n}(\Lambda_q^{\perp}(X))$. Finally, they argue that this is also an upper bound on $\lambda_{m-n}(\Lambda^{\perp}(X))$. Aggarwal and Regev in [AR16] present a more direct approach to bound $\lambda_{m-n}(\Lambda^{\perp}(X))$. In all cases, these bounds on the last minimum of $\Lambda^{\perp}(X)$ were used as a way to bound its smoothing parameter (our approach in Section 3.3 is in some sense more direct).

We shall need the following lemma, obtained by combining Theorem 3.13 with Theorem 3.4.

Lemma 3.19. Let $n \geqslant 60$, $s \geqslant 20\sqrt{n}$ and $m \geqslant 1400n \ln s$.Then, we have:

$$\Pr_{X \leftarrow (D_{\mathbb{Z}^n,s})^m}\left[\lambda_{m-n}(\Lambda^{\perp}(X)) \leqslant 48m\right] \geqslant 1 - 2^{-\Omega(n)}.$$

Proof. Let q be the smallest prime such that $q \geqslant 96sm^{3/2}$. By [HB88], there exists a prime in the range $(96sm^{3/2}, 192sm^{3/2})$, hence we have $q < 192sm^{3/2}$.[4] We apply Theorem 3.13 to conclude that $\lambda_{n+1}(\Lambda_q(X)) \geqslant \frac{q}{48}$ with overwhelming probability. From Theorem 3.4 with $i = n+1$, it follows that $\lambda_{m-n}(\Lambda_q^{\perp}(X)) \leqslant 48m$. This implies that $\Lambda_q^{\perp}(X)$ contains $m-n$ linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{m-n}$ such that $\|\mathbf{v}_1\| \leqslant \|\mathbf{v}_2\| \leqslant \cdots \leqslant \|\mathbf{v}_{m-n}\| \leqslant 48m$. As q is prime, we have that $X\mathbf{v}_j = \mathbf{0} \bmod q$ for all $j \in [m-n]$ (see the discussion after Definition 2.1).

Now, we show that $\mathbf{v}_j \in \Lambda^{\perp}(X)$ for all $j \in [m-n]$, i.e., that $X\mathbf{v}_j = \mathbf{0}$ over the integers. Thanks to Lemma 2.5 with $t = \sqrt{2\pi}$, the rows of X have norms bounded by $s\sqrt{m}$ with probability greater than $1 - 2^{-\Omega(n)}$. Therefore, for any $j \in [m-n]$, we have

$$\|X \cdot \mathbf{v}_j\|_{\infty} = \max_i |\langle \mathbf{x}_i, \mathbf{v}_j \rangle| \leqslant \max_i \|\mathbf{x}_i\| \cdot \|\mathbf{v}_j\| \leqslant 48sm^{\frac{3}{2}}$$

with overwhelming probability. Our choice of q implies that $\|X \cdot \mathbf{v}_j\|_{\infty} < q/2$, hence the equality $X \cdot \mathbf{v}_j = \mathbf{0}$ holds over $\mathbb{Z}$. The result follows. here □

We now consider the case of a wide matrix X, i.e. with very large m. We split it into t matrices of smaller dimensions $n \times m_i$ for $i \in [t]$, where m_i is independent of m and is large enough to satisfy the conditions of Lemma 3.19. For simplicity, one could think of m_i's being all equal assuming that m is divisible by m_i. In general, we may not be able to divide m into large enough and equal pieces. This is why our X_i's may have different numbers of columns. Using Lemma 3.19, we show that every orthogonal lattice defined by such small matrices has $m_i - n$ linearly independent vectors of norm at most $48m_i$. By padding these vectors with zeros appropriately (see Figure 3.1), we thus find $\sum_{i \in [t]}(m_i - n)$ short and linearly independent vectors in $\Lambda^{\perp}(X)$. To show that there are,

[4] In fact, the following stronger result is proved in [HB88]: the number of primes in the interval $(x - x^{\alpha}, x)$ is at least $\frac{x^{\alpha}}{\log x}$ for $\alpha < 7/12$. To simplify our statements, we use a looser bound.

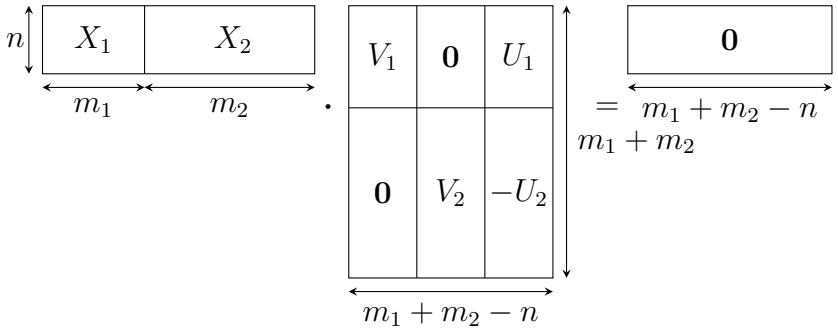

Figure 3.2: Given a wide matrix $X = (X_1|X_2) \in \mathbb{Z}^{n\times(m_1+m_2)}$, we first obtain m_1+m_2-n linearly independent short vectors in $\Lambda^\perp((X_1|X_2))$. These correspond to the columns $(V_1{}^\mathsf{t}|\mathbf{0})^\mathsf{t}$ and $(\mathbf{0}|V_2{}^\mathsf{t})^\mathsf{t}$. The n other missing short vectors are obtained via stacking U_i matrices satisfying $X_iU_i = I_n$, as depicted.

in fact, more short vectors in this lattice, we apply Lemma 3.9. We can "stack" the U-matrices from this lemma (see Figure 3.1) to obtain other short vectors orthogonal to X. Thus, in total we obtain $m-n$ short linearly independent vectors in $\Lambda^\perp(X)$ whose norms can be bounded independently from m.

Theorem 3.20. Let $n \geqslant 100$ and $s \geqslant 20\sqrt{n}$. Let m such that $2801n\ln s \leqslant m \leqslant 2^{n/2}$.Then, we have

$$\Pr_{X\leftarrow(D_{\mathbb{Z}^n,s})^m}\left[\lambda_{m-n}(\Lambda^\perp(X)) \leqslant \mathcal{O}\left(n\ln(ns)\right)\right] \geqslant 1-2^{-\Omega(n)}.$$

The "$\mathcal{O}(\cdot)$" constant can be worked out from the proof. Concretely, the term $\mathcal{O}(n\ln(ns))$ may be replaced by $134400n\ln(ns)$.

Proof. We divide our wide matrix X into smaller matrices with appropriate numbers of columns. For $m \geqslant 2801n\ln s$, we can divide the matrix X into at least two blocks of at least $m' = \lceil 1400n\ln s\rceil$ columns.

We start by splitting X into t smaller matrices $X_i \in \mathbb{Z}^{n\times m_i}$ such that $m_i \in [m', 2m']$ for all $i \in [t]$. We look at X as a block-matrix $X = [X_1|X_2|\ldots|X_t]$, where $X_i \leftarrow (D_{\mathbb{Z}^n,s})^{m_i}$ for all $i \in [t]$. We apply Lemma 3.19 to each block X_i. The lattice $\Lambda^\perp(X_i)$ has $m_i - n$ linearly independent vectors $\mathbf{v}_1^i, \ldots, \mathbf{v}_{m_i-n}^i$ such that

$$\|\mathbf{v}_1^i\| \leqslant \|\mathbf{v}_2^i\| \leqslant \ldots \leqslant \|\mathbf{v}_{m_i-n}^i\| \leqslant 48m_i \leqslant 96m',$$

with probability $1-2^{-\Omega(n)}$. It follows that we have $\sum_{i\in[t]}(m_i-n) = m-tn$ linearly independent vectors in $\Lambda^\perp(X)$ of the form:

$$\bar{\mathbf{v}}_j^i = [\mathbf{0}_{m_1+\cdots+m_{i-1}}\|\mathbf{v}_j^i\|\mathbf{0}_{m_{i+1}+\cdots+m_t}]^\mathsf{t},$$

for $j \in [m_i-n]$ and $i \in [t]$. Our goal is to have more ($m-n$, to be precise) short linearly independent vectors in $\Lambda^\perp(X)$.

Let $i \in [t]$. By Lemma 3.9, with probability greater than $1-2^{-n}$, there exists a matrix $U_i \in \mathbb{Z}^{m_i\times n}$ such that $X_iU_i = I_n$ with columns of norms $\leqslant 2\sqrt{44n\ln(ns)}$. When this event occurs, we have Let $i \in [t]$. By Lemma 3.9, with probability greater than

$1 - 2^{-n}$, there exists a matrix $U_i \in \mathbb{Z}^{m_i \times n}$ such that $X_i U_i = I_n$ with columns of norms $\leqslant 2\sqrt{44n\ln(ns)}$. When this event occurs, we have

$$X_i \cdot [V_i | U_i] = [\mathbf{0}_{n \times (m_i - n)} | I_n],$$

where V_i is the $m_i \times (m_i - n)$ matrix whose columns are $\bar{\mathbf{v}}_1^i, \bar{\mathbf{v}}_2^i \cdots, \bar{\mathbf{v}}_{m_i - n}^i$.

With probability $\geqslant 1 - t2^{-n}$ (which is $\geqslant 1 - 2^{-\Omega(n)}$ by assumption on m), we can write:

$$\left[\begin{array}{c|c|c|c|c} X_1 & X_2 & X_3 & \ldots & X_t \end{array}\right] \cdot \left[\begin{array}{c|c|c|c|c|c|c|c|c} V_1 & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & U_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \hline \mathbf{0} & V_2 & \mathbf{0} & \ldots & \mathbf{0} & -U_2 & U_2 & \ldots & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & V_3 & \ldots & \mathbf{0} & \mathbf{0} & -U_3 & \ldots & \mathbf{0} \\ \hline \vdots & & & \ddots & & & & \ddots & \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & V_t & \mathbf{0} & \mathbf{0} & \ldots & -U_t \end{array}\right] = \mathbf{0}_{m \times (m-n)}.$$

Now, we argue that the columns of the matrix built from the U_i's and V_i's are linearly independent. First, for each i, the columns of $[V_i|U_i]$ are linearly independent since they satisfy $X_i \cdot [V_i|U_i] = [\mathbf{0}_{m_1 - n}|I_{m_1}]$, and since the columns of V_i are linearly independent. This implies that for every i, the "block row" $[\mathbf{0}|\ldots|\mathbf{0}|V_i|\mathbf{0}|\ldots|\mathbf{0}|-U_i|\ldots]$ has rank exactly m_i. If one re-orders the block columns appropriately, the matrix has a "block triangular" shape. Its rank is $m_1 + \ldots + m_{t-1} + m_t - n = m - n$.

Overall, we obtain $m - n$ linearly independent vectors in $\Lambda^{\perp}(X)$, with norms $\leqslant \mathcal{O}(n\ln(ns))$, with probability $\geqslant 1 - 2^{-\Omega(n)}$. here $\square$

Chapter 4

An Anonymous Trace-and-Revoke Broadcast Encryption Scheme

This chapter is based on a joint work with O. Blazy, S. Mukherjee, H. Phan and D. Stehlé. It was accepted in ACISP 2021 [BMN+21].

4.1 Introduction

Trace-and-revoke systems, introduced in [NP01, NNL01] has been studied extensively in many works, including [DF03, KHL03, BW06, NWZ16, ABP+17].

Anonymity of receivers is important in numerous real-life applications and has been considered in multiple works, such as [BBW06, LPQ12, FP12, LG18, DPY20]. The standard notion of anonymity requires that the adversary cannot distinguish between ciphertexts of two targeted sets of its choice, even if it can corrupt any user in the intersection of these two sets or outside of the two sets. Unfortunately, it turned out to be extremely difficult to achieve this anonymity level in the general case without any restriction on the size of the target set. The state of the art constructions by Barth *et al.* [BBW06] and Libert *et al.* [LPQ12] start from a public-key encryption and result in schemes with ciphertext size which is N times larger, where N denotes the total number of users. Moreover, Kiayias and Samari [KS12] proved lower bounds in the general case that the ciphertext size has to be linear in N.

For revoke systems, the efficiency is often negatively correlated to the upper bound on the number of revoked users. One of the most important applications of broadcast encryption is Pay-TV and it can typically be in the form of a revoke system: the service broadcasts to all users except the revoked users who were detected as traitors or who unsubscribed from the system. The state of the art revoke systems [NP01,NNL01,BW06, ABP+17] have compact ciphertext sizes that grow as $\mathcal{O}(r)$ where r is the bound of the revoked users and which is not dependent on the number of users. None of these schemes is anonymous. An attempt was made to consider outsider adversaries, who can only corrupt users outside of the two targeted sets. In this limited setting, Fazio and Perera [FP12] showed that one can get key and ciphertext sizes that are sublinear in the number of users. We observe totally different situations for getting anonymity in broadcast encryption and in revoke systems: in broadcast encryption, optimal solutions exist [AY20] but one cannot get the anonymity with sublinear ciphertext size in the total number of users; in revoke systems, no impossibility result has been settled and it does not exclude the possibility to get an anonymous schemes which is as efficient as non-

anonymous ones, namely ciphertext size is $O(r)$, independent from the number of users. In this paper, we show that we can design anonymous schemes with $O(r)$ ciphertext size. Moreover, we also handle traceability to achieve anonymous trace-and-revoke systems.

4.1.1 Contributions

Our primary contribution is to develop the first symmetric-key trace-and-revoke scheme with traceability and anonymous revocation. We give two constructions of trace-and-revoke schemes, namely $\mathbf{TR}_0$ and $\mathbf{TR}_1$ from so-called linear functional encryptions. The former $\mathbf{TR}_0$ is generically constructed from inner product functional encryption (IPFE) and encrypts single bit messages. Similarly, $\mathbf{TR}_1$ is constructed from matrix multiplication functional encryption (MMFE) to support n-bit messages. Interestingly, unlike [ABP+17], our DDH instantiations do not require discrete-log evaluation for ciphertext decryption.

Our second contribution is to propose efficient constructions. We give an efficient construction of MMFE in the prime-order groups and prove that our MMFE construction is indeed tightly secure under the standard matDH assumption. Then we present IPFE construction and its security proof follow from those of MMFE. This construction can be seen as tweaking Tomida's tightly secure IPFE for the symmetric-key settings [Tom19]. However, we note that our security argument is somewhat different from Tomida's. On top of that, our tightly secure MMFE is more efficient than applying [Tom19] naively.

Our third contribution is a cryptanalysis on the LWE-based IPFE construction of [WFL19]. This justifies our choice of LWE-based IPFE to instantiate $\mathbf{TR}_0$.

Anonymous Revocation. Before describing our results, we discuss the notion of anonymous revocation in trace-and-revoke schemes. The Enc algorithm of any trace-and-revoke scheme takes a message m and a revoked user set description $\mathcal{R}$ and computes a ciphertext that can only be decrypted by users outside $\mathcal{R}$. The anonymity property intuitively means that no information on $\mathcal{R}$ should be inferred from the ciphertext. A typical multi-challenge security model is defined by polynomially many challenge phases where the adversary adaptively produces $(m^{(t)}, \mathcal{R}_0^{(t)}, \mathcal{R}_1^{(t)})$ on the t-th phase and gets an encryption of $(m^{(t)}, \mathcal{R}_\beta^{(t)})$ for the same $\beta \leftarrow \{0,1\}$ throughout the phases.

However, this security model is quite strong and there are practical scenarios that do not require such stronger definition. For example, a typical trace-and-revoke scheme revokes more and more users over time. If a revoked user wants to get access to the system again, it has to contact the broadcaster, which can give the user a new key. In such a scenario, the revoked user set increases with time, such that $\mathcal{R}^{(t-1)} \subseteq \mathcal{R}^{(t)}$ for any timestamp $t > 1$. We model this scenario by introducing the restriction that, for any t, if the adversary produces the challenge $(m^{(t)}, \mathcal{R}_0^{(t)}, \mathcal{R}_1^{(t)})$, then $\mathcal{R}_0^{(t-1)} \subseteq \mathcal{R}_0^{(t)}$ and $\mathcal{R}_1^{(t-1)} \subseteq \mathcal{R}_1^{(t)}$, and call the resulting security property multi-challenge monotonic anonymity $\mathsf{mIND\text{-}ID\text{-}CPA}$. Although this setting may suffice in many cases, this multi-challenge security model puts an additional restriction on the adversary that the challenge revocation sets must be related in a particular manner. This raises the following question: if we restrict ourselves to the single-challenge security model, can we get rid of such restriction on the monotonicity of challenge queries? We define another model for anonymity security that allows polynomially many ciphertext queries $(m^{(t)}, \mathcal{R}^{(t)})$

and a single anonymity query $(m, \mathcal{R}_0, \mathcal{R}_1)$ along with adaptively chosen key extraction queries. Note that, the revocation sets $\mathcal{R}^{(t)}$ in the ciphertext queries can be adaptively chosen, having no relation with the challenge revocation sets $\mathcal{R}_0$ and $\mathcal{R}_1$. That being said, we still had to impose a different restriction on the adversary on the post-challenge key queries. We present this model in Section 4.2.2.2 and show that our construction is secure in this model. We note that this is stronger than "insider anonymous" security where the post-challenge key queries are made only on users in $\mathcal{R}_0 \cap \mathcal{R}_1$.

4.1.2 Technical Overview

We start with a basic description of the trace-and-revoke scheme by Agrawal *et al.* [ABP+17] (in the bounded collusion model). Each user id in this scheme is associated with a vector $\mathbf{x}_{\mathsf{id}}$ and, correspondingly, a set $\mathcal{R}$ is associated with $X_{\mathcal{R}}$, the vector space spanned by $(\mathbf{x}_{\mathsf{id}})_{\mathsf{id} \in \mathcal{R}}$. Then, the predicate '$\mathsf{id} \notin \mathcal{R}$' can be emulated by testing if '$\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle = 0$' for $\mathbf{v}_{\mathcal{R}}$ orthogonal to $X_{\mathcal{R}}$. Using this relation, one encrypts a message m by encrypting $m \cdot \mathbf{v}_{\mathcal{R}}$ using an IPFE. An IPFE key for $\mathbf{x}_{\mathsf{id}}$ is used to evaluate $\mathsf{id} \notin \mathcal{R}$ in the encrypted domain. We now describe the decryption algorithm of [ABP+17] to clarify that this construction does not achieve anonymity of the revocation set. Decryption takes a ciphertext ct for $(m, \mathcal{R})$ and a secret key sk for id and runs IPFE decryption to obtain an intermediate $\mathsf{Res} = \langle \mathbf{x}_{\mathsf{id}}, m \cdot \mathbf{v}_{\mathcal{R}} \rangle$. The correctness then follows from the fact that decryption can compute $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ and divide Res by it to retrieve m. This is the reason why the description of $\mathcal{R}$ is provided as part of the ciphertext. Thus, the Agrawal *et al.* scheme does not achieve revocation set hiding.

Our constructions build on [ABP+17], but avoid the above difficulty by exploiting the fact that if we consider the message to be single bit (i.e., $m \in \{0, 1\}$), we have the following four cases:

- $m = 0$, $\mathsf{id} \in \mathcal{R}$: The value of $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ is zero.
- $m = 1$, $\mathsf{id} \in \mathcal{R}$: Same as above where the value of $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ is zero; therefore, when $\mathsf{id} \in \mathcal{R}$, the message m is hidden.
- $m = 0$, $\mathsf{id} \notin \mathcal{R}$: The value of $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle$ is again zero.
- $m = 1$, $\mathsf{id} \notin \mathcal{R}$: The value of $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle = m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle$ is non-zero.

The above list of cases shows that a secret key for $\mathbf{x}_{\mathsf{id}}$ decrypts an IPFE ciphertext for $m \cdot \mathbf{v}_{\mathcal{R}}$ and retrieves $m \in \{0, 1\}$ correctly if $\mathsf{id} \notin \mathcal{R}$. Note that the decryption algorithm no longer requires the description of the revoked set $\mathcal{R}$. Based on this observation, our constructions translate $(m, \mathcal{R})$ into a vector $m \cdot \mathbf{v}_{\mathcal{R}}$ where $\mathbf{v}_{\mathcal{R}}$ is a random vector orthogonal to $X_{\mathcal{R}}$ and id to a non-zero vector $\mathbf{x}_{\mathsf{id}}$. The monotonic anonymity (in the $\mathsf{mIND\text{-}ID\text{-}CPA}$ security model discussed above) then follows from the fact that the underlying IPFE hides the plaintext vector (here $m \cdot \mathbf{v}_{\mathcal{R}}$). For an n-bit message space, we can run independent and parallel executions of the IPFE that allow bit-by-bit retrieval of the message encrypted. We propose a more efficient alternative, namely, matrix multiplication functional encryption (MMFE). Our generic transformation above ensures that any efficient instantiation of MMFE will result in efficient trace-and-revoke scheme. We discuss constructions of MMFE in both the group-based settings and in the lattice-based

settings. We further show that our group-based construction of MMFE is tightly secure under standard assumptions. For lattice-based setting, we suggest use of [ABP+17] as we could mount a concrete attack on the state-of-the-art [WFL19] rendering it insecure. Lastly, we note that tracing is performed in a similar fashion to [ABP+17].

An attack on the Wang *et al.* IPFE. Here, we show that the IPFE construction by Wang *et al.* can be broken for the parameters chosen in [WFL19]. Our attack can be thwarted by increasing the parameters, but then the scheme does not enjoy great efficiency compared to the one from [ABP+17]. Here, we give the overview on LWE-based IPFE from [WFL19]. The dimension n of the LWE secrets is proportional to the security parameter λ, the parameters ℓ, m, p, q are polynomial in n. The master secret key is Z, uniform over $\{0, \ldots, p-1\}^{\ell \times m}$. The public key is of the form $\mathsf{pk} = (A \in \mathbb{Z}_q^{m \times n}, T = ZA \in \mathbb{Z}_q^{\ell \times n})$. The secret key for the vector $\mathbf{x} \in \mathbb{Z}_p^{\ell}$ is $\mathsf{sk}_{\mathbf{x}} = \mathbf{x}^{\mathsf{t}} \cdot Z$. The ciphertext for a vector $\mathbf{y} \in \mathbb{Z}_p^{\ell}$ is of the form $(\mathbf{c}_0 \approx A\mathbf{s}, \mathbf{c}_1 \approx T\mathbf{s} + p^{k-1} \cdot \mathbf{y})$. The authors state that under the LWE assumption, this IPFE is adaptively secure for chosen message distributions, assuming that the secret key queries are linearly independent. We will give an algorithm that can recover the master key from the public key and ciphertexts (i.e., recover $\mathbf{z}$ from X^{t} and $X^{\mathsf{t}}\mathbf{z}$, where $\mathbf{z} \leftarrow \{0, \ldots, p-1\}^{\ell}$ and $X \in \{0, \ldots, p-1\}^{\ell \times (\ell-1)}$ is chosen by the adversary). We remark that $\mathbf{z}$ belongs to a coset of the lattice orthogonal of X defined by $\mathbf{t}$. The crux of the attack is that for parameters as above, the minimum of this lattice is larger than $\|\mathbf{z}\|$. This means that we have a Bounded Distance Decoding problem instance in a lattice of dimension 1. Finally, we also explain why our attack does not extend to the schemes from [ALS16, ABP+17].

Organization of this chapter. In Section 4.2, we present some important definitions. In Section 4.3, we present black-box transformations to convert linear functional encryptions into trace-and-revoke systems with traceability and anonymity of revocation. Before we present group-based MMFE construction, in Section 4.4, we show an attack of a recent LWE-based IPFE construction [WFL19]. Then, in Section 4.5, we present a construction of MMFE in the prime-order groups.

4.2 Definitions and Preliminaries

For any two sets S and R, we define $S \Delta R = (S \setminus R) \cup (R \setminus S)$. For a dictionary $\mathsf{D} = (k, v_k)_k$, $\mathsf{D.vals}()$ gives the set $\{v_k : k \in \mathsf{D}\}$. For a vector space V over a field K, the corresponding orthogonal space is denoted by $V^{\perp}$. For a distribution D, we write $x \leftarrow D$ to say that x is sampled from D. The ppt abbreviation stands for probabilistic polynomial time. We denote $\mathcal{G}_{gen}(1^{\lambda}, p) \rightarrow (g, \mathbb{G})$ such that $\mathbb{G}$ is a cyclic group of prime order p and g generates $\mathbb{G}$. For $A = (a_{ij}) \in \mathbb{Z}_p^{\beta \times \alpha}$ we denote $[A] = (g^{a_{ij}}) \in \mathbb{Z}_p^{\beta \times \alpha}$. For $m, k \in \mathbb{N}$ for $m > k$, we use $M \leftarrow \mathcal{D}_k[m, k]$ to get a full rank matrix $M \in \mathbb{Z}^{m \times k}$ where the first k rows are linearly independent.

4.2.1 Linear Functional Encryption

A functional encryption scheme [BSW11] allows a user, having a secret key sk_f corresponding to a function f, to evaluate $f(z)$ securely given a ciphertext ct_z for a plaintext z. The inner product function, being one of the simplest functionalities, has received

a tremendous amount of exposure [ALS16, AGRW17, ACF+18, CLT18, Tom19]. We here define an extended version for IPFE in symmetric-key settings called Matrix Multiplication Functional Encryption (MMFE). Informally speaking, having a secret key $\mathsf{sk}_{\mathbf{x}}$ for $\mathbf{x} \in \mathbb{Z}_p^\ell$, given a ciphertext ct_M for for $M \in \mathbb{Z}_p^{n\times\ell}$, MMFE outputs a binary vector of length n where the i^{th} component indicates if $M_i\mathbf{x} = 0$ for $i \in [n]$ in terms of a predicate $f : \mathbb{Z}_p \to \{0, 1\}$.

We consider inner product functional encryption (IPFE) over $\mathbb{Z}_p$ in the symmetric-key settings [1] for a prime integer $p \geqslant 2$. Unlike existing IPFE definitions in [ALS16, ABP+17, ACF+18], the $\mathit{IPFE}.\mathsf{Dec}$ algorithm here retrieves an injective function of the inner product value. In particular, it may not be the inner product value itself. More precisely, the $\mathit{IPFE}.\mathsf{Dec}$ algorithm takes as input a ciphertext ct that encrypts $\mathbf{y} \in \mathbb{Z}_p^\ell$ and a secret key $\mathsf{sk}_{\mathbf{x}}$ with respect to $\mathbf{x} \in \mathbb{Z}_p^\ell$, and outputs $f(\langle \mathbf{x}, \mathbf{y}\rangle)$.

4.2.1.1 Inner Product Functional Encryption.

Definition 4.1. An inner product functional encryption (IPFE) over $\mathbb{Z}_p$ with respect to an injective map f is a tuple $\mathit{IPFE} = (\mathit{IPFE}.\mathsf{Setup}, \mathit{IPFE}.\mathsf{KeyGen}, \mathit{IPFE}.\mathsf{Enc}, \mathit{IPFE}.\mathsf{Dec})$ of four ppt algorithms.

- $\mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell, p)$ takes as input the security parameter λ and the dimension of vectors ℓ. It outputs the public parameters pp and the master secret key msk. The public parameters pp contain the description of the injective function f.
- $\mathit{IPFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$ takes as input the public parameters pp, the master secret key msk and a vector $\mathbf{x} \in \mathbb{Z}_p^\ell$ and outputs a secret key $\mathsf{sk}_{\mathbf{x}}$.
- $\mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y})$ takes as input the public parameters pp, the master secret key msk and a vector $\mathbf{y} \in \mathbb{Z}_p^\ell$ and outputs a ciphertext ct.
- $\mathit{IPFE}.\mathsf{Dec}(\mathsf{pp}, \mathsf{sk}_{\mathbf{x}}, \mathsf{ct})$ takes as input the public parameters pp, the secret key of a user $\mathsf{sk}_{\mathbf{x}}$ and a ciphertext $\mathsf{ct}_{\mathbf{y}}$, and outputs $f(\langle \mathbf{x}, \mathbf{y}\rangle)$.

The correctness requirement is that, with overwhelming probability over the randomness used by the algorithms, for $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell, p)$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_p^\ell$, for $\mathsf{sk}_{\mathbf{x}} \leftarrow \mathit{IPFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$ and $\mathsf{ct} \leftarrow \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y})$:

$$\mathit{IPFE}.\mathsf{Dec}\left(\mathsf{pp}, \mathsf{sk}_{\mathbf{x}}, \mathsf{ct}\right) = f(\langle \mathbf{x}, \mathbf{y}\rangle).$$

Security. The security ($\mathsf{IND\text{-}CPA}$) of symmetric-key IPFE is modeled as the following security game played between a challenger and an adversary $\mathcal{A}$. This security model is reminiscent of that of [SSW09].

- The challenger runs $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathsf{Setup}(1^\lambda, 1^\ell, p)$, keeps msk secret and gives the public parameters pp to the adversary $\mathcal{A}$. The challenger further samples $\beta \leftarrow \{0, 1\}$.

[1] We define IPFE in the symmetric-key settings as a stepping stone to construct trace-and-revoke in the symmetric-key settings.

- Adversary $\mathcal{A}$ adaptively issues queries of one of the following two types:

  1. **Ciphertext query**: The adversary sends two vectors $\mathbf{y}^{(0)}, \mathbf{y}^{(1)} \in \mathbb{Z}_p^\ell$, and the challenger responds with $\mathsf{ct}^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}^{(\beta)})$.
  2. **Secret key query**: The adversary sends a vector $\mathbf{x} \in \mathbb{Z}_p^\ell$, and the challenger responds with $\mathsf{sk}_\mathbf{x} \leftarrow \mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$.

  These queries can be made under the restriction that for all ciphertext query $(\mathbf{y}^{(0)}, \mathbf{y}^{(1)})$ and all secret key query $\mathbf{x}$, we must have $f(\langle \mathbf{x}, \mathbf{y}^{(0)} \rangle) = f(\langle \mathbf{x}, \mathbf{y}^{(1)} \rangle)$.

- Finally, the adversary returns its guess $\beta' \in \{0, 1\}$ for the bit β chosen by the challenger. The adversary wins this game if $\beta = \beta'$.

The advantage of the adversary is defined as $\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{FE},\mathcal{A}} = |\Pr[\beta = \beta'] - 1/2|$. A symmetric-key IPFE scheme IPFE is said secure if $\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathsf{FE},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

4.2.1.2 Matrix Multiplication Functional Encryption.

We now define matrix multiplication functional encryption (MMFE) over $\mathbb{Z}_p$ for a prime integer $p \geqslant 2$. MMFE, as the name suggests, decrypts a ciphertext for a matrix $M \in \mathbb{Z}_p^{n \times \ell}$ with a key $\mathsf{sk}_\mathbf{x}$ made of $\mathbf{x} \in \mathbb{Z}_p^\ell$ revealing only $M\mathbf{x}$ and nothing else. Due to its similarity with the definition of IPFE, MMFE can be achieved from available IPFE. In particular, one can use n-many instances of IPFE to encrypt n vectors $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$ independently and the Dec algorithm basically computes $\langle \mathbf{y}_i, \mathbf{x} \rangle$ for each $i \in [1, n]$ individually. However, such a trivial construction suffers from a degradation proportional to n. This gets worse in case of multi-challenge security.

We give a definition and propose a concrete construction with tight security in this work. A similar primitive was already introduced for predicate encryption to allow decryption based on subspace membership relation ($M\mathbf{x} = \mathbf{0}$ or not) in [BRS13]. Looking ahead, we present a symmetric-key MMFE definition here to construct symmetric-key trace-and-revoke scheme $\mathbf{TR}_1$ for arbitrary n-bit messages in Section 4.3.2.

Definition 4.2. A matrix multiplication functional encryption scheme $\mathcal{MMFE}$ over $\mathbb{Z}_p$ with respect to an injective function f is a tuple of four ppt algorithms with the following specifications:

- $\mathcal{MMFE}.\mathsf{Setup}(1^\lambda, 1^\ell, 1^n, p)$ takes as input the security parameter λ and the dimensions (n, ℓ) of matrices. It outputs the public parameters pp and the master secret key msk. Similarly to IPFE, the public parameters pp contain the description of an injective function f.

- $\mathcal{MMFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$ takes as input the public parameters pp, the master secret key msk and a vector $\mathbf{x} \in \mathbb{Z}_p^\ell$ and outputs a secret key $\mathsf{sk}_\mathbf{x}$.

- $\mathcal{MMFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M)$ takes as input the public parameters pp, the master secret key msk and a matrix $M \in \mathbb{Z}_p^{n \times \ell}$ and outputs a ciphertext ct.

- $\mathcal{MMFE}.\mathsf{Dec}(\mathsf{pp}, \mathsf{sk}_{\mathbf{x}}, \mathsf{ct})$ takes as input the public parameters pp, the secret key of a user $\mathsf{sk}_{\mathbf{x}}$ and a ciphertext ct, and outputs $(f(M_1\mathbf{x}), \ldots, f(M_n\mathbf{x}))$ where M_i is the i^{th} row of M.

The correctness requirement is that, with overwhelming probability over the randomness used by the algorithms, for $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathcal{MMFE}.\mathsf{Setup}(1^\lambda, 1^\ell, 1^n, p)$, for all $\mathbf{x} \in \mathbb{Z}_p^\ell$ and $M \in \mathbb{Z}_p^{n\times\ell}$, for $\mathsf{sk}_{\mathbf{x}} \leftarrow \mathcal{MMFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$ and $\mathsf{ct} \leftarrow \mathcal{MMFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M)$:

$$\mathcal{MMFE}.\mathsf{Dec}\left(\mathsf{pp}, \mathsf{sk}_{\mathbf{x}}, \mathsf{ct}\right) = \left(f(M_1\mathbf{x}), \ldots, f(M_n\mathbf{x})\right).$$

Security. Full security ($\mathsf{IND\text{-}CPA}$) of symmetric-key matrix multiplication functional encryption is modeled as the following security game played between a challenger and an adversary $\mathcal{A}$.

- The challenger runs $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathsf{Setup}(1^\lambda, 1^\ell, 1^n, p)$, keeps msk secret and gives the public parameters pp to the adversary $\mathcal{A}$. The challenger further samples $\beta \leftarrow \{0, 1\}$.
- Adversary $\mathcal{A}$ adaptively issues queries of one of the following two types:
  1. **Ciphertext query**: The adversary sends two matrices $M^{(0)}, M^{(1)} \in \mathbb{Z}_p^{n\times\ell}$, and the challenger responds with $\mathsf{ct}^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M^{(\beta)})$.
  2. **Secret key query**: The adversary sends a vector $\mathbf{x} \in \mathbb{Z}_p^\ell$, and the challenger responds with $\mathsf{sk}_{\mathbf{x}} \leftarrow \mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x})$.

  These queries can be made under the restriction that for all ciphertext query $(M^{(0)}, M^{(1)})$ and all secret key query $\mathbf{x}$, we must have $f(M^{(0)}\mathbf{x}) = f(M^{(1)}\mathbf{x})$.
- Finally, the adversary returns its guess $\beta' \in \{0, 1\}$ for the bit β chosen by the challenger. The adversary wins this game if $\beta = \beta'$.

The advantage of the adversary is defined as $\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathcal{MMFE},\mathcal{A}} = |\Pr[\beta = \beta'] - 1/2|$. A symmetric-key inner matrix multiplication functional encryption scheme $\mathcal{MMFE}$ is said secure if $\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathcal{MMFE},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

4.2.1.3 Mathematical Tools and Hardness Assumptions

We assume $\mathcal{G}_{gen}$ to be the group generator that generates the prime order group description. Precisely, $\mathcal{G}_{gen}(1^\lambda, p) \to (g, \mathbb{G})$ such that $\mathbb{G}$ is a cyclic group of prime order p and g generates $\mathbb{G}$. We follow the notation of [EHK+17] to denote g^a by $[a]$ for any $a \in \mathbb{Z}_p$ and for $A = (a_{ij}) \in \mathbb{Z}_p^{\beta\times\alpha}$ we denote

$$[A] = \begin{pmatrix} g^{a_{11}} & \ldots & g^{a_{1\alpha}} \\ \vdots & \ddots & \vdots \\ g^{a_{\beta 1}} & \ldots & g^{a_{\beta\alpha}} \end{pmatrix} \in \mathbb{G}^{\beta\times\alpha}.$$

Definition 4.3. Let $m, k \in \mathbb{N}$, such that $m > k$. We call $\mathcal{D}_k[m, k]$ a matrix distribution if it outputs matrices in $\mathbb{Z}_p^{m\times k}$ of full rank k in polynomial time (w.l.o.g. we assume the first k rows of $M \leftarrow \mathcal{D}_k[m, k]$ form an invertible matrix). We write $\mathcal{D}_k = \mathcal{D}_k[k + 1, k]$.

Definition 4.4. For all adversary $\mathcal{A}$, the advantage function is defined as following

$$\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda) = |\Pr[\mathcal{A}([U], [U\mathbf{x}]) = 1] - \Pr[\mathcal{A}([U], [\mathbf{z}]) = 1]|$$

where $U \leftarrow \mathcal{D}_k$, $\mathbf{x} \leftarrow \mathbb{Z}_p^k$ and $\mathbf{z} \leftarrow \mathbb{Z}_p^{k+1}$. The $\mathcal{D}_k$-matDH assumption states that $\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda)$ is negligible in λ for all ppt adversary $\mathcal{A}$.

Definition 4.5. For all adversary $\mathcal{A}$, the advantage function is defined as following

$$\mathsf{Adv}_{\mathcal{A}}^{n\text{-}\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda) = |\Pr[\mathcal{A}([U], [UX]) = 1] - \Pr[\mathcal{A}([U], [Z]) = 1]|$$

where $U \leftarrow \mathcal{D}_k$, $X \leftarrow \mathbb{Z}_p^{k\times n}$ and $Z \leftarrow \mathbb{Z}_p^{(k+1)\times n}$. The n-fold $\mathcal{D}_k$-matDH assumption (i.e. n-$\mathcal{D}_k$-matDH) states that $\mathsf{Adv}_{\mathcal{A}}^{n\text{-}\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda)$ is negligible in λ for all ppt adversary $\mathcal{A}$.

Now, [EHK+17] showed that $\mathsf{Adv}^{n\text{-}\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda) \leqslant \mathsf{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda)$ for any fixed value n that is polynomial in λ.

Definition 4.6. For all adversary $\mathcal{A}$, the advantage function is defined as following

$$\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_{2k,k}\text{-}\mathsf{matDH}}(\lambda) = |\Pr[\mathcal{A}([V], [V\mathbf{y}]) = 1] - \Pr[\mathcal{A}([V], [\mathbf{z}]) = 1]|$$

where $V \leftarrow \mathcal{D}_{2k,k}$, $\mathbf{y} \leftarrow \mathbb{Z}_p^k$ and $\mathbf{z} \leftarrow \mathbb{Z}_p^{2k}$. The $\mathcal{D}_{2k,k}$-matDH assumption states that $\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_{2k,k}\text{-}\mathsf{matDH}}(\lambda)$ is negligible in λ for all ppt adversary $\mathcal{A}$. [EHK+17] showed that given a $\mathcal{D}_k$-matDH problem instance, one can create a $\mathcal{D}_{2k,k}$-matDH problem instance with the degradation of k i.e.

$$\mathsf{Adv}^{\mathcal{D}_{2k,k}\text{-}\mathsf{matDH}}(\lambda) \leqslant k \cdot \mathsf{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda) \tag{4.1}$$

Definition 4.7. For all adversary $\mathcal{A}$, the advantage function is defined as following

$$\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_k\text{-}\mathsf{matDH}'}(\lambda) = |\Pr[\mathcal{A}([S], \left[\mathbf{u}^\top S\right]) = 1] - \Pr[\mathcal{A}([S], \left[\mathbf{z}^\top\right]) = 1]|$$

where $S \leftarrow \mathbb{Z}_p^{k\times m}$, $\mathbf{u} \leftarrow \mathbb{Z}_p^k$ and $\mathbf{z} \leftarrow \mathbb{Z}_p^m$ for a fixed value m that is polynomial in λ. The $\mathcal{D}_k$-matDH' assumption states that $\mathsf{Adv}_{\mathcal{A}}^{\mathcal{D}_k\text{-}\mathsf{matDH}'}(\lambda)$ is negligible in λ for all ppt adversary $\mathcal{A}$.

Tomida [Tom19] showed that given a $\mathcal{D}_k$-matDH' problem instance, one can create a m-fold $\mathcal{D}_k$-matDH problem instance without any degradation i.e. the advantage $\mathsf{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}'}(\lambda) \leqslant \mathsf{Adv}^{m\text{-}\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda)$. Due to the relation between $\mathcal{D}_k$-matDH and m-$\mathcal{D}_k$-matDH mentioned above, $\mathsf{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}'}(\lambda) \leqslant \mathsf{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}}(\lambda)$.We consider inner product functional encryption (IPFE) over $\mathbb{Z}_p$ in the symmetric-key settings [2] for a prime integer $p \geqslant 2$. Unlike existing IPFE definitions in [ALS16, ABP+17, ACF+18], the $\mathit{IPFE}.\mathsf{Dec}$ algorithm here retrieves an injective function of the inner product value. In particular, it may not be the inner product value itself. More precisely, the $\mathit{IPFE}.\mathsf{Dec}$ algorithm takes as input a ciphertext ct that encrypts $\mathbf{y} \in \mathbb{Z}_p^\ell$ and a secret key $\mathsf{sk}_{\mathbf{x}}$ with respect to $\mathbf{x} \in \mathbb{Z}_p^\ell$, and outputs $f(\langle \mathbf{x}, \mathbf{y} \rangle)$.

[2] We define IPFE in the symmetric-key settings as a stepping stone to construct trace-and-revoke in the symmetric-key settings.

4.2.2 Trace-and-Revoke Systems

A symmetric key traitor tracing encryption scheme is a multi-recipient encryption system in which a broadcasting office has the master secret key for encryption and there are many users with decryption capabilities, each having its own secret key. Additionally, the encryption scheme provides a feature to let the broadcaster identify at least one user from a coalition $\mathcal{T}$ of malicious users (traitors) that built an unauthorized decryption device $\mathcal{D}$. The following is the blackbox confirmation model [BF99], in which an efficient tracing algorithm Trace is given oracle access to $\mathcal{D}$, which we denote by $\mathcal{O}^{\mathcal{D}}$. The oracle $\mathcal{O}^{\mathcal{D}}$ takes as input any message-ciphertext pair (m, C) and returns 1 if $\mathcal{D}(C) = m$ and 0 otherwise. Given as input a set $\mathcal{S}$ of suspected users containing $\mathcal{T}$, the Trace algorithm should disclose the identity of at least one user from the set $\mathcal{T}$. For security, a traitor coalition should not be able to design a useful box that escapes tracing, i.e., such that the Trace algorithm replies $\perp$ or frames an innocent user in $\mathcal{S} \setminus \mathcal{T}$.

Following [ABP+17], the probability of decryption of decoder $\mathcal{D}$, can be estimated by repeatedly querying the oracle $\mathcal{O}^{\mathcal{D}}$ with plaintext-ciphertext pairs. Therefore, we assume the decryption device $\mathcal{D}$ correctly decrypts a properly generated ciphertext with significant probability. The following is a description of $\mathcal{D}$, reproduced from [ABP+17] and modified for the symmetric-key setting. Let $\mathcal{R}$ be any set of revoked users, of size $\leqslant r$. Let the message m be sampled uniformly at random from the message space M and let $C_{\mathcal{R}}$ be the output of the encryption algorithm Enc using the master secret key msk and $\mathcal{R}$ as the set of revoked users. With $C_{\mathcal{R}}$ as input, the device $\mathcal{D}$ is assumed to output m with probability significantly more than $1/|M|$:

$$\Pr_{\substack{m \leftarrow U(M) \\ C_{\mathcal{R}} \leftarrow \mathsf{Enc}(\mathsf{msk}, \mathsf{pp}, \mathcal{R}, m)}} \left[\mathcal{O}^{\mathcal{D}}(C_{\mathcal{R}}, m) = 1\right] \geqslant \frac{1}{|M|} + \frac{1}{\lambda^c}, \tag{4.2}$$

for some constant $c > 0$.

We let the identity space ID and the message space M be implicit arguments to the setup algorithm below. We let the secret key space $\mathcal{K}$, the ciphertext space $\mathcal{C}$ (along with ID and M) and the descriptions of mathematical tools that are used be part of the public parameters output by the setup algorithm. We adapt the definition from [ABP+17] to the symmetric-key setting.

Definition 4.8. A dynamic trace-and-revoke scheme $\mathbf{TR}$ in the black-box confirmation model is a tuple $\mathbf{TR} = (\mathsf{Setup}, \mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec}, \mathsf{Trace})$ of five ppt algorithms with the following specifications.

- $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$ takes as input the security parameter λ, the bound t on the size of traitor coalitions and the bound r on the number of revoked users. It outputs $(\mathsf{msk}, \mathsf{pp}, \mathsf{dir})$ containing the master secret key msk, the public parameters pp and the initially empty user directory dir. Here, unlike [ABP+17], dir is kept secret.

- $\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathsf{id})$ takes as input the public parameters pp, the master secret msk, the user directory dir and an identity $\mathsf{id} \in \mathsf{ID}$ of a user. It outputs the corresponding secret key $\mathsf{sk}_{\mathsf{id}}$ and some information u_{id} for the given identity id. It also updates dir to include u_{id}.

- $\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)$ takes as input the public parameters pp, the master secret msk, the user directory dir, a set $\mathcal{R}$ of size $\leqslant r$ which contains the u_{id} of each revoked user in dir, and a plaintext message $m \in M$. It outputs a ciphertext $C_{\mathcal{R}} \in \mathcal{C}$.

- $\mathsf{Dec}(\mathsf{pp}, \mathsf{sk}_{\mathsf{id}}, C_{\mathcal{R}})$ takes as input the public parameters pp, a secret key $\mathsf{sk}_{\mathsf{id}}$ of a user with identity id and a ciphertext $C_{\mathcal{R}} \in \mathcal{C}$. It outputs a plaintext $m' \in M$.

- $\mathsf{Trace}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, \mathcal{S}, \mathcal{O}^{\mathcal{D}})$ is a tracing algorithm in the black-box confirmation model that takes as input the public parameters pp, the master secret key msk, the user directory dir, a set $\mathcal{R}$ of $\leqslant r$ revoked users, a set $\mathcal{S}$ of $\leqslant t$ suspect users, and has black-box access to the pirate decoder $\mathcal{D}$ through the oracle $\mathcal{O}^{\mathcal{D}}$. It outputs an identity id or $\perp$.

The correctness requirement is that, with overwhelming probability over the randomness used by the algorithms, for $(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}) \leftarrow \mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$, for any set $\mathcal{R}$ of $\leqslant r$ revoked users:

$$\forall m \in M,\ \forall \mathsf{id} \in \mathsf{ID} \setminus \mathcal{R} :\ \mathsf{Dec}(\mathsf{pp}, \mathsf{sk}_{\mathsf{id}}, \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)) = m.$$

In this work, we consider three security properties for a trace-and-revoke scheme: message hiding, revocation set hiding, and traceability.

4.2.2.1 Message Hiding.

The IND-CPA security of a trace-and-revoke scheme **TR** is defined based on the following game. Informally speaking, neither a system outsider nor a revoked user must be able to get any information about the encrypted message.

- The challenger runs $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$ and gives the produced public parameters pp to the adversary $\mathcal{A}$. The adversary may ask the challenger to add polynomially many users in the system (these user addition queries can be adaptive and take place at any time in the game). The challenger updates dir accordingly.

- The adversary can adaptively make up to r secret key queries and a single challenge ciphertext query, of the following form:

  * Given a key generation query id, the challenger provides the corresponding $\mathsf{sk}_{\mathsf{id}}$ to $\mathcal{A}$.
  * Given the challenge ciphertext query $(m_0, m_1, \mathcal{R})$ with $\mathcal{R} \subset \mathsf{ID}$ of size $\leqslant r$, the challenger samples $\beta \leftarrow \{0, 1\}$ and provides $C^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m_\beta)$ to $\mathcal{A}$.

  These queries are subject to the restriction that every queried id belongs to $\mathcal{R}$.

- Finally, the adversary returns its guess $\beta' \in \{0, 1\}$ for the bit β chosen by the challenger. The adversary wins this game if $\beta = \beta'$.

The advantage of the adversary $\mathcal{A}$ is defined as

$$\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathbf{TR},\mathcal{A}} = |\Pr[\beta = \beta'] - 1/2|.$$

A trace-and-revoke scheme **TR** is said to be IND-CPA secure if $\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathbf{TR},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

4.2.2.2 Revocation Set Hiding.

The anonymity of a trace-and-revoke scheme **TR** captures the idea of hiding the *revocation set* in the ciphertext: if t^{th} challenge ciphertext is created for one of the two adversarially chosen revoked sets $(\mathcal{R}_0^{(t)}, \mathcal{R}_1^{(t)})$ on the t^{th} challenge phase, then the adversary cannot distinguish if $\mathcal{R}_0^{(t)}$ or $\mathcal{R}_1^{(t)}$ was used for the encryption for all of t.

As we already have mentioned in the Introduction, we aim for a multi-challenge security settings that properly emulates the following scenario: A typical trace-and-revoke scheme traces and revokes more and more users over the time. In such a scenario, each new ciphertext is created for growing revoked user sets. We call this setting as *monotonic anonymity* security model (mIND-ID-CPA) and define it as following.

- The challenger runs $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$ and gives the produced public parameter pp to the adversary $\mathcal{A}$. The adversary may ask the challenger to add polynomially many users in the system (these user addition queries can be adaptive and take place at any time in the game). The challenger updates dir accordingly.
- The adversary can adaptively make up to $(r + t)$ secret key queries and polynomially many anonymity challenge queries, of the following form:
  - Given a key generation query id, the challenger provides the corresponding sk_id to $\mathcal{A}$.
  - Given a challenge anonymity query $(m, \mathcal{R}_0, \mathcal{R}_1)$ with $\mathcal{R}_0, \mathcal{R}_1 \subset \mathsf{ID}$ of size $\leqslant r$, the challenger samples $\beta \leftarrow \{0,1\}$ and provides $C^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}_\beta, m)$ to $\mathcal{A}$.

  These queries are subject to the restriction that for every queried id, either $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$ or $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$. Among all the key queries that have been made, at most t of them could be satisfying $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$ and at most r of them could be satisfying $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$. The challenge anonymity queries also have a natural restriction that $\mathcal{R}_0^{(i)} \subseteq \mathcal{R}_0^{(j)}$ and $\mathcal{R}_1^{(i)} \subseteq \mathcal{R}_1^{(j)}$ for all $i \leqslant j$ where the t^{th} challenge anonymity query was made on $(m^{(t)}, \mathcal{R}_0^{(t)}, \mathcal{R}_1^{(t)})$.
- Finally, the adversary returns its guess $\beta' \in \{0,1\}$ for the bit β chosen by the challenger. The adversary wins this game if $\beta = \beta'$.

The advantage of the adversary $\mathcal{A}$ is defined as

$$\mathrm{Adv}^{\mathsf{mIND\text{-}ID\text{-}CPA}}_{\mathbf{TR},\mathcal{A}} = |\Pr[\beta = \beta'] - 1/2|.$$

A trace-and-revoke scheme **TR** is said to be mIND-ID-CPA secure if $\mathrm{Adv}^{\mathsf{mIND\text{-}ID\text{-}CPA}}_{\mathbf{TR},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

In the Introduction, we informally discussed different practical scenarios involving the anonymity of revocation set. The anonymity security model in Section 4.2.2.2 is a multi-challenge security model and captures the security requirements of a typical broadcasting agency. However, the security definition is restrictive in principle as all the revoked sets in anonymity challenge queries are related. In this section, we first give a single-challenge security definition (IND-ID-CPA) for revocation set hiding. Being a single-challenge security definition for symmetric-key settings, this new security definition (IND-ID-CPA) for revocation set hiding supports multiple ciphertext queries along with multiple secret key queries and a single challenge anonymity query.

The positive side of IND-ID-CPA is that in the security proof, we no longer put any restriction on the revoked sets $\mathcal{R}$ across multiple ciphertext queries and challenge anonymity query. However, we still need to impose some new security restrictions on the adversary here in terms of post-challenge secret key queries. Precisely, we define IND-ID*-CPA security that allows all pre-challenge queries (both key and ciphertext) and all post-challenge ciphertext queries (satisfying the natural restriction). However, for post-challenge secret key queries, IND-ID*-CPA imposes a new restriction. In the literature, similar restriction has already been put like "outsider corruption" in [FP12]. However, unlike [FP12] we can still support "insider corruption" (i.e. post-challenge key queried on $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$) completely and "outsider corruption" (i.e. post-challenge key queried on $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$) with some restriction. The restriction is a bit unusual in the sense, the adversary is not allowed to make post-challenge secret key queries on $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$ for an id that was a part of pre-challenge ciphertext query but was not queried for secret key in the pre-challenge query phase. But, this is all we require to argue our construction $\mathbf{TR}_0$ is secure. Here, note that, $\mathbf{TR}_0$ being a trace-and-revoke scheme with unbounded users, the set $\mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$ is sufficiently big and the adversary is restricted from making query on a small subset.

4.2.3 Security Definition

We first define the IND-ID-CPA security model and then weaken it to define IND-ID*-CPA security. The IND-ID-CPA security of a trace-and-revoke scheme $\mathbf{TR}$ is defined based on the following game.

- The challenger runs $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$ and gives the produced public parameter pp to the adversary $\mathcal{A}$. The adversary may ask the challenger to add polynomially many users in the system (these user addition queries can be adaptive and take place at any time in the game). The challenger updates dir accordingly.
- The adversary can adaptively make up to $(r + t)$ secret key queries, polynomially many ciphertext queries and a single anonymity challenge query, of the following form:
  - Given a key generation query id, the challenger provides the corresponding sk_id to $\mathcal{A}$.
  - Given a ciphertext query $(m, \mathcal{R})$ with $\mathcal{R} \subset \mathsf{ID}$ of size $\leqslant r$, the challenger provides $C \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)$ to $\mathcal{A}$.

* Given the challenge anonymity query $(m, \mathcal{R}_0, \mathcal{R}_1)$ with $\mathcal{R}_0, \mathcal{R}_1 \subset \mathsf{ID}$ of size $\leqslant r$, the challenger samples $\beta \leftarrow \{0,1\}$ and provides $C^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}_\beta, m)$ to $\mathcal{A}$.

These queries are subject to the restriction that for every queried id, either $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$ or $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$. Among all the key queries that have been made, at most t of them could be satisfying $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$ and at most r of them could be satisfying $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$.

- Finally, the adversary returns its guess $\beta' \in \{0,1\}$ for the bit β chosen by the challenger. The adversary wins this game if $\beta = \beta'$.

The advantage of the adversary $\mathcal{A}$ is defined as $\mathrm{Adv}^{\mathsf{IND\text{-}ID\text{-}CPA}}_{\mathbf{TR},\mathcal{A}} = |\Pr[\beta = \beta'] - 1/2|$.

We then weaken the security model a small amount to define $\mathsf{IND\text{-}ID^*\text{-}CPA}$ security, which does not allow post-challenge secret key queries on $\mathsf{id} \in \mathsf{ID} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)$ for an id that was a part of pre-challenge ciphertext query but was not queried for secret key in the pre-challenge query phase. A trace-and-revoke scheme $\mathbf{TR}$ is said to be $\mathsf{IND\text{-}ID^*\text{-}CPA}$ secure if the advantage $\mathrm{Adv}^{\mathsf{IND\text{-}ID^*\text{-}CPA}}_{\mathbf{TR},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

4.2.3.1 Traceability.

The notion of traceability considers a suspected set $\mathcal{S}$ of users who might have produced the pirate decoder $\mathcal{D}$. Then the tracing algorithm Trace outputs an $\mathsf{id} \in \mathcal{S} \setminus \mathcal{T}$ where $\mathcal{T}$ is the set of traitors who are already detected. This requirement is formalized using the following game, denoted by $\mathsf{AD\text{-}TT}$, between an adversary $\mathcal{A}$ and a challenger. We reproduce the security model from [ABP+17] for sake of completeness.[3]More precisely, the authors of [ABP+17] achieved *public-traceability*: for this purpose, the public-key Enc algorithm was used to construct so-called probe ciphertexts to query $\mathcal{O}^{\mathcal{D}}$ and identify a traitor. Our trace-and-revoke scheme relies on a symmetric key Enc algorithm, and hence tracing relies on the master secret key msk (in particular, tracing is not public).

- The challenger runs $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$ and gives pp to $\mathcal{A}$. The adversary may ask the challenger to add polynomially many users in the system (these user addition queries can be adaptive and take place at any time in the game). The challenger updates dir accordingly.

- Adversary $\mathcal{A}$ makes adaptive traitor key queries on at most t distinct users. For every id queried, the challenger checks to find $\mathsf{u_{id}} \leftarrow \mathsf{dir}[\mathsf{id}]$. If available, records id in $\mathcal{T}$ and returns $\mathsf{sk_{id}}$. Otherwise, adds $\mathsf{u_{id}}$ to $\mathsf{dir}[\mathsf{id}]$, records id in $\mathcal{T}$ and returns $\mathsf{sk_{id}} \leftarrow \mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathsf{id})$.

- Adversary $\mathcal{A}$ sends an adaptively chosen revocation set $\mathcal{R} \subset \mathsf{ID}$ of size $\leqslant r$ and gets back all the secret keys $\{\mathsf{sk_{id}} \leftarrow \mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathsf{id})\}_{\mathsf{id} \in \mathcal{R}}$.

[3]Recently, a more general model of pirate, called *pirate distinguisher*, have been introduced and considered in [NWZ16, GKW18]. However, as proven in [DPY20], in the bit-encryption setting, such a notion of pirate distinguisher is equivalent to the pirate decoder. In this section, we consider bit-encryption and in the next section about multi-bit encryption, the tracing is reduced to the tracing in the bit-encryption sub schemes. Therefore, we keep using the definition from [ABP+17] (adapted to the symmetric-key setting).

- Adversary $\mathcal{A}$ then produces a pirate decoder $\mathcal{D}$ and gives the challenger its access in terms of an oracle $\mathcal{O}^{\mathcal{D}}$. $\mathcal{A}$ also produces a suspect set $\mathcal{S}$ of size $\leqslant t$ containing $\mathcal{T}$ and sends it to the challenger.
- The challenger then runs $\mathsf{Trace}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, \mathcal{S}, \mathcal{O}^{\mathcal{D}})$. The adversary wins if both of the following hold:
  - Equation (4.2) is satisfied for the set of revoked users $\mathcal{R}$ chosen by the adversary (i.e., decoder $\mathcal{D}$ is useful),
  - the execution of Trace outputs $\perp$ or outputs an $\mathsf{id} \in \mathcal{S} \setminus \mathcal{T}$ with probability $\geqslant 1/\lambda^c$.

We define the tracing advantage $\mathrm{Adv}^{\mathsf{AD\text{-}TT}}_{\mathbf{TR},\mathcal{A}}$ as the probability of $\mathcal{A}$'s win. A trace-and-revoke scheme $\mathbf{TR}$ is said to be $\mathsf{AD\text{-}TT}$ secure if the advantage $\mathrm{Adv}^{\mathsf{AD\text{-}TT}}_{\mathbf{TR},\mathcal{A}}$ is negligible for all ppt adversary $\mathcal{A}$.

4.3 Trace-and-Revoke from Linear Functional Encryption

In this section, we construct a trace-and-revoke system from a linear functional encryption scheme that achieves traceability and anonymous revocation. This is achieved in two steps. First, a trace-and-revoke system for single-bit messages is constructed from inner product functional encryption. Then we extend such a trace-and-revoke system to support arbitrary fixed length strings.

We first define a generic transformation similar to the one of [ABP+17], which converts an $\mathsf{IND\text{-}CPA}$ secure inner product functional encryption scheme $\mathcal{IPFE}$ into a trace-and-revoke system $\mathbf{TR}_0$ for the restricted message space $M = \{0,1\}$ that enjoys anonymous revocation. Note that this transformation converts an $\mathsf{IND\text{-}CPA}$ secure $\mathcal{IPFE}$ in the bounded collusion model to a trace-and-revoke system $\mathbf{TR}_0$ that supports an exponential number of users like [ABP+17]. Then we provide another generic transformation that converts an $\mathsf{IND\text{-}CPA}$ secure matrix multiplication functional encryption scheme $\mathcal{MMFE}$ into a trace-and-revoke system $\mathbf{TR}_1$ for the message space $M = \{0,1\}^n$ for n as large as $\mathsf{poly}(\lambda)$. This transformation also ensures that $\mathbf{TR}_1$ achieves anonymous revocation along with supporting an exponential number of users.

As, our primary contribution in this chapter, is to introduce trace-and-revoke schemes with anonymous revocation, our presentation mainly focuses on the construction and the anonymity security of $\mathbf{TR}_0$ and $\mathbf{TR}_1$. Nevertheless, in Section 4.3.1, we have provided a complete description of the $\mathbf{TR}_0$ that includes an explicit description of the Trace function. For the sake of simplicity, we however have presented the general trace-and-revoke systems $\mathbf{TR}_1$ in Section 4.3.2 without a Trace. Note that, $\mathbf{TR}_1$ can use the Trace algorithm of $\mathbf{TR}_0$.

4.3.1 Trace-and-Revoke for Single Bit Messages

We construct a trace-and-revoke scheme $\mathbf{TR}_0$ following the specifications of Definition 4.8 for the message space $M = \{0,1\}$. $\mathbf{TR}_0$ relies on a user directory dir which

contains the identities of all the users that have been assigned keys in the system. This user directory is initially empty. Unlike [ABP+17], we assume that dir can only be accessed by the central authority, which is the sender as well as the key generator. $\mathbf{TR}_0$ relies on an inner product functional encryption scheme IPFE for the ℓ-dimensional vector space on $\mathbb{Z}_p$, where the value ℓ is a function of r and t. Recall that, in a typical trace-and-revoke scheme, the bound on the number of revoked users r and the bound on the number of suspected users (traitors) t are given as the system parameters. The description of IPFE comes with an injective map f whose description is included in the public parameters pp. To define the trace-and-revoke scheme $\mathbf{TR}_0$, we define a special element in the range of the map $\mathit{elem}^* = f(0)$. Concretely, in case of a group-based construction of IPFE, we take the exponentiation map $f : x \mapsto [x]$ and have $\mathit{elem}^* = [0]$. In case of a lattice-based construction, we take the identity map $f : x \mapsto x$ and have $\mathit{elem}^* = 0$.

1. $\mathsf{Setup}(1^\lambda, 1^r, 1^\mathsf{t})$. Upon input the security parameter λ, the bound t on the number of the suspected users, and the bound r on the number of revoked users, set $p = \lambda^{\omega(1)}$ and proceed as follows:
   (a) Let $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell, p)$, where we set $\ell = 2r + t + 1$. The key space $\mathcal{K}$ and ciphertext space C are the IPFE key space and ciphertext space, respectively.
   (b) Create an empty directory dir.
   (c) Output the public parameter pp, master secret key msk and the (empty) user directory dir.
2. $\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathsf{id})$. Upon input the public parameters pp, the master secret key msk, the user directory dir and a user identity $\mathsf{id} \in \mathsf{ID}$, proceed as follows:
   (a) Sample $\mathbf{x}_\mathsf{id} \leftarrow \mathbb{Z}_p^\ell$. The pair $\mathsf{u}_\mathsf{id} = (\mathsf{id}, \mathbf{x}_\mathsf{id})$ is then appended to the user directory dir.
   (b) Let $(\mathsf{sk}_\mathsf{id}, \mathbf{x}_\mathsf{id}) \leftarrow \mathit{IPFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x}_\mathsf{id})$.
   (c) Output $(\mathsf{sk}_\mathsf{id}, \mathbf{x}_\mathsf{id})$.
3. $\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)$. Upon input the public parameters pp, the master secret key msk, the user directory dir, a set of revoked users $\mathcal{R}$ of size $\leqslant r$ and a plaintext message $m \in M = \{0, 1\}$, proceed as follows:
   (a) Sample $\mathbf{v}_\mathcal{R} \leftarrow X_\mathcal{R}^\perp$ where $X_\mathcal{R} = \{\mathbf{x}_\mathsf{id} \ : \ \mathsf{id} \in \mathcal{R}\}$.
   (b) Compute $\mathbf{y}_\mathcal{R} = m \cdot \mathbf{v}_\mathcal{R}$.
   (c) Output $C_\mathcal{R} = \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_\mathcal{R})$.
4. $\mathsf{Dec}(\mathsf{pp}, (\mathsf{sk}_\mathsf{id}, \mathbf{x}_\mathsf{id}), C_\mathcal{R})$. Upon input the public parameters pp, the secret key sk_id for user id and a ciphertext $C_\mathcal{R}$, proceed as follows:
   (a) Compute $\mathsf{Res} = \mathit{IPFE}.\mathsf{Dec}(\mathsf{pp}, (\mathsf{sk}_\mathsf{id}, \mathbf{x}_\mathsf{id}), C_\mathcal{R})$.
   (b) If $\mathsf{Res} = \mathit{elem}^*$, then output 0. Otherwise output 1.

5. $\mathsf{Trace}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, \mathcal{S}, \mathcal{O}^{\mathcal{D}})$. Upon input the master secret key msk, the user directory dir, a revoked set of users $\mathcal{R}$, a suspect set of users $\mathcal{S}$ and given access to the oracle $\mathcal{O}^{\mathcal{D}}$, proceed as follows:

   (a) Suppose the users in the suspect set $\mathcal{S}$ can distinguish between the messages $m = 0$ and $m' = 1$ except with negligible probability provided these users can access the oracle $\mathcal{O}^{\mathcal{D}}$.[4]

   (b) Set $\mathcal{S}_1 = \{\mathsf{id}_1, \mathsf{id}_2, \ldots\} = \mathcal{S} \setminus \mathcal{R}$.

   (c) Sample $\mathbf{v}_{\mathcal{R}} \leftarrow X_{\mathcal{R}}^{\perp}$ where $X_{\mathcal{R}} = \{\mathbf{x}_{\mathsf{id}} \ : \ \mathsf{id} \in \mathcal{R}\}$.

   (d) For all $i = 1, 2, \ldots, t$,

   - If $i = 1$, set $\mathbf{v}_{\mathcal{S}_i} = \mathbf{0}$. If $\mathcal{S}_i = \emptyset$, set $\mathbf{v}_{\mathcal{S}_i} = (m' - m) \cdot \mathbf{v}_{\mathcal{R}}$.
   - Otherwise, sample $\mathbf{v}_{\mathcal{S}_i} \leftarrow X_{\mathcal{R} \cup \mathcal{S}_i}^{\perp} \cap \left(X_{\mathcal{S}_1 \setminus \mathcal{S}_i}^{\perp} + (m' - m) \cdot \mathbf{v}_{\mathcal{R}} \right)$ where $X_{\mathcal{R} \cup \mathcal{S}_i} = \{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \mathcal{R} \cup \mathcal{S}_i\}$ and $X_{\mathcal{S}_1 \setminus \mathcal{S}_i} = \{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \mathcal{S}_1 \setminus \mathcal{S}_i\}$.
   - Construct $\mathbf{y}_i = \mathbf{v}_{\mathcal{S}_i} + m \cdot \mathbf{v}_{\mathcal{R}}$;
   - Provide the oracle $\mathcal{O}^{\mathcal{D}}$ with $(C_{\mathcal{S}_i}, m)$ as input and get a binary value b_i as output. Suppose the probability of $b_i = 1$ is p_i.
   - The probe ciphertext is $C_{\mathcal{S}_i} = \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_i)$; We note that, the decryption result of the probe ciphertext $C_{\mathcal{S}_i}$ is m if $\mathsf{id} \in \mathcal{S}_i$ and m' if $\mathsf{id} \in \mathcal{S} \setminus \mathcal{S}_i$.
   - If $i > 1$ and $|p_i - p_{i-1}|$ is non-negligible,
     - Output id_{i-1} as the traitor identity and abort;
     - If $\mathcal{S}_i = \phi$, output $\perp$ and abort. Otherwise, set $\mathcal{S}_{i+1} = \mathcal{S}_i \setminus \{\mathsf{id}_i\}$.

We state the following theorems essential for the correctness.

Theorem 4.9. Assume that $p = \lambda^{\omega(1)}$. Then, for every set $\mathcal{R}$ of revoked users of size $\leqslant r$, every $\mathsf{id} \notin \mathcal{R}$ and every $m \in M = \{0, 1\}$, we have

$$\mathsf{Dec}(\mathsf{pp}, (\mathsf{sk}_{\mathsf{id}}, \mathbf{x}_{\mathsf{id}}), \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)) = m,$$

with probability $\geqslant 1 - \lambda^{-\omega(1)}$.

Proof. As $\mathbf{x}_{\mathsf{id}}$ is uniform in $\mathbb{Z}_p^{\ell}$, $p = \lambda^{\omega(1)}$ and $\ell > r$, we have that $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle \neq 0$, with overwhelming probability. The execution of $\mathsf{Dec}(\mathsf{pp}, (\mathsf{sk}_{\mathsf{id}}, \mathbf{x}_{\mathsf{id}}), C_{\mathcal{R}})$, with $C_{\mathcal{R}} = \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)$, on Step (a) computes (with overwhelming probability):

$$\mathsf{Dec}(\mathsf{pp}, (\mathsf{sk}_{\mathsf{id}}, \mathbf{x}_{\mathsf{id}}), C_{\mathcal{R}}) = f(\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle) = f(m \cdot \langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle)$$

by the correctness of IPFE where f is the deterministic function included in pp.

Now, observe that, if $m = 0$, then $f(\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle) = f(0) = \mathit{elem}^*$. In this case, Dec outputs 0. On the other hand, if $m = 1$, then $f(\langle \mathbf{x}_{\mathsf{id}}, \mathbf{y}_{\mathcal{R}} \rangle) = f(\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle) \neq \mathit{elem}^*$ (since $\langle \mathbf{x}_{\mathsf{id}}, \mathbf{v}_{\mathcal{R}} \rangle \neq 0$ and f is injective). In this case, Dec outputs 1. Thus, for both values of m, Dec retrieves the correct value of m with overwhelming probability. □

[4] Note that [ABP+17] used Hoeffding's inequality to ensure that one can efficiently find such distinguishable m and m'. In our case, it is simpler, as $M = \{0, 1\}$.

Theorem 4.10. Let $\mathcal{R}$ be arbitrary of size $\leqslant r$ and assume Eq. (4.2) holds for $\mathcal{O}^{\mathcal{D}}$ and $\mathcal{R}$. Then we have:

$$\left| \Pr_{C \leftarrow \mathsf{Enc}(\mathsf{pp},\mathsf{msk},\mathsf{dir},\mathcal{R},0)}[\mathcal{O}^{\mathcal{D}}(C,0)=1] - \Pr_{C \leftarrow \mathsf{Enc}(\mathsf{pp},\mathsf{msk},\mathsf{dir},\mathcal{R},1)}[\mathcal{O}^{\mathcal{D}}(C,0)=1] \right| \geqslant \frac{2}{\lambda^c},$$

with probability $\geqslant 1-\lambda^{-\omega(1)}$ and for some constant $c>0$.

Proof. By Eq. (4.2), we have

$$\Pr_{C \leftarrow \mathsf{Enc}(\mathsf{pp},\mathsf{msk},\mathsf{dir},\mathcal{R},0)}[\mathcal{O}^{\mathcal{D}}(C,0)=1] \geqslant \frac{1}{2}+\frac{1}{\lambda^c}, \qquad \Pr_{C \leftarrow \mathsf{Enc}(\mathsf{pp},\mathsf{msk},\mathsf{dir},\mathcal{R},1)}[\mathcal{O}^{\mathcal{D}}(C,1)=1] \geqslant \frac{1}{2}+\frac{1}{\lambda^c}.$$

The latter means that if $m'=1$ is encrypted as C, then $\mathcal{O}^{\mathcal{D}}(C,1)$ outputs 1 with probability non-negligibly better than a random choice. Taking the complement, we obtain that

$$\Pr_{C \leftarrow \mathsf{Enc}(\mathsf{pp},\mathsf{msk},\mathsf{dir},\mathcal{R},1)}[\mathcal{O}^{\mathcal{D}}(C,0)=1] < \frac{1}{2}-\frac{1}{\lambda^c}.$$

The result follows naturally. □

Security.

We prove that the base scheme $\mathbf{TR}_0$ enjoys message hiding, revocation set hiding and traceability. restriction. The proof of the IND-CPA security and the AD-TT security are done in a manner similar to the IND-CPA security proof and the AD-TT security proof of [ABP+17].

Theorem 4.11. If IPFE is an IND-CPA secure inner product functional encryption scheme allowing up to r key extraction queries, then $\mathbf{TR}_0$ is IND-CPA secure.

Proof. Let $\mathcal{A}_{\mathbf{TR}_0}$ be a ppt adversary that breaks the IND-CPA security of $\mathbf{TR}_0$. We construct a ppt adversary $\mathcal{A}_{\mathit{IPFE}}$ that breaks the IND-CPA security of the underlying IPFE:

- It first obtains the public parameter pp output by the IPFE challenger (which runs the $\mathit{IPFE}.\mathsf{Setup}(1^\lambda,1^\ell)$ algorithm) and relays it to $\mathcal{A}_{\mathbf{TR}_0}$. On $\mathcal{A}_{\mathbf{TR}_0}$'s request, the adversary $\mathcal{A}_{\mathit{IPFE}}$ creates dir with polynomially many $(\mathsf{id},\mathbf{x}_{\mathsf{id}})$ pairs for $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^\ell$. The IPFE challenger samples $\beta \leftarrow \{0,1\}$.
- The adversary $\mathcal{A}_{\mathbf{TR}_0}$ can make multiple secret key queries on $\mathsf{id} \in \mathsf{ID}$ and multiple challenge ciphertext queries on $(m_0,m_1,\mathcal{R})$.
  - For every secret key query on id,
    - $\mathcal{A}_{\mathit{IPFE}}$ retrieves $\mathbf{x}_{\mathsf{id}} = \mathsf{dir}[\mathsf{id}]$.
    - $\mathcal{A}_{\mathit{IPFE}}$ then sends $\mathbf{x}_{\mathsf{id}}$ to the IPFE challenger. The latter returns $\mathsf{sk}_{\mathbf{x}_{\mathsf{id}}}$, which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as $\mathsf{sk}_{\mathsf{id}}$.
  - For every challenge anonymity query on $(m_0,m_1,\mathcal{R})$,
    - It samples $\mathbf{v}_{\mathcal{R}} \leftarrow X^\perp$ where $X = \{\mathbf{x}_{\mathsf{id}} \in \mathbb{Z}_p^\ell \ : \ \mathsf{id} \in \mathcal{R}\}$.

* It sends $\mathbf{y}_0 = m_0 \cdot \mathbf{v}_{\mathcal{R}}$ and $\mathbf{y}_1 = m_1 \cdot \mathbf{v}_{\mathcal{R}}$ to the IPFE challenger. The latter encrypts $\mathbf{y}_\beta$ as $\mathsf{ct}^{(\beta)} \leftarrow \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_\beta)$ and outputs $\mathsf{ct}^{(\beta)}$.
* It forwards the received ciphertext to $\mathcal{A}_{\mathbf{TR}_0}$ as its challenge $C^{(\beta)}$.

- Finally, the $\mathcal{A}_{\mathbf{TR}_0}$ adversary outputs its guess $\beta' \in \{0,1\}$ and $\mathcal{A}_{\mathit{IPFE}}$ also outputs β' as its own guess of β.

Note that adversary $\mathcal{A}_{\mathit{IPFE}}$ behaves as an $\mathsf{IND\text{-}CPA}$ challenger in the view of $\mathcal{A}_{\mathbf{TR}_0}$. Further, it is a valid adversary against IPFE as $\langle \mathbf{y}_0, \mathbf{x}_{\mathsf{id}} \rangle = \langle \mathbf{y}_1, \mathbf{x}_{\mathsf{id}} \rangle = 0$ for every vector $\mathbf{x}_{\mathsf{id}}$ queried to the IPFE challenger (i.e., each $\mathsf{id} \in \mathcal{R}$). The advantage of $\mathcal{A}_{\mathit{IPFE}}$ is exactly the same as the advantage of $\mathcal{A}_{\mathbf{TR}_0}$. □

Theorem 4.12. If IPFE is an $\mathsf{IND\text{-}CPA}$ secure inner product functional encryption scheme allowing up to $(t+r)$ key extraction queries, then $\mathbf{TR}_0$ is $\mathsf{mIND\text{-}ID\text{-}CPA}$ secure.

Proof. Given an $\mathsf{mIND\text{-}ID\text{-}CPA}$ adversary $\mathcal{A}_{\mathbf{TR}_0}$, we produce $\mathcal{A}_{\mathit{IPFE}}$ that breaks the $\mathsf{IND\text{-}CPA}$ security of IPFE.

- $\mathcal{A}_{\mathit{IPFE}}$ first obtains the public parameter pp output by the IPFE challenger (who runs the $\mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell)$ algorithm) and relays it to $\mathcal{A}_{\mathbf{TR}_0}$. The IPFE challenger, at this point, samples $\beta \leftarrow \{0,1\}$. On $\mathcal{A}_{\mathbf{TR}_0}$'s request, $\mathcal{A}_{\mathit{IPFE}}$ creates dir with polynomially many id without the corresponding $\mathbf{x}_{\mathsf{id}}$.

- Recall that, $\mathcal{A}_{\mathbf{TR}_0}$ can make multiple secret key queries on $\mathsf{id} \in \mathsf{ID}$ and multiple challenge ciphertext queries on $(m, \mathcal{R}_0, \mathcal{R}_1)$. To accommodate such queries, $\mathcal{A}_{\mathit{IPFE}}$ first defines a set of vector $\{\mathbf{x}_i\}_{1 \leqslant i \leqslant N} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{t+2r}\}$ where $\mathbf{x}_i \leftarrow \mathbb{Z}_p^\ell$. This set is used to answer to secret key queries.

  – For every secret key query on id,
    * If $\mathsf{id} \in \mathsf{dir}$, $\mathcal{A}_{\mathit{IPFE}}$ sets $\mathbf{x} = \mathsf{dir}[\mathsf{id}]$.
    * Otherwise, $\mathcal{A}_{\mathit{IPFE}}$ samples a vector $\mathbf{x} \leftarrow \{\mathbf{x}_i\}_{1 \leqslant i \leqslant N}$ and sets $\mathsf{dir}[\mathsf{id}] = \mathbf{x}$.
    * $\mathcal{A}_{\mathit{IPFE}}$ then sends $\mathbf{x}$ to the IPFE challenger. The latter returns $\mathsf{sk}_{\mathbf{x}_{\mathsf{id}}}$, which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as $\mathsf{sk}_{\mathsf{id}}$.
  – For every challenge anonymity query on $(m, \mathcal{R}_0, \mathcal{R}_1)$,
    * For every $\mathsf{id} \in \mathcal{R}_0 \cup \mathcal{R}_1$, if $\mathsf{id} \notin \mathsf{dir}$, $\mathcal{A}_{\mathit{IPFE}}$ samples a vector $\mathbf{x} \leftarrow \{\mathbf{x}_i\}_{1 \leqslant i \leqslant N}$ without repetition and sets $\mathsf{dir}[\mathsf{id}] = \mathbf{x}$.
    * $\mathcal{A}_{\mathit{IPFE}}$ defines $\hat{\mathcal{R}} = \mathcal{R}_0 \cap \mathcal{R}_1$.
    * Then $\mathcal{A}_{\mathit{IPFE}}$ defines three matrices:
      1. $Z = \mathsf{Matrix}(\{\mathbf{x}_i\}_{1 \leqslant i \leqslant N} \setminus Z)$ where $Z = \{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \hat{\mathcal{R}}\}$.
      2. $X_0 = \mathsf{Matrix}(X_0)$ where $X_0 = \{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in R_0\}$.
      3. $X_1 = \mathsf{Matrix}(X_1)$ where $X_1 = \{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \mathcal{R}_1\}$.
    * It samples $\begin{pmatrix} \mathbf{v}_{\mathcal{R}_0} \\ \mathbf{v}_{\mathcal{R}_1} \end{pmatrix} \leftarrow V^\perp$ where

$$V = \begin{pmatrix} Z & -Z \\ X_0 & \mathbf{0} \\ \mathbf{0} & X_1 \end{pmatrix}. \tag{4.3}$$

* It sends $\mathbf{y}_{\mathcal{R}_0} = m \cdot \mathbf{v}_{\mathcal{R}_0}$ and $\mathbf{y}_{\mathcal{R}_1} = m \cdot \mathbf{v}_{\mathcal{R}_1}$ to the IPFE challenger encrypts $\mathbf{y}_{\mathcal{R}_\beta}$ as $\mathsf{ct}^{(\beta)} \leftarrow \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_{\mathcal{R}_\beta})$ and outputs $\mathsf{ct}^{(\beta)}$.
* $\mathcal{A}_{\mathit{IPFE}}$ then forwards the received ciphertext to $\mathcal{A}_{\mathbf{TR}_0}$ as its challenge $C^{(\beta)}$.

- At the end of the game, $\mathcal{A}_{\mathbf{TR}_0}$ returns β' as its guess of β which $\mathcal{A}_{\mathit{IPFE}}$ forwards to the IPFE challenger as its answer.

From Eq. (4.3), $Z(\mathbf{v}_{\mathcal{R}_0} - \mathbf{v}_{\mathcal{R}_1}) = \mathbf{0}$, $X_0\mathbf{v}_{\mathcal{R}_0} = \mathbf{0}$ and $X_1\mathbf{v}_{\mathcal{R}_1} = \mathbf{0}$. As $\mathbf{y}_{\mathcal{R}_u} \in \mathrm{Span}(\mathbf{v}_{\mathcal{R}_u})$ for $u \in \{0,1\}$, $Z(\mathbf{y}_{\mathcal{R}_1} - \mathbf{y}_{\mathcal{R}_0}) = \mathbf{0}$ and $X_0\mathbf{y}_{\mathcal{R}_0} = X_1\mathbf{y}_{\mathcal{R}_1} = \mathbf{0}$.

We now show that $\mathcal{A}_{\mathit{IPFE}}$ is a valid challenger against $\mathcal{A}_{\mathbf{TR}_0}$ in the $\mathsf{mIND\text{-}ID\text{-}CPA}$ security model. For that we show, for every i^{th} key query on id_i and j^{th} challenge ciphertext query on $(m^{(j)}, \mathcal{R}_0^{(j)}, \mathcal{R}_1^{(j)})$ from $\mathcal{A}_{\mathbf{TR}_0}$, $\mathcal{A}_{\mathit{IPFE}}$ can forward corresponding vectors to the IPFE challenger. Due to the natural restriction, note that, $\mathsf{id}_i \in (\mathcal{R}_0^{(j)} \cap \mathcal{R}_1^{(j)}) \sqcup (\mathsf{ID} \setminus (\mathcal{R}_0^{(j)} \cup \mathcal{R}_1^{(j)}))$ for all $i \in [1, t+r-1]$ and all $j \in [1, r]$.

For all queried id_i, if one of the following two holds.

- $\mathsf{id}_i \in (\mathcal{R}_0^{(j)} \cap \mathcal{R}_1^{(j)})$: This means, $\mathsf{id}_i \notin (\mathcal{R}_0^{(t)} \Delta \mathcal{R}_1^{(t)})$ for all $t \in [1, j-1]$ due to the natural restriction. The corresponding $\mathbf{x}_{\mathsf{id}_i} \in \{\mathbf{x}_i\}_{1 \leqslant i \leqslant N} \cap \hat{\mathcal{R}}^{(j)}$ and by our reduction, the $\mathbf{x}_{\mathsf{id}_i}$ vector is included in the definition of $X_0^{(j)}$ and $X_1^{(j)}$. From Eq. (4.3) above, we see that $X_0^{(j)}\mathbf{y}_{\mathcal{R}_0}^{(j)} = X_1^{(j)}\mathbf{y}_{\mathcal{R}_1}^{(j)} = \mathbf{0}$. Thus, $\mathcal{A}_{\mathit{IPFE}}$ can forward this to the IPFE challenger for key query.

- $\mathsf{id}_i \in \mathsf{ID} \setminus (\mathcal{R}_0^{(j)} \cup \mathcal{R}_1^{(j)})$: Observe that, the corresponding $\mathbf{x}_{\mathsf{id}_i} \in \{\mathbf{x}_i\}_{1 \leqslant i \leqslant N} \setminus \hat{\mathcal{R}}^{(j)}$ and such $\mathbf{x}_{\mathsf{id}_i}$ is included in the definition of $Z^{(j)}$. From Eq. (4.3) above, we see that $Z^{(j)}\mathbf{y}_{\mathcal{R}_0}^{(j)} = Z^{(j)}\mathbf{y}_{\mathcal{R}_1}^{(j)} \neq \mathbf{0}$ (w.h.p. as $\mathbf{y}_{\mathcal{R}_0}^{(j)}, \mathbf{y}_{\mathcal{R}_1}^{(j)}$ are sampled randomly). Thus, $\mathcal{A}_{\mathit{IPFE}}$ can forward this to the IPFE challenger for key query.

Next, note that, for every query on id_i from $\mathcal{A}_{\mathbf{TR}_0}$, the adversary $\mathcal{A}_{\mathit{IPFE}}$ returns a distinct random vector $\mathbf{x}_{\mathsf{id}_i}$ from $\{\mathbf{x}_i\}_{1 \leqslant i \leqslant N}$ that were sampled at the starting of the reduction. The crucial point here is $\mathcal{A}_{\mathit{IPFE}}$ faces at most $(t+2r)$ many distinct identities id, hence $\{\mathbf{x}_i\}_{1 \leqslant i \leqslant N}$ is sufficient to assign the corresponding $\mathbf{x}_{\mathsf{id}}$. Moreover, $\mathcal{A}_{\mathbf{TR}_0}$ gets encryption of either $\mathbf{y}_{\mathcal{R}_0}$ or $\mathbf{y}_{\mathcal{R}_1}$ where both the vectors are randomly sampled. Thus, from the point of view of $\mathcal{A}_{\mathbf{TR}_0}$, $Z\mathbf{y}_{\mathcal{R}_b}$ is a random vector. Thus, $\mathcal{A}_{\mathit{IPFE}}$ behaves as a valid $\mathsf{mIND\text{-}ID\text{-}CPA}$ challenger to $\mathcal{A}_{\mathbf{TR}_0}$.

As we have seen above, for every $\mathbf{x}_{\mathsf{id}_i}$ and $(\mathbf{y}_{\mathcal{R}_0}^{(j)}, \mathbf{y}_{\mathcal{R}_1}^{(j)})$ the adversary $\mathcal{A}_{\mathit{IPFE}}$ gives to the IPFE challenger, $\langle \mathbf{x}_{\mathsf{id}_i}, \mathbf{y}_{\mathcal{R}_0}^{(j)} \rangle = \langle \mathbf{x}_{\mathsf{id}_i}, \mathbf{y}_{\mathcal{R}_1}^{(j)} \rangle$ holds. Thus, $\mathcal{A}_{\mathit{IPFE}}$ behaves as a valid $\mathsf{IND\text{-}CPA}$ adversary to the IPFE challenger.

If $\mathcal{A}_{\mathbf{TR}_0}$ can distinguish between any $\mathcal{R}_0^{(j)}$ and $\mathcal{R}_1^{(j)}$, $\mathcal{A}_{\mathit{IPFE}}$ can distinguish between corresponding $\mathbf{y}_{\mathcal{R}_0}^{(j)}$ and $\mathbf{y}_{\mathcal{R}_1}^{(j)}$. Thus, the advantage of $\mathcal{A}_{\mathit{IPFE}}$ is exactly the same as the advantage of $\mathcal{A}_{\mathbf{TR}_0}$. □

Theorem 4.13. If IPFE is an $\mathsf{IND\text{-}CPA}$ secure inner product functional encryption scheme allowing $(r+t)$ queries, then $\mathbf{TR}_0$ is $\mathsf{AD\text{-}TT}$ secure.

Proof. Given an $\mathsf{AD\text{-}TT}$ adversary $\mathcal{A}_{\mathbf{TR}_0}$, we have to produce $\mathcal{A}_{\mathit{IPFE}}$ that breaks the $\mathsf{IND\text{-}CPA}$ security of IPFE. $\mathcal{A}_{\mathit{IPFE}}$ first obtains the public parameter pp output by the

IPFE challenger (who runs the $\mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell)$ algorithm) and relays it to $\mathcal{A}_{\mathbf{TR}_0}$. On $\mathcal{A}_{\mathbf{TR}_0}$'s request, $\mathcal{A}_{\mathit{IPFE}}$ creates dir with polynomially many id without the corresponding $\mathbf{x}_{\mathsf{id}}$. The IPFE challenger, being a symmetric key primitive, provides $\mathcal{A}_{\mathit{IPFE}}$ polynomially many accesses to the encryption oracle $O_{\mathsf{ct}}(\cdot)$ and to the key generation oracle $O_{\mathsf{sk}}(\cdot)$.

$\mathcal{A}_{\mathbf{TR}_0}$ adaptively chooses $\mathsf{id} \in \mathsf{ID}$, $\mathcal{A}_{\mathit{IPFE}}$ assigns a random $\mathbf{x}_{\mathsf{id}}$ to $\mathsf{dir}[\mathsf{id}]$ and makes query to O_{sk} on $\mathbf{x}_{\mathsf{id}}$. The response it gets is forwarded to $\mathcal{A}_{\mathbf{TR}_0}$ as the secret key $\mathsf{sk}_{\mathsf{id}}$. $\mathcal{A}$ can make at most t many such queries and these queries are collected as a set $\mathcal{T}$.

$\mathcal{A}_{\mathbf{TR}_0}$ then adaptively chooses $\mathcal{R} \subset \mathsf{ID}$ such that $|\mathcal{R}| \leqslant r$. For every $\mathsf{id} \in \mathcal{R}$, $\mathcal{A}_{\mathit{IPFE}}$ assigns a $\mathbf{x}_{\mathsf{id}}$ and makes query to O_{sk} on $\mathbf{x}_{\mathsf{id}}$, the response it gets is forwarded to $\mathcal{A}_{\mathbf{TR}_0}$ as the secret key $\mathsf{sk}_{\mathsf{id}}$.

Finally, $\mathcal{A}$ produces a pirate decoder $\mathcal{O}^{\mathcal{D}}$ and a suspected list of traitors $\mathcal{S}$ that includes the traitor set $\mathcal{T}$ where $|\mathcal{S}| \leqslant t$. Next, $\mathcal{A}_{\mathit{IPFE}}$ runs Trace on $\mathcal{S}$ and $\mathcal{R}$ given access to O_{ct} and $\mathcal{O}^{\mathcal{D}}$. Precisely, for all $i \in [1, |\mathcal{S}|]$, $\mathcal{A}_{\mathit{IPFE}}$ computes $\mathbf{v}_{\mathcal{S}_i}$ and asks O_{ct} to get the so-called probe-ciphertext $C_{\mathcal{S}_i}$. Finally, Trace outputs either $\perp$ or some $\mathsf{id} \in \mathcal{S}$. More specifically, the winning condition of $\mathsf{AD\text{-}TT}$ security model tells that Trace outputs either $\perp$ or some $\mathsf{id} \in \mathcal{S} \setminus \mathcal{T}$ with probability $\geqslant 1/\lambda^c$ for some constant $c > 0$.

In case, Trace outputs $\perp$, $\mathcal{A}_{\mathit{IPFE}}$ outputs a random bit. Otherwise, we assume id to be id_{i-1} for which Trace aborted on the i^{th} round for some $i < t$. Then, by the description of Trace, $|p_i - p_{i-1}|$ is non-negligible. At this point $\mathcal{A}_{\mathit{IPFE}}$ retrieves $\mathbf{v}_{\mathcal{S}_{i-1}}$ and $\mathbf{v}_{\mathcal{S}_i}$ to define $\mathbf{y}_0 = \mathbf{v}_{\mathcal{S}_{i-1}} + m \cdot \mathbf{v}_{\mathcal{R}}$ and $\mathbf{y}_1 = \mathbf{v}_{\mathcal{S}_i} + m \cdot \mathbf{v}_{\mathcal{R}}$ and makes challenge ciphertext query to the IPFE challenger where $\mathcal{S}_i = \mathcal{S}_{i-1} \setminus \{\mathsf{id}_{i-1}\}$. The IPFE challenger responds with $C^{(\beta)} \leftarrow \mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_\beta)$ for $\beta \leftarrow \{0, 1\}$. $\mathcal{A}_{\mathit{IPFE}}$ runs $\beta' \leftarrow \mathcal{O}^{\mathcal{D}}(C^{(\beta)}, m)$ and outputs $(1 - \beta')$.

Here, we first show that $\mathcal{A}_{\mathit{IPFE}}$ is a valid adversary in the $\mathsf{IND\text{-}CPA}$ security model. In other words, we show that for all secret key queries on $\mathsf{id} \in \mathcal{R} \cup \mathcal{T}$, $\langle \mathbf{y}_0, \mathbf{x}_{\mathsf{id}} \rangle = \langle \mathbf{y}_1, \mathbf{x}_{\mathsf{id}} \rangle$. This can be seen from the following:

1. $\mathsf{id} \in \mathcal{R}$: $\langle \mathbf{y}_0, \mathbf{x}_{\mathsf{id}} \rangle = \langle \mathbf{y}_1, \mathbf{x}_{\mathsf{id}} \rangle = 0$.
2. $\mathsf{id} \in \mathcal{T} \cap \mathcal{S}_{i-1}$: $\langle \mathbf{y}_0, \mathbf{x}_{\mathsf{id}} \rangle = \langle \mathbf{y}_1, \mathbf{x}_{\mathsf{id}} \rangle = m \cdot \langle \mathbf{v}_{\mathcal{R}}, \mathbf{x}_{\mathsf{id}} \rangle$.
3. $\mathsf{id} \in \mathcal{T} \cap (\mathcal{S}_1 \setminus \mathcal{S}_{i-1})$: $\langle \mathbf{y}_0, \mathbf{x}_{\mathsf{id}} \rangle = \langle \mathbf{y}_1, \mathbf{x}_{\mathsf{id}} \rangle = m' \cdot \langle \mathbf{v}_{\mathcal{R}}, \mathbf{x}_{\mathsf{id}} \rangle$.

Now, we show that $\mathcal{A}_{\mathit{IPFE}}$ wins with probability given Trace didn't output $\perp$.

If $\beta = 0$. $C^{(\beta)}$ is an encryption of $\mathbf{y}_0$ that encoded $\mathbf{v}_{\mathcal{S}_{i-1}}$. The description of the Trace tells that $\langle \mathbf{v}_{\mathcal{S}_{i-1}}, \mathbf{x}_{\mathsf{id}_{i-1}} \rangle = 0$. Thus, given $\mathcal{O}^{\mathcal{D}}$ one views $C^{(\beta)}$ as an encryption of $\mathbf{y}_0 = m \cdot \mathbf{v}_{\mathcal{R}}$. In this case, $\mathcal{O}^{\mathcal{D}}(C^{(\beta)}, m)$ gives $\beta' = 1$ with very high probability.

If $\beta = 1$. $C^{(\beta)}$ is an encryption of $\mathbf{y}_1$ that encoded $\mathbf{v}_{\mathcal{S}_i}$. The description of the Trace tells that $\langle \mathbf{v}_{\mathcal{S}_i}, \mathbf{x}_{\mathsf{id}_{i-1}} \rangle = (m' - m) \cdot \langle \mathbf{v}_{\mathcal{R}}, \mathbf{x}_{\mathsf{id}_{i-1}} \rangle$. Thus, given $\mathcal{O}^{\mathcal{D}}$ one views $C^{(\beta)}$ as an encryption of $\mathbf{y}_1 = (m' - m) \cdot \mathbf{v}_{\mathcal{R}} + m \cdot \mathbf{v}_{\mathcal{R}} = m' \cdot \mathbf{v}_{\mathcal{R}}$. In this case, $\mathcal{O}^{\mathcal{D}}(C^{(\beta)}, m)$ gives $\beta' = 0$ with very high probability due to the so-called usefulness of $\mathcal{O}^{\mathcal{D}}$.

Thus, when $\mathcal{A}_{\mathbf{TR}_0}$ gives β', $\mathcal{A}_{\mathit{IPFE}}$ just forwards $(1 - \beta')$ as its guess of β.

Now, we prove that the probability that Trace outputs $\perp$ is negligible. We mention, [ABP+17, Lemma 17] already have made this argument. However, for completeness, we overview the argument here. From Theorem 4.10, we see that $\mathcal{O}^{\mathcal{D}}$ distinguishes between

$m = 0$ and $m' = 1$ with probability $\geqslant 2/\lambda^c$ for some constant $c > 0$. The description of Trace tells that $\left|\sum_{i\in[1,t]}(p_i - p_{i-1})\right| \geqslant 2/\lambda^c$ that is non-negligible. Then, by triangle inequality, there exists an i such that $|p_i - p_{i-1}|$ is non-negligible. Thus, Trace outputs id_{i-1} with non-negligible probability and aborts. Therefore, the probability that Trace continues t many iterations and outputs $\perp$ is negligible. □

Theorem 4.14. If IPFE is an $\mathsf{IND\text{-}CPA}$ secure inner product functional encryption scheme allowing up to $(t + r - 1)$ key extraction queries, then $\mathbf{TR}_0$ is $\mathsf{IND\text{-}ID^*\text{-}CPA}$ secure.

Before we give the proof, we informally discuss the necessity of such unusual restriction of $\mathsf{IND\text{-}ID^*\text{-}CPA}$ security. Note that, in $\mathbf{TR}_0$, for every id we assign a uniformly random vector $\mathbf{x}_{\mathsf{id}}$. However, being a symmetric-key trace-and-revoke, we define such an assignment on the fly when an id is referred for the first time. Thus, in the post-challenge phase, we can say for all id in pre-challenge ciphertext queries, a corresponding $\mathbf{x}_{\mathsf{id}}$ vector has already been assigned. With overwhelming probability, such $\mathbf{x}_{\mathsf{id}} \notin \mathsf{RowSpan}(Z)$ (see Eq. (4.4).) This then creates a distributional problem while simulation. To avoid such scenario, we impose the restriction only on post-challenge "outsider corruption" queries not to include id for which $(\mathsf{id}, \mathbf{x}_{\mathsf{id}})$ relation has been fixed but has not been queried for key extraction. We now give a formal proof of the theorem.

Proof. Let $\mathcal{A}_{\mathbf{TR}_0}$ be a ppt adversary that breaks the $\mathsf{IND\text{-}ID^*\text{-}CPA}$ security of $\mathbf{TR}_0$. Note that $\mathcal{A}_{\mathbf{TR}_0}$ is allowed to corrupt at most t legitimate users and the ciphertext is created considering at most r revoked users. We construct a ppt adversary $\mathcal{A}_{\mathit{IPFE}}$ that breaks the $\mathsf{IND\text{-}CPA}$ security of the underlying IPFE.

- It first obtains the public parameter pp output by the IPFE challenger (who runs the $\mathit{IPFE}.\mathsf{Setup}(1^\lambda, 1^\ell)$ algorithm) and relays it to $\mathcal{A}_{\mathbf{TR}_0}$. On $\mathcal{A}_{\mathbf{TR}_0}$'s request, the adversary $\mathcal{A}_{\mathit{IPFE}}$ creates dir with polynomially many $(\mathsf{id}, \mathbf{x}_{\mathsf{id}})$ pairs for $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^\ell$. It then sets up two empty dictionaries $Q_{\mathsf{sk}} = \{\}$ and $Q_{\mathsf{ct}} = \{\}$. Informally speaking, Q_{sk} contains all id for which key query have been/could be made and Q_{ct} contains all id on which key query has not yet been made.
- When $\mathcal{A}_{\mathit{IPFE}}$ receives a pre-challenge secret key query for $\mathsf{id} \in \mathsf{ID}$ from $\mathcal{A}_{\mathbf{TR}_0}$, it proceeds as follows:
  - If $\mathsf{id} \in Q_{\mathsf{ct}}$, it updates $Q_{\mathsf{sk}}[\mathsf{id}] = Q_{\mathsf{ct}}[\mathsf{id}]$ and removes the id entry from Q_{ct}.
  - If $\mathsf{id} \notin Q_{\mathsf{sk}}$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^\ell$ and sets $Q_{\mathsf{sk}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.
  - If $\mathsf{id} \in Q_{\mathsf{sk}}$, it sets $\mathbf{x}_{\mathsf{id}} = Q_{\mathsf{sk}}[\mathsf{id}]$.
  - It then sends $\mathbf{x}_{\mathsf{id}}$ to the IPFE challenger. The latter returns $\mathsf{sk}_{\mathbf{x}_{\mathsf{id}}}$, which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as $\mathsf{sk}_{\mathsf{id}}$.
- When $\mathcal{A}_{\mathit{IPFE}}$ receives a ciphertext query on $(m, \mathcal{R})$, it proceeds as follows:
  - For all $\mathsf{id} \in \mathcal{R} \setminus ((\mathcal{R} \cap Q_{\mathsf{sk}}.\mathsf{vals}()) \cup (\mathcal{R} \cap Q_{\mathsf{ct}}.\mathsf{vals}()))$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^\ell$ and then adds $Q_{\mathsf{ct}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.

* It samples $\mathbf{v}_{\mathcal{R}} \leftarrow X^{\perp}$ where $X = \{\mathbf{x}_{\mathsf{id}} \ : \ \mathsf{id} \in \mathcal{R}\} \subseteq Q_{\mathsf{sk}}.\mathsf{vals}() \cup Q_{\mathsf{ct}}.\mathsf{vals}()$.
* It sends $\mathbf{y} = m \cdot \mathbf{v}_{\mathcal{R}}$ to the IPFE challenger. The latter returns $\mathsf{ct}_{\mathbf{y}}$, which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as the ciphertext response $\mathsf{ct}_{\mathcal{R}}$.

- When $\mathcal{A}_{\mathit{IPFE}}$ receives $\mathcal{A}_{\mathbf{TR}_0}$'s challenge query on $(m, \mathcal{R}_0, \mathcal{R}_1)$, it proceeds as follows:
  1. First, it sets $Q_{\mathsf{idR}} = \{\mathsf{id} : \mathsf{id} \in Q_{\mathsf{ct}} \setminus (\mathcal{R}_0 \cup \mathcal{R}_1)\}$.
  2. Then it defines $\hat{\mathcal{R}}_0 = \mathcal{R}_0 \setminus \mathcal{R}_1$, $\hat{\mathcal{R}} = \mathcal{R}_0 \cap \mathcal{R}_1$ and $\hat{\mathcal{R}}_1 = \mathcal{R}_1 \setminus \mathcal{R}_0$.
  3. For all $\mathsf{id} \in \hat{\mathcal{R}} \setminus Q_{\mathsf{sk}}$,
     - If $\mathsf{id} \notin Q_{\mathsf{ct}}$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^{\ell}$ and updates $Q_{\mathsf{sk}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.
     - Otherwise, it updates $Q_{\mathsf{sk}}[\mathsf{id}] = Q_{\mathsf{ct}}[\mathsf{id}]$ and removes the id entry from Q_{ct}.
  4. Then it defines $Z = \mathsf{Matrix}((Q_{\mathsf{sk}}.\mathsf{vals}() \setminus \hat{\mathcal{R}}) \sqcup T)$ where $T = \{\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^{\ell}\}$ is of size $t - |(Q_{\mathsf{sk}}.\mathsf{vals}() \setminus \hat{\mathcal{R}})|$.
  5. For all $\mathsf{id} \in (\hat{\mathcal{R}}_0 \setminus Q_{\mathsf{ct}}.\mathsf{vals}()) \cup (\hat{\mathcal{R}}_1 \setminus Q_{\mathsf{ct}}.\mathsf{vals}())$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^{\ell}$ and updates $Q_{\mathsf{sk}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.
  6. It sets $X_0 = \mathsf{Matrix}(\{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \mathcal{R}_0 \cap (Q_{\mathsf{sk}}.\mathsf{vals}() \cup Q_{\mathsf{ct}}.\mathsf{vals}())\})$ and $X_1 = \mathsf{Matrix}(\{\mathbf{x}_{\mathsf{id}} : \mathsf{id} \in \mathcal{R}_1 \cap (Q_{\mathsf{sk}}.\mathsf{vals}() \cup Q_{\mathsf{ct}}.\mathsf{vals}())\})$.
  7. It samples $\begin{pmatrix} \mathbf{v}_{\mathcal{R}_0} \\ \mathbf{v}_{\mathcal{R}_1} \end{pmatrix} \leftarrow V^{\perp}$ for
  $$V = \begin{pmatrix} Z & -Z \\ X_0 & \mathbf{0} \\ \mathbf{0} & X_1 \end{pmatrix}. \tag{4.4}$$
  8. It sends $\mathbf{y}_{\mathcal{R}_0} = m \cdot \mathbf{v}_{\mathcal{R}_0}$ and $\mathbf{y}_{\mathcal{R}_1} = m \cdot \mathbf{v}_{\mathcal{R}_1}$ to the IPFE challenger who samples $\beta \leftarrow \{0,1\}$ and encrypts $\mathbf{y}_{\mathcal{R}_\beta}$ as $\mathsf{ct}^{(\beta)} \leftarrow \mathit{IPFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathbf{y}_{\mathcal{R}_\beta})$ and outputs $\mathsf{ct}^{(\beta)}$.
  9. $\mathcal{A}_{\mathit{IPFE}}$ then forwards the received ciphertext to $\mathcal{A}_{\mathbf{TR}_0}$ as its challenge $C^{(\beta)}$.

- $\mathcal{A}_{\mathbf{TR}_0}$ can make queries for secret key on $\mathsf{id} \in \mathsf{ID}$ and for ciphertext queries on $\mathcal{R}$.
  * For all post-challenge key queries on id, $\mathcal{A}_{\mathit{IPFE}}$ does the following:
    1. If $\mathsf{id} \in \mathcal{R}_0 \cap \mathcal{R}_1$, it retrieves $\mathbf{x}_{\mathsf{id}} = Q_{\mathsf{sk}}[\mathsf{id}]$.
    2. If $\mathsf{id} \notin \mathcal{R}_0 \cup \mathcal{R}_1$, if $\mathsf{id} \notin Q_{\mathsf{sk}}$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathsf{RowSpan}(Z)$ and sets $Q_{\mathsf{sk}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.
    3. It then sends $\mathbf{x}_{\mathsf{id}}$ to the challenger who returns $\mathsf{sk}_{\mathbf{x}_{\mathsf{id}}}$ which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as $\mathsf{sk}_{\mathsf{id}}$.
  * For all ciphertext queries on $(m, \mathcal{R})$, $\mathcal{A}_{\mathit{IPFE}}$ does the following:
    1. For all $\mathsf{id} \in \mathcal{R} \setminus ((\mathcal{R} \cap Q_{\mathsf{sk}}.\mathsf{vals}()) \cup (\mathcal{R} \cap Q_{\mathsf{ct}}.\mathsf{vals}()))$, it samples $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^{\ell}$ and then adds $Q_{\mathsf{ct}}[\mathsf{id}] = \mathbf{x}_{\mathsf{id}}$.
    2. It samples $\mathbf{v}_{\mathcal{R}} \leftarrow X^{\perp}$ where $X = \{\mathbf{x}_{\mathsf{id}} \ : \ \mathsf{id} \in \mathcal{R}\}$.

3. Sends $\mathbf{y} = m \cdot \mathbf{v}_{\mathcal{R}}$ to the challenger who returns $\mathsf{ct}_{\mathbf{y}}$ which $\mathcal{A}_{\mathit{IPFE}}$ forwards to $\mathcal{A}_{\mathbf{TR}_0}$ as the ciphertext response $\mathsf{ct}_{\mathcal{R}}$.

- Finally, adversary $\mathcal{A}_{\mathbf{TR}_0}$ outputs its guess $\beta' \in \{0,1\}$ and $\mathcal{A}_{\mathit{IPFE}}$ also outputs β' as its own guess of β.

From Eq. (4.4), $Z(\mathbf{v}_{\mathcal{R}_0} - \mathbf{v}_{\mathcal{R}_1}) = \mathbf{0}$, $X_0\mathbf{v}_{\mathcal{R}_0} = \mathbf{0}$ and $X_1\mathbf{v}_{\mathcal{R}_1} = \mathbf{0}$. As $\mathbf{y}_{\mathcal{R}_j} \in \mathrm{Span}(\mathbf{v}_{\mathcal{R}_j})$ for $j \in \{0,1\}$, $Z(\mathbf{y}_{\mathcal{R}_1} - \mathbf{y}_{\mathcal{R}_0}) = \mathbf{0}$ and $X_0\mathbf{y}_{\mathcal{R}_0} = X_1\mathbf{y}_{\mathcal{R}_1} = \mathbf{0}$. As a result, for any $\mathbf{x}_{\mathsf{id}} \in \mathsf{RowSpan}(Z)$, $\mathbf{x}_{\mathsf{id}}^{\top}(\mathbf{y}_{\mathcal{R}_1} - \mathbf{y}_{\mathcal{R}_0}) = 0$. Thus, $\mathcal{A}_{\mathit{IPFE}}$ behaves as a valid adversary in the mIND-ID-CPA security model.

Firstly, note that, $\mathcal{A}_{\mathbf{TR}_0}$ gets encryption of either $\mathbf{y}_{\mathcal{R}_0}$ or $\mathbf{y}_{\mathcal{R}_1}$ where both the vectors are randomly sampled. Thus, from the point of view of $\mathcal{A}_{\mathbf{TR}_0}$, $Z\mathbf{y}_{\mathcal{R}_b}$ is a random vector. Then, we show that $\mathcal{A}_{\mathbf{TR}_0}$ sees $\mathbf{x}_{\mathsf{id}}$ purely random even though $\mathbf{x}_{\mathsf{id}}$ is sampled randomly from $\mathsf{RowSpan}(Z)$. This follows from the fact that $\mathcal{A}_{\mathbf{TR}_0}$ has access to all purely random vectors $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^{\ell}$ for $\mathsf{id} \in \hat{\mathcal{R}}$. Thus, from the point of view of $\mathcal{A}_{\mathbf{TR}_0}$, it has access to $|\hat{\mathcal{R}}|$ basis vectors of $\mathbb{Z}_p^{\ell}$ and the space $\mathbb{Z}_p^{\ell}$ is left with entropy of $(\ell - |\hat{\mathcal{R}}|)$ basis vectors where $\ell - |\hat{\mathcal{R}}| > t$. As $\mathcal{A}_{\mathbf{TR}_0}$ gets at most t samples of $\mathbf{x}_{\mathsf{id}} \leftarrow \mathsf{RowSpan}(Z)$, it sees $\mathbf{x}_{\mathsf{id}}$ identically distributed to vectors chosen uniformly random from $\mathbb{Z}_p^{\ell}$. The ciphertext and the secret keys are already properly distributed since $\mathcal{A}_{\mathit{IPFE}}$ has forwarded the reply of IPFE challenger. This shows that $\mathcal{A}_{\mathit{IPFE}}$ behaves as a valid challenger in the IND-ID*-CPA security model.

If $\mathcal{A}_{\mathbf{TR}_0}$ can distinguish between $\mathcal{R}_0$ and $\mathcal{R}_1$, $\mathcal{A}_{\mathit{IPFE}}$ can distinguish between $\mathbf{y}_{\mathcal{R}_0}$ and $\mathbf{y}_{\mathcal{R}_1}$. Thus, the advantage of $\mathcal{A}_{\mathit{IPFE}}$ is exactly the same as the advantage of $\mathcal{A}_{\mathbf{TR}_0}$. □

4.3.2 Efficient Trace-and-Revoke for Bit Strings

We present a trace-and-revoke scheme $\mathbf{TR}_1$ for $M = \{0,1\}^n$ that does not run parallel independent n executions of $\mathbf{TR}_0$. However, we note that, we omit the description of Trace here as it follows from the Trace algorithm of $\mathbf{TR}_0$. This scheme again assumes the existence of a user directory dir which is initialized to be empty, contains the identities of the users that have been assigned keys in the system. We assume that dir can only be modified by the central authority who is the sender as well as the key generator. Here, we assume existence of an efficient matrix multiplication functional encryption $\mathcal{MMFE}$ that encrypts matrices of $n \times \ell$ dimension. The intuitive idea here is that, we utilize n copies of inner product of ℓ dimensional vectors as a linear system of equations $M\mathbf{x}$ where $M \in \mathbb{Z}_p^{n \times \ell}$ and $\mathbf{x} \in \mathbb{Z}_p^{\ell}$. Each of the rows of M is used to encrypt each message bit.

1. $\mathsf{Setup}(1^{\lambda}, 1^n, 1^r, 1^{\mathsf{t}})$. Upon input the security parameter λ, the message bit-length n, the bound t on the number of the suspected users and the bound r on the number of revoked users, set $p = \lambda^{\omega(1)}$ and proceed as follows:
   (a) Let $(\mathsf{pp}, \mathsf{msk}) \leftarrow \mathcal{MMFE}.\mathsf{Setup}(1^{\lambda}, 1^{\ell}, 1^n, p)$, where we set $\ell = 2r + t + n + 1$.
   (b) Output the public parameter pp, master secret key msk and an empty user directory dir.

2. $\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathsf{id})$. Upon input the public parameters pp, the master secret key msk, the user directory dir and a user identity $\mathsf{id} \in \mathsf{ID}$, proceed as follows:

   (a) Sample $\mathbf{x}_{\mathsf{id}} \leftarrow \mathbb{Z}_p^\ell$. The pair $\mathsf{u}_{\mathsf{id}} = (\mathsf{id}, \mathbf{x}_{\mathsf{id}})$ is then appended to the user directory dir.

   (b) Let $(\mathbf{x}_{\mathsf{id}}, \mathsf{sk}_{\mathsf{id}}) \leftarrow \mathcal{MMFE}.\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x}_{\mathsf{id}}) \in \mathcal{MMFE}.\mathcal{K}$.

   (c) Output $\mathsf{sk}_{\mathsf{id}}$.

3. $\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, \mathsf{dir}, \mathcal{R}, m)$. Upon input the public parameter pp, the master secret key msk, the user directory dir, a set of revoked users $\mathcal{R}$ of size $\leqslant r$ and a plaintext messages $m \in M = \{0,1\}^n$, proceed as follows:

   (a) Sample $\mathbf{v}_{\mathcal{R},1}, \ldots, \mathbf{v}_{\mathcal{R},n} \leftarrow X_{\mathcal{R}}^\perp$ where $X_{\mathcal{R}} = \{\mathbf{x}_{\mathsf{id}} \in \mathbb{Z}_p^\ell \ : \ \mathsf{id} \in \mathcal{R}\}$.

   (b) Compute $\mathbf{y}_{\mathcal{R},i} = m_i \cdot \mathbf{v}_{\mathcal{R},i}$ for $i \in [1,n]$.

   (c) Define a matrix $M_{\mathcal{R}} = (\mathbf{y}_{\mathcal{R},1}, \ldots, \mathbf{y}_{\mathcal{R},n})^\top$.

   (d) Output $C_{\mathcal{R}} = \mathcal{MMFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M_{\mathcal{R}})$.

4. $\mathsf{Dec}(\mathsf{pp}, (\mathbf{x}_{\mathsf{id}}, \mathsf{sk}_{\mathsf{id}}), C_{\mathcal{R}})$. Upon input the public parameters pp, the secret key $\mathsf{sk}_{\mathsf{id}}$ for user id and a ciphertext $C_{\mathcal{R}}$ considering the revoked set $\mathcal{R}$, proceed as follows:

   (a) Compute $\mathbf{t} = \mathcal{MMFE}.\mathsf{Dec}(\mathsf{pp}, (\mathbf{x}_{\mathsf{id}}, \mathsf{sk}_{\mathsf{id}}), C_{\mathcal{R}})$.

   (b) Output $m' = (m'_1, \ldots, m'_n) \in \{0,1\}^n$ where for all $i \in [1,n]$, $m'_i = 0$ if $t_i = \mathit{elem}^*$; else $m'_i = 1$.

Correctness.

The correctness basically follows from the correctness of $\mathbf{TR}_0$ above. The main difference is that, functionally, Enc of $\mathbf{TR}_1$ is some-what n many copies of Enc of $\mathbf{TR}_0$. Thus, Dec must concatenate all the bits to get back the message. Therefore, $\mathbf{TR}_1$ is correct if Dec of $\mathbf{TR}_1$ retrieves all the bits m_i correctly. Now, if $\exists i \in [1,n]$, such that Dec of $\mathbf{TR}_1$ didn't compute m_i correctly, this can be extended to an attack on the correctness of Dec of $\mathbf{TR}_0$. This basically ensures the correctness of $\mathbf{TR}_1$.

Security

Now, we prove the general revocation scheme $\mathbf{TR}_1$ to be anonymous secure revocation scheme.

Theorem 4.15. If $\mathcal{MMFE}$ is an $\mathsf{IND\text{-}CPA}$ secure matrix multiplication functional encryption scheme, then $\mathbf{TR}_1$ is $\mathsf{IND\text{-}CPA}$ secure.

Proof Sketch. The proof is very similar to the proof of Theorem 4.11. However, the primary difference being the ciphertext generation on a challenge $(m_0, m_1, \mathcal{R})$. In particular, $\mathcal{A}_{\mathcal{MMFE}}$ finds solution of $X \cdot V = \mathbf{0}$ such that V is a full-rank matrix in $\mathbb{Z}_p^{\ell \times n}$. Precisely, $V = \begin{pmatrix} \mathbf{v}_{\mathcal{R},1} & \ldots & \mathbf{v}_{\mathcal{R},n} \end{pmatrix}$. Then $\mathcal{A}_{\mathcal{MMFE}}$ constructs the challenge as M_0 and M_1 where $M_b = (\mathbf{y}_{\mathcal{R},b,1}, \ldots, \mathbf{y}_{\mathcal{R},b,n})^\top$ such that $\mathbf{y}_{\mathcal{R},b,j} = m_{b,j} \cdot \mathbf{v}_{\mathcal{R},j}$. The rest follows naturally. □

Theorem 4.16. If $\mathcal{MMFE}$ is an IND-CPA secure matrix-multiplication functional encryption scheme allowing at most $(t + r - 1)$ key extraction queries, then $\mathbf{TR}_1$ is mIND-ID-CPA secure.

Proof Sketch. The proof is very similar to the proof of Theorem 4.12. The difference is again how we handle ciphertext generation. For, ciphertext query on $(m, \mathcal{R})$, $\mathcal{A}_{\mathcal{MMFE}}$ finds solution of $X \cdot V = \mathbf{0}$ such that V is a full-rank matrix in $\mathbb{Z}_p^{\ell \times n}$. Precisely, $V = \begin{pmatrix} \mathbf{v}_{\mathcal{R},1} & \dots & \mathbf{v}_{\mathcal{R},n} \end{pmatrix}$. Then we construct the ciphertext query to the $\mathcal{MMFE}$ challenger as M where $M = (\mathbf{y}_{\mathcal{R},1}, \dots, \mathbf{y}_{\mathcal{R},n})^\top$ such that $\mathbf{y}_{\mathcal{R},j} = m_j \cdot \mathbf{v}_{\mathcal{R},j}$. For the challenge query on $(m, \mathcal{R}_0, \mathcal{R}_1)$, $\mathcal{A}_{\mathcal{MMFE}}$ finds non-trivial solution of the following equations where both $V_{\mathcal{R}_0}, V_{\mathcal{R}_1}$ are full-rank matrices from $\mathbb{Z}_p^{\ell \times n}$.

$$\begin{pmatrix} Z & -Z \\ X_0 & \mathbf{0} \\ \mathbf{0} & X_1 \end{pmatrix} \cdot \begin{pmatrix} V_{\mathcal{R}_0} \\ V_{\mathcal{R}_1} \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \tag{4.5}$$

where $V_{\mathcal{R}_b} = \begin{pmatrix} \mathbf{v}_{\mathcal{R}_b,1} & \dots & \mathbf{v}_{\mathcal{R}_b,n} \end{pmatrix}$ for $b \in \{0, 1\}$. Then $\mathcal{A}_{\mathcal{MMFE}}$ constructs the challenge as M_0 and M_1 where $M_b = (\mathbf{y}_{\mathcal{R}_b,1}, \dots, \mathbf{y}_{\mathcal{R}_b,n})^\top$ such that $\mathbf{y}_{\mathcal{R}_b,j} = m_j \cdot \mathbf{v}_{\mathcal{R}_b,j}$. The rest of the argument follows naturally. □

Construction $\mathbf{TR}_0$ and $\mathbf{TR}_1$. Note that, available IPFE schemes [ALS16, ABP+17] suffice to construct of $\mathbf{TR}_0$ and $\mathbf{TR}_1$. In particular, withholding the public keys of available IPFE schemes, one can get symmetric-key IPFE schemes and use them to construct $\mathbf{TR}_0$. Furthermore, $\mathbf{TR}_1$ can be constructed from running n independent instances of any symmetric-key IPFE scheme. We in fact use this technique to construct $\mathbf{TR}_0$ and $\mathbf{TR}_1$ in the lattice-based settings withholding the public key of Agrawal *et al.*'s IPFE [ABP+17]. In the group-based settings, however, we can achieve more efficient constructions than naively hiding the public key of the public-key IPFE. In Section 4.5, we propose new constructions of symmetric-key IPFE and symmetric-key MMFE in the prime-order groups.

4.4 Cryptanalysis of the Wang *et al.* IPFE Construction

As we mention above, the schemes from Section 4.3 can be instantiated with the LWE-based $\mathcal{IPFE}$ scheme from [ABP+17]. Note that the latter does not enjoy IND-CPA security, but it was showed to enjoy a weaker security property that still suffices for the trace-and-revoke scheme from [ABP+17]. That weaker security property restricts the number of key requests to be significantly smaller than the dimension of the vector space, and imposes that the vectors of the key queries are uniformly sampled. This relaxation of IND-CPA security also suffices for our adaptation from Section 4.3..

$\mathcal{IPFE}$ scheme from [WFL19], note that the LWE-based $\mathcal{IPFE}$ scheme from [WFL19] is also claimed to enjoy a security property that is stronger than IND-CPA security (which the authors leverage to obtain a decentralized Attribute-Based Encryption scheme). In fact, as we will show below, this scheme can be broken for the parameters suggested in [WFL19]. Before showing an attack, we first recall the definition.

Definition 4.17. The bounded distance decoding problem BDD_γ is as follows: given a basis B of an n-rank lattice L, $\mathbf{t} \in \mathbb{R}^n$, and real $d \leqslant \frac{\lambda_1}{2}$ such that $\mathrm{dist}(\mathbf{t}, L) \leqslant d$, find the unique $\mathbf{v} \in L$ closest to $\mathbf{t}$. Note that this is equivalent to finding $\mathbf{e} \in \mathbf{t} + L$ such that $\|\mathbf{e}\| \leqslant d$.

We now describe here a simplified version of the security property that this scheme aims to achieve, and the corresponding simplified version of the scheme (this corresponds to setting $k = 1$ in the definition from [WFL19]; our attack readily extends to $k \geqslant 1$). In the challenge phase, the adversary sends to the challenger descriptions of two distributions D_0 and D_1 over plaintext vectors. The challenger chooses $\beta \leftarrow \{0, 1\}$ and samples $\mathbf{y} \leftarrow D_\beta$; it encrypts it under the public key pk and the resulting ciphertext $\mathsf{Enc}_{\mathsf{pk}}(\mathbf{y})$ is given to the adversary. The adversary can adaptively make key queries $\mathbf{x}$, before or after the challenge phase. The security property, called adaptive security for chosen message distributions, requires that the adversary cannot guess β correctly, as long as the distributions D_0 and D_1 remain indistinguishable given the replies to the key queries.

We review their construction based on LWE.

- $\mathit{IPFE}.\mathsf{Setup}(1^n, 1^\ell, p)$. Set integers $m, q = p^e$ for some integer e, and reals $\alpha, \alpha' \in (0, 1)$. Sample $A \leftarrow \mathbb{Z}_q^{m \times n}$, $Z \leftarrow \{0, \ldots, p-1\}^{\ell \times m}$,[5] compute $T = ZA \in \mathbb{Z}_q^{\ell \times n}$, define
$$\mathsf{msk} := Z \quad \text{and} \quad \mathsf{pk} := (A, T).$$
- $\mathit{IPFE}.\mathsf{KeyGen}(\mathsf{msk}, \mathbf{x})$. Given $\mathbf{x} \in \mathbb{Z}_p^\ell$, set $\mathbf{z}_\mathbf{x} = \mathbf{x}^\mathsf{t} Z \in \mathbb{Z}^m$ (interpreting each coordinate of $\mathbf{x}$ as an integer in $\{0, \ldots, p-1\}$), and output $\mathsf{sk}_\mathbf{x} = \mathbf{z}_\mathbf{x}$.
- $\mathit{IPFE}.\mathsf{Enc}(\mathsf{pk}, \mathbf{y})$. To encrypt a vector $\mathbf{y} \in \mathbb{Z}_p^\ell$, sample $\mathbf{s} \leftarrow \mathbb{Z}_q^n$, $\mathbf{e}_0 \leftarrow D_{\mathbb{Z}^m, \alpha q}$, $\mathbf{e}_1 \leftarrow D_{\mathbb{Z}^\ell, \alpha' q}$ and compute
$$\mathbf{c}_0 = A\mathbf{s} + \mathbf{e}_0 \in \mathbb{Z}_q^m, \qquad \mathbf{c}_1 = T\mathbf{s} + \mathbf{e}_1 + p^{e-1} \cdot \mathbf{y} \in \mathbb{Z}_q^\ell.$$
Then, return the ciphertext $C = (\mathbf{c}_0, \mathbf{c}_1)$.
- $\mathit{IPFE}.\mathsf{Dec}(\mathsf{sk}, C)$. Given $C = (\mathbf{c}_0, \mathbf{c}_1)$ and secret key $\mathsf{sk}_\mathbf{x} = \mathbf{z}_\mathbf{x}$, compute $\mu' = \langle \mathbf{x}, \mathbf{c}_1 \rangle - \langle \mathbf{z}_\mathbf{x}, \mathbf{c}_0 \rangle \bmod q$, and output the value $\mu \in \mathbb{Z}_p$ that minimize $|\mu' - p^{e-1}\mu|$.

In [WFL19], the dimensions n is proportional to the security parameter λ, the parameters $\ell, m, p, q, 1/\alpha, 1/\alpha'$ are polynomial in n, and e is a constant. In [WFL19, Theorem 3.5], the authors state that under the LWE assumption, the above functional encryption for inner products is adaptively secure for chosen message distributions, assuming that the secret key queries corresponding are linearly independent.
Below, we describe a cryptanalysis of the scheme above with the specified parameters. We then explain why this attack does not apply to the schemes from [ALS16] and [ABP+17].

[5] In [WFL19], the notation $\mathbb{Z}_p^{\ell \times m}$ is used instead of $\{0, \ldots, p-1\}^{\ell \times m}$. We stress that it should indeed be interpreted as $\{0, 1, \ldots, p-1\}^{\ell \times m}$. In particular, the operation $\mathbf{x}^\mathsf{t} Z$ in the $\mathit{IPFE}.\mathsf{KeyGen}$ algorithm is over $\mathbb{Z}$ and not modulo p, as otherwise decryption correctness would not hold.

We show that even for with challenge vectors rather than distributions, key queries allow to recover the master secret key msk. Concretely, we can recover Z from X^{t} and $X^{\mathsf{t}}Z$, where $Z \leftarrow \{0,\ldots,p-1\}^{\ell\times m}$ and $X \in \{0,\ldots,p-1\}^{\ell\times(\ell-1)}$ is chosen by the adversary. We let our adversary sample $X \leftarrow \{0,\ldots,p-1\}^{\ell\times(\ell-1)}$ (recall that the multiplication $X^{\mathsf{t}}Z$ is over $\mathbb{Z}$). The fact that X has only $\ell-1$ columns means that we can find distinct challenge plaintexts (which are elements of $\mathbb{Z}_p^\ell$) so that the columns of X are valid key queries.

It suffices to show how the adversary can recover the first column $\mathbf{z}$ of Z from $X^{\mathsf{t}}\mathbf{z}$, as it can proceed similarly for all columns of Z. Given $\mathbf{t} = X^{\mathsf{t}}\mathbf{z}$ and X, we know that $\mathbf{z}$ belongs to a coset of the lattice $\Lambda^\perp(X)$ defined by $\mathbf{t}$.

Let us now study the lattice $\Lambda^\perp(X)$. As $X \leftarrow \{0,\ldots,p-1\}^{\ell\times(\ell-1)}$, its columns are expected to be linearly independent with overwhelming probability and $\det(X\mathbb{Z}^{\ell-1})$ is expected to grow as $p^{\Omega(\ell)}$. These properties would be easier to prove if the entries of X were Gaussian with standard deviation p, but it can be experimentally checked that this behavior also holds for this distribution. We also expect the lattice $X\mathbb{Z}^{\ell-1}$ to be primitive, i.e., that $X^{\mathsf{t}}\mathbb{Z}^\ell = \mathbb{Z}^{\ell-1}$. By [Ngu99, p. 30], we hence have that $\det(\Lambda^\perp(X)) = \det(X\mathbb{Z}^{\ell-1})$. As X is full column-rank, we known that $\dim(\Lambda^\perp(X)) = 1$, and hence we expect that $\lambda_1(\Lambda^\perp(X)) = p^{\Omega(\ell)}$. Finally, note that the orthogonal lattice can be efficiently computed, by using a Hermite Normal Form algorithm.

Now, recall that we want to recover $\mathbf{z}$ from a known coset of $\Lambda^\perp(X)$. As $\|\mathbf{z}\| \leqslant \sqrt{\ell}p$, by the above analysis of $\Lambda^\perp(X)$, we expect to have

$$\|\mathbf{z}\| < \lambda_1(\Lambda^\perp(X))/2.$$

This implies that $\mathbf{z}$ is uniquely determined from the coset. Moreover, this is a Bounded Distance Decoding problem instance in a lattice of dimension 1, which can be solved efficiently. Concretely, if $\Lambda^\perp(X) = \mathbf{b}\mathbb{Z}$ and we are given $\mathbf{b}$ and $k\mathbf{b}+\mathbf{z}$, we can recover $k = \lfloor\langle k\mathbf{b}+\mathbf{z}, \mathbf{b}\rangle/\|\mathbf{b}\|^2\rceil$ and hence $\mathbf{z}$.

Remarks. Our proof shows that the scheme from [WFL19] is not secure with the specified parameters. We explain here why the above attack does not work for the [ALS16] and [ABP+17] schemes. First, in the mod-p scheme from [ALS16, Section 4.1], the authors take $\mathbf{z}$ from a discrete Gaussian distribution with a large standard deviation. With the parameters specified in [ALS16], we then have that $\|\mathbf{z}\|$ is significantly larger than $\lambda_1(\Lambda^\perp(X))$. This implies that there is a large amount of entropy left in $\mathbf{z}$ given $\mathbf{t} = X^{\mathsf{t}}\mathbf{z}$. Also, this attack does not work for the [ALS16] scheme over $\mathbb{Z}$, because in that case, the matrix X and hence the lattice $\Lambda^\perp(X)$ are not random at all. Indeed, the kernel lattice is forced to be $(\mathbf{y}_0 - \mathbf{y}_1)\mathbb{Z}^\ell$, where $\mathbf{y}_0$ and $\mathbf{y}_1$ are the challenge vectors. By assumption on the scheme, these challenge vectors are small. Put differently, in that setting, if we first do $(\ell-1)$ random queries, there does not exist $\mathbf{y}_0 - \mathbf{y}_1 \neq \mathbf{0}$ short anymore that allows us to create a non-trivial challenge phase. Finally, the attack does not work for the [ABP+17] scheme variant, because in that case, the matrix X has much fewer columns than rows. This increases the dimension of $\Lambda^\perp(X)$ enough to make $\lambda_1(\Lambda^\perp(X))$ much smaller, and in particular smaller than $\|\mathbf{z}\|$.

4.5 Linear Functional Encryptions in Prime-Order Groups

As outlined in 4.3, our trace-and-revoke schemes are instantiated using different linear functional encryption schemes. In this section, we give a construction of $\mathcal{MMFE}$ in the symmetric-key setting. For $n = 1$, the $\mathcal{MMFE}$ construction reduces to $\mathcal{IPFE}$. Due to space restraint, we omit the description of $\mathcal{IPFE}$ and present the $\mathcal{MMFE}$ below. The point of interest being, the Dec in our $\mathcal{MMFE}$ (and in our $\mathcal{IPFE}$) does not compute the discrete log.

We propose a construction of matrix multiplication functional encryption ($\mathcal{MMFE}$) from $\mathcal{D}_k$-matDH. Since, the complete matrix $M = (\mathbf{y}_1, \ldots, \mathbf{y}_n)^\top$ is available to Enc at once, our construction can reuse the randomness for all $\mathbf{y}_i \in \mathbb{Z}_p^\ell$. This also allows the proof to be tightly reduced to $\mathcal{D}_k$-matDH. For this, we require n matrices $W_1, \ldots, W_n$ unlike $\mathcal{IPFE}$ from $\mathcal{D}_k$-matDH that required only one. We emphasize that, similar to $\mathcal{IPFE}$ above, $\mathcal{MMFE}$ also does not need to evaluate discrete logarithm algorithm.

- $\mathsf{Setup}(1^\lambda, 1^\ell, 1^n, p)$. Run $(g, \mathbb{G}) \leftarrow \mathcal{G}_{gen}(1^\lambda, p)$. Sample $A \leftarrow \mathcal{D}_k$ and $W_1, \ldots, W_n \leftarrow \mathbb{Z}_p^{\ell \times k\ell n}$. Define $\mathsf{msk} = (W_1, \ldots, W_n)$ and $\mathsf{pp} = ([1])$.
- $\mathsf{KeyGen}(\mathsf{pp}, \mathsf{msk}, \mathbf{x} \in \mathbb{Z}_p^\ell)$. Set $\mathsf{sk}_\mathbf{x} \leftarrow (\mathbf{x}^\top W_1, \ldots, \mathbf{x}^\top W_n, \mathbf{x})$.
- $\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M = (\mathbf{y}_1, \ldots, \mathbf{y}_n)^\top \in \mathbb{Z}_p^{n \times \ell})$ proceeds as follows to encrypt the given vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{Z}_p^\ell$. Sample $\mathbf{s} \leftarrow \mathbb{Z}_p^{k\ell n}$. Set $\mathsf{ct}_M \leftarrow ([\mathbf{s}], [\mathbf{y}_1 + W_1\mathbf{s}], \ldots, [\mathbf{y}_n + W_n\mathbf{s}])$.
- $\mathsf{Dec}(\mathsf{pp}, \mathsf{sk}_\mathbf{x}, \mathsf{ct}_M)$. Parse $\mathsf{ct}_M = ([\mathbf{c}_0], [\mathbf{c}_1], \ldots, [\mathbf{c}_n])$. Return $\mathbf{t} = (t_1, \ldots, t_n)$ where $t_i = \left[\mathbf{x}^\top \mathbf{c}_i\right] \cdot [\mathsf{sk}_\mathbf{x} \cdot \mathbf{c}_0]^{-1}$.

The correctness is easy to verify.

We show a rough comparison of our scheme with [Tom19] if their scheme was used for symmetric key settings directly. 4.1 shows that the symmetric key variant resulted from hiding the public key of [Tom19] has bigger public parameters and bigger ciphertext i.e. contain more group elements than our scheme. On the other hand, our secret key contains more elements from $\mathbb{Z}_p$. Both the schemes are proven secure under same assumption $\mathcal{D}_k$-matDH with constant degradation. We further compare the result for the SXDH based instances which shows that their scheme outputs ciphertext that is 1.5 times bigger than us.

	$\lvert\mathsf{pp}\rvert$	$\lvert\mathsf{sk}\rvert$	$\lvert\mathsf{ct}\rvert$	Degradation	Assumption
[Tom19]	$(k^3(k+1)\ell^2 + k^2\ell^2)\mathbb{G}$	$(k+1)k\ell\ \mathbb{Z}_p$	$n((k+1)k\ell + \ell)\mathbb{G}$	4	$\mathcal{D}_k$-matDH
	$(2\ell^2 + \ell^2)\mathbb{G}$	$2\ell\ \mathbb{Z}_p$	$3n\ell\ \mathbb{G}$	4	SXDH
[BMN+21]	$1\mathbb{G}$	$k\ell n^2\ \mathbb{Z}_p$	$k\ell n + \ell n\mathbb{G}$	$(k+1)$	$\mathcal{D}_k$-matDH
	$1\mathbb{G}$	$n^2\ell\ \mathbb{Z}_p$	$2n\ell\ \mathbb{G}$	2	SXDH

Table 4.1: Comparison of naive application of [Tom19] with our construction in symmetric-key settings.

Security. Next, we argue the security of $\mathcal{MMFE}$ in the IND-CPA security model. Our construction is basically a modification of [Tom19] for symmetric-key settings. This improves upon the performance in terms of ciphertext size and removes the usage of public parameters completely. Note that, this modification required us to argue the security proof in a different manner. Although the overall proof strategy stayed more-or-less the same, our proof presents a completely new proof for an essential lemma. We state the security theorem next.

Theorem 4.18. For any adversary $\mathcal{A}$ of the construction $\mathcal{MMFE}$ in the IND-CPA security model that makes at most q_{sk} secret key queries (for $q_{\mathsf{sk}} < \ell$) and q_{ct} challenge ciphertext queries in an interleaved manner, there exists adversary AC such that,

$$\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathcal{MMFE},\mathcal{A}}(\lambda) \leqslant (k+1) \cdot \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}}_{AC}(\lambda).$$

Proof. The proof is done by defining a hybrid argument of a sequence of games that begins with the real protocol (called $\mathbf{Game}_0$) and ends with a so-called final game (called $\mathbf{Game}_3$) where the adversary has no advantage at all. During the sequence, we use X_i to denote the event that the adversary has won $\mathbf{Game}_i$.

- $\mathbf{Game}_0$. This is the real game. All secret key queries on $\mathbf{x} \in \mathbb{Z}_p^\ell$ are responded as the real game. For all j^{th} (such that $j \in [1, q_{\mathsf{ct}}]$) ciphertext query on two matrices $M_j^{(0)}, M_j^{(1)} \in \mathbb{Z}_p^{n\times\ell}$, for $\beta \leftarrow \{0,1\}$ the challenge ciphertext returned is $\mathsf{ct}^{(\beta)} \leftarrow \mathcal{MMFE}.\mathsf{Enc}(\mathsf{pp}, \mathsf{msk}, M_j^{(\beta)})$. More precisely, the j^{th} ciphertext query is responded as,

$$\mathbf{s}_j \leftarrow \mathbb{Z}_p^{k\ell n}, \quad [\mathbf{c}_{j,0}] = [\mathbf{s}_j], \quad [\mathbf{c}_{j,i}] = \left[\mathbf{y}_{j,i}^{(\beta)} + W_i \mathbf{s}_j\right]$$

  for $j \in [1, q_{\mathsf{ct}}]$. At the end, $\mathcal{A}$ outputs $\beta' \in \{0,1\}$ and wins the game if $\beta = \beta'$.

- $\mathbf{Game}_1$. The response of the challenge queries are defined as following. For j^{th} ciphertext query is made on $(M_j^{(0)}, M_j^{(1)})$ where $M_j^{(b)} = (\mathbf{y}_{j,1}^{(b)}, \ldots, \mathbf{y}_{j,n}^{(b)})^\top$ for $b \in \{0,1\}$,

$$\mathbf{s}_j \leftarrow \mathbb{Z}_p^{k\ell n}, \quad [\mathbf{c}_{j,0}] = [\mathbf{s}_j], \quad [\mathbf{c}_{j,i}] = \left[\mathbf{y}_{j,i}^{(\beta)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} \mathbf{u}^\top \mathbf{v}_{j,i,\iota} \cdot \mathbf{z}_{\psi_i(\iota),i}\right]$$

  where $\mathbf{u} \leftarrow \mathbb{Z}_p^k$, $\mathbf{z}_{j,i} = \mathbf{y}_{j,i}^{(1)} - \mathbf{y}_{j,i}^{(0)}$, $\phi_i(j) = \mathsf{Rank}(\mathbf{z}_{1,i} || \ldots || \mathbf{z}_{j,i})$, $\psi_i(j) = \mathsf{min}(\phi_i^{-1}(j))$, and $\mathbf{v}_{j,i,1}, \ldots, \mathbf{v}_{j,i,\ell} \leftarrow \mathbb{Z}_p^k$ where $i \in [1,n]$ and $j \in [1, q_{\mathsf{ct}}]$. In Lemma 4.19, we show that $|\Pr[X_1] - \Pr[X_0]| \leqslant \mathrm{Adv}^{\mathscr{D}_{2k,k}\text{-}\mathsf{matDH}}(\lambda)$.

- $\mathbf{Game}_2$. The response of the challenge queries are defined as following. For j^{th} ciphertext query is made on $(M_j^{(0)}, M_j^{(1)})$ where $M_j^{(b)} = (\mathbf{y}_{j,1}^{(b)}, \ldots, \mathbf{y}_{j,n}^{(b)})^\top$ for $b \in \{0,1\}$,

$$\mathbf{s}_j \leftarrow \mathbb{Z}_p^{k\ell n}, \quad [\mathbf{c}_{j,0}] = [\mathbf{s}_j], \quad [\mathbf{c}_{j,i}] = \left[\mathbf{y}_{j,i}^{(\beta)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \cdot \mathbf{z}_{\psi_i(\iota),i}\right]$$

  where $\mathbf{z}_{j,i} = \mathbf{y}_{j,i}^{(1)} - \mathbf{y}_{j,i}^{(0)}$, $\phi_i(j) = \mathsf{Rank}(\mathbf{z}_{1,i} || \ldots || \mathbf{z}_{j,i})$, $\psi_i(j) = \mathsf{min}(\phi_i^{-1}(j))$, and $v_{j,i,1}, \ldots, v_{j,i,\ell} \leftarrow \mathbb{Z}_p$ where $i \in [1,n]$ and $j \in [1, q_{\mathsf{ct}}]$. In Lemma 4.20, we show that $|\Pr[X_2] - \Pr[X_1]| \leqslant \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}'}(\lambda)$.

- **Game**$_3$. Finally, we show that, the injected entropy is sufficient to hide β in the returned ciphertexts completely. This is because, for any $j \in [1, q_{\mathsf{ct}}]$ and $i \in [1, n]$, $\sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \mathbf{z}_{\psi_i(\iota),i}$ is basically a random vector in the span of $\{\mathbf{z}_{\psi_i(\iota),i}\}_{\iota \in [1,\phi_i(j)]}$. Furthermore, by the definition of ϕ and ψ, $\{\mathbf{z}_{\psi_i(\iota),i}\}_{\iota \in [1,\phi_i(j)]}$ are the basis and therefore each $\mathbf{z}_{j,i} \in \mathrm{Span}(\{\mathbf{z}_{\psi_i(\iota),i}\}_{\iota \in [1,\phi_i(j)]})$. Then,

$$\begin{aligned}\mathbf{y}_{j,i}^{(\beta)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \mathbf{z}_{\psi_i(\iota),i} &= \beta \mathbf{z}_{j,i} + \mathbf{y}_{j,i}^{(0)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \mathbf{z}_{\psi_i(\iota),i} \\ &\equiv \mathbf{y}_{j,i}^{(0)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \mathbf{z}_{\psi_i(\iota),i}\end{aligned}$$

  As the ciphertext distribution stays the same as in **Game**$_2$, $\Pr[X_3] = \Pr[X_2]$.

  Furthermore, $[\mathbf{c}_{j,i}] = \left[\mathbf{y}_{j,i}^{(\beta)} + W_i \mathbf{s}_j + \sum_{\iota \in [1,\phi_i(j)]} v_{j,i,\iota} \mathbf{z}_{\psi_i(\iota),i}\right]$ hides β completely for all $j \in [1, q_{\mathsf{ct}}]$ and $i \in [1, n]$. Thus $\Pr[X_3] = 1/2$.

  To summarise,

$$\begin{aligned}\mathrm{Adv}^{\mathsf{IND\text{-}CPA}}_{\mathcal{MMFE},\mathcal{A}}(\lambda) &\leqslant |1/2 - \Pr[X_0]| \\ &\leqslant |\Pr[X_3] - \Pr[X_2]| + |\Pr[X_2] - \Pr[X_1]| + |\Pr[X_1] - \Pr[X_0]| \\ &\leqslant 0 + \mathrm{Adv}^{\mathscr{D}_{2k,k}\text{-}\mathsf{matDH}}(\lambda) + \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}'}(\lambda) \\ &\leqslant k \cdot \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}}(\lambda) + \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}}(\lambda) \leqslant (k+1) \cdot \mathrm{Adv}^{\mathscr{D}_k\text{-}\mathsf{matDH}}(\lambda)\end{aligned}$$

Lemma 4.19. For any efficient adversary $\mathcal{A}$ that makes at most q_{sk} secret key queries and at most q_{ct} ciphertext queries, there exists a algorithm AB such that $|\Pr[X_1] - \Pr[X_0]| \leqslant \mathrm{Adv}^{\mathscr{D}_{2k,k}\text{-}\mathsf{matDH}}_{AB}(\lambda)$.

Proof. To simulate the game, we use a $\mathscr{D}_{2k,k}$-matDH (as described in Definition 4.6) problem instance $([A], [\mathbf{t}])$ where $\mathbf{t} = \begin{pmatrix} \mathcal{A} u \mathbf{w} \\ \mathcal{A} d \mathbf{w} + \boldsymbol{\delta} \end{pmatrix}$ for $\mathbf{w} \in \mathbb{Z}_p^k$ where $\boldsymbol{\delta} = \mathbf{0}$ or chosen uniformly random vector from $\mathbb{Z}_p^k$. In fact, we use random self-reducibility property to define $q_{\mathsf{ct}} n \ell$ many problem instances $([A], [\mathbf{t}_{j,i,\iota}])$ for $j \in [1, q_{\mathsf{ct}}]$, $i \in [1, n]$ and $\iota \in [1, \ell]$. We use such problem instances to sample the $W_1, \ldots, W_n$. First, we set

$$\mathbf{s}_j = (\overline{\mathbf{t}}_{j,1,1}, \ldots, \overline{\mathbf{t}}_{j,1,\ell}, \ldots, \overline{\mathbf{t}}_{j,n,1}, \ldots, \overline{\mathbf{t}}_{j,n,\ell})^\top.$$

For all $i \in [1, n]$, we then sample

$$W_i = \widetilde{W}_i + \sum_{\iota \in [1,\phi_i(q)]} \mathbf{z}_{\psi_i(\iota),i} \mathbf{u}^\top T \left[\mathbf{0}_{k \times k(\ell(i-1)+(\iota-1))} || I_k || \mathbf{0}_{k \times k((\ell-\iota)+\ell(n-i))}\right]$$

where $\widetilde{W}_i \leftarrow \mathbb{Z}_p^{\ell \times k\ell n}$, $\mathbf{u} \leftarrow \mathbb{Z}_p^k$, $T = \mathcal{A} d \cdot (\mathcal{A} u)^{-1}$ and $\mathbf{z}_{j,i} = \mathbf{y}_{j,i}^{(1)} - \mathbf{y}_{j,i}^{(0)}$ where $i \in [1, n]$ and $j \in [1, q_{\mathsf{ct}}]$ for j^{th} ciphertext query is made on $(M_j^{(0)}, M_j^{(1)})$ where $M_j^b = (\mathbf{y}_{j,1}^b, \ldots, \mathbf{y}_{j,n}^b)^\top$ for $b \in \{0, 1\}$.

For all $j \in [1, q_{\mathsf{sk}}]$, the j^{th} secret key query on $\mathbf{x}_j$, we respond $\mathsf{sk} = (\mathbf{x}_j^\top \widetilde{W}_1, \ldots, \mathbf{x}_j^\top \widetilde{W}_i)$. Given j^{th} ciphertext query on $(M_j^{(0)}, M_j^{(1)})$ for $j \in [1, q_{\mathsf{ct}}]$, we respond $([\mathbf{c}_{j,0}], \ldots, [\mathbf{c}_{j,n}])$

where $\mathbf{c}_{j,0} = \mathbf{s}_j$ and for all $i \in [1,n]$, $\mathbf{c}_{j,i} = \mathbf{y}^b_{j,i} + \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top \underline{\mathbf{t}}_{j,i,\iota}$ where $\phi_i(j) = \mathsf{Rank}(\mathbf{z}_{1,i}||\ldots||\mathbf{z}_{j,i})$ and $\psi_i(j) = \mathsf{min}(\phi_i^{-1}(j))$. Firstly, observe that, the ciphertext generation uses $\mathbf{z}_{1,1},\ldots,\mathbf{z}_{1,n},\ldots,\mathbf{z}_{j,1},\ldots,\mathbf{z}_{j,n}$ where $\mathbf{z}_{\iota,i} = \mathbf{y}^{(1)}_{\iota,i} - \mathbf{y}^{(0)}_{\iota,i}$ where $i \in [1,n]$ and $\iota \in [1,j]$ i.e. each j^{th} ciphertext is defined using already queried matrices $(M_1^{(0)}, M_1^{(1)}), \ldots, (M_j^{(0)}, M_j^{(1)})$. Moreover, $\mathbf{x}^\top W_i = \mathbf{x}^\top \widetilde{W}_i$ for all $i \in [1,n]$ as $\mathbf{x}^\top \mathbf{z}_{j,i} = 0$ for all $j \in [1,q]$ and therefore the secret keys are simulated properly. We now show that the ciphertexts are also simulated properly.

$$\begin{aligned}
W_i\mathbf{s}_j &= \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(q)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top T \left[\mathbf{0}_{k\times k(\ell(i-1)+(\iota-1))}||I_k||\mathbf{0}_{k\times k((\ell-\iota)+\ell(n-i))}\right] \\
&\qquad\qquad \cdot \left[\overline{\mathbf{t}}_{j,1,1} \quad \cdots \quad \overline{\mathbf{t}}_{j,1,\ell} \quad \cdots \quad \overline{\mathbf{t}}_{j,n,1} \quad \cdots \quad \overline{\mathbf{t}}_{j,n,\ell}\right]^\top \\
&= \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top T\overline{\mathbf{t}}_{j,i,\iota} \\
&= \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top \cdot (\mathcal{A}d(\mathcal{A}u)^{-1}) \cdot (\mathcal{A}u\mathbf{w}_{j,i,\iota}) \\
&= \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top \cdot (\mathcal{A}d\mathbf{w}_{j,i,\iota}) \\
&Approx \widetilde{W}_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}\mathbf{u}^\top \underline{\mathbf{t}}_{j,i,\iota}
\end{aligned}$$

Now, it is clear that if $\mathbf{t} \in \mathrm{Span}(A)$, the simulation is identical to $\mathbf{Game}_0$ as $\underline{\mathbf{t}}_{j,i,\iota} = \mathcal{A}d\mathbf{w}_{j,i,\iota}$ for $j \in [1,q_{\mathsf{ct}}]$, $i \in [1,n]$ and $\iota \in [1,\ell]$. Otherwise, the simulation is identical to $\mathbf{Game}_1$. □

Lemma 4.20. For any efficient adversary $\mathcal{A}$ that makes at most q_{sk} secret key queries and at most q_{ct} ciphertext queries, there exists a ppt algorithm $\mathcal{B}$ such that $|\Pr[X_2] - \Pr[X_1]| \leqslant \mathrm{Adv}^{\mathcal{D}_k\text{-}\mathsf{matDH}'}_{AB}(\lambda)$.

Proof. Here, $\mathcal{B}$ gets an $\mathcal{D}_k$-matDH' problem instance $([T], \left[\mathbf{v}^{(\delta)}\right])$ for $\delta \leftarrow \{0,1\}$ where $T \in \mathbb{Z}_p^{k\times m}$, $\mathbf{v}^{(0)} = A^\top T$ and $\mathbf{v}^{(1)} \leftarrow \mathbb{Z}_p^{1\times m}$ (as described in Definition 4.7) where $A \leftarrow \mathbb{Z}_p^k$. Note that here we set $m = q_{\mathsf{ct}}n\ell$ and implicitly set $\mathbf{u}$ as A and set $T = \left[\underline{\mathbf{t}}_{1,1,1} \quad \cdots \quad \underline{\mathbf{t}}_{q_{\mathsf{ct}},n,\ell}\right]$.

Given the problem instance, $\mathcal{B}$ chooses $W_1,\ldots,W_n \leftarrow \mathbb{Z}_p^{\ell\times k\ell n}$ to define msk. Since, $\mathcal{B}$ knows msk completely, it can respond to the secret key queries on its own. On j^{th} ciphertext query $(M_j^{(0)}, M_j^{(1)})$, $\mathcal{B}$ samples $\mathbf{s}_j \leftarrow \mathbb{Z}_p^{k\ell n}$ and defines the ciphertext as following:

$$\mathbf{c}_{j,0} = [\mathbf{s}_j], \qquad \mathbf{c}_{j,i} = \left[W_i\mathbf{s}_j + \sum_{\iota\in[1,\phi_i(j)]} \mathbf{z}_{\psi_i(\iota),i}v^{(\delta)}_{j,i,\iota} + \mathbf{y}^{(\beta)}_{j,i}\right]$$

for all $i \in [1,n]$ where ϕ and ψ are defined as in the previous game. It is clear that the simulation is distributionally consistent. Precisely, if $\left[\mathbf{v}^{(0)}\right]$ is provided, the simulation is identical to $\mathbf{Game}_1$. Otherwise, the simulation is identical to $\mathbf{Game}_2$. If $\mathcal{A}$ distinguishes between $\mathbf{Game}_1$ and $\mathbf{Game}_2$, $\mathcal{B}$ can distinguish between $\left[\mathbf{v}^{(0)}\right]$ and $\left[\mathbf{v}^{(1)}\right]$. □

□

Chapter 5

A new integer-LWE problem

This chapter is based on a joint work with E. Kirshanova, D. Stehlé. This paper is in preparation.

5.1 Introduction

The integer Learning with Errors. The integer learning with errors (ILWE), introduced by Bootle *et al* [BDE+18], has computations without modular reduction. They consider the problem of finding a vector $\mathbf{s} \in \mathbb{Z}^n$ given polynomially many samples of the form $(\mathbf{a}, \langle \mathbf{a}, \mathbf{s} \rangle + e) \in (\mathbb{Z}^n \times \mathbb{Z})$. Then the problem becomes easy to solve, under mild conditions on the distributions of $\mathbf{a}$, $\mathbf{s}$ and e.

Our integer Learning with Errors. We introduce a new variant of the LWE problem over the integers, without any modulus q, denoted by integer $\mathrm{LWE}_{m,\sigma_X,\sigma_k,\alpha}$. More precisely, it consists in recovering $\mathbf{k}$ from $\Big(X, \pi_{\ker(X)}\big(\mathbf{k}+\mathbf{e}\big)\Big)$, where $X \leftarrow D_{\mathbb{Z}^{m\times n},\sigma_X}$, $\mathbf{k} \leftarrow D_{\mathbb{Z}^m,\sigma_k}$ and the error $\mathbf{e}$ is taken from continuous Gaussian with standard deviation α.

Our contribution. In this chapter, we define the new integer-LWE. More precisely, we consider two problems: integer-SIS, and search integer-LWE. Then we obtain the following results.

- The lattices that underlie our new definitions of SIS and LWE stem from a matrix $X \in \mathbb{Z}^{m\times n}$ whose entries are Gaussian over the integers: the integer-SIS problem translates to the one of finding short vectors in the orthogonal $(m-n)$ dimensional lattice $\Lambda^{\perp}(X) \subset \mathbb{Z}^m$. In Section 5.3, we exhibit the reduction from hard lattice problem (SIS modulo q) to integer-SIS .
- We then introduce the integer-LWE definition in Section 5.3, and propose a quantum polynomial-time reduction from integer-SIS to the search integer-LWE in Section 5.4.

5.2 Preliminaries

For an ordered set $B = \{\mathbf{b}_i\}_{i\in[n]}$ of linearly independent vectors in $\mathbb{R}^n$, we let $\widetilde{B}$ denote its Gram-Schmidt orthogonalization. The norm of a matrix is the Euclidean norm of its longest column: $\|X\| = \max_i \|\mathbf{x}_i\|$.

For a vector subspace $V \subseteq \mathbb{R}^m$ and a vector $\mathbf{x} \in \mathbb{R}^m$, we let $\pi(\mathbf{x}, V)$ denote the projection of $\mathbf{x}$ onto V. The kernel of a matrix $X \in \mathbb{R}^{m \times n}$ is denoted $\ker(X)$.

The statistical distance between two distributions D, D' over a common support X is defined as $\Delta(D, D') = \frac{1}{2} \sum_{x \in X} |D(x) - D'(x)|$.

The following lemma states that any full-rank set of vectors in a lattice can be efficiently converted to a basis of the lattice without increasing the norms of the Gram-Schmidt vectors. It is folklore and we borrow the statement from [MG02].

Lemma 5.1 ([MG02, Lemma 7.1])**.** There is a deterministic polynomial-time algorithm that, given an arbitrary basis B of an n-dimensional lattice L and a full-rank set of lattice vectors $S \in L$, outputs a basis T of L such that $\|\widetilde{T}\| \leqslant \|\widetilde{S}\|$ and $\|T\| \leqslant \sqrt{n} \cdot \|S\|$.

We recall the definition of the Bounded Distance Decoding problem.

Definition 5.2. The Bounded Distance Decoding BDD_d problem is as follows: given a lattice L, a bound d, and a coset $\mathbf{e} + L$ where $\|\mathbf{e}\| \leqslant d$, output $\mathbf{e}$.

In [GPV08], Gentry et al. showed how to use an arbitrary basis B to sample efficiently from the discrete Gaussian distribution $D_{L,s,\mathbf{c}}$, for any s sufficiently greater than $\|\widetilde{B}\|$.

Lemma 5.3 (Adapted from [BLP+13, Lemma 2.3])**.** There is a probabilistic polynomial-time algorithm that, given a basis B of a rank-n lattice $L = L(B), \mathbf{c} \in \mathrm{Span}(L)$ and an parameter $s > 0$ satisfying $\sqrt{\ln(2n+4)/\pi} \cdot \|\widetilde{B}\| \leqslant s$, outputs a sample from $D_{L,s,\mathbf{c}}$.

5.2.1 The smoothing parameter

The following lemma states that the discrete Gaussian function is essentially invariant under shifts, if the standard deviation is sufficiently large.

Lemma 5.4 ([MR04, Lemma 4.4])**.** For any lattice $L \subseteq \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^m, \varepsilon \in (0,1)$, and $s \geqslant \eta_\varepsilon(L)$, we have

$$\rho_{s,\mathbf{c}}(L) \in \left[\frac{1-\varepsilon}{1+\varepsilon}, 1\right] \cdot \rho_s(L).$$

The following result implies that a lattice Gaussian distribution has high min-entropy.

Lemma 5.5 ([PR06, Lemma 2.11])**.** For any rank-n lattice L, $\mathbf{x} \in L$, $\varepsilon \in (0, 1/3)$, and $s \geqslant \eta_\varepsilon(L)$, we have

$$D_{L,s}(\mathbf{x}) \leqslant \frac{1-\varepsilon}{1+\varepsilon} \cdot 2^{-n}$$

The next lemma is borrowed from [DRN14, Lemma 2.13] and [MR04, Lemma 4.4].

Lemma 5.6. For any lattice $L \subseteq \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^m, \varepsilon \in (0, 1/3)$, and $s \geqslant \eta_\varepsilon(L)$, we have

$$\Pr_{\mathbf{v} \leftarrow D_{L+\mathbf{c},s}}\left[\|\mathbf{v}\| > s \cdot t\sqrt{\frac{m}{2\pi}}\right] \leqslant \exp\left(-\frac{m}{2}(t-1)^2\right).$$

5.2.2 The lattice $\Lambda^{\perp}(X)$ and its dual

We recall some lemmas which will be used in this chapter. All of them are from Chapter 3. The following lemma makes explicit the dual of the $\Lambda^{\perp}(X)$ lattice. The result is well-known and can be found, for example, in [KNSW20].

Lemma 5.7. Let $X \in \mathbb{Z}^{n\times m}$ and $\Lambda^{\perp}(X)^{\star}$ be the dual lattice of $\Lambda^{\perp}(X)$. We have

$$\Lambda^{\perp}(X)^{\star} = \Big(\mathbb{Z}^m + X^{\mathsf{t}}\mathbb{R}^n\Big) \cap \ker(X) = \pi\Big(\mathbb{Z}^m, \ker(X)\Big).$$

Let $\mathbf{b}$ be a vector in $\Lambda^{\perp}(X)^{\star}$. From the lemma above, we can write $\mathbf{b} = \mathbf{v} + X\mathbf{r}$ for some $\mathbf{v} \in \mathbb{Z}^m$ and $\mathbf{r} \in \mathbb{R}^n$. As X is integral, we can always assume that $\|\mathbf{r}\|_{\infty} \leqslant 1/2$ (by rounding its coefficients to the nearest integers and updating $\mathbf{v}$ accordingly).

The following result gives a probabilistic lower bound for the first minimum of the dual of $\Lambda^{\perp}(X)$.

Lemma 5.8 ([KNSW20, Lemma 14])**.** Let $n \geqslant 60$, $\sigma_X \geqslant 20\sqrt{n}$ and $m \geqslant 1355n\ln\sigma_X$. Then we have

$$\Pr_{X\leftarrow\left(D_{\mathbb{Z}^{m\times n},\sigma_X}\right)}\left[\lambda_1^{\infty}(\Lambda^{\perp}(X)^{\star}) \geqslant \frac{1}{96\sqrt{n+\ln m}}\right] \geqslant 1-2^{-\Omega(n)},$$

$$\Pr_{X\leftarrow\left(D_{\mathbb{Z}^{m\times n},\sigma_X}\right)}\left[\lambda_1(\Lambda^{\perp}(X)^{\star}) \geqslant \frac{1}{96}\right] \geqslant 1-2^{-\Omega(n)}.$$

The following lemma is central to proving the previous results, and we will also use it directly. It gives an upper bound and a lower bound on the singular values of a random Gaussian matrix X.

Lemma 5.9 (Adapted from [AGHS13, Lemma 8])**.** There exists a universal constant $C > 1$ such that for any $m \geqslant 2n, \varepsilon > 0$ and $\sigma_X \geqslant 2C\eta_{\varepsilon}(\mathbb{Z}^m)$, we have

$$\Pr_{X\leftarrow D_{\mathbb{Z}^{m\times n},\sigma_X}}\left[\frac{\sigma_X\sqrt{2\pi m}}{C} \leqslant \sigma_n(X) \leqslant \sigma_1(X) \leqslant \sigma_X C\sqrt{2\pi m}\right] \geqslant 1-\Big(4m\varepsilon + 2^{-\Omega(n)}\Big).$$

5.2.3 Rényi divergence

Let P and Q be two discrete distributions such that $Supp(P) \subseteq Supp(Q)$ over a countable domain. The Rényi divergence of P and Q (of order 2) is defined by $R(P\|Q) = \Sigma_{x\in M}P(x)^2/Q(x)$. The following lemma gives an upper bound on the Rényi divergence for two discrete Gaussians that differ by a shift.

Lemma 5.10 ([LSS14, Lemma 4.2])**.** For any n-rank lattice $L \subseteq \mathbb{R}^n$, set $P = D_{L,s,\mathbf{u}}$ and $Q = D_{L,s,\mathbf{v}}$ for some $\mathbf{u}, \mathbf{v} \in L$ and $s > 0$. Then we have

$$R(P\|Q) \leqslant \exp(2\pi\|\mathbf{u}-\mathbf{v}\|^2/s^2).$$

In our reductions the following lemma will be applied to bound from below the success probability of an algorithm. The reader can find a proof in, e.g., [LSS14, Lemma 4.1].

Lemma 5.11. Let P, Q be two distributions with $Supp(P) \subseteq Supp(Q)$ and $E \subseteq Supp(Q)$ be an arbitrary event. Then we have $P(E) \geqslant Q(E)^2/R(P\|Q)$.

.

5.3 Integer-SIS and Integer-LWE

In this section, we introduce and study the hardness of a new version of the Short Integer Solution (SIS) problem. We call this version integer-SIS as it does not require any modulus q. We also define the integer-LWE distribution and the main computational problems associated with it.

5.3.1 SIS over the integers

First, we recall the "standard" SIS problem introduced by Ajtai in [Ajt98]. Several standard worst-case problems over Euclidean lattices reduce to SIS (see, e.g., [Ajt98, MR04, GPV08]).

Definition 5.12. Let $m > n, q \geqslant 2, \beta > 0$ be functions of n. The small integer solution problem $\mathrm{SIS}_{m,q,\beta}$ (in the ℓ_2 norm) is as follows: given an integer q, a matrix $A \leftarrow U(\mathbb{Z}_q^{m\times n})$, and a real $\beta > 0$, find a nonzero integer vector $\mathbf{e} \in \mathbb{Z}^m$ such that $A^{\intercal}\mathbf{e} = \mathbf{0} \bmod q$ and $\|\mathbf{e}\| < \beta$.

We now define the integer-SIS problem. The main difference with the above "standard" SIS is in the distribution of the matrix A: instead of being uniform modulo q, the entries of A are now Gaussian over the integers. Similarly, the success condition does not involve any reduction modulo an integer q.

Definition 5.13. Let $m > n, \sigma_X > 0, \beta > 0$ be functions of n. The small integer solution problem over the integers integer-$\mathrm{SIS}_{m,\sigma_X,\beta}$ (in the ℓ_2 norm) is as follows: given a matrix $X \leftarrow D_{\mathbb{Z}^{m\times n},\sigma_X}$, and a real $\beta > 0$, find a nonzero integer vector $\mathbf{e} \in \mathbb{Z}^m$ such that $X^{\intercal}\mathbf{e} = \mathbf{0}$ and $\|\mathbf{e}\| < \beta$.

Our main result in this section is a reduction from $\mathrm{SIS}_{m,q,\beta}$ to integer-$\mathrm{SIS}_{m,\sigma_X,\beta}$.

Theorem 5.14. Let $n < m < n^{\mathcal{O}(1)}, q \geqslant 2$ and $\beta > 0$ be functions of n. There is a probabilistic polynomial-time reduction from $\mathrm{SIS}_{m,q,\beta}$ to integer-$\mathrm{SIS}_{m,\sigma_X,\beta}$, for any $\sigma_X \geqslant \Omega(q\sqrt{n})$.

Proof. Let $A = (\mathbf{a}_1 \| \dots \| \mathbf{a}_m)$ be an $\mathrm{SIS}_{m,q,\beta}$ instance. The reduction proceeds as follows:

1. For $i \leqslant m$, sample $\mathbf{y}_i \leftarrow D_{q\mathbb{Z}^n,\sigma_X,-\mathbf{a}_i}$ (here, the vector $\mathbf{a}_i$ is seen as a vector with entries in $\{0, \dots, q-1\}$); define $Y = (\mathbf{y}_1 \| \dots \| \mathbf{y}_m)$ and set $X = A + Y \in \mathbb{Z}^{m\times n}$.
2. Call the integer-$\mathrm{SIS}_{m,\sigma_X,\beta}$ oracle with input X; let $\mathbf{e}$ be its output.
3. Return $\mathbf{e}$.

By Lemma 5.3 with parameter $\sigma_X \geqslant \Omega(q\sqrt{n})$, there exists a probabilistic polynomial-time algorithm to sample $\mathbf{y}_i$ from $D_{q\mathbb{Z}^{m\times n},\sigma_X,-\mathbf{a}_i}$. This implies that the reduction can be run in probabilistic polynomial-time. We now prove its correctness.

We first consider the distribution D of the first row $\mathbf{x}$ of X (note that the rows are independent and identically distributed). From $\mathbf{x} = \mathbf{a} + \mathbf{y}$, we have that the support of D is $\mathbb{Z}^n$. Further, for $\mathbf{x} \in \mathbb{Z}^n$, we have:

$$\Pr_{\mathbf{a},\mathbf{y}}[\mathbf{x}] = \Pr_{\mathbf{a}\leftarrow U(\mathbb{Z}_q^n)}[\mathbf{a}] \cdot \frac{\rho_{\sigma_X,-\mathbf{a}}(\mathbf{x}-\mathbf{a})}{\rho_{\sigma_X,-\mathbf{a}}(q\mathbb{Z}^n)}$$

$$\in \frac{1}{q^n} \cdot \left[1 - 2^{-\Omega(n)}, 1\right] \cdot \frac{\rho_{\sigma_X}(\mathbf{x})}{\rho_{\sigma_X}(q\mathbb{Z}^n)} \qquad \text{(from Lemma 5.4 with } \varepsilon = 2^{-\Omega(n)}\text{)}$$
$$\in \left[1 - 2^{-\Omega(n)}, 1 + 2^{-\Omega(n)}\right] \cdot \frac{\rho_{\sigma_X}(\mathbf{x})}{\rho_{\sigma_X}(\mathbb{Z}^n)}.$$

The last equation follows from the fact that $\rho_{\sigma_X}(\mathbb{Z}^n)/\rho_{\sigma_X}(q\mathbb{Z}^n) \in [1-2^{-\Omega(n)}, 1+2^{-\Omega(n)}] \cdot q^n$, using the Poisson summation formula and the fact that $\sigma_X \geqslant \eta_{2^{-\Omega(n)}}(q\mathbb{Z}^n)$. It follows that the distribution of $\mathbf{x}$ is within $2^{-\Omega(n)}$ statistical distance to $D_{\mathbb{Z}^n,\sigma_X}$. Therefore, the distribution of matrix X is within $2^{-\Omega(n)}$ statistical distance to $D_{\mathbb{Z}^{m\times n},\sigma_X}$.

From the above, we obtain that the integer-SIS$_{m,\sigma_X,\beta}$ oracle succeeds with probability $1 - 2^{-\Omega(n)}$ if X was sampled from $D_{\mathbb{Z}^{m\times n},\sigma_X}$. Assume that it does succeed, and let $\mathbf{e}$ be its output: we have that $X^\intercal\mathbf{e} = \mathbf{0}$ and $0 < \|\mathbf{e}\| \leqslant \beta$. As $X = A + Y$, we have $A^\intercal\mathbf{e} = -Y^\intercal\mathbf{e}$. Using the fact that $Y \in q\mathbb{Z}^{m\times n}$, we deduce that $A^\intercal\mathbf{e} = \mathbf{0} \bmod q$. This completes the proof. □

5.3.2 Search integer-LWE

We now define the integer-LWE distribution, and the search integer-LWE problem.

Definition 5.15. Let $m > n \geqslant 1$, be integers, $\sigma_X, \sigma_k > 0$ and $\alpha \in (0,1)$. The integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ distribution is obtained as follows: sample $X \leftarrow D_{\mathbb{Z}^{m\times n},\sigma_X}$, $\mathbf{k} \leftarrow D_{\mathbb{Z}^m,\sigma_k}$ and $\mathbf{e} \leftarrow D_{\mathbb{R}^m,\alpha}$; output $(X, \mathbf{b} = \pi_{\ker(X)}(\mathbf{k} + \mathbf{e}))$.

We observe that the integer-SIS problem is syntactically equivalent to finding a short nonzero vector in the orthogonal lattice. Inspired by the duality with the Bounded Distance Decoding problem (BDD), we define the search integer-LWE problem as a BDD problem instance on the dual lattice.

Definition 5.16. Let n be going to infinity, and $m > n$, $\sigma_X, \sigma_k > 0$, $\alpha \in (0,1)$ be functions of n. The search integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ problem is as follows: given a pair $(X, \mathbf{b})$ from the integer-LWE distribution, output $\mathbf{k}$.

5.4 Hardness of search integer-LWE

Below, we show that an efficient algorithm solving search integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ with some non-negligible probability may be used by a quantum machine to efficiently solve integer-SIS$_{m,\sigma_X,\sqrt{m-n}/\alpha}$ with non-negligible probability, when $\sigma_k = \Omega(\sigma_X\sqrt{mn})$. An important property of the reduction is that the matrix underlying the integer-SIS and integer-LWE instances is preserved.

Theorem 5.17. Let $n \geqslant 60$, $\sigma_X \geqslant \frac{m}{\sqrt{n}}$, $\sigma_k \geqslant (C+1)\sqrt{2\pi \ln 2}\sigma_X\sqrt{mn}$, where C is as in Lemma 5.9, and $\alpha < 0.001/\max(\sqrt{m-n}, \sqrt{n(n+\ln m)})$. Suppose that there exists an algorithm that solves search integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ in time T and with probability $\geqslant \delta$. Then there exists a quantum algorithm that solves integer-SIS$_{m,\sigma_X,\sqrt{m-n}/\alpha}$ in time $\mathrm{poly}(T,n)$ and with probability $\Omega(\delta^4) - 2^{-\Omega(n)}$.

For efficiency purposes, one would be interested in using a rather small m. For security purposes, one is rather interested in larger values of α. For example, the assumptions of the theorem allow to set $m = \Theta(n \log n)$, σ_X between $\Theta(\sqrt{n} \log n)$ and any $\mathrm{poly}(n)$, and α as large as $\Theta(1/\sqrt{n \log n})$.

The proof of Theorem 5.17 is adapted from the SIS to LWE quantum reduction from [SSTX09], relying on Regev's quantum reduction from Gaussian sampling to the Bounded Distance Decoding (BDD) problem [Reg05]. It starts with the observation that integer-LWE is a BDD instance, for the lattice $\pi_{\ker(X)}(\mathbb{Z}^m)$. Conveniently, Regev's reduction maps the BDD lattice to its dual, and the dual of $\pi_{\ker(X)}(\mathbb{Z}^m)$ is $\ker(X)$ (with overwhelming probability over the choice of X, by Lemma 5.7). This implies that solving the BDD instances corresponding to integer-LWE should allow to obtain short non-zero vectors in $\ker(X)$ and hence solve integer-SIS.

In [SSTX09], the authors showed that Regev's reduction can be exploited even if the success probability of the BDD oracle is only non-negligible (as opposed to Regev's proof that requires a success probability that is close to 1). This is convenient to our setup, as an adversary may solve integer-LWE with small probability only. From a technical perspective, this BDD correctness requirement weakening due to [SSTX09] assumes that the BDD oracle is Strongly Solution Independent (SSI), which means that the success probability of the oracle when given $\mathbf{t} = \mathbf{b} + \mathbf{e}$ as input is independent of the lattice vector $\mathbf{b}$: when given $\mathbf{b} + \mathbf{e}$ and $\mathbf{b}' + \mathbf{e}$ as inputs, its success probability should be identical. In the case of [SSTX09], an SSI BDD oracle is built by re-randomizing the input to the provided LWE oracle: the LWE lattice can be viewed as defined modulo q, and a given $\mathbf{b}+\mathbf{e}$ can be mapped to $\mathbf{b}'+\mathbf{e}$ by obliviously adding to it a uniformly chosen element of L/qL where L is the LWE lattice.

In the case of integer-LWE, such a strong self-reducibility seems difficult to achieve, as we are interested in lattice vectors $\mathbf{b}$ of small Euclidean norms. We re-randomize a given $\mathbf{b} + \mathbf{e}$ by adding a Gaussian lattice vector to it. We can show that if the success probability of integer-LWE for a Gaussian lattice vector is bounded from below by a non-negligible quantity, then the re-randomization above provides a BDD algorithm whose success probability is non-negligibly bounded from below for any lattice vector of sufficiently small norm. This is formalized by the notion of Weakly Solution Independence (WSI) of the BDD solver, introduced in Definition 5.19. Conveniently, a WSI BDD solver suffices for the technique from [SSTX09] to apply. The WSI property is obtained by using the boundedness of the Rényi divergence between a Gaussian distribution and the same distribution shifted by a short vector.

We now proceed to proving Theorem 5.17. Motivated by the interpretation of integer-LWE as a Bounded Distance Decoding problem, we first introduce a variant of the latter in which the lattice vector is assumed to be of bounded norm.

Definition 5.18. The variant Bounded Distance Decoding $\mathrm{BDD}_{\eta,\chi}$ with parameters a distribution χ and a real $\eta > 0$ is as follows: given as inputs a lattice L and a vector $\mathbf{t} = \mathbf{b} + \mathbf{e}$ where $\mathbf{b} \in L$ such that $\|\mathbf{b}\| \leqslant \eta$, and $\mathbf{e}$ is distributed according to χ, the goal is to find $\mathbf{b}$.

Weak solution independence is a property of a BDD oracle that states that its success probability does not significantly depend on the choice of the lattice vector $\mathbf{b}$.

Definition 5.19. Let L be an n-rank lattice, χ be a distribution, and $\eta, \delta > 0$. A randomized algorithm $\mathcal{A}$ solving $\mathrm{BDD}_{\eta,\chi}$ is said to be weakly solution-independent (WSI) with parameter δ, if for all $\mathbf{b} \in L$ with $\|\mathbf{b}\| \leqslant \eta$, the probability over the randomness of $\mathcal{A}$ and of $\mathbf{e} \leftarrow \chi$ that $\mathcal{A}$ returns $\mathbf{b}$ when given $\mathbf{t} = \mathbf{b} + \mathbf{e}$ as input, is bounded from below by δ.

The following lemma assumes the existence of an integer-LWE solver for random choices of X and $\mathbf{k}$, and extracts from it an integer-LWE solver that succeeds for all X's in a given set and for all $\mathbf{k}$ of sufficiently small norm. We use the notation integer-$\mathrm{LWE}_{m,X,\mathbf{v},\alpha}$ for fixed X and $\mathbf{v}$.

Lemma 5.20. Let $m > n \geqslant 1$, $\sigma_X, \sigma_k > 0$ and $\alpha \in (0,1)$. Assume that there exists an algorithm $\mathcal{A}$ that solves search integer-$\mathrm{LWE}_{m,\sigma_X,\sigma_k,\alpha}$ in time T and with probability $\delta > 0$. Then there is a set $\mathcal{X} \subseteq \mathbb{Z}^{m\times n}$ with $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}) \geqslant \delta/2$ and an algorithm $\mathcal{B}$ such that for any $\mathbf{v} \in \mathbb{Z}^m$ satisfying $\|\mathbf{v}\| \leqslant \sigma_k/\sqrt{2\pi \ln 2}$ and any $X \in \mathcal{X}$, algorithm $\mathcal{B}$ solves search integer-$\mathrm{LWE}_{m,X,\mathbf{v},\alpha}$, in time $T + \mathrm{poly}(n)$ and with probability $\geqslant \delta^2/8$.

Proof. As $\mathcal{A}$ solves search integer-LWE with probability δ over the random choice of $X, \mathbf{k}, \mathbf{e}$, there exists a set $\mathcal{X}$ of X's with $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}) \geqslant \delta/2$ such that for all $X \in \mathcal{X}$ we have that $\mathcal{A}$ solves search integer-LWE with probability $\geqslant \delta/2$ (over its internal randomness and the random choices of $\mathbf{k}$ and $\mathbf{e}$). We will restrict the analysis of algorithm $\mathcal{B}$ below to any fixed $X \in \mathcal{X}$.

Given as input $(X, \mathbf{b} = \pi_{\ker(X)}(\mathbf{v}+\mathbf{e}))$ with $\|\mathbf{v}\| \leqslant \sigma_k/\sqrt{2\pi \ln 2}$, algorithm $\mathcal{B}$ proceeds as follows:

1. Sample $\mathbf{k} \leftarrow D_{\mathbb{Z}^m,\sigma_k}$;
2. Call algorithm $\mathcal{A}$ on $(X, \pi_{\ker(X)}(\mathbf{b}+\mathbf{k}))$, and let $\mathbf{k}'$ denote its output;
3. Return $\mathbf{k}' - \mathbf{k}$.

Note first that if $\mathcal{A}$ succeeds then so does $\mathcal{B}$, because $\pi_{\ker(X)}(\mathbf{b}+\mathbf{k}) = \pi_{\ker(X)}((\mathbf{v}+\mathbf{k})+\mathbf{e})$. We now show that for all $X \in \mathcal{X}$, for all $\mathbf{v}$ with $\|\mathbf{v}\| \leqslant \sigma_k/\sqrt{2\pi \ln 2}$, algorithm $\mathcal{A}$ succeeds with sufficiently large probability. As $\mathbf{k}$ is distributed according to $D_{\mathbb{Z}^m,\sigma_k}$ and $\mathbf{v} \in \mathbb{Z}^m$, the vector $\mathbf{v}+\mathbf{k}$ is distributed according to $D_{\mathbb{Z}^m,\sigma_k,\mathbf{v}}$. By Lemma 5.10, we have that

$$R\Big(D_{\mathbb{Z}^m,\sigma_k,\mathbf{v}} \| D_{\mathbb{Z}^m,\sigma_k}\Big) \leqslant \exp\left(2\pi \frac{\|\mathbf{v}\|^2}{\sigma_k^2}\right).$$

By Lemma 5.11 and the definition of $\mathcal{X}$, Algorithm $\mathcal{A}$ succeeds with probability $\geqslant (\delta/2)^2/2$. This completes the proof. □

We now obtain a WSI BDD oracle from the above integer-LWE solver for arbitrary $\mathbf{k}$.

Lemma 5.21. Let $m > n \geqslant 1, \sigma_X \geqslant 1, \alpha \in (0,1)$ and $\sigma_k \geqslant (C+1)\sqrt{2\pi \ln 2}\sigma_X\sqrt{mn}$, where C is as in Lemma 5.9. Assume that there exists an algorithm $\mathcal{A}$ that solves search integer-$\mathrm{LWE}_{m,\sigma_X,\sigma_k,\alpha}$ in time T and probability $\delta > 0$. Then there is a set $\mathcal{X}' \subseteq \mathbb{Z}^{m\times n}$ with $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}') \geqslant \delta/2 - 2^{-\Omega(n)}$ and a WSI algorithm $\mathcal{C}$ with parameter $\delta^2/8$, that solves $\mathrm{BDD}_{\sigma_X\sqrt{mn},D_{\mathbb{R}^m,\alpha}}$ in $L = \Lambda^\perp(X)^\star$ in time $T + \mathrm{poly}(n)$, for all $X \in \mathcal{X}'$.

Proof. Our reduction takes as input $X \in \mathbb{Z}^{m\times n}$ and a variant $\mathrm{BBD}_{\eta,D_{\mathbb{R}^m,\alpha}}$ instance $\mathbf{t} = \mathbf{b} + \mathbf{e}$ where $\mathbf{b} \in \Lambda^{\perp}(X)^{\star}$ and $\mathbf{e} \leftarrow D_{\mathbb{R}^m,\alpha}$. Algorithm $\mathcal{C}$ proceeds as follows:

1. Call algorithm $\mathcal{B}$ from Lemma 5.20 on input $(X, \mathbf{t})$, and let $\mathbf{v}' \in \mathbb{Z}^m$ denote its output;
2. Return $\mathbf{b} = \pi_{\ker(X)}(\mathbf{v}')$.

Let $\mathcal{X} \subseteq \mathbb{Z}^{m\times n}$ be as in Lemma 5.20. Define $\mathcal{X}'$ by restricting $\mathcal{X}$ to the set of those X's in $\mathcal{X}$ such that $\sigma_1(X) \leqslant C\sigma_X\sqrt{m}$, for C as in Lemma 5.9. By Lemma 5.9, we have $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}') \geqslant \delta/2 - 2^{-\Omega(n)}$.

We now prove correctness of algorithm $\mathcal{C}$. Let $X \in \mathcal{X}'$ and $\mathbf{b} \in \Lambda^{\perp}(X)^{\star}$ with $\|\mathbf{b}\| \leqslant \sigma_X\sqrt{mn}$. From Lemma 5.7, we can write $\mathbf{b} = \mathbf{v} + X\mathbf{r}$ for some $\mathbf{v} \in \mathbb{Z}^m$ and $\mathbf{r} \in \mathbb{R}^n$ such that $\|\mathbf{r}\|_\infty \leqslant 1/2$. By the triangular inequality and the definition of $\mathcal{X}'$, we have:

$$\|\mathbf{v}\| \leqslant \|\mathbf{b}\| + \|X\mathbf{r}\| \leqslant \sigma_X\sqrt{mn} + C\sigma_X\sqrt{m}\cdot\|\mathbf{r}\| \leqslant \sigma_k/\sqrt{2\pi\ln 2}.$$

From Lemma 5.20, algorithm $\mathcal{B}$ outputs $\mathbf{v}' = \mathbf{v}$ with probability $\geqslant \delta^2/8$. □

The next lemma states that if there exists a WSI algorithm that solves variant $\mathrm{BDD}_{\eta,D_{\mathbb{R}^m,\alpha}}$ for some lattice L, then there exists a WSI algorithm that solves $\mathrm{BDD}_{\eta,D_{L/R,\alpha}}$ for L, when R is sufficiently large. Its adaptation from [SSTX09] is direct.

Lemma 5.22 (Adapted from [SSTX09, Lemma 8])**.** Let L be an n-rank lattice, B a basis of L and $\eta, \alpha > 0$. Suppose that there exists a WSI algorithm with parameter δ that solves $\mathrm{BDD}_{\eta,D_{\mathbb{R}^n,\alpha}}$ for lattice L in time T. Then there exists an R, whose bit-length is polynomial in $T, n, |\log\alpha|$, and the bit-size of B, and a WSI algorithm with parameter $\geqslant \delta - 2^{-n}$ that solves $\mathrm{BDD}_{\eta,D_{L/R,\alpha}}$ in time polynomial in $T, n, |\log\alpha|$ and the bit-size of B.

We now adapt the BDD to Gaussian sampling reduction from [SSTX09]. The weakening from the SSI property of the BDD oracle to the WSI property resides in an upper bound on the Euclidean norms of the vectors on which the BDD oracle is called. From the proof of [SSTX09, Lemma 9], it can be observed that the BDD oracle is called on vectors of the form $\mathbf{t} = \mathbf{e} \bmod L$, where $\mathbf{e}$ is a BDD error term. One can write $\mathbf{t} = \mathbf{b} + \mathbf{e}$ with $\|\mathbf{b}\| \leqslant n\cdot\max_i\|\mathbf{b}_i\|$, where $(\mathbf{b}_i)_i$ is the given basis of L. These are the vectors of L for which the WSI property should hold.

Lemma 5.23 (Adapted from [SSTX09, Lemma 9])**.** Let L be an n-rank lattice, B be a basis of L, $R > 2^{2n}\cdot\lambda_n(L)$, $\eta \geqslant n\cdot\|B\|$ and $\alpha \leqslant \lambda_1(L)/(2\sqrt{2n})$. Assume that there exists a WSI algorithm with parameter δ that solved $\mathrm{BDD}_{\eta,D_{L/R,\alpha}}$ for L, in time T. Then there exists a quantum algorithm which outputs a vector $\mathbf{u} \in L^{\star}$ whose distribution is within statistical distance $1 - \delta^2/2 + \mathcal{O}(\delta^4) + 2^{\Omega(-n)}$ of $D_{L^{\star},1/(2\alpha)}$. It finishes in time polynomial in $T + \log R$.

We observe that the lemma above also works for lattices that are not of full rank. We now conclude this section with a proof for Theorem 5.17.

Theorem 5.17. Using Lemma 5.21, there exists a set $\mathcal{X}' \subseteq \mathbb{Z}^{m\times n}$ with $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}') \geqslant \delta/2 - 2^{-\Omega(n)}$ and a WSI algorithm $\mathcal{C}$ with parameter $\delta^2/8$, that solves $\mathrm{BDD}_{\sigma_X\sqrt{mn},D_{\mathbb{R}^m,\alpha}}$ in $L = \Lambda^\perp(X)^\star$ in time $T + \mathrm{poly}(n)$, for all $X \in \mathcal{X}'$. By using Lemma 5.22, we obtain a WSI algorithm $\mathcal{C}'$ with parameter $\delta^2/8 - 2^{-n}$ that solves $\mathrm{BDD}_{\sigma_X\sqrt{mn},D_{L/R,\alpha}}$ in time polynomial in $T, n, |\log\alpha|$ (with $L = \Lambda^\perp(X)^\star$, and for all $X \in \mathcal{X}'$).

We now would like to apply Lemma 5.23 to algorithm $\mathcal{C}'$. We first check that its assumptions are satisfied. Note that the projections of the canonical vectors of $\mathbb{Z}^m$ orthogonally to $\ker(X)$ provide a generating set of vectors of $\pi_{\ker(X)}(\mathbb{Z}^m)$ of norms $\leqslant 1$. By Lemma 5.1, we can hence efficiently obtain a basis B of $\pi_{\ker(X)}(\mathbb{Z}^m)$ such that $\|B\| \leqslant \sqrt{m-n}$. By assumption on σ_X, the inequality $\sigma_X\sqrt{mn} \geqslant (m-n)\sqrt{m-n}$ holds. From the second statement in Lemma 5.8, we have that $\lambda_1(\Lambda^\perp(X)^\star) \geqslant 1/96$ holds with probability $1 - 2^{-\Omega(n)}$ over choice of X. Thanks to the assumption on α, we obtain that $\alpha \leqslant \frac{\lambda_1(\Lambda^\perp(X)^\star)}{2\sqrt{2(m-n)}}$ for a set $\mathcal{X}'' \subseteq \mathbb{Z}^{m\times n}$ of X's such that $D_{\mathbb{Z}^{m\times n},\sigma_X}(\mathcal{X}'') \geqslant \delta/2 - 2^{-\Omega(n)}$. For later use, we assume that for an $X \in \mathcal{X}''$, both lower bounds from Lemma 5.8 hold.

Suppose that $X \in \mathcal{X}''$. Lemma 5.23 applied to $L = \Lambda^\perp(X)^\star$ allows us to quantumly obtain a vector $\mathbf{u} \in \Lambda^\perp(X)$ whose distribution is within statistical distance $1 - \Omega(\delta^4) + 2^{-\Omega(n)}$ from $D_{\Lambda^\perp(X),1/(2\alpha)}$. By the Gaussian tail bound (Lemma 5.6), the vector $\mathbf{u}$ has norm $\leqslant \sqrt{m-n}/\alpha$ with probability $\Omega(\delta^4) - 2^{-\Omega(n)}$. We now show that $\mathbf{u}$ is non-zero with sufficiently high probability. By definition of $\mathcal{X}''$, we have $\lambda_1^\infty(\Lambda^\perp(X)^\star) \geqslant 1/(96\sqrt{n + \ln m})$ for all $X \in \mathcal{X}''$. The assumption on α ensures that the hypothesis of Lemma 5.5 holds. Using the latter lemma gives that $\mathbf{u}$ is non-zero with probability $\Omega(\delta^4) - 2^{-\Omega(n)}$. This completes the proof. □

Chapter 6

Conclusion and open problems

In this chapter, we will give a short conclusion of this thesis. Then we will present some open problems that are related to the works.

6.1 Conclusion

In this thesis, we have studied some cryptographic aspects of orthogonal lattice. The ultimate objective is to understand this family of lattices sufficiently well to simplify and optimize lattice-based cryptography. In some cases, the disclosed properties lead to efficient cryptanalytic algorithms, hence invalidating candidate constructions; in other cases, they help improving security proofs, hence increasing confidence in other constructions; and they also help to show relationships between different cryptographic constructions or suggest other cryptographic designs.

We first studied the most important parameter of the orthogonal lattice, the smoothing parameter. When orthogonal lattices are used in various constructions such as cryptographic multilinear maps, traitor-tracing schemes, and inner product functional encryption, improving the smoothing parameter bound implies improving the algorithm efficiency. Our other result gives a bound of the $(m-n)$-th minimum of the orthogonal lattice. This parameter is also very useful in cryptanalytic algorithms.

We then studied the uses of the orthogonal lattices in revoke and traitor-tracing schemes. Starting from inner product functional encryption (IPFE), we construct a

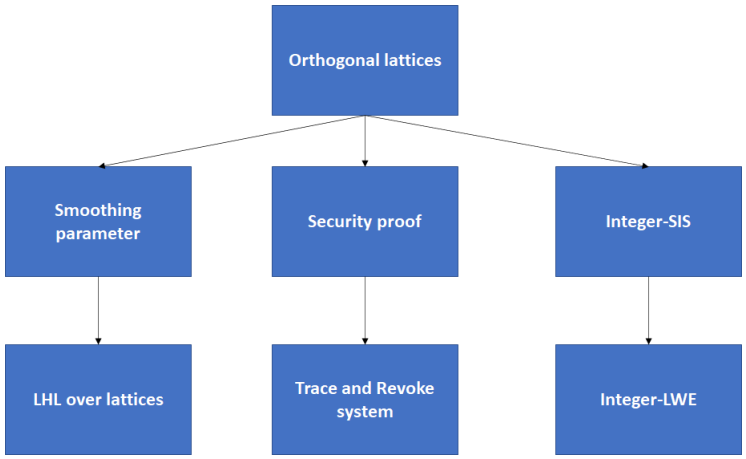

Figure 6.1: Cryptographic aspects of orthogonal lattices.

trace-and-revoke system from [ABP+17] with the novelty of achieving anonymity. And we further show that our group-based construction is tightly secure under standard assumptions. For the lattice-based setting, we suggest using the scheme of [ABP+17] as we could mount a concrete attack on the state-of-the-art [WFL19], rendering it insecure. Our attack comes from the following observation: if a vector $\mathbf{z}$ belongs to a coset of the orthogonal lattice of X and the minimum of this lattice is larger than $\|\mathbf{z}\|$ then we have a bounded distance decoding (BDD) instance in a lattice of dimension 1. It follows that we can find $\mathbf{z}$ efficiently.

Finally, we studied variants of two well-known lattice problems without any modular reduction, namely integer-SIS and integer-LWE, and put forward a reduction from one to another. First, we gave a reduction from BDD for specific lattices to integer-LWE using Rényi divergence. Then we proposed a quantum reduction from integer-SIS to BDD for those specific lattices.

6.2 Open problems

In light of the foregoing, we aim to present some open problems for future works.

Open problem 1. What is decision integer-LWE?

In our intuition, the decision integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ is as follows: given as input a pair $(X, \mathbf{b})$, decide, with non-negligible advantage, whether it was sampled from the integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ distribution or whether it was sampled from $\left(D_{\mathbb{Z}^{m\times n},\sigma_X} \times \pi_{\ker(X)}(D_{\mathbb{R}^m,\sqrt{\sigma_k^2+\alpha^2}})\right)$. Our definition comes from the fact that the integer-LWE$_{m,\sigma_X,\sigma_k,\alpha}$ distribution is statistically close to $\left(D_{\mathbb{Z}^{m\times n},\sigma_X} \times \pi_{\ker(X)}(D_{\mathbb{R}^m,\sqrt{\sigma_k^2+\alpha^2}})\right)$ where α is big enough. We will investigate it, in an effort to prove its hardness.

Open problem 2. How to construct cryptographic primitives based on the hardness of integer-LWE?

We give two suggestions to build the schemes which are based on the hardness of integer-LWE.

1. Can we construct an encryption scheme whose security is based on the hardness of the decision integer-LWE problem like [Reg05]?

2. Can we construct an encryption scheme whose security is based on the hardness of the search integer-LWE problem like [SSTX09]?

Open problem 3. What are the variants of integer-LWE?

In our work, the defining matrix X is sampled from a Gaussian distribution, leading to another family of random lattices. In efficient cryptographic constructions, the above matrix X is often randomly conditioned on a specific matrix structure, typically made of blocks that are Toeplitz matrices. In this case, the resulting lattices are algebraically richer than arbitrary lattices: they typically enjoy a module structure over an order of a number field. Other families of algebraic lattices arise in some cryptographic schemes or their proofs, such as the log-unit lattice [CDPR16] and the Schnorr-Adleman factoring lattice [Ajt98,Sch91]. Following our work on integer-LWE, I want to study a ring variant of integer-LWE and a modulo variant of integer-LWE.

Open problem 4. How to recover the encryption randomness?

Here we want to recover the encryption randomness when decrypting in lattice-based encryption schemes. In Regev's encryption scheme [Reg05] (and the NIST proposals [CJL+16]), the decryption algorithms allow to recover the inner product between the secret key vector and a noise vector. But it does not allow the recovery of the noise vector itself. J. Deneuville et al. have a heuristic way to do this in [DGGJ18]. In [HHK17], this is useful to obtain a tightly secure upgrade from CPA to CCA security in the quantum random oracle model. We will investigate the problem of the recovery of the noise vector, in an effort to improve it and to make it to be more rigorous.

Bibliography

[ABDCP15] Michel Abdalla, Florian Bourse, Angelo De-Caro, and David Pointcheval. Simple functional encryption schemes for inner products. In *International Workshop on Public Key Cryptography*, volume 9020 of *LNCS*, pages 733–751. Springer, 2015. 14, 22

[ABP+17] Shweta Agrawal, Sanjay Bhattacherjee, Duong Hieu Phan, Damien Stehlé, and Shota Yamada. Efficient public trace and revoke from standard assumptions: Extended abstract. In Bhavani M. Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu, editors, *ACM CCS 2017*, pages 2277–2293. ACM Press, October / November 2017. 14, 16, 22, 24, 49, 50, 51, 52, 53, 56, 57, 61, 62, 63, 64, 65, 68, 73, 74, 75, 92

[ACF+18] Michel Abdalla, Dario Catalano, Dario Fiore, Romain Gay, and Bogdan Ursu. Multi-input functional encryption for inner products: Function-hiding realizations and constructions without pairings. In Hovav Shacham and Alexandra Boldyreva, editors, *CRYPTO 2018, Part I*, volume 10991 of *LNCS*, pages 597–627. Springer, Heidelberg, August 2018. 53, 56

[AGHS13] Shweta Agrawal, Craig Gentry, Shai Halevi, and Amit Sahai. Discrete Gaussian leftover hash lemma over infinite domains. In *Advances in Cryptology - ASIACRYPT 2013*, pages 97–116, 2013. 12, 13, 16, 20, 21, 24, 31, 32, 33, 34, 37, 38, 39, 40, 42, 46, 83

[AGRW17] Michel Abdalla, Romain Gay, Mariana Raykova, and Hoeteck Wee. Multi-input inner-product functional encryption from pairings. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, *EUROCRYPT 2017, Part I*, volume 10210 of *LNCS*, pages 601–626. Springer, Heidelberg, April / May 2017. 53

[Ajt96] Miklós Ajtai. Generating hard instances of lattice problems (extended abstract). In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 99–108, 1996. 11, 19, 22

[Ajt98] Miklós Ajtai. The shortest vector problem in ℓ_2 is NP-hard for randomized reductions (extended abstract). In *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, pages 10–19, 1998. 11, 19, 84, 92

[ALS16] Shweta Agrawal, Benoît Libert, and Damien Stehlé. Fully secure functional encryption for inner products, from standard assumptions. In Matthew

Robshaw and Jonathan Katz, editors, *CRYPTO 2016, Part III*, volume 9816 of *LNCS*, pages 333–362. Springer, Heidelberg, August 2016. 13, 14, 21, 22, 52, 53, 56, 73, 74, 75

[APS18] Navid Alamati, Chris Peikert, and Noah Stephens-Davidowitz. New (and old) proof systems for lattice problems. In *Public-Key Cryptography - PKC 2018*, pages 619–643, 2018. 37

[AR16] Divesh Aggarwal and Oded Regev. A note on discrete Gaussian combinations of lattice vectors. *Chic. J. Theoret. Comput. Sci.*, (7), June 2016. 12, 16, 20, 24, 32, 33, 34, 35, 38, 46

[AY20] Shweta Agrawal and Shota Yamada. Optimal broadcast encryption from pairings and LWE. In Anne Canteaut and Yuval Ishai, editors, *EUROCRYPT 2020, Part I*, volume 12105 of *LNCS*, pages 13–43. Springer, Heidelberg, May 2020. 49

[Ban93] Wojciech Banaszczyk. New bounds in some transference theorems in the geometry of numbers. *Math. Ann.*, pages 625–636, 1993. 29, 34, 35, 36, 37

[BBW06] Adam Barth, Dan Boneh, and Brent Waters. Privacy in encrypted content distribution using private broadcast encryption. In Giovanni Di Crescenzo and Avi Rubin, editors, *FC 2006*, volume 4107 of *LNCS*, pages 52–64. Springer, Heidelberg, February / March 2006. 49

[BDE+18] Jonathan Bootle, Claire Delaplace, Thomas Espitau, Pierre-Alain Fouque, and Mehdi Tibouchi. LWE without modular reduction and improved side-channel attacks against bliss. In *ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security*, pages 494–524, 2018. 81

[Bel11] Jean-Claude Belfiore. Lattice codes for the compute-and-forward protocol: The flatness factor. *2011 IEEE Information Theory Workshop, ITW 2011*, 10 2011. 29

[BF99] Dan Boneh and Matthew K. Franklin. An efficient public key traitor tracing scheme. In Michael J. Wiener, editor, *CRYPTO'99*, volume 1666 of *LNCS*, pages 338–353. Springer, Heidelberg, August 1999. 57

[BLP+13] Zvika Brakerski, Adeline Langlois, Chris Peikert, Oded Regev, and Damien Stehlé. Classical hardness of learning with errors. In *STOC 2013 - Proceedings of the 2013 ACM Symposium on Theory of Computing*, pages 575–584, 2013. 16, 23, 82

[BMN+21] Olivier Blazy, Sayantan Mukherjee, Huyen Nguyen, Hieu Phan, and Damien Stehlé. An anonymous trace-and-revoke broadcast encryption scheme. In *Accepted at ACISP*, 2021. 16, 24, 49, 76

[BRS13] Dan Boneh, Ananth Raghunathan, and Gil Segev. Function-private subspace-membership encryption and its applications. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part I*, volume 8269 of *LNCS*, pages 255–275. Springer, Heidelberg, December 2013. 54

[BSW11] Dan Boneh, Amit Sahai, and Brent Waters. Functional encryption: Definitions and challenges. In Yuval Ishai, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 253–273. Springer, Heidelberg, March 2011. 52

[BW06] Dan Boneh and Brent Waters. A fully collusion resistant broadcast, trace, and revoke system. In *CCS*, page 211–220. Association for Computing Machinery, 2006. 13, 22, 49

[CDPR16] Ronald Cramer, Léo Ducas, Chris Peikert, and Oded Regev. Recovering short generators of principal ideals in cyclotomic rings. In *Advances in Cryptology – EUROCRYPT 2016*, pages 559–585, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg. 92

[CJL+16] Lidong Chen, Stephen Jordan, Yi-Kai Liu, Dustin Moody, Rene Peralta, Ray Perlner, and Daniel Smith-Tone. Report on post-quantum cryptography. NIST, 2016. 93

[CLT18] Guilhem Castagnos, Fabien Laguillaumie, and Ida Tucker. Practical fully secure unrestricted inner product functional encryption modulo p. In Thomas Peyrin and Steven Galbraith, editors, *ASIACRYPT 2018, Part II*, volume 11273 of *LNCS*, pages 733–764. Springer, Heidelberg, December 2018. 53

[CS93] John H. Conway and Neil J. A. Sloane. *Sphere packings, lattices, and groups*. Third edition, Springer-Verlag, 1993. 34

[CSV13] Jingwei Chen, Damien Stehlé, and Gilles Villard. A new view on HJLS and PSLQ: Sums and projections of lattices. In *Proceedings of the 38th International Symposium on Symbolic and Algebraic Computation*, pages 149–156, 2013. 28

[DF03] Yevgeniy Dodis and Nelly Fazio. Public key trace and revoke scheme secure against adaptive chosen ciphertext attack. In Yvo Desmedt, editor, *PKC 2003*, volume 2567 of *LNCS*, pages 100–115. Springer, Heidelberg, January 2003. 49

[DGGJ18] Jean-Christophe Deneuville, Philippe Gaborit, Qian Guo, and Thomas Johansson. Ouroboros-e: An efficient lattice-based key-exchange protocol. In *ISIT*, pages 1450–1454. IEEE, 2018. 93

[DGHV10] Marten van Dijk, Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. Fully homomorphic encryption over the integers. In *Advances in Cryptology — EUROCRYPT 2010*, pages 24–43, 2010. 13, 21

[DPY20] Xuan Thanh Do, Duong Hieu Phan, and Moti Yung. A concise bounded anonymous broadcast yielding combinatorial trace-and-revoke schemes. In Mauro Conti, Jianying Zhou, Emiliano Casalicchio, and Angelo Spognardi, editors, *ACNS 20, Part II*, volume 12147 of *LNCS*, pages 145–164. Springer, Heidelberg, October 2020. 49, 61

[DRN14] Daniel Dadush, Oded Regev, and Stephens-Davidowitz Noah. On the closest vector problem with a distance guarantee. In *Proceedings of the 2014 IEEE 29th Conference on Computational Complexity*, pages 98–109, 2014. 29, 82

[EHK+17] Alex Escala, Gottfried Herold, Eike Kiltz, Carla Ràfols, and Jorge Luis Villar. An algebraic framework for Diffie-Hellman assumptions. *Journal of Cryptology*, 30(1):242–288, January 2017. 55, 56

[FP12] Nelly Fazio and Irippuge Milinda Perera. Outsider-anonymous broadcast encryption with sublinear ciphertexts. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 225–242. Springer, Heidelberg, May 2012. 49, 60

[GGH13] Sanjam Garg, Craig Gentry, and Shai Halevi. Candidate multilinear maps from ideal lattices. In *Advances in Cryptology – EUROCRYPT 2013*, pages 1–17, 2013. 31

[GKW18] Rishab Goyal, Venkata Koppula, and Brent Waters. Collusion resistant traitor tracing from learning with errors. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *50th ACM STOC*, pages 660–670. ACM Press, June 2018. 61

[GPV08] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing*, pages 197–206, 2008. 11, 12, 15, 19, 22, 82, 84

[HB88] D.R. Heath-Brown. The number of primes in a short interval. *J. Reine Angew. Math.*, 389:22–63, 1988. 46

[HHK17] Dennis Hofheinz, Kathrin Hovelmanns, and Eike Kiltz. A modular analysis of the Fujisaki-Okamoto transformation. In *TCC*, pages 341–371. Springer, 2017. 93

[HM17] Gottfried Herold and Alexander May. LP solutions of vectorial integer subset sums – cryptanalysis of Galbraith's binary matrix LWE. In *Public-Key Cryptography – PKC 2017*, pages 3–15, Berlin, Heidelberg, 2017. Springer Berlin Heidelberg. 16, 23

[HMM98] George Havas, Bohdan S. Majewski, and Keith R. Matthews. Extended GCD and Hermite normal form algorithms via lattice basis reduction. *Exp. Math.*, 7(2):125–136, 1998. 31

[ILL89] Russell Impagliazzo, Leonid A. Levin, and Michael Luby. Pseudo-random generation from one-way functions. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing*, pages 12–24. ACM, 1989. 11, 19

[KHL03] Chong Hee Kim, Yong Ho Hwang, and Pil Joong Lee. An efficient public key trace and revoke scheme secure against adaptive chosen ciphertext attack. In Chi-Sung Laih, editor, *ASIACRYPT 2003*, volume 2894 of *LNCS*, pages 359–373. Springer, Heidelberg, November / December 2003. 49

[KNSW20] Elena Kirshanova, Huyen Nguyen, Damien Stehlé, and Alexandre Wallet. On the smoothing parameter and last minimum of random orthogonal lattices. *Designs, Codes and Cryptography*, pages 1–20, 2020. 16, 24, 31, 33, 34, 83

[KS12] Aggelos Kiayias and Katerina Samari. Lower bounds for private broadcast encryption. In *Information Hiding*, volume 7692 of *LNCS*, pages 176–190. Springer, 2012. 49

[LG18] Jiangtao Li and Junqing Gong. Improved anonymous broadcast encryptions - tight security and shorter ciphertext. In Bart Preneel and Frederik Vercauteren, editors, *ACNS 18*, volume 10892 of *LNCS*, pages 497–515. Springer, Heidelberg, July 2018. 49

[LLL82] Arjen K. Lenstra, Hendrik W. Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261:515–534, 1982. 31

[LLLS13] Fabien Laguillaumie, Adeline Langlois, Benoît Libert, and Damien Stehlé. Lattice-based group signatures with logarithmic signature size. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT 2013, Part II*, volume 8270 of *LNCS*, pages 41–61. Springer, Heidelberg, December 2013. 15, 22

[LPQ12] Benoît Libert, Kenneth G. Paterson, and Elizabeth A. Quaglia. Anonymous broadcast encryption: Adaptive security and efficient constructions in the standard model. In Marc Fischlin, Johannes Buchmann, and Mark Manulis, editors, *PKC 2012*, volume 7293 of *LNCS*, pages 206–224. Springer, Heidelberg, May 2012. 49

[LPR10] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *Proceedings of the 29th Annual International Conference on Theory and Applications of Cryptographic Techniques*, pages 1–23, 2010. 16, 23

[LPSS17] San Ling, Duong Hieu Phan, Damien Stehlé, and Ron Steinfeld. Hardness of k-LWE and applications in traitor tracing. *Algorithmica*, pages 1318–1352, 2017. 13, 21

[LS14] Adeline Langlois and Damien Stehlé. Worst-case to average-case reductions for module lattices. *Designs, Codes and Cryptography*, 75, 06 2014. 15, 22

[LSS14] Adeline Langlois, Damien Stehlé, and Ron Steinfeld. GGHLite: More efficient multilinear maps from ideal lattices. In *Advances in Cryptology – EUROCRYPT 2014*, pages 239–256, Berlin, Heidelberg, 2014. Springer Berlin Heidelberg. 83

[Lyu13] Vadim Lyubashevsky. Lattice signatures without trapdoors. In *EUROCRYPT*, volume 7237, pages 738–755. Springer, 2013. 15, 22

[MG02] Daniele Micciancio and Shafi Goldwasser. *Complexity of Lattice Problems: a cryptographic perspective*, volume 671 of *The Kluwer International Series in Engineering and Computer Science*. Kluwer Academic Publishers, 2002. 82

[MP13] Daniele Micciancio and Chris Peikert. Hardness of SIS and LWE with small parameters. In *Advances in Cryptology – CRYPTO 2013*, pages 21–39, 2013. 16, 23

[MR04] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on Gaussian measures. 37(1):267–302, 2007. Preliminary version in FOCS 2004. 11, 13, 19, 20, 29, 37, 82, 84

[Ngu99] Phong Nguyen. *La géométrie des nombres en cryptologie*. PhD thesis, Université Paris 7, 1999. 13, 20, 35, 75

[NNL01] Dalit Naor, Moni Naor, and Jeffery Lotspiech. Revocation and tracing schemes for stateless receivers. In Joe Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 41–62. Springer, Heidelberg, August 2001. 49

[NP01] Moni Naor and Benny Pinkas. Efficient trace and revoke schemes. In Yair Frankel, editor, *FC 2000*, volume 1962 of *LNCS*, pages 1–20. Springer, Heidelberg, February 2001. 49

[NS97] Phong Nguyen and Jacques Stern. Merkle-Hellman revisited: A cryptanalysis of the Qu-Vanstone cryptosystem based on group factorizations. In *Advances in Cryptology – CRYPTO 1997*, pages 198–212, 1997. 13, 21, 31

[NS99] Phong Nguyen and Jacques Stern. The hardness of the hidden subset sum problem and its cryptographic implications. In *Advances in Cryptology – CRYPTO 1999*, pages 31–46, 1999. 13, 21

[NV09] Phong Nguyen and Brigitte Vallée. *The LLL Algorithm: Survey and Applications*. Springer, 1st edition, 2009. 27

[NWZ16] Ryo Nishimaki, Daniel Wichs, and Mark Zhandry. Anonymous traitor tracing: How to embed arbitrary information in a key. In Marc Fischlin and Jean-Sébastien Coron, editors, *EUROCRYPT 2016, Part II*, volume 9666 of *LNCS*, pages 388–419. Springer, Heidelberg, May 2016. 49, 61

[Pei08] Chris Peikert. Limits on the hardness of lattice problems in ℓ_p norms. *Computational Complexity*, 17(2):300–351, 2008. 37

[Pei15] Chris Peikert. A decade of lattice cryptography. *IACR Cryptology ePrint Archive*, page 939, 2015. 11, 15, 19, 22

[PR06] Chris Peikert and Alon Rosen. Efficient collision-resistant hashing from worst-case assumptions on cyclic lattices. In *Theory of Cryptography*, pages 145–166, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. 15, 22, 82

[PRS17] Chris Peikert, Oded Regev, and Noah Stephens-Davidowitz. Pseudorandomness of ring-LWE for any ring and modulus. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing*, pages 461–473. ACM, 2017. 37

[Reg05] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 84–93, 2005. 11, 12, 15, 16, 19, 22, 23, 86, 92, 93

[RSW18] Miruna Rosca, Damien Stehlé, and Alexandre Wallet. On the Ring-LWE and Polynomial-LWE problems. In *Advances in Cryptology - EUROCRYPT 2018*, pages 146–173, 2018. 16, 23

[Sch91] C. P. Schnorr. Factoring integers and computing discrete logarithms via diophantine approximation. In *Advances in Cryptology — EUROCRYPT '91*, pages 281–293, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg. 92

[SSTX09] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In *Advances in Cryptology – ASIACRYPT 2009*, pages 617–635, 2009. 16, 23, 86, 88, 92

[SSW09] Emily Shen, Elaine Shi, and Brent Waters. Predicate privacy in encryption systems. In Omer Reingold, editor, *TCC 2009*, volume 5444 of *LNCS*, pages 457–473. Springer, Heidelberg, March 2009. 53

[Tom19] Junichi Tomida. Tightly secure inner product functional encryption: Multi-input and function-hiding constructions. In Steven D. Galbraith and Shiho Moriai, editors, *ASIACRYPT 2019, Part III*, volume 11923 of *LNCS*, pages 459–488. Springer, Heidelberg, December 2019. 50, 53, 56, 76, 77, 104

[WFL19] Zhedong Wang, Xiong Fan, and Feng-Hao Liu. FE for inner products and its application to decentralized ABE. In *PKC*, 2019. 17, 24, 50, 52, 73, 74, 75, 92

List of Figures

List of Tables