KDM and Selective Opening Secure IBE Based on the LWE Problem

Jingnan He[1,2,3], Bao Li[1,2], Xianhui Lu[1,2], Dingding Jia[1,2]*, Wenpan Jing[1,2]

[1]State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
[2]Data Assurance and Communication Security Research Center, Chinese Academy of Sciences
[3]University of Chinese Academy of Sciences, Beijing, China
{jnhe13,lb,xhlu,ddjia,wpjing}@is.ac.cn

ABSTRACT

Based on the learning with errors (LWE) problem, we construct an identity-based encryption (IBE) scheme in the standard-model which is secure against the key dependent message (KDM) attacks and the selective opening (SO) attacks.

KDM security, which requires that the scheme can still obtain secrecy even if the messages depend on the secret key, is necessary for IBE when updating the identity secret keys with the old ones. SO security, which requires that when the adversary corrupts some fraction of senders, the uncorrupted senders' messages still maintain privacy, is useful for IBE since that one user receives messages from many senders is a natural scenario. Although there is a KDM secure IBE scheme based on LWE constructed by Alperin-Sheriff and Peikert (PKC'12), and a SO secure IBE scheme based on LWE constructed by He *et al.* (IWSEC'15), those two schemes cannot achieve both of the security requirements at the same time.

In this paper, first, we propose a new transformation which is a new variant of LWE assumption. Then, we prove that a dual-style public key encryption (PKE) scheme of the LWE based PKE of Gentry *et al.* (STOC'08) is KDM secure with the help of our transformation. At last, based on the PKE scheme, we construct an IBE scheme in the standard model that is both KDM *and* SO secure.

Keywords

Key Dependent Message Security, Selective Opening Security, Learning with Errors, Identity-based Lossy Encryption

*This author is the corresponding author.

ASIA CCS '17, April 4–6, 2017, Abu Dhabi, United Arab Emirates.

DOI: http://dx.doi.org/10.1145/3055504.3055508

1. INTRODUCTION

1.1 Motivation

The learning with errors (LWE) problem, introduced by Regev in 2005 [21], is applied in a wide range of cryptography [21, 11, 14, 11, 9, 1, 2, 3, 12, 18]. It is very attractive in the last decade, because of its simple and highly parallelizable operations and its assumed hardness in the post-quantum world. Moreover, its average-case problem enjoyed worst-case hardness guarantees.

Our focus in this paper is on the construction of IBE based on LWE which can achieve two stronger security requirements, the key dependent message (KDM) security and the selective opening (SO) security. KDM secure encryption scheme, introduced by Black, Rogaway and Shrimpton [8] in 2002, guarantees the secrecy even if the encrypted messages depend on secret keys. Applebaum *et al.* [5] and Alperin-Sheriff, Peikert [4] proposed KDM secure constructions based on LWE problem. SO security is another strong security notion. It states that in the multi-party computation senario, even if the adversary corrupts a part of the senders and knows their messages and randomnesses, the unopened ciphertexts remain secure. In [6], Bellare, Hofheinz and Yilek proposed a framework of lossy encryption to achieve SO security, and the constructions based on LWE in [18, 19] also fit this framework.

In the IBE scenario, it is natrual to require the scheme to be both KDM secure *and* SO secure. As Alperin-Sheriff and Peikert [4] suggested, when the secret keys are about to expire, the private key generator (PKG) can encrypt the new secret key under the user's identity for the previous epoch; and when the user loses the old key and wants to decrypt the old ciphertext, the PKG can encrypt the old secret key under the user's identity for the current epoch. So the IBE scheme needs to be KDM secure. Besides, since that many senders send messages to one user is natural in the IBE scenario, the SO security is also necessary.

However, there is no LWE-based IBE scheme being both KDM and SO secure. On one hand, Alperin-Sheriff and Peikert [4] constructed a selective KDM secure IBE for user secret keys with the help of all-but-d trapdoor functions in 2012. Their scheme cannot achieve SO security, because the user public key in their scheme is perturbed by an extra Gaussian noise, and this extra noise results in the randomness vector drawn from the discrete Gaussian distribution rather than the uniform distribution which does not have enough entropy to achieve SO security. On the other hand,

He et al. [13] extended the notion of lossy encryption to the IBE setting and proved that the identity-based lossy encryption (IBLE) can achieve the indistinguishability-based selective opening security in the selective identity setting (IND-sID-SO) in 2015. Then they constructed an IBLE scheme from LWE. However, their scheme is not KDM secure because the trapdoor construction just allows the adversary to attack one identity, which does not fit the multi-user setting of IBE.

1.2 Our Contributions and Techniques

As mentioned above, there are two main technique hurdles to achieve KDM and SO security based on LWE at the same time: (1) the randomness' distribution of the current LWE-based KDM secure IBE scheme in [4], which is the discrete Gaussian, limits the scheme to fit the framework IBLE which can achieve SO security; (2) the trapdoor construction of the current LWE-based SO secure IBE scheme in [13], which just allows that the adversary attacks one identity, limits the scheme to be applied in the multi-user setting for KDM security. Therefore, how to achieve the proper distribution of the randomness and multi-user setting are the key problem.

In this paper, we simplified the structure of the KDM secure IBE scheme constructed by Alperin-Sheriff and Peikert [4], by eliminating the extra Gaussian noise in the user public key, and changing one Gaussian randomness vector to a uniform randomness vector. Then, we prove that our scheme not only is KDM secure, but also achieves SO security.

KDM secure IBE scheme with the uniform randomness. The key problem in proving KDM security is how to generate public keys and challenge ciphertexts whose corresponding messages are affine functions of the secret keys without knowing secret keys.

The public key of the PKE scheme in [4] is $-\mathbf{A}\mathbf{z}_1 + \mathbf{z}_0$ where $\mathbf{z}_1$ is the secret key and $\mathbf{z}_1, \mathbf{z}_0$ follow from the discrete Gaussian distribution. Since this type of the public key and the ciphertext are all Gaussian LWE instances, then it is easy to generate them from standard LWE instance by the transformation proposed by Applebaum [5], without knowing the secret key $\mathbf{z}_1$. That is why Alperin-Sheriff and Peikert [4] mentioned that the KDM security of their construction relied heavily on the structure $-\mathbf{A}\mathbf{z}_1 + \mathbf{z}_0$ of the public key which was perturbed by an extra Gaussian noise $\mathbf{z}_0$ like [14] instead of using the structure $\mathbf{A}\mathbf{z}_1$ like [11].

The problem (1) of the randomness' distribution in the IBE scheme inherits from the corresponding PKE scheme. Because Alperin-Sheriff and Peikert [4] constructed the KDM secure IBE scheme from the corresponding KDM secure PKE scheme and the all-but-d trapdoor function, the user public key $-\mathbf{A}_{id}\mathbf{z}_1 + \mathbf{z}_0$ in the IBE scheme is also perturbed by an extra Gaussian noise which inherited from the corresponding PKE scheme. However, the distribution of the randomness vector is sensitively influenced by the extra Gaussian noise $\mathbf{z}_0$. Since the randomness vector will amplify the noise $\mathbf{z}_0$ when it multiplies the user public key in the ciphertext, then to limit the amplification in the range of decryption capability, the randomness vector should be drawn from the discrete Gaussian rather than the uniform distribution. Besides, the Gaussian noise $\mathbf{z}_0$ is not necessary in the IBE scheme, since that $\mathbf{A}_{id}\mathbf{z}_1$ is statistically indistinguishable from the uniform distribution. And in fact, the classical IBE schemes based on LWE like [11, 9, 1, 2] does not have this extra Gaussian noise $\mathbf{z}_0$.

Therefore, by removing the extra Gaussian noise $\mathbf{z}_0$ from the public key of the corresponding PKE scheme, there is no amplification when the randomness multiplies the public key in the ciphertext and the randomness does not need to be drawn from the discrete Gaussian. Then the randomness can be drawn from the uniform distribution which has enough entropy to fit the IBLE framework. To do so, another problem arises that if removing the extra Gaussian noise, public keys $\mathbf{A}\mathbf{z}_1$ will be knapsack LWE instances which are different from the Gaussian LWE instances of ciphertexts and current transformations cannot generate those two instances which share the same secret key. How to generate the public keys which are knapsack LWE instances and the challenge ciphertexts which are Gaussian LWE instances, without knowing the secret keys? By extending previous results, we propose a new transformation to handle this problem, in which we can simultaneously generate knapsack LWE instances and Gaussian LWE instances from standard LWE instances. As a by-product, our transformation can derive a variant of LWE assumption, which says that even if some of the LWE instances have no errors, those instances are still computationally indistinguishable with the uniform distribution.

Finally, based on this KDM secure *PKE* scheme with the uniform randomness and the same all-but-d trapdoor function of [4], we can obtain a KDM secure LWE-based IBE scheme.

SO secure IBE scheme with the multi-user setting. In [13], He *et al.* proved that an IBLE scheme is IND-SO secure. IBLE has two indistinguishable modes, the real mode and the lossy mode. The real mode is same as the normal IBE scheme. In the lossy mode, the key point is to information-theoretically hide the plaintext message for the challenge identities. It requires that the randomness should have enough entropy in the ciphertext to information-theoretically hide the message.

Therefore, to achieve the IND-SO security, by adding a lossy mode for the above KDM secure IBE scheme, we prove that the scheme is also an IBLE. Since the randomness vector is drawn from uniform distribution in our KDM secure scheme, it is easy to achieve the property of information-theoretically hiding the message by a special structure in public key used in [13]. And the problem (2) of the multi-user setting is naturally solved based on our KDM secure IBE scheme. For specifically, the IBE scheme in [13] is selective identity IND-SO-CPA secure for just one target identity. The multi-user setting in our case means that the IBE scheme should be selective identity IND-SO-CPA secure for d target identities where d is polynomial in λ. Alperin-Sheriff and Peikert [4] proposed an all-but-d trapdoor function to handle with the multi-user setting. So we can combine their all-but-d trapdoor function with the IBLE to achieve the selective identity IND-SO-CPA secure IBE scheme for the multi-user setting. We should remark that, the all-but-d trapdoor function is used to construct a trapdoor for d identities rather than the opened ciphertexts in selective opening security.

To sum up, we propose a new transformation and combine it with the all-but-d trapdoor function in [4] with the framework IBLE [13] to construct an LWE-based IBE scheme,

which can achieve KDM security and SO security at the same time.

1.3 Organization

The rest of this paper is organized as follows. In section 2 we introduce some notations, definitions and previous results. In section 3, we describe the new transformation. In section 4, we prove a public key encryption scheme is KDM-CPA secure. Based on the KDM secure public key encryption scheme in section 4 and the all-but-d trapdoor function in [4], we just describe the construction of compact KDM secure IBE scheme in section 5. Because the proof of the security is similar with the proof in [4], we will omit the proof. At the same time, we prove that this IBE scheme is IND-SO secure.

2. PRELIMINARIES

2.1 Notations

Except special notes, all operations in this paper are under the operation of modulo q, and log means $\log_2$. Throughout, we use λ to denote our security parameter. For a positive integer l, we use $[l]$ to denote the set $\{1, ..., l\}$. We use bold lower-case letters (e.g. $\mathbf{s}$) to denote vectors, and bold upper-case letters (e.g. $\mathbf{A}$) to denote matrices. We use $x \xleftarrow{\$} X$ to denote that x is drawn uniformly at random over a set X. We use $x \leftarrow \mathcal{X}$ to denote that x is drawn according to a distribution $\mathcal{X}$. To denote the statistical distance between two distributions, we write $\Delta(\mathcal{X}, \mathcal{Y})$. For two distribution ensembles $\mathcal{X} = \mathcal{X}_\lambda, \mathcal{Y} = \mathcal{Y}_\lambda$, we write $\mathcal{X} \approx_s \mathcal{Y}$ if $\Delta(\mathcal{X}, \mathcal{Y})$ is a negligible function of λ, and we write $\mathcal{X} \approx_c \mathcal{Y}$ if for all probabilistic polynomial time (PPT) distinguishers D there is a negligible function $negl(\cdot)$ such that: $|\Pr[D(\lambda, \mathcal{X}_\lambda) = 1] - \Pr[D(\lambda, \mathcal{Y}_\lambda) = 1]| \leq negl(\lambda)$. We let $\lfloor x \rceil$ be the closest integer to x. For a matrix $\mathbf{X} \in \mathbb{R}^{n \times k}$, the largest singular value of $\mathbf{X}$ is defined as $s_1(\mathbf{X}) = \max_{\|\mathbf{u}\|=1} \|\mathbf{Xu}\|$. Let $\Lambda_{\mathbf{y}}^{\perp}(\mathbf{A}) = \{\mathbf{e} \in \mathbb{Z}^m s.t. \mathbf{Ae} = \mathbf{y} \mod q\}$ given a matrix $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$ and a vector $\mathbf{y} \in \mathbb{Z}_q^n$. We use $|x|$ to denote the absolute value of an element x, and use $\|\mathbf{s}\|$ to denote the L_2 length of the vector $\mathbf{s}$.

2.2 The Discrete Gaussians

For any $s > 0$ and $\mathbf{c} \in \mathbb{R}^n$, define the Gaussian function: $\forall \mathbf{x} \in \mathbb{R}^n, \rho_{s,\mathbf{c}}(\mathbf{x}) = \exp(-\pi \|\mathbf{x}-\mathbf{c}\|^2/s^2)$. For any $\mathbf{c} \in \mathbb{R}^n$, real $s > 0$, and n-dimensional lattice Λ, define the discrete Gaussian distribution over Λ as: $\forall \mathbf{x} \in \Lambda, \mathcal{D}_{\Lambda,s,\mathbf{c}}(\mathbf{x}) = \frac{\rho_{s,\mathbf{c}}(\mathbf{x})}{\rho_{s,\mathbf{c}}(\Lambda)}$, where $\rho_{s,\mathbf{c}}(\Lambda) = \sum_{\mathbf{y} \in \mathbf{\Lambda}} \rho_{s,\mathbf{c}}(\mathbf{y})$. We omit the parameter $\mathbf{c}$ when $\mathbf{c} = \mathbf{0}$.

The following lemma states some properties of the discrete Gaussian distribution. For $\epsilon > 0$, $\eta_\epsilon(\Lambda)$ denotes the smoothing parameter of an n-dimensional lattice introduced by Micciancio and Regev [17] which is a positive real value. The definition of it depends on the notion of dual lattice, we just need some relevant facts but not the precise definition, so for details see [17, 11, 16].

LEMMA 2.1 ([4], LEMMA 2.1). *Let $m \geq Cn \log q$ for some constant $C > 1$.*

1. For any $\omega(\sqrt{\log n})$ function, we have $\eta_\epsilon(\mathbb{Z}^n) \leq \omega(\sqrt{\log n})$ for some negligible $\epsilon(n) = negl(n)$.

2. With all but $negl(n)$ probability over the uniformly random choice of $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, the following holds:

For $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z}^m,r}$ where $r = \omega(\sqrt{\log n})$, the distribution of $\mathbf{y} = \mathbf{Ae} \mod q$ is within $negl(n)$ statistical distance of uniform, and the conditional distribution of $\mathbf{e}$ given $\mathbf{y}$ is $\mathcal{D}_{\Lambda_{\mathbf{y}}^{\perp}(\mathbf{A}),r}$.

3. For any m-dimensional lattice Λ, any $\mathbf{c} \in \mathbb{Z}^m$, any $r \geq \eta_\epsilon(\mathbf{\Lambda})$ where $\epsilon(n) = negl(n)$, and any random variable $\mathbf{x} \leftarrow \mathcal{D}_{\Lambda+\mathbf{c},r}$, we have $\|\mathbf{x}\| \leq r\sqrt{m}$ with all but $negl(n)$ probability. In addition, for $\Lambda = \mathbb{Z}$ and any random variable $x \leftarrow \mathcal{D}_{\mathbb{Z},r}$, we have $|x| \leq r \cdot \omega(\sqrt{\log n})$ except with $negl(n)$ probability.

4. For any $r > 0$, and for $\mathbf{R} \leftarrow \mathcal{D}_{\mathbb{Z},r}^{n \times k}$, we have $s_1(\mathbf{R}) \leq r \cdot O(\sqrt{n} + \sqrt{k})$ except with $negl(n)$ probability.

2.3 Hard Learning Problems

Here we recall the concepts and the hardness of LWE.

Learning with Errors (LWE).

Let $m = m(n)$, $q = q(n)$ be integers, and χ be a distribution on $\mathbb{Z}_q$. Let $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}, \mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n, \mathbf{e} \leftarrow \chi^m$, then the LWE$(m, n, q, \chi)$ problem is to find $\mathbf{s}$, given $(\mathbf{A}, \mathbf{As} + \mathbf{e})$.

When $\mathbf{s} \leftarrow \mathcal{D}_{\mathbb{Z},r}^n$, we call it Gaussian LWE. This is the search version of the LWE problem, and there is a decisional version of the LWE problem.

(Decisional) Learning with Errors (DLWE).

Let $m = m(n)$, $q = q(n)$ be integers, and χ be a distribution on $\mathbb{Z}_q$. Let $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}, \mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n, \mathbf{e} \leftarrow \chi^m$, then the DLWE$(m, n, q, \chi)$ problem is that given $(\mathbf{A}, \mathbf{b})$, decide whether $\mathbf{b}$ is distributed by $\mathbf{As} + \mathbf{e}$ or chosen uniformly at random from $\mathbb{Z}_q^m$.

Evidence for the hardness of LWE$(m, n, q, \mathcal{D}_{\mathbb{Z},\alpha q})$ follows from results of Regev [20], when $m = poly(n)$ and $\alpha q \geq 2\sqrt{n}$. The hardness of decisional version of LWE can be reduced to the search version of LWE.

(Decisional) Knapsack Learning with Errors (KLWE).

Let m, n be positive integers with $m > n$ and let χ be an error distribution on $\mathbb{Z}_q$. Let $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}, \mathbf{e} \leftarrow \chi^m, \mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^n$, then the KLWE$(m, n, q, \chi)$ problem is to distinguish the ditributions $(\mathbf{A}, \mathbf{A}^t\mathbf{e})$ and $(\mathbf{A}, \mathbf{u})$.

(Decisional) Extended Learning with Errors (ELWE).

Let m, n be positive integers with $m > n$ and let χ be an error distribution on $\mathbb{Z}_q$. Let $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m \times n}, \mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n, \mathbf{e} \leftarrow \chi^m, \mathbf{t} \leftarrow \chi^m, \mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^n$, then the ELWE$(m, n, q, \chi)$ problem is to distinguish the distributions $(\mathbf{A}, \mathbf{As} + \mathbf{e}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$ and $(\mathbf{A}, \mathbf{u}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$.

The hardnesses of these problems are as follows.

LEMMA 2.2 ([15] LEMMA 4.8). *For any $n, m \geq n + \omega(\log n)$, q and $\mathcal{D}_{\mathbb{Z},\alpha q}$, there is a polynomial time reduction from the problem of inverting LWE$(m, n, q, \mathcal{D}_{\mathbb{Z},\alpha q})$ with probability ϵ, to the problem of inverting KLWE$(m, m-n, q, \mathcal{D}_{\mathbb{Z},\alpha q}^m)$ (in the search version) with probability $\epsilon' = \epsilon + negl(\lambda)$.*

LEMMA 2.3 ([4] THEOREM 3.1). *There exists a probabilistic polynomial time oracle machine (a simulator) $\mathcal{S}$ such that for any adversary $\mathcal{A}$, $Adv_{\mathsf{LWE}(m,n,q,\chi)}(\mathcal{S}^{\mathcal{A}}) \geq \frac{1}{2p-1} \cdot Adv_{\mathsf{ELWE}(m,n,q,\chi)}(\mathcal{A}) - negl(\lambda)$, where $m \geq n + \omega(\log n)$, $|\langle \mathbf{e}, \mathbf{t} \rangle| < p$ with overwhelming probability, and p is the smallest prime divisor of the modulus q.*

2.4 Some Results about Randomness

We introduce some results about randomness which will be used as tools in the later section.

LEMMA 2.4 ([12]). *Let $\mathcal{D}$ be a distribution over $\mathbb{Z}_q^n$ with min-entropy k. For any $\varepsilon > 0$ and $l \leq (k - 2\log(1/\varepsilon) - O(1))/\log q$, the joint distribution of $(\mathbf{C}, \mathbf{C} \cdot \mathbf{s})$ where $\mathbf{C} \leftarrow \mathbb{Z}_q^{l \times n}$ is uniformly random and $\mathbf{s} \in \mathbb{Z}_q^n$ is drawn from the distribution $\mathcal{D}$ is ε-close to the uniform distribution over $\mathbb{Z}_q^{l \times n} \times \mathbb{Z}_q^l$.*

The following lemma is an extension of lemma A.2. The proof will be showed in appendix which is similar to lemma A.2 of [7].

LEMMA 2.5. *There is a distribution* Lossy *such that* $\bar{\mathbf{A}} \leftarrow \mathsf{Lossy} \approx_c \mathbf{U} \xleftarrow{\$} \mathbb{Z}_q^{2m \times n}$ *and given* $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, *and* $\mathbf{e}^{(1)} \leftarrow \mathcal{D}_{\mathbb{Z},r_1}^{m \times n}, \mathbf{e}^{(2)} \leftarrow \mathcal{D}_{\mathbb{Z},r_2}^{m \times n}$, $\tilde{H}_\infty^\epsilon(\mathbf{s}|\bar{\mathbf{A}}, \bar{\mathbf{A}}\mathbf{s} + [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t]^t) \geq n$, *where $\epsilon = negl(\lambda)$.* Lossy *is as follows.*

- *Choose* $\mathbf{B} \xleftarrow{\$} \mathbb{Z}_q^{m \times k}$, $\mathbf{C} \xleftarrow{\$} \mathbb{Z}_q^{k \times n}$, $\mathbf{Z} \leftarrow \mathcal{D}_{\mathbb{Z},r'}^{m \times n}$, *and* $\mathbf{R} \leftarrow \mathcal{D}_{\mathbb{Z},\omega(\sqrt{\log n})}^{m \times w}$, *where* $\frac{r'}{r_1} = negl(\lambda)$, $\frac{r'}{r_2} = negl(\lambda)$, $k \log q \leq n - 2\lambda + 2$, *and* $n \log q \leq m - 2\lambda + 2$.
- *Let* $\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{B} \\ -\mathbf{R}^t\mathbf{B} \end{bmatrix} \mathbf{C} + \begin{bmatrix} \mathbf{Z} \\ -\mathbf{R}^t\mathbf{Z} \end{bmatrix}$.
- *Output* $\bar{\mathbf{A}}$.

2.5 Key-Dependent Message Security

The key-dependent message security against chosen plaintext attack (KDM-CPA) is that an adversary plays a game with a challenger that answers encryption queries for functions of the users' secret keys (see figure 1). We use $\mathcal{F}$ to denote the class of functions which map l secret keys to a plaintext. The adversary can make KDM queries polynomial times with $f \in \mathcal{F}$ and $j \in \{1, ..., l\}$.

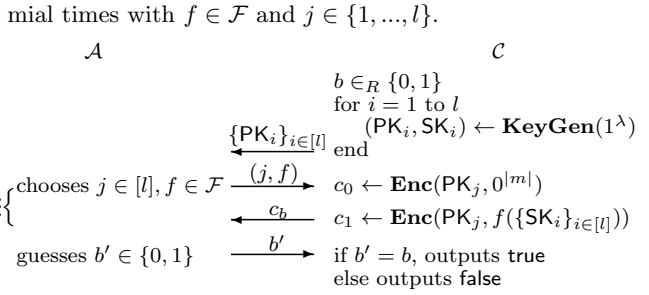

Figure 1: Game of key-dependent message security for PKE

The advantage of the adversary is $Adv_{\mathcal{A},\mathcal{PKE}}^{\mathsf{KDM\text{-}CPA}} = |2\Pr[b' = b] - 1|$. If for every probabilistic polynomial-time adversary $\mathcal{A}$, $Adv_{\mathcal{A},\mathcal{PKE}}^{\mathsf{KDM\text{-}CPA}}$ is a negligible function, then the PKE scheme is KDM-CPA secure with respect to $\mathcal{F}$.

The above definition is the KDM-CPA security for PKE, the definition for an IBE scheme (Setup, Ext, Enc, Dec) is similar. First, the challenger chooses a bit $b \leftarrow \{0,1\}$. Second, the adversary $\mathcal{A}(1^\lambda, d)$ outputs (distinct) target identities $\mathcal{I} = (id_1, ..., id_l)$ for some $l \leq d$ to the challenger. Third, the challenger runs $\mathsf{Setup}(1^\lambda, \mathsf{d})$ to generate the master public key MPK and the master secret key MSK, and sends MPK to the adversary. Then the challenger extracts secret keys for every target identities id_i. Fourth, $\mathcal{A}$ can make polynomial extraction and encryption queries. The order of these queries decided by $\mathcal{A}$. The extraction query means that $\mathcal{A}$ sends any identity $id \notin \mathcal{I}$ to the challenger, then the challenger runs $\mathsf{Ext}(\mathsf{MSK}, \mathsf{id})$ to get the user secret key sk_{id} and sends it to $\mathcal{A}$. The encryption query means that $\mathcal{A}$ sends $(f \in \mathcal{F}, 1 \leq j \leq l)$ to the challenger. If $b = 0$, the challenger computes $c_0 \leftarrow \mathsf{Enc}(id_j, 0)$. If $b = 1$, the challenger computes $c_1 \leftarrow \mathsf{Enc}(id_j, f(\{sk_{id_i}\}_{i \in [l]}))$. Then the challenger sends c_b to $\mathcal{A}$. Finally, $\mathcal{A}$ attempts to guess b and outputs $b' \in \{0,1\}$.

The advantage of the adversary is $Adv_{\mathcal{A},\mathcal{IBE}}^{\mathsf{KDM\text{-}CPA}} = |2\Pr[b' = b] - 1|$. If for every probabilistic polynomial-time adversary $\mathcal{A}$, $Adv_{\mathcal{A},\mathcal{IBE}}^{\mathsf{KDM\text{-}CPA}}$ is a negligible function, then the IBE scheme is KDM-CPA secure with respect to $\mathcal{F}$.

2.6 Selective Opening Security

We use a game to define IND-sID-SO. Let $\mathcal{D}_\mathcal{M}$ be any message sampler.

Init : The adversary outputs a list of target identities $\mathcal{I} = \{\mathbf{id}_1^*, \mathbf{id}_2^*, ..., \mathbf{id}_l^*\}$.

Setup : The challenger runs $\mathbf{Setup}(1^\lambda)$ and keeps the master secret key MSK. The challenger randomly chooses $j \in_R \{0, 1, ..., l-1\}$. The challenger samples n messages $\{\mathbf{m}_0^i\}_{i=1..n}$ from $\mathcal{D}_\mathcal{M}$ and gets n ciphertexts by using algorithm $\mathbf{Enc}(\mathbf{id}_j^*, \mathsf{MPK}, \mathbf{m}_0^i)$, $i = 1..n$. The master public key MPK and the n ciphertexts $\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_n$ are sent to the adversary.

Phase 1 : The adversary issues queries $q_1, ..., q_k$ where the i-th query q_i is a query on $\mathbf{id}_i$. We require that $\mathbf{id}_i \notin \mathcal{I}$. The challenger responds by using algorithm Extract to obtain a private key $\mathsf{SK}_{\mathbf{id}_i}$ for $\mathbf{id}_i$, and sends $\mathsf{SK}_{\mathbf{id}_i}$ to the adversary. All queries may be made adaptively, that is, the adversary may ask q_i with knowledge of the challenger's responses to $q_1, ..., q_{i-1}$.

Open & **Challenge** : Once the adversary decides that Phase 1 is over it specifies a set J and sends it to the challenger. Then the challenger resamples n messages $\{\mathbf{m}_1^i\}_{i=1..n}$ from $\mathcal{D}_\mathcal{M}$ such that $\mathbf{m}_1^{[J]} = \mathbf{m}_0^{[J]}$. The challenger picks a random bit $b \in \{0,1\}$ and sends the adversary the messages $\mathbf{m}_b$ and the randomnesses $\mathbf{r}[J]$ used in ciphertexts $\mathbf{c}[J]$.

Phase 2 : The adversary issues additional adaptive queries $q_{k+1}, ..., q_m$ where q_i is a private-key extraction query on $\mathbf{id}_i$, where $\mathbf{id}_i \notin \mathcal{I}$. The challenger responds the same as in **Phase 1**.

Guess : Finally, the adversary outputs a guess $b' \in \{0,1\}$ and wins if $b' = b$.

The advantage of $\mathcal{A}$ in attacking an IBE scheme $\mathcal{E}$ is $Adv_{\mathcal{A},\mathcal{E},\mathcal{D}_\mathcal{M},n}^{\mathsf{IND\text{-}sID\text{-}SO}}(\lambda) = |2 \cdot \Pr[b = b'] - 1|$. The probability is over the random bits used by the challenger and the adversary.

We say that an IBE system IND-sID-SO secure if for all IND-sID-SO PPT adversaries $\mathcal{A}$ we have that $Adv_{\mathcal{A},\mathcal{IBE},\mathcal{M},n}^{\mathsf{IND\text{-}sID\text{-}SO}}(\lambda)$ is a negligible function.

2.7 Identity-based lossy encryption

An IBLE scheme works in two modes. One is the real mode which is the same as an IBE scheme with standard master key generation algorithm and extraction algorithm. The other is the lossy mode with a lossy master key generation algorithm, and the corresponding extraction algorithm. The two modes share the same encryption and decryption algorithms. Let $\mathcal{I}$ be the list of the lossy identities. Then for identities $\mathbf{id} \notin \mathcal{I}$, encryptions with the lossy master public key $\mathsf{MPK}_{\mathsf{lossy}}$ are committing as the same in the real mode. For $\mathbf{id}_{\mathsf{lossy}} \in \mathcal{I}$, encryptions are not committing.

Formally, the real mode is a tuple of PPT algorithms $\{\mathsf{Setup}_{\mathsf{real}}, \mathsf{Ext}_{\mathsf{real}}, \mathsf{Enc}, \mathsf{Dec}\}$:

- $\mathbf{Setup_{real}}(1^\lambda) \rightarrow (\mathsf{MPK}_{\mathsf{real}}, \mathsf{MSK})$
- $\mathbf{Ext_{real}}(\mathbf{id}, \mathsf{MPK}_{\mathsf{real}}, \mathsf{MSK}) \rightarrow \mathsf{SK}_{\mathbf{id}}$
- $\mathbf{Enc}(\mathbf{id}, \mathsf{MPK}, \mu) \rightarrow \mathsf{C}$
- $\mathbf{Dec}(\mathbf{id}, \mathsf{SK}_{\mathbf{id}}, \mathsf{C})$ outputs either a message μ or $\perp$ in the case of failure.

The lossy mode is a tuple of PPT algorithms $\{\mathsf{Setup}_{\mathsf{lossy}}, \mathsf{Ext}_{\mathsf{lossy}}, \mathsf{Enc}, \mathsf{Dec}\}$:

- $\mathbf{Setup_{lossy}}(1^\lambda, \mathcal{I}) \rightarrow (\mathsf{MPK}_{\mathsf{lossy}}, \mathsf{MSK})$
- $\mathbf{Ext_{lossy}}(\mathbf{id}, \mathsf{MPK}_{\mathsf{lossy}}, \mathsf{MSK})$ outputs either a user secret key $\mathsf{SK}_{\mathbf{id}}$ when $\mathbf{id} \notin \mathcal{I}$ or $\perp$ when $\mathbf{id} \in \mathcal{I}$.
- **Enc** and **Dec** algorithms are the same as those in the real mode.

An Identity-based Lossy Encryption Scheme should have the properties as below.

(1) *Correctness on keys for all* $\boldsymbol{id} \notin \mathcal{I}$. For any $(\mathsf{MPK}, \mathsf{MSK})$ generated by $\mathbf{Setup_{real}}(1^k)$ or $\mathbf{Setup_{lossy}}(1^\lambda, \mathcal{I})$, any $\mathsf{SK}_{\mathbf{id}}$ generated by $\mathbf{Ext_{real/lossy}}(\mathbf{id}, \mathsf{MPK}, \mathsf{MSK})$, and any message μ, $\mathbf{Dec}(\mathbf{id}, \mathsf{SK}_{\mathbf{id}}, \mathbf{Enc}(\mathbf{id}, \mathsf{MPK}, \mu)) = \mu$.

(2) *Lossiness of encryption with lossy keys for* $\boldsymbol{id} \in \mathcal{I}$. For any lossy keys $\mathsf{MPK}_{\mathsf{lossy}}$ generated by $\mathbf{Setup_{lossy}}(1^\lambda, \mathcal{I})$ and any two messages $\mu_0 \neq \mu_1$, there is $\mathbf{Enc}(\mathbf{id}, \mathsf{MPK}_{\mathsf{lossy}}, \mu_0) \approx_s \mathbf{Enc}(\mathbf{id}, \mathsf{MPK}_{\mathsf{lossy}}, \mu_1)$.

(3) *Indistinguishability between real keys and lossy keys.* We use a game to describe this property.

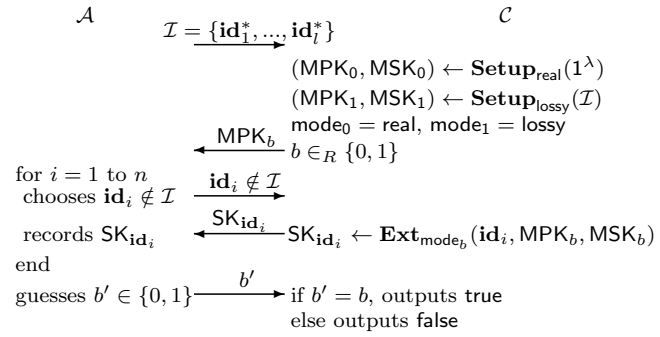

Figure 2: Game of indistinguishability between real keys and lossy keys

The advantage of the adversary is $Adv^{\mathsf{lossy\text{-}keys}}_{\mathcal{A},\mathcal{IBLE}} = |2Pr[b' = b] - 1|$. If for all PPT adversaries $\mathcal{A}$ we have that $Adv^{\mathsf{lossy\text{-}keys}}_{\mathcal{A},\mathcal{IBLE}}$ is a negligible function, then we say that the real keys generated in the real mode is indistinguishable with the lossy keys generated in the lossy mode.

LEMMA 2.6 ([13]). *Let λ be a security parameter. If $\mathcal{IBLE}$ is any identity-based lossy encryption scheme, then for all* IND-sID-SO *PPT adversaries $\mathcal{A}$, $Adv^{\mathsf{IND\text{-}sID\text{-}SO}}_{\mathcal{A},\mathcal{IBLE}}(\lambda)$ is a negligible function.*

3. TRANSFORMATION

Before the discussion of the KDM security, we will describe a transformation which is a useful tool in our security proof. It is based on two previous transformations. Theorem 1 of our transformation shows that from the same standard LWE instances $(\mathbf{A}'^{m\times(m-n)}, \mathbf{A}'\mathbf{s}' + \mathbf{t}^{m\times 1})$, we can get the knapsack LWE instances $(\mathbf{A}^{m\times n}, \mathbf{A}^t\mathbf{t})$ and polynomial many fresh Gaussian LWE instances $(\bar{\mathbf{a}}^{m\times 1}, \bar{\mathbf{a}}^t\mathbf{t} + e')$ involved with the same $\mathbf{t}$. Micciancio and Mol [15] proposed a transformation from standard LWE instances $(\mathbf{A}'^{m\times(m-n)}, \mathbf{A}'\mathbf{s}' + \mathbf{t}^{m\times 1})$ to knapsack LWE instances $(\mathbf{A}^{m\times n}, \mathbf{A}^t\mathbf{t})$. Applebaum et al. [5] proposed a transformation from standard LWE instances $(\mathbf{A}_2'^{m\times m}, \mathbf{A}_2'\mathbf{s}_2' + \mathbf{t}_2^{m\times 1})$ to Gaussian LWE instances $(\bar{\mathbf{a}}_2^{m\times 1}, \bar{\mathbf{a}}_2^t\mathbf{t}_2 + e)$. However, we can not get the knapsack LWE instances and the Gaussian LWE instances which share the same $\mathbf{t}$ by using the two transformations straightly. The reason is that in the transformation of Applebaum et al. [5], the m-dimensional Gaussian LWE is generated with the help of the *invertible square matrix* version of standard LWE instances $(\mathbf{A}_2'^{m\times m}, \mathbf{A}_2'\mathbf{s}_2' + \mathbf{t}_2^{m\times 1})$. But we should get the Gaussian LWE instances with the help of the standard LWE instances $(\mathbf{A}'^{m\times(m-n)}, \mathbf{A}'\mathbf{s}' + \mathbf{t}^{m\times 1})$ where $\mathbf{A}'$ is not an invertible square matrix. So we need to expand the standard LWE instances $(\mathbf{A}'^{m\times(m-n)}, \mathbf{A}'\mathbf{s}' + \mathbf{t}^{m\times 1})$ to the instances of an invertible square matrix, and keep $\mathbf{t}$ unchanged meanwhile.

THEOREM 1. *Let $q = p^e$ be a prime power. For any $n, m \geq n + \omega(\log n)$, $\chi = \mathcal{D}_{\mathbb{Z},r}$,*

(1) There is a polynomial-time transformation that, for arbitrary $\mathbf{s}' \in \mathbb{Z}_q^n$, maps $(\mathbf{A}', \mathbf{A}'\mathbf{s}' + \mathbf{t})$ which follows from $\mathsf{LWE}(m, m-n, q, \chi)$ to $(\mathbf{A}, \mathbf{A}^t\mathbf{t})$ which follows from knapsack $\mathsf{LWE}(m, n, q, \chi)$*, and maps $(\mathbf{A}', \mathbf{u}')$ which follows from $\mathcal{U}(\mathbb{Z}_q^{m\times(m-n)} \times \mathbb{Z}_q^m)$ to $(\mathbf{A}, \mathbf{u})$ which follows from $\mathcal{U}(\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^n)$.*

(2) With the help of $(\mathbf{A}', \mathbf{A}'\mathbf{s}' + \mathbf{t})$, there is a polynomial-time transformation that, maps $(\bar{\mathbf{a}}', \bar{\mathbf{a}}'^t\mathbf{s}' + e')$ which follows from $\mathsf{LWE}(1, m-n, q, \chi)$ to $(\bar{\mathbf{a}}, \bar{\mathbf{a}}^t\mathbf{t}+e') \leftarrow$ Gaussian $\mathsf{LWE}(1, m, q, \chi)$, and maps $(\bar{\mathbf{a}}', \bar{u}')$ which follows from $\mathcal{U}(\mathbb{Z}_q^{m-n} \times \mathbb{Z}_q)$ to $(\bar{\mathbf{a}}, \bar{u})$ which follows from $\mathcal{U}(\mathbb{Z}_q^m \times \mathbb{Z}_q)$.

PROOF. Draw m samples $\{(\mathbf{a}_i', b_i')\}_{i\in[m]}$ from some distribution $\mathcal{D}$ over $\mathbb{Z}_q^{(m-n)} \times \mathbb{Z}_q$, where $\mathcal{D}$ is either $\mathsf{LWE}(1, m-n, q, \chi)$ or $\mathcal{U}(\mathbb{Z}_q^{m-n} \times \mathbb{Z}_q)$. We use $(\mathbf{A}', \mathbf{b}')$ to denote these m samples, i.e.$(\mathbf{A}', \mathbf{b}') = ([\mathbf{a}_1', \mathbf{a}_2', ..., \mathbf{a}_m']^t, [b_1', b_2', ..., b_m']^t) \in \mathbb{Z}_q^{m\times(m-n)} \times \mathbb{Z}_q^m$.

(1) This can be obtained by lemma 2.2 which is proved by Micciancio and Mol [15]. We will simply state the method used in lemma 2.2. Because $\mathbf{A}' \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(m-n)})$, $\mathbf{A}'$ is nonsingular except with probability at most $1/p^{m-(m-n)-1} = 1/p^{n-1} = negl(n)$, where p is the smallest prime factor of q. A matrix $\mathbf{A} \in \mathbb{Z}_q^{m\times n}$ can be efficiently computed using linear algebra from $\mathbf{A}'$ such that $\mathbf{A}^t\mathbf{A}' = \mathbf{0} \mod q$. $\mathbf{A}^t$ can further be randomized by left-multiplying a random unimodualar matrix. If $\mathbf{A}'$ is chosen at random among all nonsingular matrices, then the distribution of the randomized $\mathbf{A}$ is within negligible statistical distance from $\mathcal{U}(\mathbb{Z}_q^{m\times n})$. Then we can get $(\mathbf{A}, \mathbf{A}^t\mathbf{b}')$.

When $\mathcal{D}$ is $\mathsf{LWE}(1, m-n, q, \chi)$, i.e. $(\mathbf{A}', \mathbf{b}')$ is $(\mathbf{A}', \mathbf{A}'\mathbf{s}' + \mathbf{t})$, we can get knapsack $\mathsf{LWE}(m, n, q, \chi)$ instances $(\mathbf{A}, \mathbf{A}^t\mathbf{t})$.

When $\mathcal{D}$ is $\mathcal{U}(\mathbb{Z}_q^{m-n} \times \mathbb{Z}_q)$, i.e. $(\mathbf{A}', \mathbf{b}')$ is $(\mathbf{A}', \mathbf{u}')$. Then we can get $(\mathbf{A}, \mathbf{A}^t\mathbf{u}')$. Because $m \geq n + \omega(\log n)$ and the min-entropy of $\mathbf{u}'$'s distribution is $m \log q$, by lemma 2.4, the distribution of $(\mathbf{A}, \mathbf{A}^t\mathbf{u}')$ is statistically close to the uniform distribution over $\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^n$.

(2) Because $\mathbf{A}' \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(m-n)})$, $\mathbf{A}'$ is nonsingular except with probability at most $1/p^{m-(m-n)-1} = 1/p^{n-1} = negl(n)$, where p is the smallest prime factor of q. Then choose another matrix $\mathbf{A}'' \in \mathbb{Z}_q^{m\times n}$ such that $[\mathbf{A}', \mathbf{A}''] \in$

$\mathbb{Z}_q^{m\times m}$ is invertible. Sample an instance $(\bar{\mathbf{a}}', \bar{b}')$ from $\mathcal{D}$. Choose $\bar{\mathbf{a}}'' \xleftarrow{\$} \mathbb{Z}_q^n, \mathbf{s}'' \xleftarrow{\$} \mathbb{Z}_q^n$. Let $\bar{\mathbf{a}}^t = -[\bar{\mathbf{a}}'^t, \bar{\mathbf{a}}''^t][\mathbf{A}', \mathbf{A}'']^{-1} \in \mathbb{Z}_q^m$ and $\bar{b} = \bar{\mathbf{a}}^t(\mathbf{b}' + \mathbf{A}''\mathbf{s}'') + \bar{b}' + \bar{\mathbf{a}}''^t\mathbf{s}''$. Because $[\bar{\mathbf{a}}'^t, \bar{\mathbf{a}}''^t]$ is uniform and $[\mathbf{A}', \mathbf{A}'']$ is invertible modulo q, $\bar{\mathbf{a}} \in \mathbb{Z}_q^m$ is uniform as well.

When $\mathcal{D}$ is $\mathsf{LWE}(1, m-n, q, \chi)$, i.e. $(\mathbf{A}', \mathbf{b}')$ is $(\mathbf{A}', \mathbf{A}'\mathbf{s}' + \mathbf{t})$, and $(\bar{\mathbf{a}}', \bar{b}')$ is $(\bar{\mathbf{a}}', \bar{\mathbf{a}}'^t\mathbf{s}' + e')$. Then

$$\begin{aligned}\bar{b} &= \bar{\mathbf{a}}^t(\mathbf{b}' + \mathbf{A}''\mathbf{s}'') + \bar{b}' + \bar{\mathbf{a}}''^t\mathbf{s}'' \\ &= \bar{\mathbf{a}}^t(\mathbf{A}'\mathbf{s}' + \mathbf{t} + \mathbf{A}''\mathbf{s}'') + \bar{\mathbf{a}}'^t\mathbf{s}' + e' + \bar{\mathbf{a}}''^t\mathbf{s}'' \\ &= -[\bar{\mathbf{a}}'^t, \bar{\mathbf{a}}''^t][\mathbf{A}', \mathbf{A}'']^{-1}[\mathbf{A}', \mathbf{A}''] \begin{bmatrix}\mathbf{s}' \\ \mathbf{s}''\end{bmatrix} + \bar{\mathbf{a}}^t\mathbf{t} + \\ &\quad [\bar{\mathbf{a}}'^t, \bar{\mathbf{a}}''^t] \begin{bmatrix}\mathbf{s}' \\ \mathbf{s}''\end{bmatrix} + e' \\ &= \bar{\mathbf{a}}^t\mathbf{t} + e'.\end{aligned}$$

Therefore, $(\bar{\mathbf{a}}, \bar{b})$ is distributed according to gaussian $\mathsf{LWE}(1, m, q, \chi)$.

When $\mathcal{D}$ is $\mathcal{U}(\mathbb{Z}_q^{m-n} \times \mathbb{Z}_q)$, i.e. $(\mathbf{A}', \mathbf{b}')$ is $(\mathbf{A}', \mathbf{u}')$, and $(\bar{\mathbf{a}}', \bar{b}')$ is $(\bar{\mathbf{a}}', \bar{u}')$. Then $\bar{b} = \bar{\mathbf{a}}^t(\mathbf{b}' + \mathbf{A}''\mathbf{s}'') + \bar{b}' + \bar{\mathbf{a}}''^t\mathbf{s}'' = \bar{\mathbf{a}}^t(\mathbf{u}' + \mathbf{A}''\mathbf{s}'') + \bar{u}' + \bar{\mathbf{a}}''^t\mathbf{s}''$. Because $\bar{u}'$ is uniform and independent with $\bar{\mathbf{a}}^t(\mathbf{u}' + \mathbf{A}''\mathbf{s}'') + \bar{\mathbf{a}}''^t\mathbf{s}''$, $\bar{b}$ is uniform too. Therefore, $(\bar{\mathbf{a}}, \bar{b})$ is distributed according to $\mathcal{U}(\mathbb{Z}_q^m \times \mathbb{Z}_q)$. □

Obviously, this theorem inducts a variant of DLWE assumption that $(\mathbf{A}, \mathbf{At}, \mathbf{B}, \mathbf{Bt} + \mathbf{e})$ is computationally indistinguishable with $(\mathbf{A}, \mathbf{u}_1, \mathbf{B}, \mathbf{u}_2)$ based on the DLWE assumption, where $\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{n\times m}$, $\mathbf{B} \xleftarrow{\$} \mathbb{Z}_q^{l\times m}$, $\mathbf{t} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m$, $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^l$, $\mathbf{u}_1 \xleftarrow{\$} \mathbb{Z}_q^n, \mathbf{u}_2 \xleftarrow{\$} \mathbb{Z}_q^l$, and $r \geq 2\sqrt{n}, q, m, n$ satisfy the parameters in theorem 1. It means that the Gaussian LWE instances even if some of which has no errors, they are still computationally indistinguishable with the uniform.

4. KDM-CPA SECURE PKE SCHEME

In this section, we will prove that the dual Regev type PKE scheme like [11] proposed by Gentry, Peikert and Vaikuntanathan is KDM-CPA secure for affine functions of the secret keys.

4.1 Construction

The structure of this PKE scheme is the same as the scheme of [11]. The message space $\mathcal{M}$ is $\mathbb{Z}_p$, and the concrete construction is as follows.

$\mathbf{KeyGen}(1^\lambda)$	$\mathbf{Enc}(\mathsf{PK} = (\mathbf{A}, \mathbf{A}^t\mathbf{t}), \mu)$	$\mathbf{Dec}(\mathsf{SK} = \mathbf{t}, \mathsf{C})$
$\mathbf{A} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$	$\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$	$(\mathbf{c}_1, c_2) = \mathsf{C}$
$\mathbf{t} \longleftarrow \mathcal{D}_{\mathbb{Z},r}^m$	$\mathbf{e} \longleftarrow \mathcal{D}_{\mathbb{Z},r}^m, e' \longleftarrow \mathcal{D}_{\mathbb{Z},r}$	$\mu' = c_2 - \mathbf{t}^t\mathbf{c}_1$
$\mathsf{PK} = (\mathbf{A}, \mathbf{A}^t\mathbf{t})$	$\mathbf{c_1} = \mathbf{As} + \mathbf{e}$	$\mu = decode(\mu')$
$\mathsf{SK} = \mathbf{t}$	$c_2 = \mathbf{t}^t\mathbf{As} + e' + \mu \cdot p$	return μ
return $(\mathsf{PK}, \mathsf{SK})$	return $(\mathbf{c}_1, c_2)$	

Figure 3: Construction of the PKE scheme

$decode(\mu')$ means that for μ', outputs the number in $\mathbb{Z}_p$ which is closer to $(\mu' \pm \lfloor \frac{p}{2} \rfloor)$ *mod* q.

Parameters. Let $n = poly(\lambda)$, $m \geq n + \omega(\log n)$, $q = p^2$, $p > 2(r\omega(\sqrt{\log n})$
$+ r^2\sqrt{m} \cdot \omega(\sqrt{\log m}))$, $r \geq 2\sqrt{n}$.

Correctness. For all $(\mathsf{PK}, \mathsf{SK})$ generated by $\mathbf{KeyGen}(1^\lambda)$ and all message μ, $(\mathbf{c}_1, c_2) \leftarrow \mathsf{Enc}(\mathbf{A}, \mathbf{A}^t\mathbf{t}, \mu)$, $c_2 - \mathbf{t}^t\mathbf{c}_1 = \mathbf{t}^t\mathbf{As} + e' + \mu \cdot p - \mathbf{t}^t\mathbf{As} - \mathbf{t}^t\mathbf{e} = \mu \cdot p + e' - \mathbf{t}^t\mathbf{e}$. By Cauchy-Schwarz and lemma 2.1, $|e' - \mathbf{t}^t\mathbf{e}| \leq r\omega(\sqrt{\log n}) + r^2\sqrt{m} \cdot \omega(\sqrt{\log m}) < \frac{p}{2}$. Therefore, the algorithm *decode*() will get the correct message with overwhelming probability.

4.2 KDM-CPA Security.

We will prove that the scheme mentioned above is KDM-CPA secure.

THEOREM 2. *Parameters are the same in section 4.1. $l \geq 1$ denotes the number of users in the definition of* KDM-CPA *security. And the adversary $\mathcal{A}$ makes at most $d = poly(n)$ times KDM queries. Then the scheme in section 4.1 is* KDM-CPA *secure for affine functions, assuming that the* extended LWE *problem and the* LWE *problem is hard. Specifically,* $Adv_{\mathcal{A},\text{instantiation}}^{\mathsf{KDM\text{-}CPA}} \leq ld \cdot Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})} +$
$Adv_{\mathcal{S}_2}^{\mathsf{LWE}(d+lm,m-n,q,\mathcal{D}_{\mathbb{Z},r})} + negl(\lambda)$.

PROOF. We use f to denote affine functions over $\mathbb{Z}_p$ of the l users' secret keys $\mathbf{t}_1, \mathbf{t}_2, ..., \mathbf{t}_l$. f is defined as $f_{\mathbf{K},w}(\{\mathbf{t}_i\}_{i\in[l]}) = \sum_{i\in[l]} < \mathbf{k}_i, \mathbf{t}_i > + w \mod p$, where $\mathbf{K} = [\mathbf{k}_1, ..., \mathbf{k}_l]$. For simplicity, we will omit $\mathbf{K}$ and w, and use f to denote the affine function.

We will use a game sequence to prove this theorem. We start with the game $\mathbf{G}_0$, which is the real game of the definition of KDM-CPA security in section 2.5. We generate the public keys and the corresponding secret keys.

Then, we move to the game $\mathbf{G}_1$. In this game, instead of using public keys to generate ciphertexts, we use the secret key and uniformly randomly chosen vector $\mathbf{u}$ to generate ciphertexts. By the extended learning with errors (ELWE) assumption and the hybrid technique, we can obtain that this game is computationally indistinguishable from the game $\mathbf{G}_0$.

Next is the game $\mathbf{G}_2$, in which, we do not choose public keys and secret keys, but generate public keys and the challenge ciphertexts from LWE instances using the transformation (theorem 3) without knowing secret keys. This transformation guarantees that the distributions of public keys and challenge ciphertexts are identical as the distributions in game $\mathbf{G}_1$.

In the final game $\mathbf{G}_3$, public keys and challenge ciphertexts are all uniformly random vectors generated from uniform instances by the transformation (theorem 3). Then based on the LWE assumption, we can get that this game is computationally indistinguishable from the game $\mathbf{G}_2$.

Since the challenge ciphertext in $\mathbf{G}_3$ is uniformly random, there is no information about the message.

Next we will describe every game and the indistinguishability between them in detail.

$\boxed{\mathbf{G}_0:}$ This is the original game from the definition of KDM-CPA security in section 2.5. At first, we randomly choose a bit $b \xleftarrow{R} \{0, 1\}$. Next, we choose l users' public key $\{(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)\}_{i\in[l]}$ and secret keys $\{\mathbf{t}_i\}_{i\in[l]}$ like the KeyGen algorithm, and send the public keys to the adversary $\mathcal{A}$. Then the adversary makes the KDM query of the form (j_z, f_z) at most $d = poly(\lambda)$ times, where $1 \leq z \leq d$ and $1 \leq j_z \leq l$. The adversary can make this query at most $d = poly(\lambda)$ times. The query (j_z, f_z) means that it is the z-th query and the adversary requires that the challenger uses j_z-th public key to encrypt "$\mathbf{0}$" if $b = 0$, and $f_z(\{\mathbf{t}_i\}_{i\in[l]})$ if $b = 1$. More concretely, the challenge ciphertext sent to the adversary for the KDM query (j_z, f_z) is $C_0 = (\mathbf{A}_{j_z}\mathbf{s} + \mathbf{e}, \mathbf{t}_{j_z}^t\mathbf{A}_{j_z}\mathbf{s} + e')$

if $b = 0$, or $C_1 = (\mathbf{A}_{j_z}\mathbf{s} + \mathbf{e}, \mathbf{t}_{j_z}^t\mathbf{A}_{j_z}\mathbf{s} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$, where $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m, e' \leftarrow \mathcal{D}_{\mathbb{Z},r}$ chosen by us. Then the adversary guesses the value of b. Figure 6 in section B shows the process.

$\boxed{\mathbf{G}_1:}$ In this game, we generate the public keys $\{(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)\}_{i\in[l]}$ and secret keys $\{\mathbf{t}_i\}_{i\in[l]}$ as in $\mathbf{G}_0$. The difference of this game from $\mathbf{G}_0$ is the method of generating challenge ciphertexts. For every KDM query (j_z, f_z) from the adversary, we choose $\mathbf{u} \xleftarrow{\$} \mathbb{Z}_q^m$, $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m$, $e' \leftarrow \mathcal{D}_{\mathbb{Z},r}$, and generate the challenge ciphertext is, $C_0 = (\mathbf{u}, \mathbf{t}_{j_z}^t\mathbf{u} - \mathbf{t}_{j_z}^t\mathbf{e} + e')$ if $b = 0$, or $C_1 = (\mathbf{u}, \mathbf{t}_{j_z}^t\mathbf{u} - \mathbf{t}_{j_z}^t\mathbf{e} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$.

Then, we show that the above two games $\mathbf{G}_0$ and $\mathbf{G}_1$ are computationally indistinguishable based on the extended LWE assumption.

CLAIM 1. $|\Pr[\mathbf{G}_0(\mathcal{A}) = 1] - \Pr[\mathbf{G}_1(\mathcal{A}) = 1]| \leq ld \cdot Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})}$.

PROOF. In the definition of KDM-CPA security, the adversary in a game can make at most d times KDM query (j_z, f_z), so it is reasonable that the adversary queries the same user's public key many times, i.e. $j_i = j_k$ for the i-th query and the k-th query, where $1 \leq i \leq d, 1 \leq k \leq d$ and $i \neq k$. To show this situation clearly, we will use a sequence of hybrid games.

$\mathbf{H}_0$: The same as $\mathbf{G}_0$.

$\mathbf{H}_k$: When the challenger receives the first k queries $(j_z, f_z), z \in \{1, ..., k\}$ from the adversary, he uses the method of encryption in $\mathbf{G}_1$ to generate the challenge ciphertext $C_0 = (\mathbf{u}, \mathbf{t}_{j_z}^t\mathbf{u} - \mathbf{t}_{j_z}^t\mathbf{e} + e')$ if $b = 0$, or $C_1 = (\mathbf{u}, \mathbf{t}_{j_z}^t\mathbf{u} - \mathbf{t}_{j_z}^t\mathbf{e} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$.

When the challenger receives the rest $(d-k)$ queries (j_z, f_z), $z \in \{k+1, ..., d\}$ from the adversary, he uses the method of encryption in $\mathbf{G}_0$ to generate the challenge ciphertext $C_0 = (\mathbf{A}_{j_z}\mathbf{s} + \mathbf{e}, \mathbf{t}_{j_z}^t\mathbf{A}_{j_z}\mathbf{s} + e')$ if $b = 0$, or $C_1 = (\mathbf{A}_{j_z}\mathbf{s} + \mathbf{e}, \mathbf{t}_{j_z}^t\mathbf{A}_{j_z}\mathbf{s} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$.

$\mathbf{H}_d$: The same as $\mathbf{G}_1$.

Suppose there is a PPT adversary $\mathcal{A}$ has non-negligible advantage in distinguishing $\mathbf{H}_{k-1}$ and $\mathbf{H}_k$, for any $k \in \{1, ..., d\}$. Then we use $\mathcal{A}$ to construct an algorithm $\mathcal{S}_1$ as figure 7 in section B to distinguish $\mathsf{ELWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instance with the uniform instance. The oracle $\mathcal{O}_{\mathsf{ELWE}}$ uniform randomly outputs the ELWE instances $(\mathbf{A}, \mathbf{As} + \mathbf{e}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$ or the uniform instances $(\mathbf{A}, \mathbf{u}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$.

At first, $\mathcal{S}_1$ randomly chooses a bit $b \xleftarrow{R} \{0, 1\}$. Then, after receiving an instance $(\mathbf{A}, \mathbf{b}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$ sent by the oracle $\mathcal{O}_{\mathsf{ELWE}}$, $\mathcal{S}_1$ chooses $j^* \xleftarrow{\$} \{1, ..., l\}$ uniformly at random. Next, $(\mathbf{A}, \mathbf{A}^t\mathbf{t})$ is the public key and $\mathbf{t}$ is the secret key for the j^*-th user. And for other users, $\mathcal{S}_1$ chooses the public key $(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)$ and secret key $\mathbf{t}_i$ as the KeyGen algorithm where $i \neq j^*$.

$\mathcal{S}_1$ sends the l public keys to the adversary $\mathcal{A}$ and invoke $\mathcal{A}$. Then $\mathcal{S}_1$ will receive KDM query and sends back the challenge ciphertext. This process can be looped at most d times. For the first $k-1$ KDM queries from the adversary $\mathcal{A}$, $\mathcal{S}_1$ generates the challenge ciphertext as the game $\mathbf{G}_1$. For the k-th KDM query (j_k, f_k), if $j_k \neq j^*$, then $\mathcal{S}_1$ outputs 0 or 1 randomly, and ends this algorithm. If else, i.e. $j_k = j^*$, then $\mathcal{S}_1$ chooses $e' \leftarrow \mathcal{D}_{\mathbb{Z},r}$ and uses the instance sent by $\mathcal{O}_{\mathsf{ELWE}}$ to generate the challenge ciphertext as follows, $C_0 = (\mathbf{b}, \mathbf{t}^t\mathbf{b} - \mathbf{t}^t\mathbf{e} + e')$ if $b = 0$, or $C_1 = (\mathbf{b}, \mathbf{t}^t\mathbf{b} - \mathbf{t}^t\mathbf{e} + e' + f_k(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$. And for the rest $d-k$ KDM queries, $\mathcal{S}_1$ simulates as the game $\mathbf{G}_0$.

Specifically, let's analyze the case of $j_k = j^*$. When the oracle $\mathcal{O}_{\mathsf{ELWE}}$ outputs the $\mathsf{ELWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instance $(\mathbf{A}, \mathbf{b} = \mathbf{As} + \mathbf{e}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$, the j^*-th public key is $(\mathbf{A}, \mathbf{A}^t\mathbf{t})$, and the challenge ciphertext is $C_0 = (\mathbf{As} + \mathbf{e}, \mathbf{t}^t(\mathbf{As} + \mathbf{e}) - \mathbf{t}^t\mathbf{e} + e') = (\mathbf{As} + \mathbf{e}, \mathbf{t}^t\mathbf{As} + e')$ if $b = 0$, or $C_1 = (\mathbf{As} + \mathbf{e}, \mathbf{t}^t\mathbf{As} + e' + f_k(\{\mathbf{t}_i\}_{i\in[l]})p)$if $b = 1$, which is the same as the game $\mathbf{G}_0$. So in the case that $j_k = j^*$ and $\mathcal{O}_{\mathsf{ELWE}}$ outputs $\mathsf{ELWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instance, $\mathcal{S}_1$ simulates the game $\mathbf{H}_{k-1}$.

When the oracle $\mathcal{O}_{\mathsf{ELWE}}$ outputs uniform random instances $(\mathbf{A}, \mathbf{b} = \mathbf{u}, \mathbf{t}, \mathbf{t}^t\mathbf{e})$, the j^*-th public key is $(\mathbf{A}, \mathbf{A}^t\mathbf{t})$, and the challenge ciphertext is $C_0 = (\mathbf{u}, \mathbf{t}^t\mathbf{u} - \mathbf{t}^t\mathbf{e} + e')$ if $b = 0$, or $C_1 = (\mathbf{u}, \mathbf{t}^t\mathbf{u} - \mathbf{t}^t\mathbf{e} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$ if $b = 1$, which is the same as the game $\mathbf{G}_1$. So in the case that $j_k = j^*$ and $\mathcal{O}_{\mathsf{ELWE}}$ outputs a uniform random instance, $\mathcal{S}_1$ simulates the game $\mathbf{H}_k$.

Since the probability of $j_k = j^*$ is $\frac{1}{l}$, so $\frac{1}{l}|\Pr[\mathbf{H}_{k-1}(\mathcal{A}) = 1] - \Pr[\mathbf{H}_k(\mathcal{A}) = 1]| = Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})}$. $\mathbf{H}$ is a hybrid sequence, consequently, $\frac{1}{l}|\Pr[\mathbf{H}_0(\mathcal{A}) = 1] - \Pr[\mathbf{H}_d(\mathcal{A}) = 1]| \leq d \cdot Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})}$. Since $\mathbf{H}_0$ is $\mathbf{G}_0$ and $\mathbf{H}_d$ is $\mathbf{G}_1$, $|\Pr[\mathbf{G}_0(\mathcal{A}) = 1] - \Pr[\mathbf{G}_1(\mathcal{A}) = 1]| \leq ld \cdot Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})}$ holds. □

$\boxed{\mathbf{G}_2:}$ In this game, we use the transformation from LWE instances to $\mathsf{knapsack\ LWE}$ instances by theorem 1(1) to generate public keys without knowing secret keys. Specifically, we can generate $\mathsf{KLWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instances $\{(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)\}_{i\in[l]}$, which are used as public keys, from l $\mathsf{LWE}(m, m-n, q, \mathcal{D}_{\mathbb{Z},r})$ instances $\{(\mathbf{A}'_i, \mathbf{A}'_i\mathbf{s}' + \mathbf{t}_i)\}_{i\in[l]}$.

It should be noted that we do not know the secret keys $\{\mathbf{t}_i\}_{i\in[l]}$. But we can construct the relation of the secret keys for all users by the transformation. In detail, by theorem 1(1), the public key of i-th user is the knapsack LWE instance $(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)$ generated by the LWE instance $(\mathbf{A}'_i, \mathbf{b}_i = \mathbf{A}'_i\mathbf{s}' + \mathbf{t}_i)$, and the public key of j-th user is the knapsack LWE instance $(\mathbf{A}_j, \mathbf{A}_j^t\mathbf{t}_j)$ generated by the LWE instance $(\mathbf{A}'_j, \mathbf{b}_j = \mathbf{A}'_j\mathbf{s}' + \mathbf{t}_j)$ for any $i, j \in [l]$ and $i \neq j$. Because $\mathbf{A}'_i, \mathbf{A}'_j \leftarrow \mathcal{U}(\mathbb{Z}_q^{m\times(m-n)})$, $\mathbf{A}'$ is nonsingular except with probability at most $1/p^{m-(m-n)-1} = 1/p^{n-1} = negl(\lambda)$, we can choose another two matrices $\mathbf{C}'_i, \mathbf{C}'_j \in \mathbb{Z}_q^{m\times n}$ such that $[\mathbf{A}'_i, \mathbf{C}'_i]$ and $[\mathbf{A}'_j, \mathbf{C}'_j]$ are nonsingular. Then we can get a matrix $\mathbf{V}_{ji} \in \mathbb{Z}_q^{m\times m}$ such that $\mathbf{V}_{ji}[\mathbf{A}'_j, \mathbf{C}'_j] = [\mathbf{A}'_i, \mathbf{C}'_i]$ using linear algebra. That is, $\mathbf{V}_{ji}\mathbf{A}'_j = \mathbf{A}'_i$. Therefore, $\mathbf{V}_{ji}\mathbf{b}_j - \mathbf{b}_i = \mathbf{V}_{ji}\mathbf{A}'_j\mathbf{s}' + \mathbf{V}_{ji}\mathbf{t}_j - \mathbf{A}'_i\mathbf{s}' - \mathbf{t}_i = \mathbf{A}'_i\mathbf{s}' + \mathbf{V}_{ji}\mathbf{t}_j - \mathbf{A}'_i\mathbf{s}' - \mathbf{t}_i = \mathbf{V}_{ji}\mathbf{t}_j - \mathbf{t}_i$.

Then,we can get $\mathbf{t}_i = \mathbf{b}_i - \mathbf{V}_{ji}\mathbf{b}_j + \mathbf{V}_{ji}\mathbf{t}_j$. We define $\mathbf{V}_{jj} = \mathbf{I}$. Then the affine function of secret keys can be written as an affine function of one secret key, as follows

$$
\begin{aligned}
&f(\{\mathbf{t}_i\}_{i\in[l]}) = \sum_{i\in[l]} < \mathbf{k}_i, \mathbf{t}_i > + w \mod p \\
&= \sum_{i\in[l]} < \mathbf{k}_i, \mathbf{b}_i - \mathbf{V}_{ji}\mathbf{b}_j + \mathbf{V}_{ji}\mathbf{t}_j > + w \mod p \\
&= \sum_{i\in[l]} (< \mathbf{V}_{ji}^t\mathbf{k}_i, \mathbf{t}_j > + < \mathbf{k}_i, \mathbf{b}_i - \mathbf{V}_{ji}\mathbf{b}_j >) + w \mod p \\
&= < \tilde{\mathbf{k}}_j, \mathbf{t}_j > + \tilde{w}_j \mod p,
\end{aligned}
$$

where $\tilde{\mathbf{k}}_j = \sum_{i\in[l]} (\mathbf{V}_{ji}^t\mathbf{k}_i) \mod p$ and $\tilde{w}_j = \sum_{i\in[l]} (< \mathbf{k}_i, \mathbf{b}_i - \mathbf{V}_{ji}\mathbf{b}_j >) + w \mod p$. Then obviously, $\tilde{\mathbf{k}}_j$ and $\tilde{w}_j$ are known to us.

For every KDM query (j_z, f_z), with the help of the LWE

instance $(\mathbf{A}'_{j_z}, \mathbf{A}'_{j_z}\mathbf{s}' + \mathbf{t}_{j_z})$ which is used to generate the public key $(\mathbf{A}_{j_z}, \mathbf{A}_{j_z}^t\mathbf{t}_{j_z})$, we can get a fresh gaussian LWE instance $(\bar{\mathbf{a}}, \bar{\mathbf{a}}^t\mathbf{t}_{j_z} + e')$ by theorem 1(2). We chooses $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m$ and generate the challenge ciphertext as below, $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e')$ if b=0, or $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e' + \tilde{w}_{j_z}p)$ if b=1.

Then, we show that the two games $\mathbf{G}_1$ and $\mathbf{G}_2$ are statistically indistinguishable.

CLAIM 2. *$\mathbf{G}_1$ and $\mathbf{G}_2$ are statistically indistinguishable.*

PROOF. Firstly, in $\mathbf{G}_1$, public keys are $\mathsf{KLWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instances. In $\mathbf{G}_2$, public keys are generated from LWE instances by theorem 1 which are also $\mathsf{KLWE}(m, n, q, \mathcal{D}_{\mathbb{Z},r})$ instances.

Then, in $\mathbf{G}_2$, for the KDM query (j_z, f_z) the challenge ciphertext is $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e' + \tilde{w}_{j_z}p)$ if b=1.

Define $\mathbf{y} = \bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p$, then $C_1 = (\mathbf{y}, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e' + \tilde{w}_{j_z}p) = (\mathbf{y}, \mathbf{t}_{j_z}^t(\mathbf{y} - \mathbf{e} + \tilde{\mathbf{k}}_{j_z}p) + e' + \tilde{w}_{j_z}p) = (\mathbf{y}, \mathbf{t}_{j_z}^t\mathbf{y} - \mathbf{t}_{j_z}^t\mathbf{e} + e' + f_z(\{\mathbf{t}_i\}_{i\in[l]})p)$. Since $\bar{\mathbf{a}}$ is uniform by theorem 1(2) and independent of $\mathbf{e} - \tilde{\mathbf{k}}_{j_z}p$, so the distribution of $\mathbf{y}$ is uniform. Consequently, the challenge ciphertext in $\mathbf{G}_2$ is distributed exactly as in $\mathbf{G}_1$ when $b = 1$. The case of $b = 0$ follows symmetrically.

Therefore $\mathbf{G}_1$ and $\mathbf{G}_2$ are statistically indistinguishable. □

$\mathbf{G}_3$: In this game, instead of using LWE instances to generate public keys and challenge ciphertexts in $\mathbf{G}_2$, we use uniform instances to generate public keys and challenge ciphertexts by theorem 1. In detail, by theorem 1(1), from l many $\mathcal{U}(\mathbb{Z}_q^{m\times(m-n)} \times \mathbb{Z}_q^m)$ instances $\{(\mathbf{A}'_i, \mathbf{u}'_i)\}_{i\in[l]}$, we can generate $\mathcal{U}(\mathbb{Z}_q^{m\times n} \times \mathbb{Z}_q^m)$ instances $(\mathbf{A}_i, \mathbf{u}_i)$ which are used as public keys.

For every KDM query (j_z, f_z), with the help of $(\mathbf{A}'_{j_z}, \mathbf{u}'_{j_z})$ which is used to generate the public key $(\mathbf{A}_{j_z}, \mathbf{u}_{j_z})$, we can get a fresh uniform instance $(\bar{\mathbf{a}}, \bar{u})$ which follows $\mathcal{U}(\mathbb{Z}_q^m \times \mathbb{Z}_q)$ by theorem 1(2). Then we choose $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m$ and generate the challenge ciphertext as below, $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \bar{u})$if $b = 0$, or $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \bar{u} + \tilde{w}_{j_z}p)$ if $b = 1$.

Then we will show that the two games $\mathbf{G}_2$ and $\mathbf{G}_3$ are computationally indistinguishable based on the $\mathsf{DLWE}(d + lm, m - n, q, \mathcal{D}_{\mathbb{Z},r})$ assumption.

CLAIM 3. $|\Pr[\mathbf{G}_2(\mathcal{A}) = 1] - \Pr[\mathbf{G}_3(\mathcal{A}) = 1]| = Adv_{\mathcal{S}_2}^{\mathsf{DLWE}(d+lm,m-n,q,\mathcal{D}_{\mathbb{Z},r})}$.

PROOF. Suppose there is a PPT adversary $\mathcal{A}$ has non-negligible advantage in distinguishing $\mathbf{G}_2$ and $\mathbf{G}_3$. Then we use $\mathcal{A}$ to construct an algorithm $\mathcal{S}_2$ as Figure 8 in section B to solve the decisional LWE problem. The oracle $\mathcal{O}_{\mathsf{DLWE}}$ uniform randomly outputs the LWE instances or the uniform instances.

At first, $\mathcal{S}_2$ randomly chooses a bit $b \xleftarrow{R} \{0, 1\}$. Then, after receiving instances $(\{(\mathbf{A}'_i, \mathbf{b}'_i)\}_{i\in[l]})$ sent by the oracle $\mathcal{O}_{\mathsf{DLWE}}$, $\mathcal{S}_2$ gets $\{(\mathbf{A}_i, \mathbf{b}_i)\}_{i\in[l]}$ by theorem 1(1) as the public keys.

Then $\mathcal{S}_2$ sends the l public keys $\{(\mathbf{A}_i, \mathbf{b}_i)\}_{i\in[l]}$ to the adversary $\mathcal{A}$ and invoke $\mathcal{A}$. Then $\mathcal{S}_2$ will receive KDM query and sends back the challenge ciphertext. This process can be looped at most d times. For each KDM query (j_z, f_z) from the adversary $\mathcal{A}$, $\mathcal{S}_2$ invokes the oracle $\mathcal{O}_{\mathsf{DLWE}}$ to get an instance $(\bar{\mathbf{a}}', \bar{b}')$. Then by theorem 1(2), $\mathcal{S}_2$ gets an instance $(\bar{\mathbf{a}}, \bar{b})$ from the instance $(\bar{\mathbf{a}}', \bar{b}')$. Next, $\mathcal{S}_2$ draws $\mathbf{e} \leftarrow \mathcal{D}_{\mathbb{Z},r}^m$ and generates the challenge ciphertext is $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \bar{b}^t)$ if $b = 0$, or $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \bar{b}^t + \tilde{w}_{j_z}p)$ if $b = 1$, where $\tilde{\mathbf{k}}_{j_z}$ and $\tilde{w}_{j_z}$ can be computed by f_z and $\{\mathbf{A}_i\}_{i\in[l]}$.

When the oracle $\mathcal{O}_{\mathsf{DLWE}}$ outputs LWE instances $\{\mathbf{A}'_i, \mathbf{b}'_i = \mathbf{A}'_i\mathbf{s}' + \mathbf{t}_i\}_{i\in[l]}$ and $(\bar{\mathbf{a}}', \bar{b}' = \bar{\mathbf{a}}'^t\mathbf{s}' + e')$, the public keys are $\{(\mathbf{A}_i, \mathbf{A}_i^t\mathbf{t}_i)\}_{i\in[l]}$ (by theorem 1(1)), and for every KDM query (j_z, f_z), $(\bar{\mathbf{a}}, \bar{b})$ is $(\bar{\mathbf{a}}, \bar{\mathbf{a}}^t\mathbf{t}_{j_z} + e')$ (by theorem 1(2)). Then the challenge ciphertext is $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e')$ if $b = 0$ or $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \mathbf{t}_{j_z}^t\bar{\mathbf{a}} + e' + \tilde{w}_{j_z}p)$ if $b = 1$. Therefore, in this case, $\mathcal{S}_2$ simulates $\mathbf{G}_2$.

When the oracle $\mathcal{O}_{\mathsf{DLWE}}$ outputs uniform random instances $\{\mathbf{A}'_i, \mathbf{b}'_i = \mathbf{u}'_i\}_{i\in[l]}$ and $(\bar{\mathbf{a}}', \bar{b}' = \bar{u}')$, the public keys are $\{(\mathbf{A}_i, \mathbf{u}_i)\}_{i\in[l]}$ (by theorem 1(1)), and for every KDM query (j_z, f_z), $(\bar{\mathbf{a}}, \bar{b})$ is $(\bar{\mathbf{a}}, \bar{u})$ (by theorem 1(2)). Then the challenge ciphertext is $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \bar{u})$ if $b = 0$, or $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p), \bar{u} + \tilde{w}_{j_z}p)$ if $b = 1$. So in this case $\mathcal{S}_2$ simulates $\mathbf{G}_3$.

Consequently, $|\Pr[\mathbf{G}_2(\mathcal{A}) = 1] - \Pr[\mathbf{G}_3(\mathcal{A}) = 1]| = Adv_{\mathcal{S}_2}^{\mathsf{DLWE}(d+lm,m-n,q,\mathcal{D}_{\mathbb{Z},r})}$. □

CLAIM 4. $\Pr[\mathbf{G}_3(\mathcal{A})] = \frac{1}{2} + negl(\lambda)$.

PROOF. In $\mathbf{G}_3$, for every KDM query (j_z, f_z), when $b = 0$, the ciphertext is $C_0 = (\bar{\mathbf{a}} + \mathbf{e}, \bar{u})$. When $b = 1$, the ciphertext is $C_1 = (\bar{\mathbf{a}} + \mathbf{e} - \tilde{\mathbf{k}}_{j_z}p, \bar{u} + \tilde{w}_{j_z}p)$.

Since $(\bar{\mathbf{a}}, \bar{u})$ is uniform by theorem 1 and independent of everything else in the ciphertext, so the distributions of C_0 and C_1 are uniform. Therefore, the distributions of the ciphertext in the case of $b = 0$ and in the case of $b = 1$ are statistically indistinguishable. So the adversary can just randomly guess b. That is, $\Pr[\mathbf{G}_3(\mathcal{A})] = \frac{1}{2} + negl(\lambda)$. □

Above all, $Adv_{\mathcal{A}}^{\mathsf{KDM\text{-}CPA}} = |2\Pr[\mathbf{G}_0(\mathcal{A})] - 1| \le ld \cdot Adv_{\mathcal{S}_1}^{\mathsf{ELWE}(m,n,q,\mathcal{D}_{\mathbb{Z},r})} + Adv_{\mathcal{S}_2}^{\mathsf{DLWE}(d+lm,m-n,q,\mathcal{D}_{\mathbb{Z},r})} + \quad negl(\lambda) = negl(\lambda)$ states that the scheme is KDM-CPA secure.

We should note that when q is super-polynomial in n, the above reduction is still valid. And the hardness of the corresponding extended LWE problem is under an LWE problem with super-polynomial modulus q and inverse error rate $1/\alpha$. In the later IBE construction, q will be super-polynomial in n.

5. KDM SECURE AND SO SECURE IBE

In this section, we will describe the construction of the IBE scheme as figure 4. Based on the all-but-d trapdoor function proposed by Alperin-Sheriff and Peikert [4] and the KDM secure PKE scheme in section 4, this IBE scheme is KDM secure. Because the proof is similar as [4], we will simply describe it. Compared with the KDM secure IBE scheme in [4], the user public key in Ext algorithm does not have the extra Gaussian noise, and the randomness vector $\mathbf{s}$ in Enc algorithm can be drawn from the uniform distribution instead of the discrete Gaussian. Thanks to these modifications, we can construct the lossy mode to the scheme by the method of [13] and prove it is an IBLE scheme. Then, we can get this scheme is IND-SO secure.

5.1 Construction

The following lemma is the efficient trapdoor construction and associated sampling algorithm of Micciancio and Peikert [16] which will be used in the extraction of user secret keys.

As in [16], we say that $\mathbf{R}$ is a strong trapdoor for $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$ if $\mathbf{A}\begin{bmatrix}\mathbf{R}\\ \mathbf{I}\end{bmatrix} = H(\mathbf{G})$ for some efficiently computable and invertible linear transformation H over $\mathbb{Z}_q^n$, which is applied column-wise to $\mathbf{G}$ where $\mathbf{G}$ is a universal pubic "gadget" matrix.

LEMMA 5.1 ([16],THEOREM 5.1). *Let $\mathbf{R}$ be a strong trapdoor for $\mathbf{A} \in \mathbb{Z}_q^{n\times m}$. There is an efficient randomized algorithm that, given $\mathbf{R}$, any $\mathbf{u} \in \mathbb{Z}_q^n$, and any $r \geq s_1(\mathbf{R}) \cdot \omega(\sqrt{\log n}) \geq \eta_\epsilon(\Lambda^\perp(\mathbf{A}))$ (for some $\epsilon(\lambda) = negl(\lambda)$), samples from a distribution within $negl(\lambda)$ distance of $\mathcal{D}_{\Lambda_{\mathbf{u}}^\perp(\mathbf{A}),r}$.*

$\mathbf{Setup}(1^\lambda, d)$

$\mathbf{R} \leftarrow \mathcal{D}^{md\times w}_{\mathbb{Z},\omega(\sqrt{\log n})}$
for $i = 0$ to $d-1$
  $\mathbf{A}_i \xleftarrow{\$} \mathbb{Z}_q^{n\times md}$, $\tilde{\mathbf{A}}_i = -\mathbf{A}_i\mathbf{R} \in \mathbb{Z}_q^{n\times w}$, $\mathbf{y}_i \xleftarrow{\$} \mathbb{Z}_q^n$
end
$\mathbf{A} = \begin{bmatrix}\mathbf{A}_0\\ ...\\ \mathbf{A}_{d-1}\end{bmatrix}$, $\tilde{\mathbf{A}} = \begin{bmatrix}\tilde{\mathbf{A}}_0\\ ...\\ \tilde{\mathbf{A}}_{d-1}\end{bmatrix} = -\mathbf{AR}$, $\mathbf{y} = \begin{bmatrix}\mathbf{y}_0\\ ...\\ \mathbf{y}_{d-1}\end{bmatrix}$
$\mathsf{MPK} = (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y})$, $\mathsf{MSK} = \mathbf{R}$
return $(\mathsf{MPK}, \mathsf{MSK})$

$\mathbf{Ext}(\mathsf{MPK} = (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y}), \mathsf{MSK} = \mathbf{R}, u)$

$\vec{u}^t = (u^0, u^1, ..., u^{d-1})$
$\bar{\mathbf{A}}_u = \vec{u}^t \cdot \mathbf{A}$, $\mathbf{y}_u = \vec{u}^t \cdot \mathbf{y}$
$\mathbf{A}_u = [\vec{u}^t \cdot \mathbf{A} | u^d\mathbf{G} + \vec{u}^t \cdot \tilde{\mathbf{A}}] = [\bar{\mathbf{A}}_u | u^d\mathbf{G} - \bar{\mathbf{A}}_u\mathbf{R}]$
$\mathbf{t}_u \leftarrow \mathcal{D}_{\Lambda_{\mathbf{y}_u}^\perp(\mathbf{A}_u),r}$ (by lemma 5.1)
$\mathsf{SK}_u = \mathbf{t}_u$
return (SK_u)

$\mathbf{Enc}(\mathsf{MPK} = (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y}), u, \mu)$

$\vec{u}^t = (u^0, u^1, ..., u^{d-1})$, $\mathbf{A}_u = [\vec{u}^t \cdot \mathbf{A} | u^d\mathbf{G} + \vec{u}^t \cdot \tilde{\mathbf{A}}]$, $\mathbf{y}_u = \vec{u}^t \cdot \mathbf{y}$
$\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, $\mathbf{e}^{(1)} \leftarrow \mathcal{D}^{md}_{\mathbb{Z},r}, \mathbf{e}^{(2)} \leftarrow \mathcal{D}^{w}_{\mathbb{Z},\gamma}, e' \leftarrow \mathcal{D}_{\mathbb{Z},\gamma}$
$\mathbf{c}_1^t = \mathbf{s}^t\mathbf{A}_u + [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t]$, $c_2 = \mathbf{s}^t\mathbf{y}_u + e' + p \cdot \mu$
return $(\mathbf{c}_1, c_2)$

$\mathbf{Dec}(\mathsf{MPK}, \mathsf{SK}_u, (\mathbf{c}_1, c_2))$

$\mu = decode(c_2 - \mathbf{c}_1^t\mathsf{SK}_u)$
return μ

Figure 4: Construction of identity-based encryption

The following lemma is a useful tool in the proof of KDM security and the construction of the lossy mode.

LEMMA 5.2 ([4]). *Let $\mathbf{A}^* \in \mathbb{Z}_q^{nl\times\bar{m}}$, and let $V \in \mathcal{R}^{l\times d}$ be the Vandermonde matrix whose rows are the vectors $\vec{u}_i^t = (u_i^0, u_i^1, ..., u_i^{d-1})$ where $l \leq d$ and $\bar{m} \geq 1$. There is a polynomial-time algorithm $\mathsf{MapUser}(\mathbf{A}^*, \mathcal{I} = \{u_1, ..., u_l\})$ that outputs $\mathbf{A} \in \mathbb{Z}_q^{nd\times\bar{m}}$ such that $V \cdot \mathbf{A} = \mathbf{A}^*$. And $\mathbf{A}$ is uniformly random over the uniformly random choice of $\mathbf{A}^*$.*

Figure 4 is the construction of IBE scheme. As in [4], let $\mathcal{R}$ denote any commutative ring (with efficiently computable operations, including inversion of multiplicative units) such that the additive group $\mathbb{G} = \mathbb{Z}_q^n$ is an $\mathcal{R}$-module, and such that there are at least $d+1$ known elements $U = \{u_0 = 0, u_1, u_2, ...\} \subseteq \mathcal{R}$ where $u_i - u_j$ is invertible in $\mathcal{R}$ for every $i \neq j$. Then the identity space is $U\backslash\{0\} \subseteq \mathcal{R}$. Let u denote the identity and $\vec{u}^t = (u^0, u^1, ..., u^{d-1})$. $\vec{u}^t \cdot \mathbf{A}$ means that $\sum_{i=0}^{d-1} u^i\mathbf{A}_i$ where $\mathbf{A} = [\mathbf{A}_0^t | \mathbf{A}_1^t | ... | \mathbf{A}_{d-1}^t]^t \in \mathbb{Z}_q^{nd\times md}$. We take $\mathbf{G} = \mathbf{I}_n \otimes [1, 2, 2^2, ..., 2^{\frac{w}{n}-1}] \in \mathbb{Z}_q^{n\times w}$ where $\frac{w}{n} = \lceil \log q \rceil$ as in [16]. decode() is same in section 4.

Parameters. Let $n = poly(\lambda)$, $m \geq n\log q + 2\lambda - 2$, $q = p^2$, $p = \gamma \cdot poly(n)$ for a sufficiently large $poly(n)$ term to ensure correctness, $\gamma = r^{\omega(1)}$, and $r \geq O(\sqrt{md}) \cdot \omega(\sqrt{\log n})^2$.

Correctness. Let $\mathbf{e}^t = [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t]$. For $(\mathbf{c}_1, c_2) \leftarrow \mathsf{Enc}(\mathsf{MPK}, u, \mu)$ and $\mathbf{t}_u \leftarrow \mathsf{Ext}(\mathsf{MPK}, \mathsf{MSK}, u)$ under identity u, $c_2 - \mathbf{c}_1^t\mathbf{t}_u = \mathbf{s}^t\mathbf{y}_u + e' + p \cdot \mu - \mathbf{s}^t\mathbf{A}_u\mathbf{t}_u - \mathbf{e}^t\mathbf{t}_u = p \cdot \mu + e' - \mathbf{e}^t\mathbf{t}_u$. By Cauchy-Schwarz and lemma 2.1, $|e' + \mathbf{e}^t\mathbf{t}_u| \leq |e'| + |\mathbf{t}^t\mathbf{e}| \leq \gamma \cdot (\omega(\sqrt{\log n}) + w) + r^2md < \frac{p}{2}$. Therefore, the algorithm *decode*() will get the correct message with overwhelming probability.

5.2 KDM security

THEOREM 3. *Parameters are the same in section 5.1. The IBE scheme in section 5.1 is selective identity* KDM-CPA *secure for affine functions, under the* LWE *assumption and the* KDM-CPA *security of the PKE scheme in section 4.*

Proof sketch. The proof is same as the proof in [4], so we just give a brief introduction. Game 0 is the actual KDM-CPA security game from section 4.2.

In Game 1, the challenger uses an all-but-d trapdoor construction to construct the master public key, and this game is statistically indistinguishable from the Game 0. In detail, the adversary sends a list of target identities $\mathcal{I} = \{u_1^*, u_2^*, ..., u_l^*\}$. The challenger chooses d uniform random matrices $\mathbf{A}_i^* \in \mathbb{Z}_q^{n\times md}$ and master secret key $\mathbf{R} \leftarrow \mathcal{D}^{md\times w}_{\mathbb{Z},\omega(\sqrt{\log n})}$. And for $i \in [d]$, the challenger chooses the secret key for the identity u_i to be $\mathbf{z}_i \leftarrow \mathcal{D}^{md+w}_{\mathbb{Z},r}$ and sets $\mathbf{y}_i^* = [\mathbf{A}_i^* | -\mathbf{A}_i^*\mathbf{R}]\mathbf{z}_i$. By lemma 2.1, this is statistically close to choosing $\mathbf{y}_i^*$ uniformly an random and then sampling $\mathbf{z}_i$ from $\mathcal{D}_{\Lambda_{\mathbf{y}_i^*}^\perp([\mathbf{A}_i^* | -\mathbf{A}_i^*\mathbf{R}]),r}$. Let $\mathbf{A}^*$ denote the stack of matrices $\mathbf{A}_i^*$, and $\mathbf{y}^*$ denote the stack of vectors $\mathbf{y}_i^*$. By the algorithm $\mathsf{MapUser}(\mathbf{A}^*, \mathcal{I})$ and $\mathsf{MapUser}(\mathbf{y}^*, \mathcal{I})$ (lemma 5.2), the challenger computes $\mathbf{A}$ and $\mathbf{y}$ such that $\mathbf{A}_i^* = \vec{u}_i^{*t} \cdot \mathbf{A}$ and $\mathbf{y}_i^* = \vec{u}_i^{*t} \cdot \mathbf{y}$. Like in [4], we define the monic degree-d polynomial as $f(x) = x^{d-l} \cdot \prod_{i\in[l]}(x - u_i^*) = c_0 + c_1x + \cdots + c_{d-1}x^{d-1} + x^d$, when $x \notin \mathcal{I}$, $f(x)$ is invertible, and when $x \in \mathcal{I}$, $f(x)$ is 0. And let $\tilde{\mathbf{A}} = -\mathbf{AR} + [c_0, ..., c_{d-1}]^t\mathbf{G}$. Then $(\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y})$ as the master public key is sent to the adversary. Now the user public key is $pk_u = [\vec{u}^t \cdot \mathbf{A} | -\vec{u}^t \cdot \mathbf{AR} + f(u)\mathbf{G}]$, and this is an all-but-d trapdoor function in which $\mathbf{R}$ is the trapdoor if $f(u)$ is invertible, because $pk_u \cdot \begin{bmatrix}\mathbf{R}\\ \mathbf{I}\end{bmatrix} = f(u) \cdot \mathbf{G}$. When the adversary makes extraction queries, the challenger uses the Ext algorithm normally if the identity $u \notin \mathcal{I}$. If the identity $u \in \mathcal{I}$, $f(u) = 0$ then $\mathbf{R}$ is not the trapdoor for pk_u. So Ext can not sample the user secret key.

In Game 2, the challenger uses the KDM secure PKE scheme discussed in section 4 to statistically simulate Game 1. The challenger uses the outputs of the KDM secure PKE scheme and the all-but-d trapdoor function to construct the keys and the ciphertexts. Because the dimensions of keys and ciphertexts in the IBE scheme are larger than the dimensions in the PKE scheme, it should be handled carefully. Specifically, when the challenger constructs IBE ciphertexts from the PKE ciphertexts, a super-polynomial extra noise should be used to disguise the noise introduced from the

PKE ciphertext. Then the IBE ciphertext will be statistically distinguishable from the real IBE ciphertext created by the encryption algorithm. It will lead to a super-polynomial modulus q. Since all the operations are same as [4], we omit them here. In this game, the KDM security of this IBE scheme is reduced to the KDM security of the PKE scheme.

5.3 SO security

THEOREM 4. *Parameters are the same in section 5.1, and* $r = r'^{\omega(1)}, r' \geq 2\sqrt{n}, k \log q \leq n - 2\lambda + 2$. *Then the IBE scheme in section 5.1 is* IND-sID-SO *secure, assuming that the* LWE *problem is hard.*

PROOF. First, we construct a lossy mode to the IBE scheme of section 5.1, and prove it is an IBLE scheme. Then by lemma 2.6, this scheme is IND-sID-SO secure.

Now, we will construct the IBLE scheme. The $\mathbf{Setup_{lossy}}$ algorithm is as figure 5, besides, $\mathbf{Setup_{real}}$ is same as $\mathbf{Setup}$, $\mathbf{Ext_{real}}$ and $\mathbf{Ext_{lossy}}$ are same as $\mathbf{Ext}$ in figure 4.

Then, we show this scheme fulfills the properties of IBLE.

1. *Correctness on keys for all* $\boldsymbol{id} \notin \mathcal{I}$. This is the same as the correctness property in section 5.1.

2. *Lossiness of encryption with lossy keys for* $\boldsymbol{id} = u_j^* \in \mathcal{I} = \{u_1^*, ..., u_l^*\}$ for some $j \in \{1, ..., l\}$.

$$\begin{aligned}
&\mathbf{Enc}(\mathbf{id}, \mathsf{MPK}_{\mathsf{lossy}}, \mu) = \mathbf{Enc}(u_j^*, (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y}), \mu) \\
&= (\mathbf{s}^t[\vec{u}_j^{*t} \cdot \mathbf{A}| - \vec{u}_j^{*t} \cdot \mathbf{AR} + (\sum_{i=0}^{d-1} c_i (u_j^*)^i + (u_j^*)^d)\mathbf{G})] + \\
&\quad [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t], \mathbf{s}^t(\vec{u}_j^{*t} \cdot \mathbf{y}) + e' + \mu \cdot p) \\
&= (\mathbf{s}^t[(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t | (-(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t\mathbf{R} + f(u_j^*)\mathbf{G})] + \\
&\quad [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t], \mathbf{s}^t\mathbf{y}_j^* + e' + \mu \cdot p) \\
&= (\mathbf{s}^t[(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t | (-(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t\mathbf{R})] + [(\mathbf{e}^{(1)})^t | (\mathbf{e}^{(2)})^t], \\
&\quad \mathbf{s}^t\mathbf{y}_j^* + e' + \mu \cdot p) \\
&= ((\begin{bmatrix} \mathbf{B}_j\mathbf{C} + \mathbf{Z}_j \\ -\mathbf{R}^t(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j) \end{bmatrix} \mathbf{s} + \begin{bmatrix} \mathbf{e}^{(1)} \\ \mathbf{e}^{(2)} \end{bmatrix})^t, \mathbf{s}^t\mathbf{y}_j^* + e' + \mu \cdot p) \\
&= (((\begin{bmatrix} \mathbf{B}_j \\ -\mathbf{R}^t\mathbf{B}_j \end{bmatrix} \mathbf{C} + \begin{bmatrix} \mathbf{Z}_j \\ -\mathbf{R}^t\mathbf{Z}_j \end{bmatrix})\mathbf{s} + \begin{bmatrix} \mathbf{e}^{(1)} \\ \mathbf{e}^{(2)} \end{bmatrix})^t, \mathbf{s}^t\mathbf{y}_j^* + e' + \mu \cdot p).
\end{aligned}$$

Let $\mathbf{A}' = \begin{bmatrix} \mathbf{B}_j \\ -\mathbf{R}^t\mathbf{B}_j \end{bmatrix} \mathbf{C} + \begin{bmatrix} \mathbf{Z}_j \\ -\mathbf{R}^t\mathbf{Z}_j \end{bmatrix}, \mathbf{e} = \begin{bmatrix} \mathbf{e}^{(1)} \\ \mathbf{e}^{(2)} \end{bmatrix}$. Because the parameters satisfy the requirements of lemma 2.5, we know that $\tilde{H}_\infty(\mathbf{s}|\mathbf{A}'\mathbf{s} + \mathbf{e}) \geq n$. Then because $\log q \leq k - 2\log(1/\varepsilon) - O(1)$, and by lemma 2.4, given $\mathbf{A}'\mathbf{s} + \mathbf{e}$, $\mathbf{s}^t\mathbf{y}_j^*$ is ε-close to $\mathcal{U}(\mathbb{Z}_q)$. When $\varepsilon = negl(\lambda)$, $\mathbf{s}^t\mathbf{y}_j^* \approx_s \mathcal{U}(\mathbb{Z}_q)$ given $\mathbf{A}'\mathbf{s} + \mathbf{e}$. Therefore, for any $\mu \in \mathcal{M}$, $(\mathbf{s}^t\mathbf{y}_j^* + e' + \mu \cdot p)$ is statistically close to $\mathcal{U}(\mathbb{Z}_q)$, given $\mathbf{A}'\mathbf{s} + \mathbf{e}$, i.e. for the lossy identity u_j^*, any lossy keys $\mathsf{MPK}_{\mathsf{lossy}}$ generated by $\mathbf{Setup_{lossy}}(1^n, u_j^*)$ and any two messages $\mu_0 \neq \mu_1$, there is $\mathbf{Enc}(u_j^*, \mathsf{MPK}_{lossy}, \mu_0) \approx_s \mathbf{Enc}(u_j^*, \mathsf{MPK}_{\mathsf{lossy}}, \mu_1)$.

3. *Indistinguishability between real keys and lossy keys.* We use a game sequence to prove this property.

$\mathbf{Game}_0$: This is the actual game from the definition of the third property of identity-based lossy encryption described as figure 2. In the process of $\mathbf{Setup_{lossy}}$, we choose $\mathbf{B}_j, \mathbf{C}, \mathbf{Z}_j$ where $j = 1$ to l to construct $\left[\mathbf{B}_1^t, ..., \mathbf{B}_l^t\right]^t \mathbf{C} + \left[\mathbf{Z}_1^t, ..., \mathbf{Z}_l^t\right]^t$ as figure 5.

$\mathbf{Game}_1$: In this game, we use LWE instances $(\mathbf{H}_j^*, \mathbf{A}_j^*) = (\mathbf{B}_j, (\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t)$ where $j = 1$ to l and process as the algorithm $\mathbf{Setup_{lossy}}$ (figure 5) to generate the lossy keys $(\mathsf{MPK}_1, \mathsf{MSK}_1)$. The remainder of the game is unchanged. So this game is identical to $\mathbf{Game}_0$.

$\mathbf{Setup_{lossy}}(1^\lambda, d, \mathcal{I} = \{u_1^*, u_2^*, ..., u_l^*\})$

$f(x) = x^{d-l} \cdot \prod_{i \in [l]} (x - u_i^*) = c_0 + c_1 x + \cdots + c_{d-1}x^{d-1} + x^d$

$\mathbf{R} \leftarrow \mathcal{D}^{md \times w}_{\mathbb{Z}, \omega(\sqrt{\log n})}$

for $j = 1$ to l

$\quad \mathbf{B}_j \xleftarrow{\$} \mathbb{Z}_q^{md \times k}, \mathbf{C} \xleftarrow{\$} \mathbb{Z}_q^{k \times n}, \mathbf{Z}_j \leftarrow \mathcal{D}^{md \times n}_{\mathbb{Z}, r'}$

$\quad \mathbf{A}_j^* = (\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t$, $\mathbf{y}_j^* \xleftarrow{\$} \mathbb{Z}_q^n$

end

$\mathbf{A}^* = \begin{bmatrix} \mathbf{A}_1^* \\ ... \\ \mathbf{A}_l^* \end{bmatrix}, \mathbf{y}^* = \begin{bmatrix} \mathbf{y}_1^* \\ ... \\ \mathbf{y}_l^* \end{bmatrix}, \mathbf{A} = \begin{bmatrix} \mathbf{A}_0 \\ ... \\ \mathbf{A}_{d-1} \end{bmatrix} \leftarrow \mathsf{MapUser}(\mathbf{A}^*, \mathcal{I})$

$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ ... \\ \mathbf{y}_{d-1} \end{bmatrix} \leftarrow \mathsf{MapUser}(\mathbf{y}^*, \mathcal{I}), \tilde{\mathbf{A}} = \begin{bmatrix} -\mathbf{A}_0\mathbf{R} + c_0\mathbf{G} \\ ... \\ -\mathbf{A}_{d-1}\mathbf{R} + c_{d-1}\mathbf{G} \end{bmatrix}$

$\mathsf{MPK} = (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y}), \mathsf{MSK} = \mathbf{R}$

return $(\mathsf{MPK}, \mathsf{MSK})$

Figure 5: Construction of the lossy mode

$\mathbf{Game}_2$: In this game, we use uniformly random matrices $\mathbf{X}_j^*$ to replace $(\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t$ where $j = 1$ to l and process as the algorithm $\mathbf{Setup_{lossy}}$ (figure 5) to generate the lossy keys $(\mathsf{MPK}_1, \mathsf{MSK}_1)$. The remainder of the game is unchanged. Based on the hardness of LWE assumption, $\mathbf{A}_j^* = (\mathbf{B}_j\mathbf{C} + \mathbf{Z}_j)^t \approx_c \mathbf{X}_j^*$, so this game is computationally indistinguishable from $\mathbf{Game}_1$.

In $\mathbf{Game}_2$, $\mathsf{MPK}_1 = (\mathbf{A}', \tilde{\mathbf{A}}', \mathbf{y}')$ where $\vec{u}_j^{*t} \cdot \mathbf{A}' = \mathbf{X}_j^*$, $\tilde{\mathbf{A}}' = -\mathbf{A}'\mathbf{R} + [c_0, ..., c_{d-1}]^t\mathbf{G}$, $\vec{u}_j^{*t} \cdot \mathbf{y}' = \mathbf{y}_j^*$ and $j = 1$ to l. Then, by lemma 5.2, $\mathbf{A}'$ and $\mathbf{y}'$ are uniformly random because $\mathbf{X}^*$ which is the stack of matrices $\mathbf{X}_j^*$ and $\mathbf{y}^*$ which is the stack of vectors $\mathbf{y}_j^*$ are uniformly random. And by lemma 2.1, $(\mathbf{A}', \mathbf{A}'\mathbf{R})$ is statistically indistinguishable from uniform, i.e. $(\mathbf{A}', \tilde{\mathbf{A}}')$ is statistically indistinguishable from uniform. Therefore, $\mathsf{MPK}_1 = (\mathbf{A}', \tilde{\mathbf{A}}', \mathbf{y}')$ is statistically distinguishable from $\mathsf{MPK}_0 = (\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y})$ where $\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{y}$ are uniformly random.

Furthermore, $\mathbf{Ext_{real}}$ and $\mathbf{Ext_{lossy}}$ are same as the algorithm $\mathbf{Ext}$ in figure 4. Since the distributions of MPK_0 and MPK_1 are statistically indistinguishable, and the distributions of MSK_0 and MSK_1 are the same, then the distributions between the output of $\mathbf{Ext_{real}}(u, \mathsf{MPK}_0, \mathsf{MSK}_0) = \mathbf{Ext}(\mathsf{u}, \mathsf{MPK}_0, \mathsf{MSK}_0)$ and the output of $\mathbf{Ext_{lossy}}(u, \mathsf{MPK}_1, \mathsf{MSK}_1) = \mathbf{Ext}(\mathsf{u}, \mathsf{MPK}_1, \mathsf{MSK}_1)$ are statistically indistinguishable, i.e. any SK_u generated by $\mathbf{Ext_{real}}(u, \mathsf{MPK}_0, \mathsf{MSK}_0)$ is statistically indistinguishable with any SK_u generated by $\mathbf{Ext_{lossy}}(u, \mathsf{MPK}_1, \mathsf{MSK}_1)$ for all $u \notin \mathcal{I}$.

Therefore, the advantage of the adversary of $\mathbf{Game}_2$ is $negl(\lambda)$. So the advantage of the adversary of $\mathbf{Game}_0$ is $negl(\lambda)$. This completes the proof. □

6. ACKNOWLEDGMENTS

This research is supported by the National Nature Science Foundation of China (No.61572495, No.61379137 and No.61502484), and the National Basic Research Program of China (973 project) (No.2013CB338002).

7. REFERENCES

[1] Shweta Agrawal, Dan Boneh, and Xavier Boyen. Efficient lattice (h) ibe in the standard model. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 553–572. Springer, 2010.

[2] Shweta Agrawal, Dan Boneh, and Xavier Boyen. Lattice basis delegation in fixed dimension and shorter-ciphertext hierarchical ibe. In *Annual Cryptology Conference*, pages 98–115. Springer, 2010.

[3] Adi Akavia, Shafi Goldwasser, and Vinod Vaikuntanathan. Simultaneous hardcore bits and cryptography against memory attacks. In *Theory of Cryptography Conference*, pages 474–495. Springer, 2009.

[4] Jacob Alperin-Sheriff and Chris Peikert. Circular and kdm security for identity-based encryption. In *Public Key Cryptography–PKC 2012*, pages 334–352. Springer, 2012.

[5] Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *Advances in Cryptology-CRYPTO 2009*, pages 595–618. Springer, 2009.

[6] Mihir Bellare, Dennis Hofheinz, and Scott Yilek. Possibility and impossibility results for encryption and commitment secure under selective opening. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 1–35. Springer, 2009.

[7] Alexandra Berkoff and Feng-Hao Liu. Leakage resilient fully homomorphic encryption. In *Theory of Cryptography*, pages 515–539. Springer, 2014.

[8] John Black, Phillip Rogaway, and Thomas Shrimpton. Encryption-scheme security in the presence of key-dependent messages. In *Selected Areas in Cryptography*, volume 2595, pages 62–75. Springer, 2002.

[9] David Cash, Dennis Hofheinz, Eike Kiltz, and Chris Peikert. Bonsai trees, or how to delegate a lattice basis. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 523–552. Springer, 2010.

[10] Yevgeniy Dodis, Leonid Reyzin, and Adam Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In *Advances in cryptology-Eurocrypt 2004*, pages 523–540. Springer, 2004.

[11] Craig Gentry, Chris Peikert, and Vinod Vaikuntanathan. Trapdoors for hard lattices and new cryptographic constructions. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 197–206. ACM, 2008.

[12] Shafi Goldwasser, Yael Tauman Kalai, Chris Peikert, and Vinod Vaikuntanathan. Robustness of the learning with errors assumption. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 5-7, 2010. Proceedings*, pages 230–240, 2010.

[13] Jingnan He, Bao Li, Xianhui Lu, Dingding Jia, Haiyang Xue, and Xiaochao Sun. Identity-based lossy encryption from learning with errors. In *International Workshop on Security*, pages 3–20. Springer, 2015.

[14] Richard Lindner and Chris Peikert. Better key sizes (and attacks) for lwe-based encryption. In *Topics in Cryptology–CT-RSA 2011*, pages 319–339. Springer, 2011.

[15] Daniele Micciancio and Petros Mol. Pseudorandom knapsacks and the sample complexity of lwe search-to-decision reductions. In *Advances in Cryptology–CRYPTO 2011*, pages 465–484. Springer, 2011.

[16] Daniele Micciancio and Chris Peikert. Trapdoors for lattices: Simpler, tighter, faster, smaller. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 700–718. Springer, 2012.

[17] Daniele Micciancio and Oded Regev. Worst-case to average-case reductions based on gaussian measures. *SIAM Journal on Computing*, 37(1):267–302, 2007.

[18] Chris Peikert, Vinod Vaikuntanathan, and Brent Waters. A framework for efficient and composable oblivious transfer. In *Annual International Cryptology Conference*, pages 554–571. Springer, 2008.

[19] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. *SIAM Journal on Computing*, 40(6):1803–1844, 2011.

[20] Oded Regev. On Lattices, Learning with Errors, Random Linear Codes, and Cryptography. In *STOC*, pages 84–93, 2005.

[21] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *Journal of the ACM (JACM)*, 56(6):34, 2009.

APPENDIX

A. PROOF OF LEMMA 5

The **min-entropy** of a random variable X is $\tilde{H}_\infty(X) = -\log(\max_x \Pr[X = x])$, and the **average min-entropy** of X conditioned on Y, defined by [10], is

$$\begin{aligned}\tilde{H}_\infty(X|Y) &= -\log(\mathbf{E}_{y\leftarrow Y}[\max_x \Pr[X = x|Y = y]])\\ &= -\log(\mathbf{E}_{y\leftarrow Y}[2^{-\tilde{H}_\infty(X|Y=y)}]).\end{aligned}$$

DEFINITION 1 ([10]). *For two random variables X and Y, the ϵ-**smooth average min-entropy** of X conditioned on Y, denoted $\tilde{H}^\epsilon_\infty(X|Y)$ is*

$$\tilde{H}^\epsilon_\infty(X|Y) = \max_{(X',Y'):\Delta((X,Y),(X',Y'))<\epsilon} \tilde{H}_\infty(X'|Y').$$

LEMMA A.1 ([7]). *For any random variable X, given distributions $\mathcal{D}_Y \approx_s \mathcal{D}_Z$ where $Y \leftarrow \mathcal{D}_Y$ and $Z \leftarrow \mathcal{D}_Z$, there exists some ϵ such that $\Delta(Y, Z) < \epsilon = negl(\lambda)$, and $\tilde{H}^\epsilon_\infty(X|Y) \geq \tilde{H}^\epsilon_\infty(X|Z)$.*

Next we prove lemma 2.5. The method of proof is similar to lemma A.2 in [7].

PROOF. 1. $\bar{\mathbf{A}} \approx_c \mathbf{U} \xleftarrow{\$} \mathbb{Z}_q^{2m\times n}$:

$(\mathbf{BC} + \mathbf{Z}, \mathbf{R}^t(\mathbf{BC} + \mathbf{Z})) \overset{(1)}{\approx}_c (\mathbf{U}_1, \mathbf{R}^t\mathbf{U}_1) \overset{(2)}{\approx}_s (\mathbf{U}_1, \mathbf{U}_2)$,

where $\mathbf{U}_1, \mathbf{U}_2 \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$

- Under the hardness of LWE assumption, approximate formula (1) holds.

- Let $\mathbf{r}_i$ be the i-th column of $\mathbf{R}$ where $\mathbf{r}_i \leftarrow \mathcal{D}^m_{\mathbb{Z},\omega(\sqrt{\log q})}$. Because $n\log q \le m - 2\lambda + 2$, by lemma 2.1, $(\mathbf{U}_1^t, \mathbf{U}_1^t\mathbf{r}_i)$ is statistically close to the uniform distribution over $\mathbb{Z}_q^{n\times m} \times \mathbb{Z}_q^n$. Because the columns of $\mathbf{R} = [\mathbf{r}_1\ \mathbf{r}_2\ ...\ \mathbf{r}_m]$ are sampled independently, $(\mathbf{U}_1^t, \mathbf{U}_1^t\mathbf{R})$ is statistically close to the uniform distribution over $\mathbb{Z}_q^{n\times m} \times \mathbb{Z}_q^{n\times m}$. Taking the transpose, (2) holds.

2. $\tilde{H}^\epsilon_\infty(\mathbf{s}|\bar{\mathbf{A}}, \bar{\mathbf{A}}\mathbf{s}+\mathbf{x}) \ge n$, where $\epsilon = negl(\lambda)$: Let $\mathbf{s}_0 \xleftarrow{\$} \{0,1\}^n$, $\mathbf{s}_1 \xleftarrow{\$} \mathbb{Z}_q^n$, then, think of $\mathbf{s} = \mathbf{s}_0 + \mathbf{s}_1$. Because $\tilde{H}^\epsilon_\infty(\mathbf{s}|\bar{\mathbf{A}}\mathbf{s}+\mathbf{e}) \ge \tilde{H}^\epsilon_\infty(\mathbf{s}_0|\bar{\mathbf{A}}\mathbf{s}+\mathbf{e})$, we will consider $\tilde{H}^\epsilon_\infty(\mathbf{s}_0|\bar{\mathbf{A}}\mathbf{s}+\mathbf{e})$.

$$
\begin{aligned}
&\bar{\mathbf{A}}\mathbf{s}+\mathbf{e}\\
&= \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_0 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix}\\
&\overset{(1)}{\approx}_s \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix}\\
&\overset{(2)}{\approx}_s \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{u}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix}
\end{aligned}
$$

- Since $\frac{r'}{r_1} = negl(\lambda)$, each element of $\mathbf{Z}\mathbf{s}_0$ is negligibly small compared to the corresponding element of $\mathbf{e}^{(1)}$. $\frac{r'}{r_2} = negl(\lambda)$ and $\mathbf{R}^t\mathbf{Z}\mathbf{s}_0$ is polynomial number of operations on elements of $\mathbf{Z}\mathbf{s}_0$, so each element of $\mathbf{R}^t\mathbf{Z}\mathbf{s}_0$ is negligibly small compared to the corresponding element of $\mathbf{e}^{(2)}$. Therefore, $\begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix} + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_0 \approx_s \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix}$, and the approximate formula (1) holds. It means that their statistical distance is some $\epsilon_1 = negl(\lambda)$.
- Since $\mathbf{s}_0 \xleftarrow{\$} \{0,1\}^n$ and $k\log q \le n - 2\lambda + 2$, by lemma 2.4, the approximate formula (2) holds where $\mathbf{u}_0 \xleftarrow{\$} \mathbb{Z}_q^k$. It means that their statistical distance is some $\epsilon_2 = negl(\lambda)$.

Then, for $\epsilon = \epsilon_1 + \epsilon_2 = negl(\lambda)$,

$$
\begin{aligned}
&\tilde{H}^\epsilon_\infty(\mathbf{s}_0|\begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_0 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix})\\
&\overset{(3)}{\ge} \tilde{H}^\epsilon_\infty(\mathbf{s}_0|\begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix})\\
&\ge \tilde{H}^\epsilon_\infty(\mathbf{s}_0|\begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{u}_0 + \begin{bmatrix}\mathbf{B}\\ \mathbf{R}^t\mathbf{B}\end{bmatrix}\mathbf{C}\mathbf{s}_1 + \begin{bmatrix}\mathbf{Z}\\ \mathbf{R}^t\mathbf{Z}\end{bmatrix}\mathbf{s}_1 + \begin{bmatrix}\mathbf{e}^{(1)}\\ \mathbf{e}^{(2)}\end{bmatrix})\\
&\overset{(4)}{=} \tilde{H}_\infty(\mathbf{s}_0)\\
&= n
\end{aligned}
$$

By lemma A.1, (3) holds. Because each of $\mathbf{B}, \mathbf{R}, \mathbf{C}, \mathbf{Z}, \mathbf{u}_0, \mathbf{s}_1, \mathbf{e}$ is independent of $\mathbf{s}_0$, (4) holds. □

LEMMA A.2 ([7]). *There exists a distribution* Lossy *such that* $\bar{\mathbf{A}} \leftarrow \mathsf{Lossy} \approx_c \mathbf{U} \xleftarrow{\$} \mathbb{Z}_q^{m\times n}$ *and given* $\mathbf{s} \xleftarrow{\$} \mathbb{Z}_q^n$, *and* $\mathbf{e} \leftarrow \mathcal{D}^{m\times n}_{\mathbb{Z},\beta q}$, $\tilde{H}^\epsilon_\infty(\mathbf{s}|\bar{\mathbf{A}}, \bar{\mathbf{A}}\mathbf{s}+\mathbf{x}) \ge n$, *where* $\epsilon = negl(\lambda)$. Lossy *is as follows,*

- *Choose* $\mathbf{B} \xleftarrow{\$} \mathbb{Z}_q^{m\times k}$, $\mathbf{C} \xleftarrow{\$} \mathbb{Z}_q^{k\times n}$, *and* $\mathbf{Z} \leftarrow \mathcal{D}^{m\times n}_{\mathbb{Z},\alpha q}$, *where* $\frac{\alpha}{\beta} = negl(\lambda)$ *and* $k\log q \le n - 2\lambda + 2$.
- *Let* $\bar{\mathbf{A}} = \mathbf{B}\mathbf{C} + \mathbf{Z}$.
- *Output* $\bar{\mathbf{A}}$.

B. EXPLANATION IN FIGURES

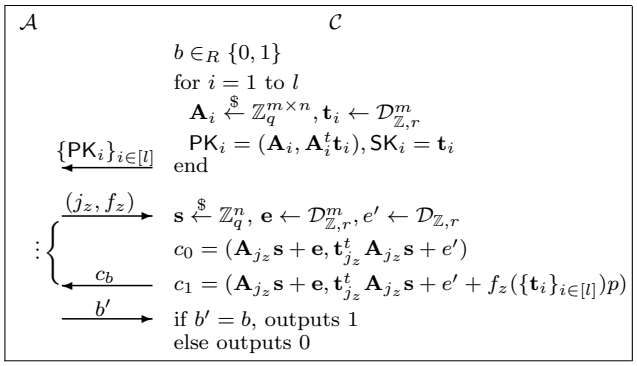

Figure 6: The process of $\mathbf{G}_0$

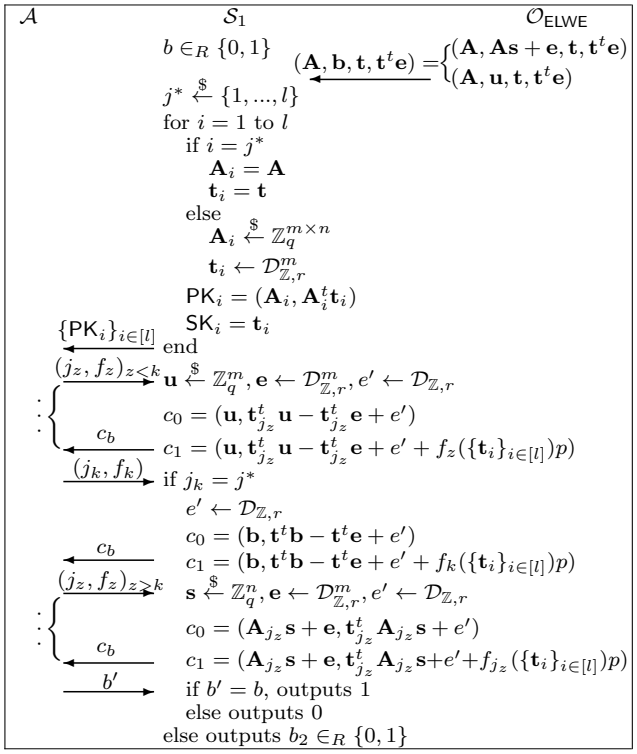

Figure 7: Game of distinguishing $\mathbf{H}_{k-1}$ and $\mathbf{H}_k$

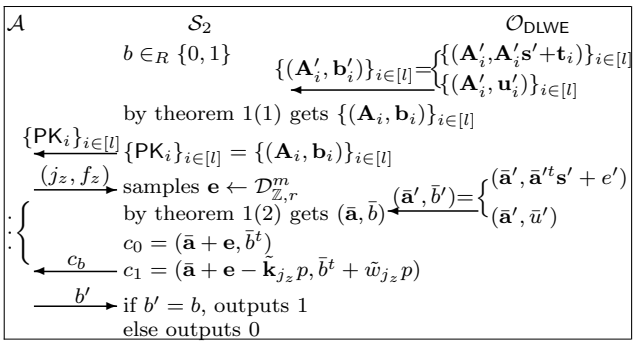

Figure 8: Game of distinguishing $\mathbf{G}_2$ and $\mathbf{G}_3$