Circular chosen-ciphertext security with compact ciphertexts

Dennis Hofheinz*

January 19, 2013

Abstract

A key-dependent message (KDM) secure encryption scheme is secure even if an adversary obtains encryptions of messages that depend on the secret key. Such key-dependent encryptions naturally occur in scenarios such as harddisk encryption, formal cryptography, or in specific protocols. However, there are not many provably secure constructions of KDM-secure encryption schemes. Moreover, only one construction, due to Camenisch, Chandran, and Shoup (Eurocrypt 2009) is known to be secure against active (i.e., CCA) attacks.

In this work, we construct the first public-key encryption scheme that is KDM-secure against active adversaries and has compact ciphertexts. As usual, we allow only circular key dependencies, meaning that encryptions of arbitrary *entire* secret keys under arbitrary public keys are considered in a multi-user setting.

Technically, we follow the approach of Boneh, Halevi, Hamburg, and Ostrovsky (Crypto 2008) to KDM security, which however only achieves security against passive adversaries. We explain an inherent problem in adapting their techniques to active security, and resolve this problem using a new technical tool called "lossy algebraic filters" (LAFs). We stress that we significantly deviate from the approach of Camenisch, Chandran, and Shoup to obtain KDM security against active adversaries. This allows us to develop a scheme with compact ciphertexts that consist only of a constant number of group elements.

Keywords: key-dependent messages, chosen-ciphertext security, public-key encryption.

1 Introduction

KDM security. An encryption scheme is key-dependent message (KDM) secure if it is secure even against an adversary who has access to encryptions of messages that depend on the secret key. Such a setting arises, e.g., in harddisk encryption [12], computational soundness results in formal methods [8, 3], or specific protocols [15]. KDM security does not follow from standard security [2, 17], and there are indications [23, 6] that KDM security (at least in its most general form) cannot be proven using standard techniques; it seems that dedicated constructions and proof techniques are necessary.[1]

The BHHO approach to KDM-CPA security. Boneh, Halevi, Hamburg, and Ostrovsky [12] (henceforth BHHO) were the first to construct and prove a public-key encryption (PKE) scheme that is KDM secure under chosen-plaintext attacks (KDM-CPA-secure) in the standard model, under the Decisional Diffie-Hellman (DDH) assumption. While they did not prove their scheme secure under messages that *arbitrarily* depend on the secret key, their result encompasses the important case of *circular (CIRC-CPA) security.* Loosely speaking, a PKE scheme is circular secure if it is secure even in a multi-user setting where encryptions of arbitrary secret keys under arbitrary public keys are known. This notion is sufficient for certain applications [15], and can often be extended to stronger forms of KDM security [6, 14]. Inspired by BHHO, KDM-CPA-secure PKE schemes from other computational assumptions followed [5, 13, 28].

*Supported by DFG grant GZ HO 4534/2-1.

[1] We mention, however, that there are semi-generic transformations that *enhance* the KDM security of an already "slightly" KDM-secure scheme [6, 14, 4].

Since we will be using a similar approach, we give a high-level intuition of BHHO's approach. The crucial property of their scheme is that it is *publicly* possible to construct encryptions of the secret key (under the corresponding public key). Thus, encryptions of the secret key itself do not harm the (IND-CPA) security of that scheme. Suitable homomorphic properties of both keys and ciphertexts allow to extend this argument to circular security (for arbitrarily many users/keys), and to affine functions of all keys.

Why the BHHO approach fails to achieve KDM-CCA security. When considering an *active* adversary, we require a stronger form of KDM security. Namely, KDM-CCA, resp. CIRC-CCA security requires security against an adversary who has access to key-dependent encryptions *and* a decryption oracle. (Naturally, to avoid a trivial notion, the adversary is not allow to submit any of those given KDM encryptions to its decryption oracle.) Now if we want to extend BHHO's KDM-CPA approach to an adversary with a decryption oracle, the following problem arises: since it is publicly possible to construct (fresh) encryptions of the secret key, an adversary can generate such an encryption and then submit it to its decryption oracle, thus obtaining the full secret key. Hence, the very property that BHHO use to prove KDM-CPA security seemingly contradicts chosen-ciphertext security.

Our technical tool: lossy algebraic filters (LAFs). Before we describe our approach to KDM-CCA security, let us present the core technical tool we use. Namely, a *lossy algebraic filter* (LAF) is a family of functions, indexed by a public key and a tag. A function from that family takes a vector $X = (X_i)_{i=1}^{\mathtt{n}}$ as input. Now if the tag is *lossy*, then the output of the function reveals only a linear combination of the X_i. If the tag is *injective*, however, then so is the function. We require that there are many lossy tags, which however require a special trapdoor to be found. On the other hand, lossy and injective tags are computationally indistinguishable. This concept is very similar to (parameterized) lossy trapdoor functions [30], and in particular to all-but-many lossy trapdoor functions (ABM-LTFs [24]). However, we do not require efficient inversion, but we do require that lossy functions always reveal *the same* linear combination about the input. In particular, evaluating the same input under many lossy tags will still leave the input (partially) undetermined.

We give a construction of LAFs under the Decision Linear (DLIN) assumption in pairing-friendly groups. Similar to ABM-LTFs, lossy tags correspond to suitably blinded signatures. (This in particular allows to release many lossy tags, while still making the generation of a fresh lossy tag hard for an adversary.) However, unlike with ABM-LTFs, functions with lossy tags always release the same information about its input. Our construction has compact tags with $\mathbf{O}(1)$ group elements, which will be crucial for our KDM-CCA secure encryption scheme.[2]

Our approach to KDM-CCA security. We can now describe our solution to the KDM-CCA dilemma explained above. We will start from a hybrid between the BHHO-like PKE schemes of Brakerski and Goldwasser [13], resp. Malkin et al. [28]. This scheme has compact ciphertexts ($\mathbf{O}(1)$ group elements), and its KDM-CPA security can be proved under the Decisional Composite Residuosity (DCR) assumption. As with the BHHO scheme, the scheme's KDM-CPA security relies on the fact that encryptions of its secret key can be publicly generated. Essentially, our modification consists of adding a suitable authentication tag to each ciphertext. This authentication tag comprises the (encrypted) image of the plaintext message under an LAF. During decryption, a ciphertext is rejected in case of a wrong authentication tag.

In our security proof, all authentication tags for the key-dependent encryptions the adversary gets are made with respect to lossy filter tags. This means that information-theoretically, little information about the secret key is released (even with many key-dependent encryptions, resp. LAF evaluations). However, any decryption query the adversary makes must refer (by the LAF properties) to an injective tag. Hence, in order to place a valid key-dependent decryption query, the

[2]The size of the LAF public key depends on the employed signature scheme. In our main construction, we use Waters signatures, which results in very compact tags, but public keys of $\mathbf{O}(k)$ group elements, where k is the security parameter. Alternatively, at the end of Section 3.2, we sketch an LAF with constant-size (but larger than in our main construction) tags *and* constant-size public keys.

adversary would have to guess the whole (hidden) secret key.[3]

Thus, adding a suitable authentication tag allows us to leverage the techniques by BHHO, resp. Brakerski and Goldwasser, Malkin et al. to chosen-ciphertext attacks. In particular, we obtain a CIRC-CCA-secure PKE scheme with compact ciphertexts (of $\mathbf{O}(1)$ group elements). We prove security under the conjunction of the following assumptions: the DCR assumption (in $\mathbb{Z}_{N^3}^*$), the DLIN assumption (in a pairing-friendly group), and the DDH assumption (somewhat curiously, in the subgroup of order $(P-1)(Q-1)/4$ of $\mathbb{Z}_{N^3}^*$, where $N = PQ$).[4]

Relation to Camenisch et al.'s CIRC-CCA-secure scheme. Camenisch, Chandran, and Shoup [16] present the only other known CIRC-CCA-secure PKE scheme in the standard model. They also build upon BHHO techniques, but instead use a Naor-Yung-style double encryption technique [29] to achieve chosen-ciphertext security. As an authentication tag, they attach to each ciphertext a non-interactive zero-knowledge proof that *either* the encryption is consistent (in the usual Naor-Yung sense), *or* that they know a signature for the ciphertext. Since they build on the original, DDH-based BHHO scheme, they can use Groth-Sahai proofs [22] to prove consistency. Compared to our scheme, their system is less efficient: they require $\mathbf{O}(k)$ group elements per ciphertext, and the secret key can only be encrypted bitwise. However, their sole computational assumption to prove circular security is the DDH (or, more generally, k-Linear) assumption in pairing-friendly groups. One thing to point out is their implicit use of a signature scheme. Their argument is conceptually not unlike our LAF argument. However, since they can apply a hybrid argument to substitute all key-dependent encryptions with random ciphertexts, they only require one-time signatures. Furthermore, the meaning of "consistent ciphertext" and "proof" in our case is very different. (Unlike Camenisch et al., we apply an argument that rests on the *information* that the adversary has about the secret key.)

Note about concurrent work. In a work concurrent to ours, Galindo, Herranz, and Villar [21] define and instantiate a strong notion of KDM security for identity-based encryption (IBE) schemes. Using the IBE→PKE transformation of Boneh, Canetti, Halevi, and Katz [11], they derive a KDM-CCA-secure PKE scheme. Their concrete construction is entropy-based and achieves only a bounded form of KDM security, much like the KDM-secure SKE scheme from [26]. Thus, while their ciphertexts are very compact, they can only tolerate a number of (arbitrary) KDM queries that is linear in the size of the secret key. In particular, it is not clear how to argue that the encryption of a full secret key in their scheme is secure.

2 Preliminaries

Notation. For $n \in \mathbb{N}$, let $[n] := \{1, \ldots, n\}$. Throughout the paper, $k \in \mathbb{N}$ is the security parameter. For a finite set $\mathcal{S}$, $s \leftarrow \mathcal{S}$ denotes the process of sampling s uniformly from $\mathcal{S}$. For a probabilistic algorithm A, $y \leftarrow A(x; R)$ denotes the process of running A on input x and with randomness R, and assigning y the result. We write $y \leftarrow A(x)$ for $y \leftarrow A(x; R)$ with uniformly chosen R. If A's running time is polynomial in k, A is called probabilistic polynomial-time (PPT).

Key-unique SKE schemes. A secret-key encryption (SKE) scheme (E, D) consists of two PPT algorithms. Encryption $\mathsf{E}(K, M)$ takes a key K and a message M, and outputs a ciphertext C. Decryption $\mathsf{D}(K, C)$ takes a key K and a ciphertext C, and outputs a message M. For correctness, we want $\mathsf{D}(K, C) = M$ for all M, all K, and all $C \leftarrow \mathsf{E}(K, M)$. We say that (E, D) is *key-unique* if for every ciphertext C, there is at most one key K with $\mathsf{D}(K, C) \neq \perp$. For instance, ElGamal

[3] We will also have to protect against a re-use of (lossy) authentication tags, and "ordinary", key-independent chosen-ciphertext attacks. This will be achieved by a combination of one-time signatures and 2-universal hash proof systems [19, 27, 25].

[4] Very roughly, we resort to the DDH assumption since we release *partial* information about our secret keys. Whereas the argument of [13, 28] relies on the fact that the secret key sk is completely hidden modulo N, where computations take place in $\mathbb{Z}_N$, we cannot avoid to leak some information about $sk \bmod N$ by releasing LAF images of sk. However, using a suitable message encoding, we *can* argue that sk is completely hidden modulo the coprime order $(P-1)(Q-1)/4$ of quadratic residues modulo N, which enables a reduction to the DDH assumption.

encryption can be interpreted as a key-unique SKE scheme through $\mathsf{E}(x, M) := (g^x, g^y, g^{xy} \cdot M)$ (and the obvious D). This example assumes a publicly known group $\mathbb{G} = \langle g \rangle$ in which the DDH assumption holds.[5] If a larger message space (e.g., $\{0,1\}^*$) is desired, hybrid encryption techniques (which are easily seen to preserve key-uniqueness) can be employed.

PKE schemes. A public-key encryption (PKE) scheme PKE consists of four[6] PPT algorithms $(\mathsf{Pars}, \mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$. The parameter generator $\mathsf{Pars}(1^k)$ outputs public parameters pp such as a group description. Key generation $\mathsf{Gen}(pp)$ outputs a public key pk and a secret key sk. Encryption $\mathsf{Enc}(pp, pk, M)$ takes parameters pp, a public key pk, and a message M, and outputs a ciphertext C. Decryption $\mathsf{Dec}(pp, sk, C)$ takes public parameters pp, a secret key sk, and a ciphertext C, and outputs a message M. For correctness, we want $\mathsf{Dec}(pp, sk, C) = M$ for all M, all $pp \leftarrow \mathsf{Pars}(1^k)$, all $(pk, sk) \leftarrow \mathsf{Gen}(pp)$, and all $C \leftarrow \mathsf{Enc}(pp, pk, M)$.

IND-CPA security of SKE and PKE schemes. An SKE scheme (E, D) is IND-CPA secure iff no efficient adversary A wins the following game with probability non-negligibly away from $1/2$. First, A selects two equal-length messages M_0, M_1, then gets an encryption $\mathsf{E}(K, M_b)$ (for random K and $b \leftarrow \{0,1\}$), and then takes a guess $b' \in \{0,1\}$. During this, A gets access to an encryption oracle $\mathsf{E}(K, \cdot)$. We say that A wins iff $b = b'$. For concrete security analyses, let $\mathsf{Adv}^{\mathsf{ind\text{-}cpa}}_{(\mathsf{E},\mathsf{D}),A}(k)$ denote the probability that A wins this game. This definition can be adapted to the PKE setting by initially giving A the public parameters pp and the public key pk instead of access to an encryption oracle.

(Chameleon) hashing. A hash function H is collision-resistant iff the probability $\mathsf{Adv}^{\mathsf{cr}}_{\mathsf{H},C}(k)$ that C, upon input H, finds $X \neq X'$ with $\mathsf{H}(X) = \mathsf{H}(X')$ is negligible for every PPT C. A chameleon hash function CH is a keyed and randomized hash function in which key generation outputs a keypair (Hpk, Htd). Given a preimage X and randomness R_{CH}, the evaluation key Hpk allows to efficiently evaluate CH, written $\mathsf{CH}_{Hpk}(X; R_{\mathsf{CH}})$. We require collision-resistance in the sense that it is infeasible to find $(X, R_{\mathsf{CH}}) \neq (X', R'_{\mathsf{CH}})$ with $\mathsf{CH}_{Hpk}(X; R_{\mathsf{CH}}) = \mathsf{CH}_{Hpk}(X'; R'_{\mathsf{CH}})$. However, the trapdoor Htd allows to produce collisions, in the following sense: given arbitrary X, R_{CH}, X', Htd allows to efficiently find R'_{CH} with $\mathsf{CH}_{Hpk}(X; R_{\mathsf{CH}}) = \mathsf{CH}_{Hpk}(X'; R'_{\mathsf{CH}})$ for the corresponding Hpk. We require that the distribution of R'_{CH} is uniform given only Hpk and X'.

Signature schemes. A signature scheme Sig consists of three PPT algorithms $(\mathsf{SGen}, \mathsf{Sig}, \mathsf{Ver})$. Key generation $\mathsf{SGen}(1^k)$ outputs a verification key vk and a signing key $sigk$. The signature algorithm $\mathsf{Sig}(sigk, M)$ takes a signing key $sigk$ and a message M and outputs a signature σ. Verification $\mathsf{Ver}(vk, M, \sigma)$ takes a verification key vk, a message M and a potential signature σ and outputs a verdict $b \in \{0,1\}$. For correctness, we require that $\mathsf{Ver}(vk, M, \sigma) = 1$ for all M, all $(vk, sigk) \leftarrow \mathsf{SGen}(1^k)$, and all $\sigma \leftarrow \mathsf{Sig}(sigk, M)$.

(One-time, strong) existential unforgeability. A signature scheme Sig is existentially unforgeable (EUF-CMA secure) iff no PPT forger F wins the following game with non-negligible probability. First, F gets a verification key vk as well as access to a signature oracle $\mathsf{Sig}(sigk, \cdot)$. A wins iff it finally outputs a valid signature σ for a fresh message M that has not yet been queried to $\mathsf{Sig}(sigk, \cdot)$. Let $\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{\mathsf{Sig},A}(k)$ denote the probability that A wins this game. Sig is called one-time existentially unforgeable (OT-EUF-CMA secure) iff no PPT forger F that makes at most one signature query wins the above game with non-negligible probability. Sig is strongly (OT-)EUF-CMA secure iff it is (OT-)EUF-CMA secure as above, but in a game in which an adversary already wins already if it generates a fresh signature for a (perhaps already signed) message. We let $\mathsf{Adv}^{\mathsf{seuf\text{-}cma}}_{\mathsf{Sig},A}(k)$ denote the probability that A wins this strong EUF-CMA security game.

DCR assumption. The Decisional Composite Residuosity (DCR) assumption over a group $\mathbb{Z}^*_{N^{s+1}}$ (for $N = PQ$ with primes P, Q, and $s \geq 1$) states that for every PPT adversary A,

$$\mathsf{Adv}^{\mathsf{dcr}}_{\mathbb{Z}^*_{N^{s+1}},A}(k) \ := \ \Pr\left[A(N, g) = 1\right] - \Pr\left[A(N, g \cdot h) = 1\right],$$

[5]In our application, $\mathbb{G}$ can be made part of the public parameters.
[6]We will only use public parameters for PKE schemes, but not, e.g., for signature schemes.

is negligible, where $g = \tilde{g}^{N^s}$ for uniform $\tilde{g} \in \mathbb{Z}^*_{N^{s+1}}$ is a uniformly chosen N^s-th power, and $h := 1 + N \in \mathbb{Z}^*_{N^{s+1}}$ is a fixed element of order N^s. Damgård and Jurik [20] have shown that the DCR assumptions over $\mathbb{Z}^*_{N^{s+1}}$ and $\mathbb{Z}^*_{N^{s'}+1}$ are equivalent for any s, s'.

DDH and DLIN assumptions. The Decisional Diffie-Hellman (DDH), resp. Decision Linear [9] (DLIN) assumptions over a group $\mathbb{G}$ of (not necessarily prime) order q state that for every PPT adversary A, the respective following functions are negligible:

$$\begin{aligned} \mathsf{Adv}^{\mathsf{ddh}}_{\mathbb{G},A}(k) &:= \Pr\left[A(g, g^x, g^y, g^{xy}) = 1\right] - \Pr\left[A(g, g^x, g^y, g^z) = 1\right], \\ \mathsf{Adv}^{\mathsf{dlin}}_{\mathbb{G},A}(k) &:= \Pr\left[A(g, U_1, U_2, g^{s_0}, U_1^{s_1}, U_2^{s_0+s_1}) = 1\right] - \Pr\left[A(g, U_1, U_2, g^{s_0}, U_1^{s_1}, U_2^{s_2}) = 1\right], \end{aligned}$$

where g is a uniform generator of $\mathbb{G}$, and $U_1, U_2 \leftarrow \mathbb{G}$ and $x, y, z, s_0, s_1, s_2 \leftarrow \mathbb{Z}_q$ are uniform.

Pairings. A (symmetric) pairing is a map $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$ between two cyclic groups $\mathbb{G}$ and $\mathbb{G}_T$ that satisfies $e(g, g) \neq 1$ and $e(g^a, g^b) = e(g, g)^{ab}$ for all generators g of $\mathbb{G}$ and all $a, b \in \mathbb{Z}$.

Waters signatures. In [31], Waters proves the following signature scheme EUF-CMA secure:[7]

- $\mathsf{Gen}(1^k)$ chooses groups $\mathbb{G}, \mathbb{G}_T$ of prime order p, along with a pairing $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$, a generator $g \in \mathbb{G}$, and uniform group elements $g^\omega, H_0, \ldots, H_k \in \mathbb{G}$. Output is

$$vk = (\mathbb{G}, \mathbb{G}_T, e, p, g, (H_i)_{i=0}^k, e(g, g)^\omega), \qquad\qquad sigk = (vk, g^\omega).$$

- $\mathsf{Sig}(sigk, M)$, for $M = (M_i)_{i=1}^k \in \{0,1\}^k$, uniformly chooses $r \leftarrow \mathbb{Z}_p$, and outputs $\sigma := (g^r, g^\omega \cdot (H_0 \prod_{i=1}^k H_i^{M_i})^r)$.
- $\mathsf{Ver}(vk, M, (\sigma_0, \sigma_1))$, outputs 1 iff $e(g, \sigma_1) = e(g, g)^\omega \cdot e(\sigma_0, H_0 \prod_{i=1}^k H_i^{M_i})$.

KDM-CCA and CIRC-CCA security. Let $n = n(k)$ and let PKE be a PKE scheme with message space $\mathcal{M}$. PKE is chosen-ciphertext secure under key-dependent message attacks (n-KDM-CCA secure) iff

$$\mathsf{Adv}^{\mathsf{kdm\text{-}cca}}_{\mathsf{PKE},n,A}(k) := \Pr\left[\mathsf{Exp}^{\mathsf{kdm\text{-}cca}}_{\mathsf{PKE},n,A}(k) = 1\right] - 1/2$$

is negligible for all PPT A, where experiment $\mathsf{Exp}^{\mathsf{kdm\text{-}cca}}_{\mathsf{PKE},n,A}$ is defined as follows. First, the experiment tosses a coin $b \leftarrow \{0,1\}$, and samples public parameters $pp \leftarrow \mathsf{Pars}(1^k)$ and n keypairs $(pk_i, sk_i) \leftarrow \mathsf{Gen}(pp)$. Then A is invoked with input pp and $(pk_i)_{i=1}^n$, and access to two oracles:

- a KDM oracle $\mathcal{KDM}_b(\cdot,\cdot)$ that maps $i \in [n]$ and a function $f : (\{0,1\}^*)^n \to \{0,1\}^*$ to a ciphertext $C \leftarrow \mathsf{Enc}(pp, pk_i, M)$. If $b = 0$, then $M = f((sk_i)_{i=1}^n)$; else, $M = 0^{|f((sk_i)_{i=1}^n)|}$.
- a decryption oracle $\mathcal{DEC}(\cdot,\cdot)$ that takes as input an index $i \in [n]$ and a ciphertext C, and outputs $\mathsf{Dec}(pp, sk_i, C)$.

When A finally generates an output $b' \in \{0,1\}$, the experiment outputs 1 if $b = b'$ (and 0 else). We require that (a) A never inputs a ciphertext C to $\mathcal{DEC}$ that has been produced by $\mathcal{KDM}_b$ (for the same index i), and (b) A only specifies PPT-computable functions f that always output messages of the same length. As a relevant special case, PKE is *n-CIRC-CCA-secure* if it is n-KDM-CCA secure against all A that only query $\mathcal{KDM}_b$ with functions $f \in \mathcal{F}$ for

$$\mathcal{F} := \{f_j : f_j((sk_i)_{i=1}^n) = sk_j\}_{j\in[n]} \cup \{f_M : f_M((sk_i)_{i=1}^n) = M\}_{M\in\mathcal{M}}.$$

(Technically, what we call "circular security" is called "clique security" in [12]. However, our notion of circular security implies that of [12].) Our main result will be a PKE scheme that is n-CIRC-CCA-secure for all polynomials $n = n(k)$.

[7] In fact, our description is a slight folklore variant of Waters's scheme. The original scheme features elements g^α, g^β in vk, so that $e(g^\alpha, g^\beta)$ takes the role of $e(g, g)^\omega$.

3 Lossy algebraic filters

3.1 Definition

Informal description. An $(\ell_{\mathsf{LAF}}, \mathfrak{n})$-lossy algebraic filter (LAF) is a family of functions indexed by a public key Fpk and a tag t. A function $\mathsf{LAF}_{Fpk,t}$ from the family maps an input $X = (X_i)_{i=1}^{\mathfrak{n}} \in \mathbb{Z}_p^{\mathfrak{n}}$ to an output $\mathsf{LAF}_{Fpk,t}(X)$, where p is an ℓ_{LAF}-bit prime contained in the public key.

The crucial property of an LAF is its lossiness. Namely, for a given public key Fpk, we distinguish *injective* and *lossy* tags.[8] For an injective tag t, the function $\mathsf{LAF}_{Fpk,t}(\cdot)$ is injective, and thus has an image of size $p^{\mathfrak{n}}$. However, if t is lossy, then $\mathsf{LAF}_{Fpk,t}(\cdot)$ only depends on a linear combination $\sum_{i=1}^{\mathfrak{n}} \omega_i X_i \bmod p$ of its input. In particular, different X with the same value $\sum_{i=1}^{\mathfrak{n}} \omega_i X_i \bmod p$ are mapped to the same image. Here, the coefficients $\omega_i \in \mathbb{Z}_p$ only depend on Fpk (but not on t). For a lossy tag t, the image of $\mathsf{LAF}_{Fpk,t}(\cdot)$ is thus of size at most p. Note that the modulus p is public, while the coefficients ω_i may be (and in fact will have to be) computationally hidden.

For this concept to be useful, we require that (a) lossy and injective tags are computationally indistinguishable, (b) lossy tags can be generated using a special trapdoor, but (c) new lossy (or, rather, non-injective) tags cannot be found efficiently without that trapdoor, even when having seen polynomially many lossy tags before. In view of our application, we will work with structured tags: each tag $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ consists of a *core tag* t_{c} and an *auxiliary tag* t_{a}. The auxiliary tag will be a ciphertext part that is authenticated by a filter image.

Definition 3.1. *An* $(\ell_{\mathsf{LAF}}, \mathfrak{n})$-lossy algebraic filter (LAF) LAF *consists of three PPT algorithms:*

Key generation. $\mathsf{FGen}(1^k)$ *samples a keypair* (Fpk, Ftd)*. The public key* Fpk *contains an* ℓ_{LAF}*-bit prime* p *and the description of a tag space* $\mathcal{T} = \mathcal{T}_{\mathsf{c}} \times \{0,1\}^*$ *for efficiently samplable* $\mathcal{T}_{\mathsf{c}}$*. A tag* $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ *consists of a core tag* $t_{\mathsf{c}} \in \mathcal{T}_{\mathsf{c}}$ *and an auxiliary tag* $t_{\mathsf{a}} \in \{0,1\}^*$*. A tag may be injective, or lossy, or neither.* Ftd *is a trapdoor that will allow to sample lossy tags.*

Evaluation. $\mathsf{FEval}(Fpk, t, X)$*, for a public key* Fpk *and a tag* $t = (t_{\mathsf{c}}, t_{\mathsf{a}}) \in \mathcal{T}$*, maps an input* $X = (X_i)_{i=1}^{\mathfrak{n}}) \in \mathbb{Z}_p^{\mathfrak{n}}$ *to a unique output* $\mathsf{LAF}_{Fpk,t}(X)$*.*

Lossy tag generation. $\mathsf{FTag}(Ftd, t_{\mathsf{a}})$*, for a trapdoor* Ftd *and* $t_{\mathsf{a}} \in \{0,1\}^*$*, samples a core tag* t_{c} *such that* $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ *is lossy.*

We require the following:

Lossiness. *The function* $\mathsf{LAF}_{Fpk,t}(\cdot)$ *is injective if* t *is injective. If* t *is lossy, then* $\mathsf{LAF}_{Fpk,t}(X)$ *depends only on* $\sum_{i=1}^{\mathfrak{n}} \omega_i X_i \bmod p$ *for* $\omega_i \in \mathbb{Z}_p$ *that only depend on* Fpk*.*

Indistinguishability. *Lossy tags are indistinguishable from random tags:*

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathsf{LAF},A}(k) := \Pr\left[A(1^k, Fpk)^{\mathsf{FTag}(Ftd,\cdot)} = 1\right] - \Pr\left[A(1^k, Fpk)^{\mathcal{O}_{\mathcal{T}_{\mathsf{c}}}(\cdot)} = 1\right]$$

is negligible for all PPT A*, where* $(Fpk, Ftd) \leftarrow \mathsf{FGen}(1^k)$*, and* $\mathcal{O}_{\mathcal{T}_{\mathsf{c}}}(\cdot)$ *is the oracle that ignores its input and samples a random core tag* t_{c}*.*

Evasiveness. *Non-injective (and in particular lossy) tags are hard to find, even given multiple lossy tags:*

$$\mathsf{Adv}^{\mathsf{eva}}_{\mathsf{LAF},A}(k) := \Pr\left[t \text{ non-injective} \;\middle|\; t \leftarrow A(1^k, Fpk)^{\mathsf{FTag}(Ftd,\cdot)}\right]$$

is negligible with $(Fpk, Ftd) \leftarrow \mathsf{FGen}(1^k)$*, and for any PPT algorithm* A *that never outputs a tag obtained through oracle queries (i.e.,* A *never outputs* $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ *when* t_{c} *has been obtained by an oracle query* $\mathsf{FTag}(Ftd, t_{\mathsf{a}})$*).*

3.2 Construction

Intuition. We present a construction based on the DLIN problem in a group $\mathbb{G}$ of order p with symmetric pairing $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$. Essentially, each tag corresponds to $\mathfrak{n}$ DLIN-encrypted

[8] Technically, there may also be tags that are neither injective nor lossy.

Waters signatures. If the signatures are valid, the tag is lossy. The actual filter maps an input $X = (X_i)_{i=1}^{\mathfrak{n}} \in \mathbb{Z}_p^{\mathfrak{n}}$ to the tuple

$$\mathsf{LAF}_{Fpk,t}(X) \;:=\; \mathbf{M} \circ X \;:=\; (\prod_{j=1}^{\mathfrak{n}} \mathbf{M}_{i,j}^{X_j})_{j=1}^{\mathfrak{n}} \;\in\; \mathbb{G}_T^{\mathfrak{n}}, \tag{1}$$

where the matrix $\mathbf{M} = (\mathbf{M}_{i,j})_{i,j \in [\mathfrak{n}]} \in \mathbb{G}_T^{\mathfrak{n} \times \mathfrak{n}}$ is computed from public key and tag. Note that this mapping is lossy if and only if the matrix

$$\widetilde{\mathbf{M}} \;:=\; (\widetilde{\mathbf{M}}_{i,j}) \;:=\; (\mathrm{dlog}_{e(g,g)}(\mathbf{M}_{i,j}))_{i,j} \;\in\; \mathbb{Z}_p^{\mathfrak{n} \times \mathfrak{n}} \tag{2}$$

of discrete logarithms (to some arbitrary basis $e(g, g) \in \mathbb{G}_T$) is non-invertible.

For a formal description, let $\ell_{\mathsf{LAF}}(k), \mathfrak{n}(k)$ be two functions.

Key generation. $\mathsf{FGen}(1^k)$ generates cyclic groups $\mathbb{G}, \mathbb{G}_T$ of prime order p (where p has bitlength $\lfloor \log_2(p) \rfloor = \ell_{\mathsf{LAF}}(k)$), and a symmetric pairing $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$. Then FGen chooses

- a generator $g \in \mathbb{G}$ and a uniform exponent $\omega \leftarrow \mathbb{Z}_p$,
- uniform group elements $U_1, \ldots, U_{\mathfrak{n}} \leftarrow \mathbb{G}$, $H_0, \ldots, H_k \leftarrow \mathbb{G}$, and
- a keypair (Hpk, Htd) for a chameleon hash function$\mathsf{CH} : \{0,1\}^* \to \{0,1\}^k$.

FGen finally outputs

$$\begin{aligned} Fpk \;&:=\; (\mathbb{G}, \mathbb{G}_T, e, p, g, (H_i)_{i=0}^k, (U_i)_{i=1}^{\mathfrak{n}}, W := e(g,g)^{\omega}, Hpk) \\ Ftd \;&:=\; (Fpk, g^{\omega}, Htd). \end{aligned}$$

For convenience, write $U_i = g^{u_i}$ for suitable (unknown) exponents u_i.

Tags. (Core) tags are of the form

$$t_{\mathsf{c}} \;:=\; (R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}, R_{\mathsf{CH}}) \;\in\; \mathbb{G} \times \mathbb{G} \times \mathbb{G}^{\mathfrak{n} \times \mathfrak{n}} \times \mathcal{R}_{\mathsf{CH}}$$

(for CH's randomness space $\mathcal{R}_{\mathsf{CH}}$), where we require $e(U_{j'}, S_{i,j}) = e(U_j, S_{i,j'})$ whenever $i \notin \{j, j'\}$. This means we can write

$$R = g^r, \qquad\qquad \widetilde{S}_i = g^{\widetilde{s}_i}, \qquad\qquad S_{i,j} = U_j^{s_i} \qquad (i \neq j)$$

for suitable $r, s_i, \widetilde{s}_i$. To a tag $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ (with auxiliary part $t_{\mathsf{a}} \in \{0,1\}^*$), we associate the matrix $\mathbf{M} = (\mathbf{M}_{i,j})_{i,j=1}^{\mathfrak{n}} \in \mathbb{G}_T^{\mathfrak{n} \times \mathfrak{n}}$ with

$$\begin{aligned} \mathbf{M}_{i,j} \;&=\; e(U_j, \widetilde{S}_i) \cdot e(g, S_{i,j}) \;=\; e(g,g)^{u_j(\widetilde{s}_i + s_i)} \qquad (i \neq j) \\ \mathbf{M}_{i,i} \;&=\; \frac{e(g, S_{i,i})}{W \cdot e(H_0 \prod_{i=1}^k H_i^{T_i}, R)} \end{aligned} \tag{3}$$

for $(T_i)_{i=1}^k := \mathsf{CH}_{Hpk}(R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}, t_{\mathsf{a}}; R_{\mathsf{CH}})$. If the matrix $\widetilde{\mathbf{M}}$ of discrete logarithms (see (2)) is invertible, we say that t is injective; if $\widetilde{\mathbf{M}}$ has rank 1, then t is lossy. Thus, for lossy tags, $\mathbf{M}_{i,j} = e(g,g)^{u_j(\widetilde{s}_i + s_i)}$ for *all* i, j.

Evaluation. $\mathsf{FEval}(Fpk, t, X)$, for $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$, $t_{\mathsf{a}} \in \{0,1\}^*$, $X = (X_i)_{i=1}^{\mathfrak{n}} \in \mathbb{Z}_p^{\mathfrak{n}}$, and Fpk and t_{c} as above, computes $\mathbf{M}$ as in (3) and then $(Y_i)_{i=1}^{\mathfrak{n}} := \mathsf{LAF}_{Fpk,t}(X) \in \mathbb{G}_T^{\mathfrak{n}}$ as in (1).

Lossiness. If we write $Y_i = e(g,g)^{y_i}$, the definition of FEval implies $(y_i)_{i=1}^{\mathfrak{n}} = \widetilde{\mathbf{M}} \cdot X$. Since injective tags satisfy that $\widetilde{\mathbf{M}}$ is invertible, they lead to injective functions $\mathsf{LAF}_{Fpk,t}(\cdot)$. But for a lossy tag, $\widetilde{\mathbf{M}}_{i,j} = u_j(\widetilde{s}_i + s_i)$, so that

$$y_i \;=\; \sum_{j=1}^{\mathfrak{n}} u_j(\widetilde{s}_i + s_i) X_j \;=\; (\widetilde{s}_i + s_i) \cdot \sum_{j=1}^{\mathfrak{n}} u_j X_j \mod p.$$

Specifically, $\mathsf{LAF}_{Fpk,t}(X)$ depends only on $\sum_i \omega_i X_i \bmod p$ for $\omega_i := u_i$.

Lossy tag generation. $\mathsf{FTag}(Ftd, t_\mathsf{a})$, for Ftd as above and $t_\mathsf{a} \in \{0,1\}^*$, first chooses a random CH-image $T = (T_i)_{i=1}^k \in \{0,1\}^k$ that can later be explained, using Htd, as the CH-image of an arbitrary preimage. FTag then chooses uniform $r, s_i, \widetilde{s_i} \leftarrow \mathbb{Z}_p$ and sets (for $i \neq j$)

$$R := g^r, \quad \widetilde{S_i} := g^{\widetilde{s_i}}, \quad S_{i,j} := U_j^{s_i}, \quad S_{i,i} := U_i^{\widetilde{s_i}+s_i} \cdot g^\omega \cdot \left(H_0 \prod_{i=1}^k H_i^{T_i}\right)^r. \tag{4}$$

Finally, FTag chooses R_CH with $\mathsf{CH}_{Hpk}(R, (\widetilde{S_i})_{i=1}^\mathfrak{n}, (S_{i,j})_{i,j=1}^\mathfrak{n}, t_\mathsf{a}; R_\mathsf{CH}) = T$ and outputs $t_\mathsf{c} = (R, (\widetilde{S_i})_{i=1}^\mathfrak{n}, (S_{i,j})_{i,j=1}^\mathfrak{n}, R_\mathsf{CH})$. Intuitively, t_c consists of $\mathfrak{n}$ DLIN encryptions (with correlated randomness $s_i, \widetilde{s_i}$) of Waters signatures $(g^r, g^\omega \cdot (H_0 \prod_{i=1}^k H_i^{T_i})^r)$ for message T. Indeed, substituting into (3) yields

$$\mathbf{M}_{i,i} := \frac{e(g,g)^{u_i(\widetilde{s_i}+s_i)} \cdot W \cdot e(g, (H_0 \prod_{i=1}^k H_i^{T_i})^r)}{W \cdot e(g^r, H_0 \prod_{i=1}^k H_i^{T_i})} = e(g,g)^{u_i(\widetilde{s_i}+s_i)}.$$

Hence, $\widetilde{\mathbf{M}}_{i,j} = u_j(\widetilde{s_i} + s_i)$ for *all* i, j, and thus the resulting tag $t = (t_\mathsf{c}, t_\mathsf{a})$ is lossy.

A generalization with constant-size evaluation keys. The LAF LAF above inherits a rather large public key of $\mathbf{O}(k)$ group elements from Waters signatures. We now sketch how to generalize LAF to any structure-preserving signature scheme; plugging in, e.g., the DLIN-based signature scheme of Abe et al. [1] yields an LAF with constant-size tags and keys. (Compared to LAF, tags will be larger, however.) The idea is to have tags that directly contain a matrix $\mathbf{M} \in \mathbb{G}_T^{\mathfrak{n}\times\mathfrak{n}}$ as above, along with a DLIN ciphertext C, and a Groth-Sahai [22] proof π. The statement proved by π is that either $\mathbf{M}$ is injective (e.g., in the sense that there exist $V_1, \ldots, V_\mathfrak{n}$ with $\mathbf{M}_{i,i} = e(U_i, V_i) \cdot e(g,g)$ and $\mathbf{M}_{i,j} = e(U_j, V_i)$ for fixed elements U_i from the public key and all $i \neq j$), or that C contains a fresh signature (e.g., of a chameleon hash value T as above). Evaluation of this LAF takes place as in (1). Lossy tags can be generated using a signing key and proving the "or" branch of the statement. The soundness of Groth-Sahai proofs ensures that any adversarially produced lossy tag (with lossy $\mathbf{M}$) would imply a fresh forged signature.

Other instances and further applications of LAFs. Since LAFs can be seen as "disguised signature schemes", it seems interesting to try to convert other signature schemes (and in particular schemes that do not require pairing-friendly groups) to LAFs. Besides, LAFs would seem potentially useful in other settings, specifically in settings with inherently many challenges (e.g., the selective-opening setting [7]).

3.3 Security proof

Theorem 3.2. *If the DLIN assumption holds in* $\mathbb{G}$*, and* CH *is a chameleon hash function, then the LAF construction* LAF *from Section 3.2 satisfies Definition 3.1.*

The lossiness of LAF has already been discussed in Section 3.2. We prove indistinguishability and evasiveness separately.

Lemma 3.3. *For every adversary A on* LAF*'s indistinguishability, there exists a DLIN distinguisher B such that*

$$\mathsf{Adv}^{\mathsf{ind}}_{\mathsf{LAF},A}(k) \;=\; \mathfrak{n} \cdot \mathsf{Adv}^{\mathsf{dlin}}_B(k). \tag{5}$$

Intuitively, to see Lemma 3.3, observe that lossy tags differ from random tags only in their $S_{i,i}$ components, and in how the CH randomness R_CH is generated. For lossy tags, the $S_{i,i}$ are (parts of) DLIN ciphertexts, which are pseudorandom under the DLIN assumption. Furthermore, the uniformity property of CH guarantees that the distribution of R_CH is the same for lossy and random tags.

Proof. Assume a PPT adversary A. We proceed in games. In Game i, A gets an input Fpk and interacts with an oracle $\mathcal{O}_i$. Let out_i denote A's output in Game i.

In **Game** 1, we let $\mathcal{O}_1(\cdot) := \mathsf{FTag}(Ftd, \cdot)$, where Ftd is the trapdoor initially sampled alongside Fpk. Thus, $\mathcal{O}_1(t_{\mathsf{a}})$ outputs core tags $t_{\mathsf{c}} = (R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}, R_{\mathsf{CH}})$ generated as in (4).

In **Game** $2.i^*$ (for $0 \leq i^* \leq \mathfrak{n}$), we let $\mathcal{O}_2$ generate core tags as in Game 1, but with independently and uniformly chosen $S_{i,i} \in \mathbb{G}$ for $i \leq i^*$. Note that Game 2.0 is equivalent to Game 1. Let furthermore Game 2 be defined as Game $2.\mathfrak{n}$. We claim

$$\Pr[out_1 = 1] - \Pr[out_2 = 1] \;=\; \Pr[out_{2.0} = 1] - \Pr[out_{2.\mathfrak{n}} = 1] \;=\; \mathfrak{n} \cdot \mathsf{Adv}_B^{\mathsf{dlin}}(k) \qquad (6)$$

for a suitable DLIN distinguisher B. Namely, B uniformly chooses $i^* \in [\mathfrak{n}]$ and parses its DLIN challenge as $(g, U, U_{i^*}, g^{\widetilde{s_{i^*}}}, U^{s_{i^*}}, C)$, where $C = U_{i^*}^{\widetilde{s_{i^*}}+s_{i^*}}$ or $C \in \mathbb{G}$ is uniform. B then first re-randomizes its input to obtain many tuples $(g^{\widetilde{s_{i^*,\ell}}}, U^{s_{i^*,\ell}}, C_\ell)$, where (a) the $\widetilde{s_{i^*,\ell}}, s_{i^*,\ell}$ are independently and uniformly random, and (b) $C_\ell = U_{i^*}^{\widetilde{s_{i^*,\ell}}+s_{i^*,\ell}}$ iff $X = U_{i^*}^{\widetilde{s_{i^*}}+s_{i^*}}$ (otherwise, all C_ℓ are independently and uniformly random). Next, B simulates Game $2.(i^*-1)$ or Game $2.i^*$, depending on its own challenge C. Concretely, to prepare a key Fpk for A, B sets $U_j = U^{\alpha_j}$ for all $j \neq i^*$ and uniform $\alpha_j \leftarrow \mathbb{Z}_p$. (Like Game $2.i^*$, B chooses $\omega \leftarrow \mathbb{Z}_p$ and a CH keypair (Hpk, Htd) on its own.) When answering A's ℓ-th oracle query, B proceeds as in Game $2.i^*$, but sets up (a) $\widetilde{S_{i^*}} = g^{\widetilde{s_{i^*,\ell}}}$, (b) $S_{i,i}$ as in Game 1 for $i > i^*$, (c) $S_{i,i} \leftarrow \mathbb{G}$ uniformly (as in Game 2) for $i < i^*$, (d) $S_{i^*,j} = (U^{s_{i^*,\ell}})^{\alpha_j} = U_j^{s_{i^*,\ell}}$ for $j \neq i^*$, (e) $S_{i^*,i^*} = C_\ell \cdot g^\omega \cdot \left(H_0 \prod_{i=1}^k H_i^{T_i}\right)^r$. This implicitly sets $s_{i^*} = s_{i^*,\ell}$ and $\widetilde{s_{i^*}} = \widetilde{s_{i^*,\ell}}$. (All other $s_i, \widetilde{s}_i$ are chosen by B.) Furthermore, if $C = U_{i^*}^{\widetilde{s_{i^*}}+s_{i^*}}$, this setting of S_{i,i^*} yields Game $2.(i^*-1)$; but if C is uniform, then all C_i are independently uniform, and we obtain Game $2.i^*$. We get (6).

In **Game** 3, we choose the hash values R_{CH} in the core tags output by $\mathcal{O}_3$ uniformly and independently. Recall that up to Game 2, R_{CH} was instead chosen as follows: first choose a random CH-output T, and later select R_{CH} such that $\mathsf{CH}_{Hpk}(R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}}) = T$ holds. By definition of chameleon hashing, this induces a uniform distribution of R_{CH}. Moreover, T is not used in Game 2 or Game 3. Hence, the change in Game 3 is merely conceptual, and we obtain

$$\Pr[out_3 = 1] \;=\; \Pr[out_2 = 1].$$

Now note that in Game 3, the tags t_{c} output by $\mathcal{O}_3$ are random tags. Taking things together, (5) follows as desired. □

Lemma 3.4. *For every adversary A on* LAF*'s evasiveness, there exist adversaries B, C, and F with*

$$\mathsf{Adv}_{\mathsf{LAF},A}^{\mathsf{eva}}(k) \leq \left|\mathsf{Adv}_{\mathsf{LAF},B}^{\mathsf{ind}}(k)\right| + \mathsf{Adv}_{\mathsf{CH},C}^{\mathsf{cr}}(k) + \mathsf{Adv}_{\mathsf{Sig_{Wat}},F}^{\mathsf{euf\text{-}cma}}(k). \qquad (7)$$

Intuitively, Lemma 3.4 holds because lossy (or, rather, non-injective) tags correspond to DLIN-encrypted Waters signatures. Hence, even after seeing many lossy tags (i.e., encrypted signatures), an adversary cannot produce a fresh encrypted signature. We note that the original Waters signatures from [31] are re-randomizable and thus not *strongly* unforgeable. To achieve evasiveness, we have thus used a chameleon hash function, much like Boneh et al. [10] did to make Waters signatures strongly unforgeable.

Proof. Assume a PPT adversary A. Again, we proceed in games. Let $\mathsf{bad_i}$ denote the event that A's output in Game i is a fresh non-injective tag. In **Game** 1, A gets input Fpk and interacts with an $\mathsf{FTag}(Ftd, \cdot)$ oracle. By definition,

$$\Pr[\mathsf{bad_1}] \;=\; \mathsf{Adv}_{\mathsf{LAF},A}^{\mathsf{eva}}(k).$$

To describe **Game** 2, denote A's output by $t^* = (t_{\mathsf{c}}^*, t_{\mathsf{a}}^*)$, for $t_{\mathsf{c}}^* = (R^*, (\widetilde{S}_i^{\,*})_{i=1}^{\mathfrak{n}}, (S_{i,j}^*)_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}}^*)$ Denote by $\mathsf{bad_{coll}}$ the event that t^* induces a CH-collision in the sense that

$$T^* = \mathsf{CH}_{Hpk}(R^*, (\widetilde{S}_i^{\,*})_{i=1}^{\mathfrak{n}}, (S_{i,j}^*)_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}}^*) = \mathsf{CH}_{Hpk}(R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}}) = T$$

for some hash value T associated with an FTag-output $t_\mathsf{c} = (R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}})$ (and the corresponding query t_a). In Game 2, we abort (and do not raise event bad_2) if $\mathsf{bad}_{\mathsf{coll}}$ occurs. Intuitively, we would expect to use CH's collision resistance directly to argue that $\mathsf{bad}_{\mathsf{coll}}$ occurs only negligibly often. However, both in Game 1 and Game 2, we use CH's trapdoor Htd to construct lossy tags for A.

Hence, we first argue that $\mathsf{bad}_{\mathsf{coll}}$ occurs with essentially the same probability in a modified **Game** $1'$, in which A gets random tags instead of lossy tags as oracle answers. Indeed, since lossy and random tags are indistinguishable by Lemma 3.3, and $\mathsf{bad}_{\mathsf{coll}}$ is efficiently recognizable from A's view, we obtain

$$\Pr\left[\mathsf{bad}_{\mathsf{coll}} \text{ in Game } 1'\right] - \Pr\left[\mathsf{bad}_{\mathsf{coll}} \text{ in Game } 1\right] \;=\; \mathsf{Adv}^{\mathsf{ind}}_{\mathsf{LAF},B}(k)$$

for a suitable adversary B on LAF's indistinguishability. Furthermore, since in Game $1'$, the CH-trapdoor Htd is not required, we have

$$\Pr\left[\mathsf{bad}_{\mathsf{coll}} \text{ in Game } 1'\right] \;=\; \mathsf{Adv}^{\mathsf{cr}}_{\mathsf{CH},C}(k)$$

for a suitable collision-finder C. However, Game 1 and Game 2 only differ when $\mathsf{bad}_{\mathsf{coll}}$ occurs, and so we finally get

$$\left|\Pr\left[\mathsf{bad}_2\right] - \Pr\left[\mathsf{bad}_1\right]\right| \;\leq\; \Pr\left[\mathsf{bad}_{\mathsf{coll}} \text{ in Game } 1\right] \;\leq\; \left|\mathsf{Adv}^{\mathsf{ind}}_{\mathsf{LAF},B}(k)\right| + \mathsf{Adv}^{\mathsf{cr}}_{\mathsf{CH},C}(k).$$

The final reduction. Now that CH-collisions are excluded, we can finally conclude that any occurence of bad_2 means that A has forged a Waters signature. Concretely, we show that

$$\Pr\left[\mathsf{bad}_2\right] \;=\; \mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{\mathsf{Sig}_{\mathsf{Wat}},F}(k) \tag{8}$$

for a suitable forger F that attacks $\mathsf{Sig}_{\mathsf{Wat}}$ and internally simulates Game 2 with A. Namely, F gets as input a $\mathsf{Sig}_{\mathsf{Wat}}$ public key $(\mathbb{G}, \mathbb{G}_T, e, p, g, (H_i)_{i=0}^k, W := e(g,g)^\omega)$. F extends this public key to an LAF public key Fpk by picking $U_i = g^{u_i}$ and Hpk. (In particular, F knows all u_i and Htd.) Upon an FTag-query from A, F constructs elements $\widetilde{S}_i$ and $S_{i,j}$ (for $i \neq j$) exactly as in (4); note, however, that F cannot directly compute the $S_{i,i}$, since F does not know g^ω. Instead, F requests a $\mathsf{Sig}_{\mathsf{Wat}}$ signature for the message $T \in \{0,1\}^k$ (as derived in (4)). Such a signature is of the form

$$(g^r, g^\omega \cdot \left(H_0 \prod_{i=1}^k H_i^{T_i}\right)^r),$$

from which F can compute the elements R and $S_{i,i}$ as in (4). Since F also knows the CH-trapdoor Htd, this allows to construct lossy tags exactly as FTag would do in Game 2.

It remains to describe how F extracts a $\mathsf{Sig}_{\mathsf{Wat}}$-signature out of a lossy tag $t = (t_\mathsf{c}, t_\mathsf{a})$ that A finally outputs. By our definition of tags, we may assume that $t_\mathsf{c} = (R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}, R_{\mathsf{CH}})$ is of the form $R = g^r$, $\widetilde{S}_i = g^{\widetilde{s}_i}$, and $S_{i,j} = U_j^{s_i}$ for suitable $r, s_i, \widetilde{s}_i$ and all $i \neq j$. Furthermore, since t_c is lossy,

$$\operatorname{rank}(\widetilde{\mathbf{M}}) < \mathfrak{n} \implies \exists i : \widetilde{\mathbf{M}}_{i,i} = u_i(\widetilde{s}_i + s_i) \implies \exists i : S_{i,i} = U_i^{\widetilde{s}_i + s_i} \cdot g^\omega \cdot \left(H_0 \prod_{i=1}^k H_i^{T_i}\right)^r. \tag{9}$$

Since F knows all u_i, it can compute

$$\sigma_i \;:=\; \frac{S_{i,i}}{\widetilde{S}_i^{\,u_i} \cdot S_{i,j}^{u_i/u_j}} \;=\; \frac{S_{i,i}}{U_i^{\widetilde{s}_i + s_i}}$$

for all i (and some $j \neq i$). By (9), for some i, the pair (R, σ_i) forms a valid $\mathsf{Sig}_{\mathsf{Wat}}$ signature for $T = \mathsf{CH}_{Hpk}(R, (\widetilde{S}_i)_{i=1}^{\mathfrak{n}}, (S_{i,j})_{i,j=1}^{\mathfrak{n}}; R_{\mathsf{CH}})$. Because Game 2 aborts in case of a CH-collision, we may further assume that T is a message for which F has not yet requested a signature. Consequently, F can output a forged signature for a fresh message whenever bad_2 occurs. This yields (8). Putting things together finally gives (7). □

Combining Lemma 3.3, Lemma 3.4, and the fact that Waters signatures are EUF-CMA secure already under the CDH assumption, we obtain Theorem 3.2.

4 CIRC-CCA-secure encryption scheme

4.1 The scheme

Intuition. The PKE scheme we are about to present borrows ideas from the KDM-CPA-secure PKE schemes from [13, 28]. Both of these schemes build upon BHHO [12], and in particular allow the public generation of key-dependent ciphertexts; besides, both schemes work in rings $\mathbb{Z}_{N^i}$ and are proven secure under the DCR assumption. However, [13] use several $\mathbb{Z}_{N^i}$-elements to encrypt a "blinding element" (which is used to hide the actual payload message), whereas [28] use only one $\mathbb{Z}_{N^i}$-element in a larger ring (i.e., with larger i) for this purpose. One consequence is that the secret key in [13] consists of several smaller "$\mathbb{Z}_N$-compartments", while the secret key of [28] consists of one large $\mathbb{Z}_{N^j}$-element (for $j > 1$). In our scheme, we will combine both concepts: our ciphertext encrypts a "blinding element" in several $\mathbb{Z}_{N^i}$-elements, each of which encodes a $\mathbb{Z}_{N^j}$-component. Our secret key thus consists of several larger $\mathbb{Z}_{N^j}$-components. This allows to use ideas from [13] in a setting in which some parts of the key are leaked (through key-dependent encryptions).

As we have mentioned, the schemes of [13, 28] allow to publicly generate ciphertexts which decrypt to a secret key component. This property enables a simulation that shows KDM-CPA security; namely, a simulation can now generate key-dependent challenge ciphertexts for an internally simulated adversary. However, this property is also highly dangerous in combination with a decryption oracle (as in KDM-CCA security), since now also an adversary can construct key-dependent ciphertexts and ask for their decryption. Our extension thus essentially consists in adding an "authentication tag" that causes the rejection of key-dependent decryption queries. The technical difficulty in this is that our simulation needs to add valid authentication tags to all simulated key-dependent ciphertexts. However, at the same time, the adversary must not be able to generate a valid authentication for a fresh key-dependent ciphertext. (Thus, the simulation will require some kind of leverage over the adversary.)

We resolve this problem with lossy algebraic filters. Essentially, we add an (encrypted) filter image of the encrypted message to each ciphertext as authentication tag. In the proof, all these filter images will be generated with respect to a lossy filter tag, so the image reveals little information about the encrypted secret keys. (Furthermore, even with different lossy filter tags, *always the same* information about the encrypted secret key is revealed.) At the same time, each decryption query of the adversary will refer to an injective filter tag (unless the adversary breaks the filter's evasiveness property). Thus, the adversary would have to predict the full secret key to generate a valid authentication tag for a key-dependent ciphertext. Since public key and (simulated) KDM ciphertexts reveal little information about the secret key, the adversary will fail to generate the right authentication tag with high probability.

The above strategy ensures that an adversary cannot (successfully) submit any key-dependent ciphertexts for decryption. However, we still need to ensure that our scheme withstands ordinary (i.e., key-independent) chosen-ciphertext attacks. To this end, we will additionally authenticate ciphertexts using a 2-universal hash proof system [19, 27, 25]. (Roughly speaking, this hash proof system ensures that the adversary does not gain any new information through key-independent decryption queries.) Our use of the 2-universal hash proof system is rather implicit – in our scheme below, hash proofs consist of the group elements $(G_1, G_2) = (g_1^r, g_2^r)$ and $(\widetilde{G}_1, \widetilde{G}_2) = (g_1^{\widetilde{r}}, g_2^{\widetilde{r}})$.

Setting and ingredients. First, we assume an algorithm GenN that outputs ℓ_N-bit Blum integers $N = PQ$ along with their prime factors P and Q. If N is clear from the context, we write $\mathbb{G}_{\mathsf{rnd}}$ and $\mathbb{G}_{\mathsf{msg}}$ for the unique subgroups of $\mathbb{Z}_{N^3}^*$ of order $(P-1)(Q-1)/4$, resp. N^2. We also write $h := 1 + N \bmod N^3$, so $\langle h \rangle = \mathbb{G}_{\mathsf{msg}}$. Note that it is efficiently possible to compute $\mathrm{dlog}_h(X) := x$ for $X := h^x \in \mathbb{G}_{\mathsf{msg}}$ and $x \in \mathbb{Z}_{N^2}$. Specifically, it is efficiently possible to test for membership in $\mathbb{G}_{\mathsf{msg}}$. In our scheme, $\mathbb{G}_{\mathsf{msg}}$ will be used to embed a suitably encoded message, and $\mathbb{G}_{\mathsf{rnd}}$ will be used for blinding. We require that

- P and Q are safe primes of bitlength between $\ell_N/2 - k$ and $\ell_N/2 + k$,
- $\gcd((P-1)(Q-1)/4, N) = 1$ (as, e.g., for uniform P, Q of a certain length),

- $\ell_N \geq 25k + 8$ (e.g., $k = 80$ and $\ell_N = 2048$)[9]
- the DCR assumption holds in $\mathbb{Z}^*_{N^3}$, and the DDH assumption holds in $\mathbb{G}_{\mathsf{rnd}}$.

We also assume an $(\ell_{\mathsf{LAF}}, \mathfrak{n})$-lossy algebraic filter LAF for $\mathfrak{n} = 6$ and $\ell_{\mathsf{LAF}} = (\ell_N + k + 1)/(\mathfrak{n} - 2)$. Our scheme will encrypt messages from the domain

$$\mathcal{M} := \mathbb{Z}_{2^{3k}} \times \mathbb{Z}_{p \cdot 2^k} \times \mathbb{Z}_{N \cdot 2^{k-2}},$$

where p is the modulus of the used LAF. (The reason for this weird-looking message space will become clearer in the proof.) During encryption, we will have to treat a message $M = (a, b, c) \in \mathcal{M}$ both as an element of $\mathbb{Z}_{N^2}$ and as an LAF-input from $\mathbb{Z}_p^{\mathfrak{n}}$. In these cases, we can encode

$${}_{\mathbb{Z}} := a + 2^{3k} \cdot b + p \cdot 2^{4k} \cdot c \in \mathbb{Z}, \quad [M]_{\mathbb{Z}_p^{\mathfrak{n}}} := (a, b \bmod p, c_0, \ldots, c_{\mathfrak{n}-3}) \in \mathbb{Z}_p^{\mathfrak{n}} \tag{10}$$

for the natural interpretation of $\mathbb{Z}_i$-elements as integers between 0 and $i - 1$, and c's p-adic representation $(c_i)_{i=0}^{\mathfrak{n}-3} \in \mathbb{Z}_p^{\mathfrak{n}-2}$ with $c = \sum_{i=0}^{\mathfrak{n}-3} c_i \cdot p^i$. By our choice of ℓ_N and ℓ_{LAF}, we have $0 \leq [M]_{\mathbb{Z}} < N^2 - 2^k$. However, the encoding $[M]_{\mathbb{Z}_p^{\mathfrak{n}}}$ is *not* injective, since it only depends on $b \bmod p$ (while $0 \leq b < p \cdot 2^k$).

Finally, we assume a strongly OT-EUF-CMA secure signature scheme $\mathsf{Sig} = (\mathsf{SGen}, \mathsf{Sig}, \mathsf{Ver})$ with k-bit verification keys, and a key-unique IND-CPA secure symmetric encryption scheme (E, D) (see Section 2) with k-bit symmetric keys K and message space $\{0,1\}^*$.

Now consider the following PKE scheme PKE:

Public parameters. $\mathsf{Pars}(1^k)$ first runs $(N, P, Q) \leftarrow \mathsf{GenN}(1^k)$. Recall that this fixes the groups $\mathbb{G}_{\mathsf{rnd}}$ and $\mathbb{G}_{\mathsf{msg}}$. Then, Pars selects two generators g_1, g_2 of $\mathbb{G}_{\mathsf{rnd}}$. Finally, Pars runs $(Fpk, Ftd) \leftarrow \mathsf{FGen}(1^k)$ and outputs

$$pp = (N, g_1, g_2, Fpk).$$

In the following, we denote with p the LAF modulus contained in Fpk.

Key generation. $\mathsf{Gen}(pp)$ uniformly selects four messages $s_j = (a_j, b_j, c_j) \in \mathcal{M}$ (for $1 \leq j \leq 4$) as secret key, and sets

$$pk := \left(u := g_1^{[s_1]_{\mathbb{Z}}} g_2^{[s_2]_{\mathbb{Z}}}, v := g_1^{[s_3]_{\mathbb{Z}}} g_2^{[s_4]_{\mathbb{Z}}}\right), \qquad sk := (s_j)_{j=1}^4.$$

Encryption. $\mathsf{Enc}(pp, pk, M)$, for pp and pk as above, and $M \in \mathcal{M}$, uniformly selects exponents $r, \widetilde{r} \leftarrow \mathbb{Z}_{N/4}$, a random filter core tag t_{c}, a Sig-keypair $(vk, sigk) \leftarrow \mathsf{SGen}(1^k)$, and a random symmetric key $K \in \{0,1\}^k$ for (E, D), and computes

$$(G_1, G_2) := (g_1^r, g_2^r) \qquad Z := (u^{vk} v)^{r \cdot N^2}$$
$$(\widetilde{G}_1, \widetilde{G}_2) := (g_1^{\widetilde{r}}, g_2^{\widetilde{r}}) \qquad \widetilde{Z} := (u^{vk} v)^r \cdot u^{\widetilde{r}} \cdot h^{K + 2^k \cdot [M]_{\mathbb{Z}}},$$
$$C_{\mathsf{E}} \leftarrow \mathsf{E}(K, \mathsf{LAF}_{Fpk,t}([M]_{\mathbb{Z}_p^{\mathfrak{n}}})), \qquad \sigma \leftarrow \mathsf{Sig}(sigk, ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}}))$$
$$C := ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}}, vk, \sigma)$$

for the auxiliary tag $t_{\mathsf{a}} := vk$, and the resulting filter tag $t := (t_{\mathsf{c}}, t_{\mathsf{a}})$.

Decryption. $\mathsf{Dec}(pp, sk, C)$, for pp, sk and C as above, first checks the signature σ and rejects with $\perp$ if $\mathsf{Ver}(vk, ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}}), \sigma) = 0$, or if

$$Z \neq \left(G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}}\right)^{N^2}.$$

Then Dec computes

$$Z' := G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}} \widetilde{G}_1^{[s_1]_{\mathbb{Z}}} \widetilde{G}_2^{[s_2]_{\mathbb{Z}}}$$

[9] Depending on $\mathfrak{n}$ below, shorter ℓ_N are possible. The relevant inequality that must hold is (17).

and then $K \in \{0,1\}^k, M \in \mathcal{M}$ with

$$K + 2^k \cdot [M]_{\mathbb{Z}} := \mathrm{dlog}_h(\widetilde{Z}/Z').$$

If $\widetilde{Z}/Z' \notin \mathbb{G}_{\mathsf{msg}}$, or no such M exists, or $\mathsf{D}(K, C_{\mathsf{E}}) \neq \mathsf{LAF}_{Fpk,t}([M]_{\mathbb{Z}_p^{\mathfrak{n}}})$ (for $t = (t_{\mathsf{c}}, t_{\mathsf{a}})$ computed as during encryption), then Dec rejects with $\perp$. Else, Dec outputs M.

Secret keys as messages. Our scheme has secret keys $s = (s_j)_{j=1}^4 \in \mathcal{M}^4$; hence, we can only encrypt one quarter s_j of a secret key at a time. In the security proof below, we will thus only consider KDM queries that ask to encrypt a specific secret key *part*. Alternatively, we can change our scheme, so that 4-tuples of $\mathcal{M}$-elements are encrypted. To avoid malleability (which would destroy CCA security), we of course have to use only one LAF tag for this. Our CIRC-CCA proof below applies to such a changed scheme with minor syntactic changes.

Efficiency. When instantiated with our DLIN-based LAF construction from Section 3, and taking $\mathfrak{n} = 6$ as above, our scheme has ciphertexts with 43 $\mathbb{G}$-elements, 6 $\mathbb{Z}_{N^3}$-elements, plus chameleon hash randomness, a one-time signature and verification key, and a symmetric ciphertext (whose size could be in the range of one $\mathbb{Z}_{N^2}$-element plus some encryption randomness). The number of group elements in the ciphertext is constant, and does not grow in the security parameter. The public parameters contain $\mathbf{O}(k)$ group elements[10] (most of them from $\mathbb{G}$), and public keys contain two $\mathbb{Z}_{N^3}$-elements; secret keys consist of four $\mathbb{Z}_{N^2}$-elements. While these parameters are not competitive with current non-KDM-secure schemes, they are significantly better than those from the circular-secure scheme of Camenisch et al. [16].[11]

4.2 Security proof (single-user case)

It is instructive to first treat the single-user case. Here, we essentially only require that PKE is IND-CCA secure, even if encryptions of its secret key are made public.

Theorem 4.1. *Assume the DCR assumption holds in $\mathbb{Z}_{N^3}$, the DDH assumption holds in $\mathbb{G}_{\mathsf{rnd}}$, LAF is an LAF, Sig is a strongly OT-EUF-CMA secure signature scheme, H is collision-resistant, and (E, D) is a key-unique IND-CPA secure SKE scheme. Then PKE is 1-CIRC-CCA-secure.*

Proof. Assume a PPT adversary A on PKE's 1-CIRC-CCA security. Say that A always makes $q = q(k)$ KDM queries. We proceed in games. Let out_i denote the output of Game i.

Game 1 is the 1-KDM-CCA experiment with PKE and A. By definition,

$$\Pr[out_1 = 1] - 1/2 = \mathsf{Adv}^{\mathsf{kdm\text{-}cca}}_{\mathsf{PKE},A}(k).$$

In **Game** 2, we modify the way KDM queries are answered. Namely, in each ciphertext prepared for A, we set up Z and $\widetilde{Z}$ up as

$$\begin{aligned} Z &:= \left(G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}}\right)^{N^2} \\ \widetilde{Z} &:= G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}} \widetilde{G}_1^{[s_1]_{\mathbb{Z}}} \widetilde{G}_2^{[s_2]_{\mathbb{Z}}} \cdot h^{K + 2^k \cdot [M]_{\mathbb{Z}}}. \end{aligned} \tag{11}$$

for the already prepared $(G_j, \widetilde{G}_j) = (g_j^r, g_j^{\widetilde{r}})$. This change is only conceptual by our setup of u, v, so

$$\Pr[out_2 = 1] = \Pr[out_1 = 1].$$

In **Game** 3, we again change how KDM ciphertexts are prepared. Intuitively, our goal is now to prepare the G_j and $\widetilde{G}_j$ with additional $\mathbb{G}_{\mathsf{msg}}$-components, such that $\widetilde{Z}$, as computed in (11), is

[10] Using the generalized LAF mentioned at the end of Section 3.2, public parameters with $\mathbf{O}(1)$ group elements are possible, at the cost of a (constant) number of extra group elements per tag.

[11] For instance, Section 7 of the full version of [16] implies that their scheme has a public key, resp. ciphertext of about 500, resp. 1000 $\mathbb{G}$-elements (for $\log_2(|\mathbb{G}|) = 160$).

of the form $g \cdot h^K$ for some $g \in \mathbb{G}_{\mathsf{rnd}}$. (That is, we want the $\mathbb{G}_{\mathsf{msg}}$-components of the $G_j, \widetilde{G}_j$ to cancel out the $h^{2^k \cdot [M]_\mathbb{Z}}$ term in (11).) To do so, we prepare

$$G_j = g_j^r / h^{\alpha_j \cdot 2^k} \qquad\qquad \widetilde{G}_j = g_j^{\widetilde{r}} / h^{\widetilde{\alpha}_j \cdot 2^k}$$

for $j \in \{1, 2\}$ and suitable $\alpha_j, \widetilde{\alpha}_j$ to be determined. $\widetilde{Z}$ is still computed as in (11), so

$$\widetilde{Z} \;=\; g \cdot h^{K + 2^k \cdot [M]_\mathbb{Z} - 2^k (\alpha_1([s_1]_\mathbb{Z} \cdot vk + [s_3]_\mathbb{Z}) + \alpha_2([s_2]_\mathbb{Z} \cdot vk + [s_4]_\mathbb{Z}) + \widetilde{\alpha}_1 [s_1]_\mathbb{Z} + \widetilde{\alpha}_2 [s_2]_\mathbb{Z})}$$

for

$$g \;=\; g_1^{r \cdot ([s_1]_\mathbb{Z} \cdot vk + [s_3]_\mathbb{Z}) + \widetilde{r}[s_1]_\mathbb{Z}} g_2^{r \cdot ([s_2]_\mathbb{Z} \cdot vk + [s_4]_\mathbb{Z}) + \widetilde{r}[s_2]_\mathbb{Z}} \;=\; \left(u^{vk} v\right)^r u^{\widetilde{r}} \;\in\; \mathbb{G}_{\mathsf{rnd}}.$$

So to prepare a KDM encryption of s_{j^*} with a $\widetilde{Z}$ of the form$\widetilde{Z} = g \cdot h^K$, we can set

$$(\alpha_1, \alpha_2, \widetilde{\alpha}_1, \widetilde{\alpha}_2) \;:=\; \begin{cases} (0, 0, 1, 0) & \text{for } j^* = 1 \\ (0, 0, 0, 1) & \text{for } j^* = 2 \\ (1, 0, -vk, 0) & \text{for } j^* = 3 \\ (0, 1, 0, -vk) & \text{for } j^* = 4. \end{cases}$$

(vk can be chosen independently in advance.) The remaining parts of C are prepared as in Game 2. We claim

$$\Pr[out_3 = 1] - \Pr[out_2 = 1] \;\leq\; 4 \cdot \mathsf{Adv}^{\mathsf{dcr}}_{\mathbb{Z}^*_{N^3}, B}(k) + \mathbf{O}(2^{-k}) \tag{12}$$

for a suitable DCR distinguisher B that simulates Game 2, resp. Game 3. Concretely, B gets as input a value $\widetilde{W} \in \mathbb{Z}^*_{N^3}$ of the form $\widetilde{W} = \tilde{g}^{N^2} \cdot h^b$ for $b \in \{0, 1\}$. Note that if we set $W := \widetilde{W}^{-2^k}$, we have $W = g^{\widehat{r}} / h^{b \cdot 2^k} \in \mathbb{Z}^*_{N^3}$, with uniform $g^{\widehat{r}} \in \mathbb{G}_{\mathsf{rnd}}$. First, B guesses a value of $j^* \in [4]$. (This gives a very small hybrid argument, in which in the j^*-th step, only encryptions of s_{j^*} are changed.) We only detail B's behavior for the case $j^* = 3$; the other cases are easier or analogous. First, B sets up $g_1 := W^{N^2}$ and $g_2 := W^{\gamma N^4}$ for uniform $\gamma \in \mathbb{Z}_{N/4}$. To prepare an encryption of s_3, B chooses uniform $\rho, \widetilde{\rho} \in \mathbb{Z}_{N^2/4}$ and sets

$$G_1 := W^{\rho \cdot (\rho^{-1})} \qquad G_2 := W^{\gamma \cdot \rho \cdot (\rho^{-1}) \cdot N^2} \qquad \widetilde{G}_1 := W^{vk \cdot \widetilde{\rho} \cdot (\widetilde{\rho}^{-1})} \qquad \widetilde{G}_2 := W^{\gamma \cdot vk \cdot \widetilde{\rho} \cdot (\widetilde{\rho}^{-1}) \cdot N^2},$$

where the values $\rho^{-1}, \widetilde{\rho}^{-1}$ are computed modulo N^2. This implicitly sets $r = \rho \cdot (\rho^{-1}) / N^2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$ and $\widetilde{r} = vk \cdot \widetilde{\rho} \cdot (\widetilde{\rho}^{-1}) / N^2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$, both of which are statistically close to uniform. Furthermore, $G_j = g_j^r / h^{b \cdot \alpha_j \cdot 2^k}$ and $\widetilde{G}_j = g_j^{\widetilde{r}} / h^{b \cdot \widetilde{\alpha_j} \cdot 2^k}$; so, depending on B's challenge, encryptions of s_3 are prepared as in Game 2 or Game 3. Similar arguments work for $j^* = 1, 2, 4$, and (12) follows. (The $\mathbf{O}(2^{-k})$ term in (12) accounts for the statistical defect caused by choosing $\mathbb{G}_{\mathsf{rnd}}$-exponents from $\mathbb{Z}_{N/4}$, resp. $\mathbb{Z}_{N^2/4}$.)

Using the definition of u and v, our change in Game 3 implies $\widetilde{Z} = (u^{vk} v)^r \cdot u^{\widetilde{r}} \cdot h^K$ when a key part s_j is to be encrypted. (However, note that we still have $Z = (u^{vk} v)^{r \cdot N^2}$ in any case.) This means that A still obtains information about the s_j (beyond what is public from pk) from its KDM queries, but this information is limited to values $\mathsf{LAF}_{Fpk,t}([s_j]_{\mathbb{Z}_p^n})$. We will now further cap this leaked information by making $\mathsf{LAF}_{Fpk,t}(\cdot)$ lossy. Namely, in **Game** 4, we use the LAF trapdoor Ftd initially sampled along with Fpk. Concretely, when preparing a ciphertext C for A, we sample t_c using $t_\mathsf{c} \leftarrow \mathsf{FTag}(Ftd, t_\mathsf{a})$ for the corresponding auxiliary tag $t_\mathsf{a} = vk$. A simple reduction shows

$$\Pr[out_4 = 1] - \Pr[out_3 = 1] \;=\; \mathsf{Adv}^{\mathsf{ind}}_{\mathsf{LAF}, C_2}(k)$$

for a suitable adversary C_2 on LAF's indistinguishability.

In **Game** 5, we reject all decryption queries of A that re-use a verification key vk from one of the KDM ciphertexts. To show that this change does not significantly affect A's view, assume a decryption query C that re-uses a key $vk = vk^*$ from a KDM ciphertext C^*. Recall that C contains a signature σ of $X := ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}})$ under an honestly generated Sig-verification-key $vk = t_{\mathsf{a}} = t_{\mathsf{a}}^* = vk^*$. Since we assumed $t = (t_{\mathsf{c}}, t_{\mathsf{a}}) = (t_{\mathsf{c}}^*, t_{\mathsf{a}}^*) = t^*$, and A is not allowed to query unchanged challenge ciphertexts for decryption, we must have $(X, \sigma) \neq (X^*, \sigma^*)$ for the corresponding message X^* and signature σ^* from C^*. Hence, Game 4 and Game 5 only differ when A manages to forge a signature. A straightforward reduction to the strong OT-EUF-CMA security of Sig yields

$$\Pr[out_5 = 1] - \Pr[out_4 = 1] \ = \ q(k) \cdot \mathsf{Adv}_{\mathsf{LAF},F}^{\mathsf{seuf\text{-}cma}}(k)$$

for a forger F against Sig that makes at most one signature query.

In **Game** 6.i (for $0 \leq i \leq q$), the first i challenge ciphertexts are prepared using $Z = \widehat{g}^{N^2}$ and $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^K$ (if a key component s_j is to be encrypted), resp. $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^{K+2^k[M]_{\mathbb{Z}}}$ (if a constant $M \in \mathcal{M}$ is to be encrypted) for an independently uniform $\widehat{g} \leftarrow \mathbb{G}_{\mathsf{rnd}}$ drawn freshly for each ciphertext. Obviously, Game 6.0 is identical to Game 5:

$$\Pr[out_{6.0} = 1] \ = \ \Pr[out_5 = 1].$$

We will move from Game 6.i to Game 6.$(i+1)$ in several steps. During these steps, we denote with $C = ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}}, vk, \sigma)$ the $(i+1)$-st KDM ciphertext.

In **Game** 6.i.1, we change the $\mathbb{G}_{\mathsf{rnd}}$ parts of G_1, G_2 from a Diffie-Hellman tuple (with respect to g_1, g_2) to a random tuple. Concretely, if an s_j is to be encrypted, we set $(G_1, G_2) = (g_1^{r_1}/h^{\alpha_1 \cdot 2^k}, g_2^{r_2}/h^{\alpha_2 \cdot 2^k})$; if a constant M is encrypted, we set $(C_1, C_2) = (g_1^{r_1}, g_2^{r_2})$, in both cases for independently uniform $r_1, r_2 \leftarrow \mathbb{Z}_{N/4}$. The $\mathbb{G}_{\mathsf{msg}}$ parts of G_1, G_2 are thus unchanged compared to Game 6.i. Note that the $\widetilde{G}_j$ are still prepared as $\widetilde{G}_j = g_j^{\widetilde{r}}/h^{\widetilde{\alpha} \cdot 2^k}$, resp. $\widetilde{G}_j = g_j^{\widetilde{r}}$. A straightforward reduction to the DDH assumption in $\mathbb{G}_{\mathsf{rnd}}$ yields

$$\sum_{i=1}^{q(k)} \left(\Pr[out_{6.i} = 1] - \Pr[out_{6.i.1} = 1]\right) \ = \ q(k) \cdot \mathsf{Adv}_{\mathbb{G}_{\mathsf{rnd}},D_1}^{\mathsf{ddh}}(k) + \mathbf{O}(2^{-k})$$

for a suitable D_1. The $\mathbf{O}(2^{-k})$ error term accounts for the statistical difference caused by the choice of exponents $r_j \leftarrow \mathbb{Z}_{N/4}$, which induces an only almost-uniform distribution on group elements g^{r_j}. Note that at this point, Z and $\widetilde{Z}$ are still computed as in (11), even if an s_j is to be encrypted.

In **Game** 6.i.2, we compute Z and $\widetilde{Z}$ as $Z = \widehat{g}^{N^2}$ and $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^K$, resp. $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^{K+2^k[M]_{\mathbb{Z}}}$ for a fresh $\widehat{g} \leftarrow \mathbb{G}_{\mathsf{rnd}}$. Thus, the difference to Game 6.i.1 is that we substitute a $\mathbb{G}_{\mathsf{rnd}}$-element computed as $G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}}$ with a fresh random $\widehat{g}$. To show that this change affects A's view only negligibly, it suffices to show that A's statistical information about

$$X := \mathrm{dlog}_g \left(G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}} \right) = \gamma_1 r_1([s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}) + \gamma_2 r_2([s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}) \bmod |\mathbb{G}_{\mathsf{rnd}}|$$

(for some arbitrary generator g of $\mathbb{G}_{\mathsf{rnd}}$ and $\gamma_j = \mathrm{dlog}_g(g_j)$) is negligible. This part will be rather delicate, since we will have to argue that both A's KDM queries and A's decryption queries yield (almost) no information about X.

First, observe that A gets the following information about the s_j:

- pk reveals (through u and v) precisely the two linear equations $\gamma_1[s_1]_{\mathbb{Z}} + \gamma_2[s_2]_{\mathbb{Z}} \bmod |\mathbb{G}_{\mathsf{rnd}}|$ and $\gamma_1[s_3]_{\mathbb{Z}} + \gamma_2[s_4]_{\mathbb{Z}} \bmod |\mathbb{G}_{\mathsf{rnd}}|$ about the s_j, where the γ_j are as above. For $r_1 \neq r_2$, these equations are linearly independent of the equation that defines X. Hence, for uniform r_1, r_2, X is (almost) independent of pk.
- By LAF's lossiness, KDM queries yield (through $C_{\mathsf{E}} = \mathsf{E}(K, \mathsf{LAF}_{Fpk,t}([s_j]_{\mathbb{Z}_p^{\mathsf{n}}}))$ in total at most one equation $\omega_1 a_j + \omega_2 b_j + \sum_{i=0}^{\mathsf{n}-2} \omega_{3+i} c_{j,i} \bmod p$ for each j, where $(a_j, b_j, c_{j,0}, \ldots, c_{j,\mathsf{n}-3}) := [s_j]_{\mathbb{Z}_p^{\mathsf{n}}}$, and the ω_i are the (fixed) coefficients from LAF's lossiness property. (Recall the encodings $[s_j]_{\mathbb{Z}}, [s_j]_{\mathbb{Z}_p^{\mathsf{n}}}$ of the $s_j = (a_j, b_j, c_j)$ from (10).) Hence, the $b_j \in \mathbb{Z}_{p \cdot 2^k}$ fully blind the information released about the $c_j \in \mathbb{Z}_{2^{k-2}N}$ through the KDM ciphertexts. Thus, KDM ciphertexts reveal no information about $c_j \bmod |\mathbb{G}_{\mathsf{rnd}}|$ and hence also about $[s_j]_{\mathbb{Z}} \bmod |\mathbb{G}_{\mathsf{rnd}}|$.

Consequently, even given pk and the KDM ciphertexts, X is statistically close to independently uniform. This already shows that our change from Game 6.i.2 affects A's view only negligibly if A makes no decryption queries. It remains to show that decryption queries yield no additional information about the s_j.

To do so, let us say that a ciphertext C is *consistent* iff there exist $r, \widetilde{r}$ with $(G_j, \widetilde{G}_j) = (g_j^r, g_j^{\widetilde{r}})$ for both $j \in \{1, 2\}$. Note that the decryption of a consistent ciphertext yields no information about the s_j beyond pk. (pk and $r, \widetilde{r}$ determine the values Z, Z' computed during decryption; everything else follows from Z' and C.) So it suffices to prove the following lemma (which we do after the main proof):

Lemma 4.2. *In the situation of Game 6.$i.\ell$ (for $\ell \in \{1, 2\}$), let $\mathsf{bad}_{\mathsf{query.i.}\ell}$ be the event that A places an inconsistent decryption query that is not rejected. Then*

$$\sum_{i=1}^{q(k)} \left(\Pr\left[\mathsf{bad}_{\mathsf{query.i.1}}\right] + \Pr\left[\mathsf{bad}_{\mathsf{query.i.2}}\right]\right) \;\leq\; 2 \cdot q(k) \cdot \mathsf{Adv}^{\mathsf{eva}}_{\mathsf{LAF},F}(k) + \mathbf{O}(2^{-3k}).$$

for a suitable evasiveness adversary F on LAF.

By our discussion above and Lemma 4.2, we obtain that

$$\sum_{i=1}^{q(k)} \left|\Pr\left[out_{6.i.2} = 1\right] - \Pr\left[out_{6.i.1} = 1\right]\right| \;\leq\; 2 \cdot q(k) \cdot \mathsf{Adv}^{\mathsf{eva}}_{\mathsf{LAF},F}(k) + \mathbf{O}(2^{-3k}).$$

In **Game** 6.i.3, we reverse the change from Game 6.i.1. Concretely, we prepare the G_j as $G_j = g_j^r / h^{\alpha_j \cdot 2^k}$, resp. $G_j = g_j^r$ for a single $r \leftarrow \mathbb{Z}_{N/4}$. Another straightforward reduction to the DDH assumption in $\mathbb{G}_{\mathsf{rnd}}$ yields that

$$\sum_{i=1}^{q(k)} \left(\Pr\left[out_{6.i.3} = 1\right] - \Pr\left[out_{6.i.2} = 1\right]\right) \;=\; q(k) \cdot \mathsf{Adv}^{\mathsf{ddh}}_{\mathbb{G}_{\mathsf{rnd}},D_2}(k) + \mathbf{O}(2^{-k})$$

for a suitable D_2. To close the hybrid argument, note that Games 6.i.3 and 6.(i+1) are identical.

In **Game** 7, we clear the $\mathbb{G}_{\mathsf{msg}}$-component of $\widetilde{Z}$ in all ciphertexts prepared for A. That is, instead of computing $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^K$, resp. $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}} \cdot h^{K + [M]_{\mathbb{Z}}}$ for a freshly uniform $\widehat{g} \leftarrow \mathbb{G}_{\mathsf{rnd}}$, we set $\widetilde{Z} = \widehat{g} \cdot u^{\widetilde{r}}$. (We stress that we still compute $Z = \widehat{g}^{N^2}$.) Since all $\widetilde{Z}$ already have an independently uniform $\mathbb{G}_{\mathsf{rnd}}$-component, a straightforward reduction to the DCR assumption yields

$$\Pr\left[out_{6.q} = 1\right] - \Pr\left[out_7 = 1\right] \;=\; \mathsf{Adv}^{\mathsf{dcr}}_{\mathbb{Z}^*_{N^3},E}(k) + \mathbf{O}(2^{-k})$$

for a DCR distinguisher E. Note that because of the re-randomizability of DCR, there is no factor of $q(k)$, even though we substitute many group elements at once. However, since the precise order of $\mathbb{G}_{\mathsf{rnd}}$ is not known, this re-randomization costs us an error term of $\mathbf{O}(2^{-k})$.

In **Game** 8, we substitute the symmetric ciphertexts C_{E} in all KDM ciphertexts by encryptions of random messages. By our change in Game 7, we do not use the symmetric keys K used to produce C_{E} anywhere else. Thus, a reduction to the IND-CPA security of (E, D) gives

$$\Pr\left[out_7 = 1\right] - \Pr\left[out_8 = 1\right] \;=\; q(k) \cdot \mathsf{Adv}^{\mathsf{ind\text{-}cpa}}_{(\mathsf{E},\mathsf{D}),G}(k)$$

for an IND-CPA adversary G. Note that in Game 8, A's view is independent of the challenge bit b initially selected by the KDM challenger. Hence, we have

$$\Pr\left[out_8 = 1\right] \;=\; 1/2.$$

Taking things together yields the theorem. □

It remains to prove Lemma 4.2, which we do now:

Proof. Let $\mathsf{bad}_{\mathsf{tag.i.}\ell}$ be the event that in Game $6.i.\ell$, A submits a decryption query that refers to a lossy tag t. By our change from Game 5, we may assume that vk and thus t is fresh, i.e., has not been generated through FTag by the experiment. Thus, by LAF's evasiveness, $\mathsf{bad}_{\mathsf{tag.i.}\ell}$ can occur only with negligible probability. Concretely, it is easy to construct an evasiveness adversary F with

$$\sum_{i=1}^{q(k)} \left(\Pr\left[\mathsf{bad}_{\mathsf{tag.i.1}}\right] + \Pr\left[\mathsf{bad}_{\mathsf{tag.i.2}}\right]\right) \;\leq\; 2 \cdot q(k) \cdot \mathsf{Adv}^{\mathsf{eva}}_{\mathsf{LAF},F}(k). \tag{13}$$

Now suppose that we are in Game $6.i.\ell$, and say that $\mathsf{bad}_{\mathsf{tag.i.}\ell}$ does not occur. Consider an inconsistent decryption query $C = ((G_j, \widetilde{G}_j)_{j=1}^2, Z, \widetilde{Z}, C_{\mathsf{E}}, t_{\mathsf{c}}, vk, \sigma)$ from A. Write $(G_j, \widetilde{G}_j) = (g_j^{r_j} \cdot h^{\delta_j}, \widetilde{g}_j^{\widetilde{r}_j} \cdot h^{\widetilde{\delta}_j})$ (for $j \in \{1, 2\}$) and $\widetilde{Z} = \widetilde{g} \cdot h^{\widetilde{\delta}}$ for $\widetilde{g} \in \mathbb{G}_{\mathsf{rnd}}$. Recall that decryption first checks

$$Z \;\stackrel{?}{=}\; \left(G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}}\right)^{N^2}, \tag{14}$$

then computes

$$Z' \;=\; G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}} \widetilde{G}_1^{[s_1]_{\mathbb{Z}}} \widetilde{G}_2^{[s_2]_{\mathbb{Z}}}, \tag{15}$$

and finally attempts to find $K \in \{0,1\}^k, M \in \mathcal{M}$ with

$$K + 2^k \cdot [M]_{\mathbb{Z}} \;=\; \mathrm{dlog}_h(\widetilde{Z}/Z') \;=\; \widetilde{\delta} - \delta' \bmod N^2 \tag{16}$$

for $\delta' = \delta_1 \left([s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}\right) + \delta_2 \left([s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}\right) + \widetilde{\delta}_1 [s_1]_{\mathbb{Z}} + \widetilde{\delta}_2 [s_2]_{\mathbb{Z}}$. As usual, we write $s_j = (a_j, b_j, c_j) \in \mathcal{M} = \mathbb{Z}_{2^{3k}} \times \mathbb{Z}_{p \cdot 2^k} \times \mathbb{Z}_{N \cdot 2^{k-2}}$ for $i \in \{1, 2\}$.

h-inconsistent ciphertexts. First, consider the case that there is a $j^* \in \{1, 2\}$ with $\delta_{j^*} \neq 0 \bmod N^2$ or $\widetilde{\delta_{j^*}} \neq 0 \bmod N^2$. (In that case, we may say that C is *h-inconsistent.*) Then, we claim that either C is rejected, or A has (information-theoretically) successfully narrowed down the value of δ' to a set of size at most 2^k. Indeed, C_{E} determines K and thus $\mathsf{LAF}_{Fpk,t}([M]_{\mathbb{Z}_p^{\mathfrak{n}}}) = \mathsf{D}(K, C_{\mathsf{E}})$ by (E, D)'s key-uniqueness. Moreover, since we assumed $\neg\mathsf{bad}_{\mathsf{tag.i.}\ell}$, the used tag t is injective, and so $\mathsf{LAF}_{Fpk,t}([M]_{\mathbb{Z}_p^{\mathfrak{n}}})$ determines M up to $\lfloor b/p \rfloor \in \mathbb{Z}_{2^k}$. (Recall that the encoding $[M]_{\mathbb{Z}_p^{\mathfrak{n}}}$ only depends on $b \bmod p$.) Thus, a non-rejected ciphertext allows to infer (a 2^k-candidate set for) δ' by substituting K, M, and $\widetilde{\delta}$ (as defined by $\widetilde{Z}$) into (16).

However, we will now argue that δ' has min-entropy at least $5k$, even given pk and all KDM ciphertexts. Hence, A cannot predict a correct 2^k-candidate set for δ' (and thus cannot supply an h-inconsistent decryption query that is not rejected) with non-negligible probability. To prove our claim, we need some preparations. For concreteness, say that $\delta_1 \neq 0 \bmod N^2$ (the other cases are similar). Then either $P^2 \nmid \delta_1$ or $Q^2 \nmid \delta_1$ (or both) for the factors P, Q of N. Without losing generality, say that $P^2 \nmid \delta_1$, so that the subterm $\delta_1 \cdot [s_3]_{\mathbb{Z}} \bmod N^2$ of δ' reveals $[s_3]_{\mathbb{Z}} \bmod P$. Furthermore,

$$\begin{aligned}
[s_3]_{\mathbb{Z}} &\stackrel{(10)}{<} 2^{5k-2} \cdot p \cdot N \;\leq\; 2^{5k+1} \cdot 2^{\ell_{\mathsf{LAF}}} \cdot |\mathbb{G}_{\mathsf{rnd}}| \stackrel{\ell_{\mathsf{LAF}} = \frac{\ell_N + k + 1}{\mathfrak{n}-2}}{<} 2^{(5+1/(\mathfrak{n}-2))k+2} \cdot 2^{\ell_N/(\mathfrak{n}-2)} \cdot |\mathbb{G}_{\mathsf{rnd}}| \\
&\stackrel{P \geq 2^{(\ell_N/2)-k}}{\leq} 2^{(6+1/(\mathfrak{n}-2))k+2-(1/2-1/(\mathfrak{n}-2))\ell_N} \cdot |\mathbb{G}_{\mathsf{rnd}}| \cdot P \stackrel{\substack{\ell_N \geq 25k+8 \\ \mathfrak{n}=6}}{\leq} |\mathbb{G}_{\mathsf{rnd}}| \cdot P.
\end{aligned} \tag{17}$$

Using $\gcd(P, |\mathbb{G}_{\mathsf{rnd}}|) = 1$, the Chinese Remainder Theorem hence gives that $[s_3]_{\mathbb{Z}} \bmod |\mathbb{G}_{\mathsf{rnd}}|$ and $[s_3]_{\mathbb{Z}} \bmod P$ uniquely determine $[s_3]_{\mathbb{Z}}$. Thus, since $[s_3]_{\mathbb{Z}}$ initially has min-entropy at least $5k - 2 + \ell_{\mathsf{LAF}} + \ell_N$, even revealing $[s_3]_{\mathbb{Z}} \bmod |\mathbb{G}_{\mathsf{rnd}}|$ (through pk) leaves at least $5k + \ell_{\mathsf{LAF}}$ bits of min-entropy in $[s_3]_{\mathbb{Z}} \bmod P$. The KDM ciphertexts reveal no more than ℓ_{LAF} bits of entropy about $[s_3]_{\mathbb{Z}} \bmod P$, so that $[s_3]_{\mathbb{Z}} \bmod P$ has min-entropy at least $5k$.

However, C implies 2^k candidates for δ' which, given s_1, s_2, s_4, in turn determine 2^k candidates for $[s_3]_{\mathbb{Z}} \bmod P$. So, assuming $\neg\mathsf{bad}_{\mathsf{tag.i.j}}$, the probability that a given h-inconsistent C implies "the correct $[s_3]_{\mathbb{Z}} \bmod P$" (which is a prerequisite for non-rejection), is at most 2^{-4k}. The case for

$\delta_2 \neq 0 \bmod N^2$ makes the analogous argument about $[s_4]_{\mathbb{Z}}$, and the cases $\widetilde{\delta_{j^*}} \neq 0 \bmod N^2$ consider $[s_{j^*}]_{\mathbb{Z}}$, then using $\delta_1, \delta_2 = 0 \bmod N^2$.

g-inconsistent ciphertexts. Now assume that $\delta_1 = \delta_2 = \widetilde{\delta_1} = \widetilde{\delta_2} = 0 \bmod N^2$. Since C is inconsistent, $r_1 \neq r_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$ or $\widetilde{r}_1 \neq \widetilde{r}_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$. We may call such ciphertexts *g-inconsistent.*

Let us first assume $r_1 \neq r_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$. Recall that $|\mathbb{G}_{\mathsf{rnd}}| = (P-1)(Q-1)/4$, where P, Q are safe primes. Hence, without loss of generality, we can assume that $r_1 \neq r_2 \bmod (P-1)/2$, where $(P-1)/2$ is prime. We claim that the value

$$X' := \mathrm{dlog}_g \left(G_1^{[s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}} G_2^{[s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}} \right)$$
$$= \gamma_1 r_1 ([s_1]_{\mathbb{Z}} \cdot vk + [s_3]_{\mathbb{Z}}) + \gamma_2 r_2 ([s_2]_{\mathbb{Z}} \cdot vk + [s_4]_{\mathbb{Z}}) \bmod |\mathbb{G}_{\mathsf{rnd}}|$$

is (up to a small statistical defect) independently and uniformly random modulo $(P-1)/2$ from A's point of view. Hence, also the value $g^{X' \cdot N^2}$ from (14) to which Z is compared is unpredictable for A. This can be seen as in the discussion after Game $6.i.2$, where a similar value X is seen as essentially uniform. In particular, pk contains two linear equations (in the s_j) that are independent of X', and the information about X' from the KDM challenges is suitably blinded by the b_j-components of the s_j. There are two differences to Game $6.i.2$: first, the r_i in our case are adversarially chosen, and so could be equal modulo a factor of $|\mathbb{G}_{\mathsf{rnd}}|$. Thus, we can only conclude linear independence modulo $(P-1)/2$. Second, if $\ell = 1$, then A additionally receives one g-inconsistent challenge ciphertext that reveals another linear equation

$$\gamma_1 r_1^* ([s_1]_{\mathbb{Z}} \cdot vk^* + [s_3]_{\mathbb{Z}}) + \gamma_2 r_2^* ([s_2]_{\mathbb{Z}} \cdot vk^* + [s_4]_{\mathbb{Z}}) \bmod |\mathbb{G}_{\mathsf{rnd}}|$$

about the s_j. However, by our change from Game 5, we may assume $vk \neq vk^*$, so that this linear equation is also independent of X'. Hence, X' looks (almost) independently uniform modulo $(P-1)/2$ to A, so that $r_1 \neq r_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$ implies rejection (because of the check (14)) except with probability at most $\mathbf{O}(2^{-4k})$.[12]

Let us now assume $r_1 = r_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$ but $\widetilde{r}_1 \neq \widetilde{r}_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$. This case is similar to the case $r_1 \neq r_2 \bmod |\mathbb{G}_{\mathsf{rnd}}|$, but simpler. A similar analysis as above yields that (the $\mathbb{G}_{\mathsf{rnd}}$-component of) Z' is unpredictable for A. Specifically, C will be rejected except with probability $\mathbf{O}(2^{-4k})$.

Summarizing, and using a union bound, we obtain that

$$\Pr\left[\mathsf{bad}_{\mathsf{query.i.}\ell} \mid \neg \mathsf{bad}_{\mathsf{tag.i.}\ell}\right] \leq q'(k) \cdot 2^{-4k} = \mathbf{O}(2^{-3k})$$

for the number $q'(k)$ of A's decryption queries. Combining with (13) shows the lemma. We stress that in this proof, it may appear that several bounds have been chosen too conservatively. In particular, we arrive at an error bound that is significantly smaller than, e.g., $\mathbf{O}(2^{-k})$. These extra "entropy cushions" are used in the multi-user case. □

4.3 Security proof (multi-user case)

Theorem 4.3. *Assume the DCR assumption holds in $\mathbb{Z}_{N^3}$, the DDH assumption holds in $\mathbb{G}_{\mathsf{rnd}}$,* LAF *is an LAF,* Sig *is a strongly OT-EUF-CMA secure signature scheme,* H *is collision-resistant, and* (E, D) *is a key-unique IND-CPA secure SKE scheme. Then* PKE *is n-CIRC-CCA-secure for every polynomial $n = n(k)$.*

Proof sketch. The proof is very similar to the proof of Theorem 4.1. The way we achieve multi-user KDM security is to have n "virtual" secret keys s^i that are set up as

$$s^i = (s_1^i, s_2^i, s_3^i, s_4^i) = (s_1, s_2, s_3, s_4) + (\hat{s}_1^i, \hat{s}_2^i, \hat{s}_3^i, \hat{s}_4^i) \quad (18)$$

[12] This argument follows in the footsteps of IND-CCA security proofs of PKE schemes based on hash proof systems [18, 19]. Specifically, the knowledge that an adversary receives through pk and one inconsistent ciphertext in our case is essentially the same as in the analysis of the Kurosawa-Desmedt scheme [27].

(with component-wise addition, also in each sum $s_j + \hat{s}^i_j$) for uniformly chosen $\hat{s}^i = (\hat{s}^i_j)_{j=1}^4 \leftarrow \mathcal{M}^4$. Intuitively, the $\hat{s}^i$ blind a single $s = (s_j)_{j=1}^4 \in \mathcal{M}^4$ in several instances. While the $\hat{s}^i$ are all uniform, however, we choose the $s_j = (a_j, b_j, c_j) \in \mathcal{M}$ with "small" components. Concretely, we pick $(a_j, b_j, c_j) \leftarrow \mathbb{Z}_{3k} \times \mathbb{Z}_p \times \mathbb{Z}_{N/4}$ and embed s_j into $\mathcal{M}$ in the natural way. This choice guarantees that $[s^i_j]_\mathbb{Z} = [s_j]_\mathbb{Z} + [\hat{s}^i_j]_\mathbb{Z}$ and $[s^i_j]_{\mathbb{Z}_p^\mathfrak{n}} = [s_j]_{\mathbb{Z}_p^\mathfrak{n}} + [\hat{s}^i_j]_{\mathbb{Z}_p^\mathfrak{n}}$, except with probability $\mathbf{O}(2^{-k})$. Intuitively, the $\hat{s}^i$ can be known to A at all times, while we will try to argue that the information A has about s is very limited.

We will now go through the proof of Theorem 4.1, and sketch the necessary modifications for the multi-user case. Generally, we assume a setup of keys as in (18) (which guarantees independently uniform s^i). **Games** 1 **to** 6.i.2 are as with Theorem 4.1, where the changes apply of course to KDM queries under all public keys. The corresponding reductions to DCR, DDH, the indistinguishability of LAF, and the security of Sig and H apply almost verbatim. The only noteworthy change occurs in the justification of the change from Game 6.i.2.

Here, we have to argue that A obtains no useful information about the $s^i_j \bmod |\mathbb{G}_\mathsf{rnd}|$ from all public keys pk^i, all challenge ciphertexts, and all decryption queries. First, each pk^i reveals (through the corresponding u, v) exactly two linear equations

$$\begin{aligned}\gamma_1[s^i_1]_\mathbb{Z} + \gamma_2[s^i_2]_\mathbb{Z} &= (\gamma_1[s_1]_\mathbb{Z} + \gamma_2[s_2]_\mathbb{Z}) + \big(\gamma_1[\hat{s}^i_1]_\mathbb{Z} + \gamma_2[\hat{s}^i_2]_\mathbb{Z}\big) \bmod |\mathbb{G}_\mathsf{rnd}| \\ \gamma_1[s^i_3]_\mathbb{Z} + \gamma_2[s^i_4]_\mathbb{Z} &= (\gamma_1[s_3]_\mathbb{Z} + \gamma_2[s_4]_\mathbb{Z}) + \big(\gamma_1[\hat{s}^i_3]_\mathbb{Z} + \gamma_2[\hat{s}^i_4]_\mathbb{Z}\big) \bmod |\mathbb{G}_\mathsf{rnd}|\end{aligned}$$

about $s = (s_1, s_2, s_3, s_4)$. These equations only depend on $\gamma_1[s_1]_\mathbb{Z} + \gamma_2[s_2]_\mathbb{Z} \bmod |\mathbb{G}_\mathsf{rnd}|$ and $\gamma_1[s_3]_\mathbb{Z} + \gamma_2[s_4]_\mathbb{Z} \bmod |\mathbb{G}_\mathsf{rnd}|$ (but not on other information about the s_j), just like in the single-user case. Similarly, since $[s^i_j]_{\mathbb{Z}_p^\mathfrak{n}} = [s_j]_{\mathbb{Z}_p^\mathfrak{n}} + [\hat{s}^i_j]_{\mathbb{Z}_p^\mathfrak{n}}$, all challenge ciphertexts depend only on

$$\omega_1 a_j + \omega_2 b_j + \sum_{i=0}^{\mathfrak{n}-2} \omega_{3+i} c_{j,i} \bmod p$$

(for $(a_j, b_j, c_{j,0}, \ldots, c_{j,\mathfrak{n}-3}) := [s_j]_{\mathbb{Z}_p^\mathfrak{n}}$) and the $\hat{s}^i_j$. This equation is fully blinded by $b_j \in \mathbb{Z}_p$. Next, carefully considering the (slightly reduced) entropy in the s_j, we can prove an analog of Lemma 4.2 for the multi-user case.

As in Lemma 4.2, we will have to argue that the additional linear equation released about the s_j by the $(i+1)$-st challenge ciphertext in Game 6.i.1 does not help in producing g-inconsistent decryption queries. The corresponding analysis is the same as that in Lemma 4.2, but only considers the s_j-terms (and not the $\hat{s}^i_j$) in the exponent. Furthermore, because of the reduced entropy, the $\mathbf{O}(2^{-3k})$ bound from the lemma will become $\mathsf{poly} \cdot 2^{-k}$.

Finally, to justify the change from Game 6.i.2, it suffices to note that hence, A's view is essentially independent of

$$\gamma_1 r_1([s_1]_\mathbb{Z} \cdot vk + [s_3]_\mathbb{Z}) + \gamma_2 r_2([s_2]_\mathbb{Z} \cdot vk + [s_4]_\mathbb{Z}) \bmod |\mathbb{G}_\mathsf{rnd}|,$$

where $r_1 \neq r_2$ are the $\mathbb{G}_\mathsf{rnd}$-exponents of the considered C_1, C_2.

The remaining **Games** 6.i.3 **to Game** 8 are again as with Theorem 4.1, of course again applied to KDM queries under all public keys. The corresponding reductions to DDH, DCR, and the IND-CPA security of (E, D) apply verbatim. □

Acknowledgements

I would like to thank Hoeteck Wee for pointing out to me the connection between hash proof systems and KDM security. His interpretation of the schemes of [12] and [13] was the starting point for this work. I am grateful to Tibor Jager for very useful feedback and discussions. I am indebted to Baodong Qin, who pointed out two flaws (along with very helpful suggestions on how to fix them). Finally, I am thankful to the Crypto 2012 referees who found an annoying cyclic dependency in the PKE construction.

References

[1] Masayuki Abe, Melissa Chase, Bernardo David, Markulf Kohlweiss, Ryo Nishimaki, and Miyako Ohkubo. Constant-size structure-preserving signatures: Generic constructions and simple assumptions. In Xiaoyun Wang and Kazue Sako, editors, *ASIACRYPT 2012*, volume 7658 of *LNCS*, pages 4–24. Springer, December 2012.

[2] Tolga Acar, Mira Belenkiy, Mihir Bellare, and David Cash. Cryptographic agility and its relation to circular encryption. In Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 403–422, French Riviera, May 30 – June 3, 2010. Springer, Berlin, Germany.

[3] Pedro Adão, Gergei Bana, Jonathan Herzog, and Andre Scedrov. Soundness of formal encryption in the presence of key-cycles. In Sabrina De Capitani di Vimercati, Paul F. Syverson, and Dieter Gollmann, editors, *ESORICS 2005*, volume 3679 of *LNCS*, pages 374–396, Milan, Italy, September 12–14, 2005. Springer, Berlin, Germany.

[4] Benny Applebaum. Key-dependent message security: Generic amplification and completeness. In Kenneth G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 527–546, Tallinn, Estonia, May 15–19, 2011. Springer, Berlin, Germany.

[5] Benny Applebaum, David Cash, Chris Peikert, and Amit Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In Shai Halevi, editor, *CRYPTO 2009*, volume 5677 of *LNCS*, pages 595–618, Santa Barbara, CA, USA, August 16–20, 2009. Springer, Berlin, Germany.

[6] Boaz Barak, Iftach Haitner, Dennis Hofheinz, and Yuval Ishai. Bounded key-dependent message security. In Henri Gilbert, editor, *EUROCRYPT 2010*, volume 6110 of *LNCS*, pages 423–444, French Riviera, May 30 – June 3, 2010. Springer, Berlin, Germany.

[7] Mihir Bellare, Dennis Hofheinz, and Scott Yilek. Possibility and impossibility results for encryption and commitment secure under selective opening. In Antoine Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 1–35, Cologne, Germany, April 26–30, 2009. Springer, Berlin, Germany.

[8] John Black, Phillip Rogaway, and Thomas Shrimpton. Encryption-scheme security in the presence of key-dependent messages. In Kaisa Nyberg and Howard M. Heys, editors, *SAC 2002*, volume 2595 of *LNCS*, pages 62–75, St. John's, Newfoundland, Canada, August 15–16, 2003. Springer, Berlin, Germany.

[9] Dan Boneh, Xavier Boyen, and Hovav Shacham. Short group signatures. In Matthew Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 41–55, Santa Barbara, CA, USA, August 15–19, 2004. Springer, Berlin, Germany.

[10] Dan Boneh, Emily Shen, and Brent Waters. Strongly unforgeable signatures based on computational Diffie-Hellman. In Moti Yung, Yevgeniy Dodis, Aggelos Kiayias, and Tal Malkin, editors, *PKC 2006*, volume 3958 of *LNCS*, pages 229–240, New York, NY, USA, April 24–26, 2006. Springer, Berlin, Germany.

[11] Dan Boneh, Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from identity-based encryption. *SIAM Journal on Computing*, 36(5):1301–1328, 2007.

[12] Dan Boneh, Shai Halevi, Michael Hamburg, and Rafail Ostrovsky. Circular-secure encryption from decision diffie-hellman. In David Wagner, editor, *CRYPTO 2008*, volume 5157 of *LNCS*, pages 108–125, Santa Barbara, CA, USA, August 17–21, 2008. Springer, Berlin, Germany.

[13] Zvika Brakerski and Shafi Goldwasser. Circular and leakage resilient public-key encryption under subgroup indistinguishability - (or: Quadratic residuosity strikes back). In Tal Rabin, editor, *CRYPTO 2010*, volume 6223 of *LNCS*, pages 1–20, Santa Barbara, CA, USA, August 15–19, 2010. Springer, Berlin, Germany.

[14] Zvika Brakerski, Shafi Goldwasser, and Yael Tauman Kalai. Black-box circular-secure encryption beyond affine functions. In Yuval Ishai, editor, *TCC 2011*, volume 6597 of *LNCS*, pages 201–218, Providence, RI, USA, March 28–30, 2011. Springer, Berlin, Germany.

[15] Jan Camenisch and Anna Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In Birgit Pfitzmann, editor, *EUROCRYPT 2001*, volume 2045 of *LNCS*, pages 93–118, Innsbruck, Austria, May 6–10, 2001. Springer, Berlin, Germany.

[16] Jan Camenisch, Nishanth Chandran, and Victor Shoup. A public key encryption scheme secure against key dependent chosen plaintext and adaptive chosen ciphertext attacks. In Antoine Joux, editor, *EUROCRYPT 2009*, volume 5479 of *LNCS*, pages 351–368, Cologne, Germany, April 26–30, 2009. Springer, Berlin, Germany.

[17] David Cash, Matthew Green, and Susan Hohenberger. New definitions and separations for circular security. Cryptology ePrint Archive, Report 2010/144, 2010. `http://eprint.iacr.org/`.

[18] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive

chosen ciphertext attack. In Hugo Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 13–25, Santa Barbara, CA, USA, August 23–27, 1998. Springer, Berlin, Germany.

[19] Ronald Cramer and Victor Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 45–64, Amsterdam, The Netherlands, April 28 – May 2, 2002. Springer, Berlin, Germany.

[20] Ivan Damgård and Mats Jurik. A generalisation, a simplification and some applications of Paillier's probabilistic public-key system. In Kwangjo Kim, editor, *PKC 2001*, volume 1992 of *LNCS*, pages 119–136, Cheju Island, South Korea, February 13–15, 2001. Springer, Berlin, Germany.

[21] David Galindo, Javier Herranz, and Jorge L. Villar. Identity-based encryption with master key-dependent message security and leakage-resilience. In Sara Foresti, Moti Yung, and Fabio Martinelli, editors, *ESORICS 2012*, volume 7459 of *LNCS*, pages 627–642, Pisa, Italy, September 10–12, 2012. Springer, Berlin, Germany.

[22] Jens Groth and Amit Sahai. Efficient non-interactive proof systems for bilinear groups. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 415–432, Istanbul, Turkey, April 13–17, 2008. Springer, Berlin, Germany.

[23] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In Omer Reingold, editor, *TCC 2009*, volume 5444 of *LNCS*, pages 202–219. Springer, Berlin, Germany, March 15–17, 2009.

[24] Dennis Hofheinz. All-but-many lossy trapdoor functions. In David Pointcheval and Thomas Johansson, editors, *EUROCRYPT 2012*, volume 7237 of *LNCS*, pages 209–227, Cambridge, UK, April 15–19, 2012. Springer, Berlin, Germany.

[25] Dennis Hofheinz and Eike Kiltz. Secure hybrid encryption from weakened key encapsulation. In Alfred Menezes, editor, *CRYPTO 2007*, volume 4622 of *LNCS*, pages 553–571, Santa Barbara, CA, USA, August 19–23, 2007. Springer, Berlin, Germany.

[26] Dennis Hofheinz and Dominique Unruh. Towards key-dependent message security in the standard model. In Nigel P. Smart, editor, *EUROCRYPT 2008*, volume 4965 of *LNCS*, pages 108–126, Istanbul, Turkey, April 13–17, 2008. Springer, Berlin, Germany.

[27] Kaoru Kurosawa and Yvo Desmedt. A new paradigm of hybrid encryption scheme. In Matthew Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 426–442, Santa Barbara, CA, USA, August 15–19, 2004. Springer, Berlin, Germany.

[28] Tal Malkin, Isamu Teranishi, and Moti Yung. Efficient circuit-size independent public key encryption with KDM security. In Kenneth G. Paterson, editor, *EUROCRYPT 2011*, volume 6632 of *LNCS*, pages 507–526, Tallinn, Estonia, May 15–19, 2011. Springer, Berlin, Germany.

[29] Moni Naor and Moti Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *22nd ACM STOC*, Baltimore, Maryland, USA, May 14–16, 1990. ACM Press.

[30] Chris Peikert and Brent Waters. Lossy trapdoor functions and their applications. In Richard E. Ladner and Cynthia Dwork, editors, *40th ACM STOC*, pages 187–196, Victoria, British Columbia, Canada, May 17–20, 2008. ACM Press.

[31] Brent R. Waters. Efficient identity-based encryption without random oracles. In Ronald Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 114–127, Aarhus, Denmark, May 22–26, 2005. Springer, Berlin, Germany.