Key-Dependent Message Security: Generic Amplification and Completeness Theorems

Benny Applebaum*

October 7, 2010

Abstract

Key-dependent message (KDM) secure encryption schemes provide secrecy even when the attacker sees encryptions of messages related to the secret-key sk. Namely, the scheme should remain secure even when messages of the form $f(\mathsf{sk})$ are encrypted, where f is taken from some function class $\mathcal{F}$. A KDM *amplification* procedure takes an encryption scheme which satisfies basic level of KDM security with respect to some simple function class $\mathcal{F}_{\mathsf{basic}}$ and converts it into a new encryption scheme which achieves high level of KDM security with respect to a richer function class $\mathcal{F}_{\mathsf{ext}}$. Such procedures were recently studied by Brakerski et al. (ePrint 2009/485) and Barak et al. (EUROCRYPT 2010), who showed that a strong form of amplification is sometimes possible, provided that the underlying encryption scheme satisfies some special additional properties (i.e., entropic-KDM security and simulatable-KDM security).

In this work, we prove the first *generic* KDM amplification theorem which relies solely on the KDM security of the underlying scheme without making any other assumptions. Specifically, we show that an elementary form of KDM security against functions in which each output bit either copies or flips a single bit of the key (aka *projections*) can be amplified into KDM security with respect to any function family that can be computed in arbitrary fixed polynomial-time. Furthermore, our amplification theorem and its proof are insensitive to the exact setting of KDM security, and they hold in the presence of multiple-keys and in the symmetric-key/public-key and the CPA/CCA cases. As a result, we can amplify the security of all known KDM constructions, including ones that could not be amplified before.

Our main result is proven by combining ideas from the previous amplification theorems of Brakerski et al., and Barak et al., together with a novel use of the machinery of *randomized encoding* (Ishai and Kushilevitz, FOCS 2000). The resulting proof is very simple and, in some sense, illuminates a new connection between the previous approaches.

Finally, we study the minimal conditions under which full-KDM security (with respect to all functions) can be achieved. We show that under strong notion of KDM security, the existence of cyclic-secure fully-homomorphic encryption is not only sufficient for full-KDM security, as shown by Barak et al., but also necessary. On the other hand, we observe that for standard KDM security, this condition can be relaxed by adopting Gentry's bootstrapping technique (STOC 2009) to the KDM setting.

*School of Electrical Engineering, Tel-Aviv University, benny.applebaum@gmail.com. Work done while a postdoc at the Weizmann Institute of Science, supported by Koshland Fellowship.

1 Introduction

The study of secure encryption scheme is perhaps the most central subject in cryptography. Since the discovery of semantic security [24] till the formulation of CCA-security [31, 33, 17], modern cryptography have successfully developed increasingly stronger notions of security providing secrecy in highly adversarial settings. Still, all these strong notions of security guarantee secrecy only as long as the encrypted messages are independent of the secret key. This limitation dates back to the seminal work of Goldwasser and Micali [24] who observed that semantic security may not hold if the adversary gets to see an encryption of the secret key. For many years, such usage scenarios were considered as "security bugs" that should be prevented by system designers.

A decade ago, the assumption of independency between the secret key and the encrypted data was challenged by Camenisch and Lysyanskaya [15] and Black et al. [11]. Specifically, Camenisch and Lysyanskaya considered schemes that remain secure under a "key cycle" usage, where we have t keys organized in a cycle and each key is encrypted under its left neighbor. A generalization of this notion, called *key-dependent message* (KDM) security, was suggested by Black et al. Informally, an encryption is $\mathsf{KDM}^{(t)}$ secure with respect to a function class $\mathcal{F}$ if security holds even when the adversary can ask for an encryption of the message $M = f(\mathsf{sk}_1, \ldots, \mathsf{sk}_t)$ under the i-th public-key, where $\mathsf{sk}_1, \ldots, \mathsf{sk}_t$ are the secret keys present in the system and f is an arbitrary function in $\mathcal{F}$. This notion of security implies cyclic-security if $\mathcal{F}$ is expressive enough (e.g., contains all "projections" functions), and it becomes strictly stronger when the function class $\mathcal{F}$ grows. Hence, one would like to achieve KDM security while making the function class $\mathcal{F}$ as large as possible.

The notion of KDM security was extensively studied in the past few years in several flavors including the symmetric/public-key and the CPA/CCA settings [15, 11, 26, 8, 12, 16, 7, 27, 25, 4, 14, 2, 9, 13]. These works were motivated by the fundamental nature of the question as well as by concrete applications including encrypted storage systems (e.g., BitLocker [12]), anonymous credentials [15], and realization of security proofs at the framework of axiomatic security [1, 11, 3]. (See [12] for more motivations and details.)

Although much is known today about KDM security both on the positive and negative sides, it is still unclear whether a standard encryption scheme can be transformed into a scheme which provides $\mathsf{KDM}^{(t)}$ security, even with respect to a single key (i.e., $t = 1$) and simple non-trivial function families (e.g., projections).[1] Hence, it is natural to move forward and explore the possibility of building strong KDM security given a weak form of KDM security as a primitive. This makes sense as today, following the seminal work of Boneh et al. [12] and its follow-ups [16, 4, 13], it is known that a basic form of KDM security (with respect to the family of "affine functions") can be achieved in several settings under various concrete cryptographic assumptions. Therefore, following [14] we ask:

> Is there a *generic* transformation which *amplifies* KDM security from a weak family of functions $\mathcal{F}_{\mathsf{basic}}$ to a larger family of functions $\mathcal{F}_{\mathsf{ext}}$?

The two main features of such a procedure are *generality* – the transformation should work with any scheme which satisfies $\mathcal{F}_{\mathsf{basic}}$-KDM security without relying on any other additional property –

[1] It is known that KDM security with respect to sufficiently rich families of functions cannot be based on standard assumptions via fully black-box reductions [25]. However, this impossibility result (and its extension in [9]) does not hold for simple function class (e.g., projections). Moreover, this result is restricted to reductions which also treats the function family in a black-box manner, a restriction which can be bypassed as shown in [9] and in the current paper.

and large *amplification gap* – ideally, $\mathcal{F}_{\mathsf{basic}}$ is a very simple function class whereas $\mathcal{F}_{\mathsf{ext}}$ is as rich as possible. The question of KDM amplification was recently addressed by Brakerski et al. [14] and Barak et al. [9], who made an important progress by showing how to amplify the KDM security of several existing schemes. While these works achieve relatively large amplification gap, they fall short of providing full generality as they strongly rely on additional properties of the underlying scheme (i.e., *simulatable*-KDM security and *entropic*-KDM security). As a concrete example, it is unknown how to use any of these techniques to amplify the KDM-security of the symmetric-key encryption scheme of [4] which is based on the Learning Parity With Noise (LPN) assumption. (See Section 1.3 for more details about these works and their relation to our approach.)

1.1 Our Results

We give an affirmative answer to the above question by providing the first generic KDM amplification procedure. In particular, we consider the *projection* function class Proj^t of all functions $f : (\mathsf{sk}_1, \ldots, \mathsf{sk}_t) \mapsto v$ in which each output bit depends on (at most) single bit of the input. Namely, each output bit v_j is either fixed to a constant or copies/flips an original bit of one of the keys. We show that this elementary function family is *complete* in the following sense:

Theorem 1.1 (Completeness of projections (Informal))**.** *Let $\mathcal{F}_{\mathsf{ext}}$ be any function family which can be computed in some fixed polynomial time. Then, any encryption scheme which satisfies $\mathsf{KDM}^{(t)}$ security with respect to projections can be transformed into an encryption scheme which is $\mathsf{KDM}^{(t)}$-secure with respect to $\mathcal{F}_{\mathsf{ext}}$.*

Generality. Theorem 1.1 assumes nothing but KDM security regarding the underlying scheme. Furthermore, the theorem (and its surprisingly simple proof) is insensitive to the exact setting of KDM security: it holds for any number of keys (t), and in both symmetric-key/public-key and CPA/CCA settings. In all these cases, the new scheme is proven to be secure exactly in the same setting as the original scheme. This allows us, for example, to amplify the security of the affine-KDM secure scheme of [4], and obtain the first symmetric-key encryption scheme with strong KDM security based on the Learning Parity with Noise assumption.

Large gap. Theorem 1.1 provides a large amplification gap. In fact, this gap can be further expanded as follows. First, we can achieve *length-dependent* KDM security [9], which means that the target family $\mathcal{F}_{\mathsf{ext}}$ can be taken to be the family of all polynomial-size circuits whose size grows with their input and output lengths via a fixed polynomial rate (e.g., the circuit size is quadratic in the input and output lengths). This family is very powerful and it was shown to be rich enough for most known applications of KDM security [9].[2] (See Section 3 for details.) In addition, in the case of CPA security (both in the public-key and symmetric-key settings), we can weaken the requirement from the underlying scheme and ask for KDM security with respect to projections with a *single output*: namely, all Boolean functions $f(\mathsf{sk}_1, \ldots, \mathsf{sk}_t) \mapsto b$ which output a single bit of one of the keys, or its negation. This can be extended to the CCA setting via the transformations of [8, 16] (though in the public-key setting one has to assume, in addition, the existence of non-interactive zero-knowledge proofs for **NP**).

[2] Most of the statements in [9] refer to the slightly weaker notion of *Bounded KDM security* in which the circuit size grows only as a function of the input via a fixed polynomial rate. However, as observed in [9, Sec. 6] their construction actually satisfy the stronger definition of *length-dependent* KDM security.

The relaxation to single-output projections also enables a liberal interface to which we can easily plug previous constructions. For example, one can instantiate our reduction with schemes that enjoy KDM security with respect to affine functions, while almost ignoring technical details such as the underlying field and its representation. (These details required some effort in previous works. See the appendices in [14, 9, 13].) This, together with the simple proof of our main theorem, allows to simplify the proofs of [9, 13] for the existence of length-dependent KDM secure encryption scheme under the Decisional Diffie-Hellman (DDH) assumption [12], the Learning With Errors assumptions (LWE) [4], and the Quadratic Residuosity (QR) assumption [13].

Given this completeness theorem, the current status of KDM security resembles the status of other "complete" primitives in cryptography such as one-way functions or oblivious transfer [32, 18]: We do not know how to build these primitives from generic weaker assumptions, however, any instantiation of them suffices for an entire world of applications (i.e., all symmetric-key primitives in the case of one-way functions, and generic secure-computation in the case of oblivious transfer, cf. [22, 23]).

Beyond length-dependent security. Although length-dependent KDM security seems to suffice for most applications, one can strive for an even stronger notion of security in which the KDM function class contains all functions (or equivalently all functions computable by circuits of *arbitrary* polynomial size). It is somewhat likely that any length-dependent secure scheme actually achieves *full-KDM* security (see the discussion in [9]). Still, one may want to construct such a scheme in a provably secure way. As a basic feasibility result, it was shown in [9] that any fully homomorphic encryption scheme [20] which allows to encrypt the secret-key (i.e., "cyclic-secure") is also full-KDM secure. In light of the small number of FHE candidates [20, 34], and our little understanding of this notion, one may ask whether it is possible to relax this requirement and achieve full-KDM security under weaker assumptions.

We make two simple observations regarding this question. First, we consider the case of simulatable KDM security [9], in which it should be possible to simulate an encryption of $f(\mathsf{sk})$ given only the corresponding public-key in a way that remains indistinguishable even to someone who knows the secret-key. We show that in this setting the two notions: circular-secure FHE and full-KDM are equivalent. Hence, achieving full-KDM security under a relaxed assumption requires to use non-simulatable constructions.

Our second observation asserts that the bootstrapping technique of Gentry [20] can be used in the KDM setting as well (even for the case of non-simulatable constructions). That is, if one can construct an encryption scheme which guarantees KDM security with respect to circuits whose depth is only slightly larger than the depth of the decryption algorithm, then this scheme is actually fully KDM secure. Unfortunately, all known amplification techniques [9, 14] including the ones in this paper, amplify KDM security at the cost of making the decryption algorithm "deeper". Still, we view this observation as an interesting direction for future research.

1.2 Our Techniques

To formalize the question of KDM amplification, we define the notion of *reduction* between KDM function families $\mathcal{F}_{\mathsf{ext}} \leq_{\mathsf{KDM}} \mathcal{F}_{\mathsf{basic}}$ which means that any scheme that provides KDM security with respect to $\mathcal{F}_{\mathsf{basic}}$ can be transformed (via a fully black-box reduction) to a scheme that satisfies

KDM security with respect to $\mathcal{F}_{\mathsf{ext}}$. We describe a novel way to derive such KDM reductions based on the machinery of *randomized encoding* of functions [29, 6]. Before we explain this notion, let us start with the simpler case of *deterministic encoding*.

Say that a function f deterministically encodes a function g if for every x the output of $f(x)$ "encodes" the output of $g(x)$ in the sense that $g(x)$ can be efficiently computed based on $f(x)$ and vice versa. That is, there are two efficiently computable mappings S and R such that $S(g(x)) = f(x)$, and $R(f(x)) = g(x)$. Suppose that we are given a scheme which provides KDM security with respect to the encoding f, and we would like to immunize it against the function g. This can be easily achieved by modifying the encryption scheme as follows: to encrypt a message M we first translate it into the f-encoding by computing $S(M)$, and then encrypt the result under the original encryption scheme. Decryption is done by applying the original decryption algorithm, and then applying the recovery algorithm R to translate the result back to its original form. Observe that an encryption of $g(\mathsf{sk})$ in the new scheme is the same as an encryption of $S(g(\mathsf{sk})) = f(\mathsf{sk})$ under the original scheme. Hence, the KDM security of the new scheme with respect to g reduces to the KDM security of the original scheme with respect to f.

This simple idea provides a direct reduction with very nice structure: any KDM query for the new scheme is translated into a single KDM query for the original scheme. This simple single-query-to-single-query translation leads to high level of generality: the transformation is insensitive to the exact KDM setting (symmetric-key/public-key and CPA/CCA), to the number of keys, and it can be used with respect to large function families $\mathcal{F}_{\mathsf{ext}}$ and $\mathcal{F}_{\mathsf{basic}}$ as long as every function in $\mathcal{F}_{\mathsf{ext}}$ is encoded by some function in $\mathcal{F}_{\mathsf{basic}}$ via a pair of universal mappings S and R. On the down side, one may complain that security was not really *amplified*, as the function g and its encoding f are essentially equivalent. It turns out that this drawback can be easily fixed by letting f be a *randomized* encoding of g.

In the case of randomized encoding (RE), the function $f(x; r)$ depends not only on x but also on an additional random input r. For every fixed x, the output of $f(x; r)$ is now viewed as a distribution (induced by a random choice of r) which should encode the value of $g(x)$. Namely, there are two efficiently computable randomized mappings S and R such that for every x: (1) the distribution $S(g(x))$ is indistinguishable from $f(x; r)$, and (2) with high probability over the choice of r (or even with probability one) $R(f(x; r)) = g(x)$. One can view these conditions as saying that $g(x)$ is encoded by a *collection* of functions $\{h_r(x) = f(x; r)\}_r$.

Now suppose that our scheme is KDM secure with respect to the family $\{h_r(x) = f(x; r)\}_r$, then we can apply the above approach and get a scheme which satisfies KDM security with respect to g. The only difference is that now the message preprocessing step is randomized: To encrypt a message M first encode it by the randomized mapping $S(M)$, and then use the original encryption function. The security reduction is essentially the same except that a KDM query for g in the new scheme is emulated by an old KDM query for a *randomly chosen* function h_r. This idea can be easily extended to the case where all functions in $\mathcal{F}_{\mathsf{ext}}$ are encoded by functions in $\mathcal{F}_{\mathsf{basic}}$:

Theorem 1.2 (Informal)**.** *If $\mathcal{F}_{\mathsf{basic}}$ is an RE of $\mathcal{F}_{\mathsf{ext}}$, then $\mathcal{F}_{\mathsf{ext}} \leq_{\mathsf{KDM}} \mathcal{F}_{\mathsf{basic}}$.*

The crux of this theorem, is that, unlike deterministic encoding, randomized encoding can represent complicated functions by collections of very simple functions [29, 30, 6, 5]. Specifically, by combining the above theorem with the REs of [5], which, in turn, are based on Yao's garbled circuit [35], we obtain our main results (Thm. 1.1).

1.3 Comparison with BGK and BHHI

Our techniques are inspired by both [14] (BGK) and [9] (BHHI). We believe that our approach inherits the positive features of each of these works, and sheds new light on the way they relate to each other. Let us review the main ideas behind these constructions and explain how they compare to our solution.

1.3.1 The BGK reduction

The starting point in [14] is an encryption scheme which satisfies entropic KDM security with respect to $\mathcal{F}_{\mathsf{basic}}$. Roughly speaking, this means that KDM security should hold not only when sk is chosen uniformly from the key space $\mathcal{K} = \{0,1\}^k$ but also when it is chosen uniformly from a smaller domain $\mathcal{K}'$, e.g., $\mathcal{K}' = \{0,1\}^{k^\varepsilon}$. By relying on this notion, BGK shows that for every efficiently computable injective mapping $\alpha : \mathcal{K}' \to \mathcal{K}$, one can amplify security from $\mathcal{F}_{\mathsf{basic}}$ to the class $\mathcal{F}_{\mathsf{basic}} \circ \alpha$, i.e., with respect to functions $f(\alpha(\mathsf{sk}))$ for every $f \in \mathcal{F}_{\mathsf{basic}}$. The idea is to choose the key sk' from $\mathcal{K}'$ and employ the original scheme with the key $\mathsf{sk} = \alpha(\mathsf{sk}')$. This allows to translate a KDM query $f(\alpha(\mathsf{sk}'))$ for the new scheme into an entropic-KDM query $f(\mathsf{sk})$ for the old scheme.

The deterministic encoding (DE) approach is inspired by the BGK approach, and can be seen as a complementary solution. BGK extends a function $f : \mathcal{K} \to \mathcal{M}$ to $f \circ \alpha : \mathcal{K}' \to \mathcal{M}$ by shrinking the key space (from $\mathcal{K}$ to $\mathcal{K}'$), whereas in the DE approach $f : \mathcal{K} \to \mathcal{M}$ is extended to $R \circ f : \mathcal{K} \to \mathcal{M}'$ by padding messages which effectively shrinks the message space (from $\mathcal{M}$ to $\mathcal{M}' = R(\mathcal{M})$).

As a result BGK enjoys a similar attractive security reduction with single-query-to-single-query translation. This leads to flexibility with resect to the KDM *setting*. Indeed, although the BGK approach is not fully general due to its use of entropic-KDM security (a notion which seems stronger than standard KDM security), it immediately generalizes to the CCA and the symmetric-key settings, as long as the underlying scheme provides entropic-KDM security.

It should be mentioned that in our approach the amplification is achieved by modifying the encryption algorithm, rather than the key-generation algorithm as in BGK. This minor difference turns to have a considerable effect on the amplification-gap. First, it allows to use fresh randomness in every application of the encryption algorithm, and so the linkage between functions in $\mathcal{F}_{\mathsf{ext}}$ to functions in $\mathcal{F}_{\mathsf{basic}}$ can be *randomized*. Indeed, this is exactly what allows us to exploit the power of randomized encoding. In contrast, the BGK approach tweaks the key-generation algorithm and so the relation between $\mathcal{F}_{\mathsf{ext}}$ to $\mathcal{F}_{\mathsf{basic}}$ is bounded to be deterministic. In addition, since our modification happens in the encryption (and decryption) phases, we can let the function class $\mathcal{F}_{\mathsf{ext}}$ grow not only with the security parameter but also with the size of the messages. This leads to the strong notion of length-dependent security, and in addition allows to achieve $\mathsf{KDM}^{(t)}$ where the number of keys t grows both with the message length and the security parameter.

In contrast, the family $\mathcal{F}_{\mathsf{ext}}$ of BGK cannot grow with the message length, and it can only contain a polynomial number of functions. This limitation prevents it from being used in applications which require KDM security wrt larger functions classes (e.g., secure realization of symbolic protocols with axiomatic proofs of security). Furthermore, amplification for large number of keys can be achieved only at the expense of putting more restrictions on the underlying scheme (i.e., simulatable KDM security). On the other hand, assuming these additional properties, the BGK approach can get $\mathsf{KDM}^{(t)}$ for arbitrary unbounded t with respect to some concrete function families (e.g., constant degree polynomials), whereas in our approach t is always bounded by some fixed polynomial (in

the security parameter and message length).[3] Finally, it is important to mention that the BGK reduction treats $\mathcal{F}_{\mathsf{ext}}$ in a black-box way, while the randomized encoding approach treats this class in a non-black-box way.

1.3.2 The BHHI reduction

The BHHI approach relies on a novel connection between homomorphic encryptions and KDM security. First, it is observed that in order to obtain KDM security with respect to $\mathcal{F}_{\mathsf{ext}}$ it suffices to construct a scheme which provides both cyclic-security (i.e., KDM security with respect to the identity function) and homomorphism with respect to a function family $\mathcal{F}_{\mathsf{ext}}$, i.e., it should be possible to convert a ciphertext $C = \mathsf{Enc}_{\mathsf{pk}}(M)$ into $C' = \mathsf{Enc}_{\mathsf{pk}}(f(M))$ for every $f \in \mathcal{F}_{\mathsf{ext}}$. Indeed, the homomorphism property can be used to convert a ciphertext $\mathsf{Enc}_{\mathsf{pk}}(\mathsf{sk})$ into a the ciphertext $\mathsf{Enc}_{\mathsf{pk}}(f(\mathsf{sk}))$, and so cyclic-security is amplified to $\mathcal{F}_{\mathsf{ext}}$-KDM security.

BHHI construct such an encryption scheme by combining a two-party secure computation protocol with two messages (i.e., Yao's garbled circuit technique [35]) with a strong version of oblivious transfer which satisfies an additional *cyclic-security* property. The latter primitive is referred to as *targeted encryption* (TE). The basic idea is to view the homomorphic property as a secure-computation task in which the first party holds the message M and the second party holds the function f. The cyclic nature of the TE primitive allows to implement this homomorphism even when the input M is the secret-key. Finally, BHHI show that TE can be constructed based on affine-KDM secure encryption scheme which satisfies a strong notion of simulation: There exists a simulator which given the public-key pk can simulate a ciphertext $\mathsf{Enc}_{\mathsf{pk}}(f(\mathsf{sk}))$ in a way which is indistinguishable even for someone who holds the secret-key.

The BHHI construction seems conceptually different from our RE approach (i.e., homomorphism vs. encoding). Moreover, the construction itself is not only syntactically different, but also relies on different building blocks (e.g., TE). Still, the RE construction share an important idea with BHHI: The use of secure-computation techniques. It is well known that REs are closely related to secure multiparty-computation (MPC) protocols[4], and, indeed, the role of REs in our reduction resembles the role of MPC in BHHI. In both solutions at some point the security reduction applies the RE/MPC to the function f in $\mathcal{F}_{\mathsf{ext}}$. Furthermore, both works achieve strong KDM security by instantiating the RE/MPC with Yao's garbled circuit (GC) — a tool which leads to both: stand-alone RE construction [5] and, when equipped with an OT, to a two-party secure-computation protocol.

It should be emphasized, however, that the actual constructions differ in some important aspects. While we essentially encrypt the whole GC-based encoding under the underlying KDM encryption scheme, BHHI tweak the GC protocol with a cyclic-secure OT (i.e., TE). Pictorially, our underlying KDM-secure scheme "wraps" the GC encoding, whereas in BHHI the KDM-secure primitive is "planted inside" the GC protocol. This difference affects both generality and simplicity as follows.

First, BHHI are forced to implement a KDM-secure OT, a primitive which seems much stronger than standard KDM secure encryption schemes. For example, KDM-secure symmetric-key en-

[3] In fact, we can achieve a slightly stronger notion. Assuming that the underlying scheme satisfies $\mathsf{KDM}^{(t)}$ security for arbitrary t's (as in [12, 4]), we get a $\mathsf{KDM}^{(t)}$ secure scheme where there exists an unbounded number of keys in the system, but the arity of the KDM functions available to the adversary is polynomially bounded (in the security parameter and message length). Still, these functions can be applied to arbitrary subsets of the keys.

[4] In fact, REs were originally defined as a strong form of non-interactive reductions for MPC [29].

cryption schemes can be constructed at the presence of a random oracle [11] while OT protocols cannot [28].[5] Moreover, as we already mentioned, although TE can be based on several known affine-secure KDM schemes (i.e., ones which enable strong simulation), the LPN assumption (with constant error-rate) is a concrete example under which symmetric-key encryption scheme with KDM-security wrt affine functions exist, yet OT is not known to exist. Furthermore, since BHHI send the garbled circuit in the clear, it is not hard to show that the resulting scheme is not CCA-secure even if the TE provides CCA security. Finally, the modification of the GC protocol leads to a relatively complicated security proof.

2 Preliminaries

For a positive integer $n \in \mathbb{N}$, denote by $[n]$ the set $\{1, \ldots, n\}$. A function $\varepsilon(n)$ is *negligible* if it tends to zero faster than $1/n^c$ for every constant $c > 0$. The term *efficient* refers to probabilistic polynomial time.

Efficient functions and randomized functions. A *randomized function* $f : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ is a function whose second input is treated as a random input. We write $f(x;r)$ to denote the evaluation of f on deterministic input x and random input r, and typically assume length regularity and efficient evaluation as follows: there are efficiently computable polynomials $m(n)$ and $\ell(n)$ and an efficiently computable circuit family $\left\{f_n : \{0,1\}^n \times \{0,1\}^{m(n)} \to \{0,1\}^{\ell(n)}\right\}$ which computes the restriction of f to n-bit deterministic inputs. If the function is *not* length regular, we assume that the circuit family is indexed by a pair of input and output parameters (n, ℓ), and require evaluation in time $\mathrm{poly}(n, \ell)$. Finally, a *deterministic* function corresponds to the special case where $m(n) = 0$.

Function ensembles. A *function ensemble* is a collection of functions $\{f_z\}_{z \in Z}$ indexed by an index set $Z \subseteq \{0,1\}^*$, where for each z the function f_z has a finite domain $\{0,1\}^{n(z)}$ and a finite range $\{0,1\}^{\ell(z)}$, where $n, \ell : \{0,1\}^* \to \mathbb{N}$. (This means that different functions may have different domains but each fixed function f_z is regular.) By default, we assume that ensembles are efficiently computable, that is, the functions $n(z), \ell(z)$, as well as the function $F(z,x) = f_z(x)$ are computable in time $\mathrm{poly}(|z|)$. Hence $n(z), \ell(z) < \mathrm{poly}(|z|)$. We also assume that $|z| < \mathrm{poly}(n(z), \ell(z))$.

Randomized encoding of functions. Intuitively, a randomized encoding of a function $f(x)$ is a randomized mapping $\hat{f}(x,r)$ whose output distribution depends only on the output of f. We formalize this intuition via the notion of *computationally-private randomized encoding* of [5], while adopting the original definition from a non-uniform adversarial setting to the uniform setting where adversaries are modeled by probabilistic polynomial-time Turing machines. Let $f = \left\{f_n : \{0,1\}^n \to \{0,1\}^{\ell(n)}\right\}$ be a function and $\hat{f} = \left\{\hat{f}_n : \{0,1\}^n \times \{0,1\}^{m(n)} \to \{0,1\}^{s(n)}\right\}$ be a randomized function, which are both efficiently computable. We say that $\hat{f}$ *encodes* f, if there exist an efficient recovery algorithms Rec and an efficient simulator Sim that satisfy the following:

[5] It seems that a similar statement holds even for public-key KDM-secure schemes. See [11, 21].

- **perfect correctness.** For every $x \in \{0,1\}^n$, the error probabilities $\Pr[\mathsf{Rec}(1^n, \hat{f}(x, \mathsf{U}_{m(n)})) \neq f(x)]$ and $\Pr[\mathsf{Rec}(1^n, \mathsf{Sim}(1^n, f(x))) \neq f(x)]$ are both zero.[6]
- **computational privacy.** For every efficient adversary $\mathcal{A}$ we have that

$$\Pr[\mathcal{A}^{\hat{f}(\cdot;\mathsf{U})}(1^n) = 1] - \Pr[\mathcal{A}^{\mathsf{Sim}(f(\cdot))}(1^n) = 1] < \mathrm{neg}(n),$$

  where the oracles are defined as follows: Given x the first oracle returns a sample from $\hat{f}(x; \mathsf{U}_{m(|x|)})$ and the second oracle returns a sample from $\mathsf{Sim}(1^{|x|}, f(x))$.

This notion is naturally extended to functions $f_{n,\ell}$ which are not length-regular and are indexed by both input and output lengths. However, we always assume that privacy is parameterized *only* with the input length (i.e., the adversary's running-time/distiunguishing-probability should be polynomial/negligible in the input length.) Note that, without loss of generality, we can assume that the relevant output length ℓ, is always known to the decoder and simulator.

Encryption schemes (syntax). An encryption scheme consists of three probabilistic-polynomial time algorithms $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$, where Gen is a key generation algorithm which given a security parameter 1^k outputs a pair $(\mathsf{sk}, \mathsf{pk})$ of decryption and encryption keys; Enc is an encryption algorithm that takes a message $M \in \{0,1\}^*$ and an encryption key pk and outputs a ciphertext C; and Dec is a randomized decryption algorithm that takes a ciphertext C and a decryption key sk and outputs a plaintext M'. We also assume that both algorithms take the security parameter 1^k as an additional input, but typically omit this dependency for simplicity. Correctness requires that the decryption error

$$\max_{M \in \{0,1\}^*} \Pr_{(\mathsf{pk},\mathsf{sk}) \xleftarrow{R} \mathsf{Gen}(1^k)} [\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{pk}(M)) \neq M],$$

should be negligible in k, where the probability is taken over the randomness of $\mathsf{Gen}, \mathsf{Enc}$ and Dec. For security parameter k, let $\mathcal{K}_k$ denote the space from which decryption keys are chosen. Without loss of generality, we always assume that $\mathcal{K}_k = \{0,1\}^k$.

Following Goldreich [23], we note that the above definition corresponds to both public-key and symmetric-key encryption schemes where the latter correspond to the special case in which the decryption key sk and encryption key pk are equal. As we will see, the difference between the two settings will be part of the security definitions.

3 KDM-Security

Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme with key space $\mathcal{K} = \{\mathcal{K}_k\}$. Let $t : \mathbb{N} \to \mathbb{N}$ be a function. A t-ary KDM function ensemble is an efficient ensemble of functions $\mathcal{F} = \left\{ f_{k,z} : \mathcal{K}_k^{t(k)} \to \{0,1\}^* \right\}_{(k,z)}$. We let $\mathcal{F}_k$ denote the set $\left\{ f_{k,z} : \mathcal{K}_k^{t(k)} \to \{0,1\}^* \right\}_z$. We define KDM-CCA security in the public-key setting with respect to $\mathcal{F}$ via a game that takes place between a challenger and an adversary $\mathcal{A}$, and is indexed by the security parameter k. The game is defined in Figure 1.

[6] Previous definitions require only that the first quantity is zero, however, all known constructions (of perfectly-correct randomized encoding) satisfy the current (stronger) definition.

- **Initialization.** The challenger randomly chooses a bit $b \xleftarrow{R} \{0,1\}$ and $t = t(k)$ key-pairs $(\mathsf{sk}_1, \mathsf{pk}_1) \ldots, (\mathsf{sk}_t, \mathsf{pk}_t)$ by invoking $\mathsf{Gen}(1^k)$ for t times. The adversary $\mathcal{A}$ can send a "public-key" query and ask to see all the encryption keys $(\mathsf{pk}_1, \ldots, \mathsf{pk}_t)$.
- **Queries.** The adversary $\mathcal{A}$ may adaptively make polynomially-many queries of the following types:
  - **Encryption (KDM) queries.** $\mathcal{A}$ makes a query of the form (i, f) where $i \in [t]$ and $f \in \mathcal{F}_k$. The challenger computes $M = f(\mathsf{sk}_1, \ldots, \mathsf{sk}_t)$ and outputs
  $$C \xleftarrow{R} \begin{cases} \mathsf{Enc}(\mathsf{pk}_i, M) & \text{if } b = 1, \\ \mathsf{Enc}(\mathsf{pk}_i, 0^{|M|}) & \text{if } b = 0. \end{cases}$$
  - **Decryption queries.** $\mathcal{A}$ can also make queries of the form (i, C) where $i \in [t]$, as long as C was not given as an answer of a previous encryption query of the form (i, f) for some $f \in \mathcal{F}_k$. The challenger responds with $M = \mathsf{Dec}_{\mathsf{sk}_i}(C)$ regardless of the value of b.
- **Final phase.** The adversary attempts to guess b and outputs a bit $b' \in \{0,1\}$.

Figure 1: The KDM game with respect to the function ensemble $\mathcal{F} = \{\mathcal{F}_k\}$. The notion of public-key query captures the public-key setting (as opposed to the symmetric-key setting).

By restricting the power of the adversary in the above game we get other KDM settings. Specifically, the symmetric-key setting corresponds to adversaries of type sym who do not ask public-key queries, and the CPA setting corresponds to adversaries of type CPA who do not make decryption queries. Hence, we can classify KDM adversaries into one of the following four *types*: $(\mathsf{pub}, \mathsf{CCA})$, $(\mathsf{pub}, \mathsf{CPA})$, $(\mathsf{sym}, \mathsf{CCA})$, and $(\mathsf{sym}, \mathsf{CPA})$.

Definition 3.1. (KDM-secure encryption) *Let $T \in \{\mathsf{pub}, \mathsf{sym}\} \times \{\mathsf{CCA}, \mathsf{CPA}\}$. An encryption scheme is* T-KDM secure *with respect to a function ensemble $\mathcal{F}$ if every polynomial-time attacker $\mathcal{A}$ of type T has at most negligible advantage in guessing the value of the bit b in the KDM game, where the running time and the advantage are measured as functions of the security parameter k.*

Interesting KDM functions ensembles. For every $t = t(k)$ and for every type T we consider the following ensembles:

- **Constants, selectors, and projections**. If $\mathcal{F}_k$ contains all constant functions $\{f_M : (\mathsf{sk}_1, \ldots, \mathsf{sk}_t) \mapsto M\}_M$, then, as observed in [12], KDM security implies standard security (with respect to the type T) as fixed KDM functions can emulate a standard encryption oracle. If the ensemble contains in addition all selector functions $\{f_j : (\mathsf{sk}_1, \ldots, \mathsf{sk}_t) \mapsto \mathsf{sk}_j\}_{j \in [t]}$, we get the notion of *clique* security [12] (which is stronger than *circular security* [15]), that is, the scheme is secure even if the adversary sees encryptions of the form $\mathsf{Enc}_{\mathsf{pk}_i}(\mathsf{sk}_j)$ for every $i, j \in [t]$. Another elementary class, that slightly generalizes the previous ones, is the class of all functions $f : (\vec{\mathsf{sk}}) \mapsto v$ in which each output bit depends on (at most) single bit of the input $\vec{\mathsf{sk}}$. Namely, the j-th output bit v_j is either fixed to a constant or copies/flips an

original bit of one of the keys, i.e., $v_j \in \{0, 1, \mathsf{sk}_{i,q}, 1 - \mathsf{sk}_{i,q}\}$, where $\mathsf{sk}_{i,q}$ is the q-th bit of the i-th secret key. We refer to this class as the class of *projections* and let $\mathsf{Proj}^t_{k,\ell}$ denote the restriction of this class to functions of input length kt and output length $\ell(k)$. Note that this is a subclass of the class of affine functions $L : \mathbb{F}_2^{kt} \to \mathbb{F}_2^{\ell(k)}$.

- **Polynomial-size circuits [9].** For polynomials $p(\cdot)$ and $\ell(\cdot)$, let $\mathcal{C}^t_{k,\ell,p}$ denote the class of all circuits $C : \{0,1\}^{kt} \to \{0,1\}^{\ell(k)}$ of size at most $p(k) + p(\ell)$. Security with respect to this class is denoted by (p, ℓ)-*bounded circuit-size* KDM security. A slightly stronger notion of security is *p-length-dependent* KDM security which means that the scheme is KDM secure with respect to $\mathcal{C}^t_{k,\ell,p}$ for *every* polynomial ℓ. While, ultimately one would like to have KDM security with respect to all polynomial-size circuits (for arbitrary polynomial), it seems that p length-dependent security, say for quadratic p, may be considered to be almost as powerful since it allows the adversary to use larger circuits by encrypting longer messages. In particular, one can represent essentially any polynomial-time computable function via padding. That is, if a function f is not in the class since its circuit is too large, then a "padded" version f' of f in which the output is padded with zeroes does fall into the ensemble. Furthermore, in [9] it was shown that if p is sufficiently large (e.g., the quadratic polynomial) then length-dependent security is sufficient for axiomatic-security applications (i.e., it gives the ability to securely instantiate symbolic protocols with axiomatic proofs of security).

The above definitions become stronger when the arity t grows. At one extreme, one may consider a single scheme which satisfies any of the above definitions for an arbitrary polynomial $t(k)$, and at the other extreme one may consider the case of $t = 1$, which is still non-trivial even for projection functions.

Reductions among KDM-ensembles. We say that a KDM function ensemble $\mathcal{F}_{\mathsf{ext}}$ KDM-reduces to another KDM function ensemble $\mathcal{F}_{\mathsf{basic}}$ (in symbols $\mathcal{F}_{\mathsf{ext}} \leq_{\mathsf{KDM}} \mathcal{F}_{\mathsf{basic}}$) if there exists a transformation which converts an encryption scheme PKC that is KDM secure wrt to $\mathcal{F}_{\mathsf{basic}}$ into an encryption scheme $\widehat{\mathsf{PKC}}$ which is KDM secure wrt to $\mathcal{F}_{\mathsf{ext}}$. Formally, such a (black-box) reduction is composed of (1) (construction) an encryption scheme $\widehat{\mathsf{PKC}}$ which is given an oracle access to the scheme PKC; and (2) (security reduction) an efficient algorithm $\mathcal{B}$ such that for any adversary $\mathcal{A}$ that α-breaks the KDM-security of the scheme PKC wrt to $\mathcal{F}_{\mathsf{basic}}$, the adversary $\mathcal{B}^{\mathcal{A},\mathsf{PKC}}$ breaks the scheme $\widehat{\mathsf{PKC}}$ wrt to $\mathcal{F}_{\mathsf{ext}}$ with similar probability (up to a negligible loss). This definition can be instantiated with respect to all four different types. We say that the reduction is *type-preserving* if $\mathcal{B}^{\mathcal{A},\mathsf{PKC}}$ always ask the same type of queries that $\mathcal{A}$ asks in the KDM game. Type preserving reduction extend KDM-security while being insensitive to the concrete setting which is being used. Formally,

Lemma 3.2 (KDM-reductions)**.** *Suppose that the KDM function ensemble $\mathcal{F}_{\mathsf{ext}}$ KDM-reduces to the ensemble $\mathcal{F}_{\mathsf{basic}}$ via a type-preserving reduction $(\widehat{\mathsf{PKC}}, \mathcal{B})$. Then, for every $T \in \{\mathsf{pub}, \mathsf{sym}\} \times \{\mathsf{CCA}, \mathsf{CPA}\}$, if the encryption scheme PKC is T-KDM secure with respect to $\mathcal{F}_{\mathsf{basic}}$, then the encryption scheme $\widehat{\mathsf{PKC}}^{\mathsf{PKC}}$ is T-KDM secure with respect to $\mathcal{F}_{\mathsf{ext}}$.*

4 Reductions and Completeness results

4.1 KDM reductions via randomized encoding

Let $\mathcal{F} = \{f_{k,z}\}$ and $\mathcal{G} = \{g_{k,w}\}$ be a pair of KDM function ensembles with the same arity $t = t(k)$. We say that $\mathcal{F}$ is *encoded* by $\mathcal{G}$ if functions in $\mathcal{F}$ have a randomized encoding, such that for every function $f(x) \in \mathcal{F}_{\mathsf{ext}}$ the encoding $\hat{f}(x;r)$, restricted to any fixed random string r, is in $\mathcal{F}_{\mathsf{basic}}$. More formally, the evaluation function $F_k(z,x)$ of $\mathcal{F}$ should have a randomized encoding $\hat{F}_k((z,x);r)$ such that for every fixing of r and index z, the function $\hat{F}_{k,z,r}(x) = \hat{F}(k,z,x;r)$ corresponds to a function $g_{k,w}$ in $\mathcal{G}$, where the mapping from (z,r) to w should be efficiently computable in $\mathrm{poly}(k)$ time. Note that this means that the simulator and decoder are *universal* for all indices z, and depend only in the value of k.[7]

Theorem 4.1 (main theorem)**.** *Suppose that the KDM function ensemble $\mathcal{F}_{\mathsf{basic}}$ encodes the KDM function ensemble $\mathcal{F}_{\mathsf{ext}}$. Then, $\mathcal{F}_{\mathsf{ext}}$ KDM-reduces to $\mathcal{F}_{\mathsf{basic}}$ via a type-preserving reduction.*

To prove the theorem we need to describe a construction and a security reduction. From now on, let Sim and Rec be the universal simulator and recovery algorithm which establish the encoding of $\mathcal{F}_{\mathsf{ext}}$ by $\mathcal{F}_{\mathsf{basic}}$.

Construction 4.2. *Given an oracle access to the encryption scheme $\mathsf{PKC} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$, we define the scheme $\widehat{\mathsf{PKC}}$ as follows*

$$\widehat{\mathsf{Gen}}(1^k) = \mathsf{Gen}(1^k) \qquad \widehat{\mathsf{Enc}}_{\mathsf{pk}}(M) = \mathsf{Enc}_{\mathsf{pk}}(\mathsf{Sim}(M)) \qquad \widehat{\mathsf{Dec}}_{\mathsf{sk}}(C) = \mathsf{Rec}(\mathsf{Dec}_{\mathsf{sk}}(C)),$$

where all algorithms (i.e., encryption, decryption, simulator and recovery) get the security parameter 1^k as an additional input.

It is not hard to show that $\widehat{\mathsf{PKC}}$ satisfies the syntactic requirements of encryption schemes, namely correctness.

Lemma 4.3 (correctness)**.** *The decryption error of the scheme $\widehat{\mathsf{PKC}}$ is the same as the decryption error of PKC, and so it is negligible.*

Proof. The probability that a message M is incorrectly decrypted is bounded by

$$\Pr_{(\mathsf{pk},\mathsf{sk}) \xleftarrow{R} \mathsf{Gen}(1^k), M' \xleftarrow{R} \mathsf{Sim}(M)} [\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{pk}}(M')) \neq M'] + \Pr[\mathsf{Rec}(M') \neq M],$$

since the second term is 0, due to the (perfect) correctness of the encoding, we can bound the above by $\max_{M' \in \{0,1\}^*} \Pr[\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{pk}}(M')) \neq M']$, where $(\mathsf{pk},\mathsf{sk}) \xleftarrow{R} \mathsf{Gen}(1^k)$. □

Next, We show that the security of $\widehat{\mathsf{PKC}}$ can be based on those of PKC.

[7] In fact, the encoding itself may access in a black-box manner to the underlying encryption scheme PKC (or any cryptographic primitive which can be based on it via a black-box reduction, e.g., one-way function). More precisely, our results hold (i.e., lead to black-box KDM reduction/construction) even if the simulator and decoder use such a BB access, as well as in the case that such access is needed to compute the mapping from an index/randomness pair (z,r) to the index w of the function $g_{k,w} = \hat{F}_{k,z,r}(x)$.

Reduction 4.4. *Given an oracle access to an adversary $\mathcal{A}$ that KDM-attacks PKC wrt $\mathcal{F}_{\mathsf{basic}}$ we define an adversary $\mathcal{B}$ that KDM-attacks $\widehat{\mathsf{PKC}}$ wrt $\mathcal{F}_{\mathsf{ext}}$ as follows:*

- ***Initialization:*** *$\mathcal{B}$ invokes $\mathcal{A}$. If $\mathcal{A}$ asks for the encryption keys then $\mathcal{B}$ makes a similar query and passes the answer to $\mathcal{A}$.*
- ***Encryption query:*** *If $\mathcal{A}$ makes an encryption query (i, f), for $i \in [t]$ and $f \in \mathcal{F}_{\mathsf{ext}}$, then $\mathcal{B}$ does the following: She uniformly chooses randomness r for the randomized encoding $\hat{f}$ of f, and asks the encryption query (i, g) where $g(\cdot) = \hat{f}(\cdot, r)$ which, by our assumption, is in $\mathcal{F}_{\mathsf{basic}}$. The answer of the challenger is being sent to $\mathcal{A}$.*
- ***Decryption query:*** *If $\mathcal{A}$ makes a decryption query (i, C), then $\mathcal{B}$ checks that it is legal (by inspecting all previous encryption queries), and if so, (1) passes the same decryption query to the challenger, (2) applies the recovery algorithm Rec to the result, and (3) sends it back to $\mathcal{A}$.*
- ***Termination:*** *$\mathcal{B}$ terminates with the same output of $\mathcal{A}$.*

Note that the above reduction is indeed type-preserving. Before we formally prove the correctness of the reduction, observe that, intuitively, the difference between the emulated view of $\mathcal{A}$ and the view of $\mathcal{A}$ in the actual game, is only due to the difference in the way encryption queries are answered when the challenger is in the "real-mode", i.e., where the challenge bit b equal to 1. In the real game, encryptions are computed properly as $\widehat{\mathsf{Enc}}_{\mathsf{pk}_i}(f(\vec{\mathsf{sk}})) = \mathsf{Enc}_{\mathsf{pk}_i}(\mathsf{Sim}(f(\vec{\mathsf{sk}})))$, whereas in the emulated game they are computed by $\mathsf{Enc}_{\mathsf{pk}_i}(\hat{f}(\vec{\mathsf{sk}}, \mathsf{U}))$. However, this difference should not be noticeable due to the privacy of the randomized encoding. Formally, we prove the following lemma.

Lemma 4.5. *If $\mathcal{A}$ is an efficient adversary that breaks $\widehat{\mathsf{PKC}}$ wrt $\mathcal{F}_{\mathsf{ext}}$ with advantage $\alpha(k)$, then the adversary $\mathcal{B}^{\mathcal{A},\mathsf{PKC}}$ breaks PKC wrt $\mathcal{F}_{\mathsf{basic}}$ with advantage $\alpha(k) - \mathrm{neg}(k)$.*

Proof. We show that if the claim does not hold then the privacy of the randomized encoding can be broken. Formally, let $F(z, x) = f_z(x)$ be the uniform evaluation function of $\mathcal{F}_{\mathsf{ext}} = \{f_z\}$ and let $\hat{F}(z, x; r)$ be the encoding of F. We define the following distinguisher $\mathcal{D}$ which, given an oracle access to either $\hat{F}(\cdot; \mathsf{U})$ or to $\mathsf{Sim}(F(\cdot))$, attempts to distinguish between the two. The adversary $\mathcal{D}$ emulates the challenger: It tosses a coin b, and generates a key vector $(\mathsf{sk}_i, \mathsf{pk}_i)_{i\in[t]}$ by executing the key-generation algorithm $\mathsf{Gen}(1^k)$ for t times. Then $\mathcal{D}$ invokes $\mathcal{A}$. If $\mathcal{A}$ asks for the public-keys, $\mathcal{D}$ passes them to him. If $\mathcal{A}$ makes an encryption query (i, f_z) then $\mathcal{D}$ calls its oracle with the value $F(z, \vec{\mathsf{sk}})$. Let M denote the answer of the oracle. The distinguisher computes the ciphertext $C = \mathsf{Enc}_{\mathsf{pk}_i}(M)$ if $b = 1$, and $C = \mathsf{Enc}_{\mathsf{pk}_i}(0^{|M|})$ otherwise. Then $\mathcal{D}$ sends the ciphertext C to $\mathcal{A}$. If $\mathcal{A}$ asks for decryption query (i, C), the distinguisher checks that it is legal by inspecting all previous encryption queries, and if so, sends $\mathsf{Dec}_{\mathsf{sk}_i}(C)$. The distinguisher halts with output 1 if and only if $\mathcal{A}$ guesses the bit b correctly.

Note that: (1) If $\mathcal{D}$ gets an oracle access to $\mathsf{Sim}(F(\cdot))$ then the view of $\mathcal{A}$ is distributed exactly as in the real game and so in this case $\mathcal{D}$ outputs 1 with probability $\alpha(k)$; (2) If $\mathcal{D}$ gets an oracle access to $\hat{F}(\cdot; \mathsf{U})$ then the view of $\mathcal{A}$ is distributed exactly as in the above reduction (when $\mathcal{B}$ emulates the game). Hence, by the privacy of the encoding, the distinguisher $\mathcal{D}^{\hat{F}(\cdot;\mathsf{U})}$ outputs 1 with probability at most $\alpha(k) - \mathrm{neg}(k)$, and so the claim follows. □

4.2 Completeness of projections

In [5], it is shown, based on Yao's garbled circuit technique, that efficiently computable functions can be encoded by decomposable encoding in which every bit depends on at most single bit of the deterministic input. By combining this fact with Thm 4.1, we get the following:

Proposition 4.6 (Completeness of projections)**.** *For every polynomials $p(\cdot), t(\cdot), \ell(\cdot)$, there exists a polynomial $q(\cdot)$ for which*

$$\mathcal{C}^t_{k,\ell,p} \leq_{\mathsf{KDM}} \mathsf{Proj}^t_{k,q}, \qquad \mathcal{C}^t_{k,p} \leq_{\mathsf{KDM}} \mathsf{Proj}^t_{k}, \tag{1}$$

where $\mathcal{C}^t_{k,\ell,p}$ denotes the t-ary ensemble of p-bounded circuits of output length ℓ, $\mathsf{Proj}^t_{k,q}$ denotes the t-ary ensemble of projections of output length q, $\mathcal{C}^t_{k,p} = \bigcup_{a\in\mathbb{N}} \mathcal{C}^t_{k,k^a,p}$, and $\mathsf{Proj}^t_{k} = \bigcup_{a\in\mathbb{N}} \mathsf{Proj}^t_{k,k^a}$. Furthermore, the reductions are type-preserving.

Hence, one can upgrade KDM security from (almost) the weakest KDM function ensemble to the very powerful notion of p-length dependent KDM security.

Proof. In [5], it was shown that, based on Yao's garbled circuit technique [35], any efficiently computable circuit family $\{g_k(x)\}$ of circuit complexity $a(k)$ can be encoded by a uniform computationally-private perfectly-correct encoding $\{\hat{g}_k(x;r)\}$ with the following properties: (1) The simulator and decoder use a black-box access to a one-time symmetric encryption (equivalently, to a one-way function); (2) For every fixed randomness r, the resulting function $\hat{g}_{k,r}(x) = \hat{g}_k(x;r)$ is a single-bit operation function of output length $a(k)^{1+\varepsilon}$, where $\varepsilon > 0$ is an arbitrary small constant. (3) Furthermore, the mapping from the circuit of g_k to the circuit of $\hat{g}_{k,r}$ is efficiently computable given a black-box access to the one-time symmetric encryption scheme.

Let $\{F_k\}$ be the universal (and uniform) circuit family for the mapping $(x,z) \mapsto y$ where $x \in (\{0,1\}^k)^t$, the string z is a description of a circuit $C_z : (\{0,1\}^k)^t \to \{0,1\}^{\ell(k)}$ of size $p(k) + p(\ell(k))$, and the string $y \in \{0,1\}^{\ell(k)}$ is $C_z(x)$. By applying the encoding from [5] to $\{F_k\}$ it follows that $\mathcal{C}^t_{k,\ell,p}$ is encoded by $\mathsf{Proj}^t_{k,q}$ where q is polynomial in the circuit size of F_k. The first part of the proposition now follows from Thm 4.1.

The second part follows similarly, except that now we consider the (non-regular) function $\{G_{k,\ell}\}$ which computes the same mapping of F_k but for circuits C_z whose output length ℓ is given as an additional index, and not as a fixed polynomial in k. Again, by applying the encoding from [5] to $\{G_k\}$ it follows that $\mathcal{C}^t_{k,p}$ is encoded by Proj^t_{k}. Hence, the second part of the proposition follows from Thm 4.1.[8] □

In the case of CPA KDM security, one can actually derive KDM-security with respect to projections of arbitrary output length (i.e., Proj^t_k) from single-output projections $\mathsf{Proj}^t_{k,1}$.

Claim 4.7 (Completeness of single-output projections for CPA-KDM)**.** *For every polynomial $t(\cdot)$, we have $\mathsf{Proj}^t_k \leq_{\mathsf{KDM}} \mathsf{Proj}^t_{k,1}$, where the reduction holds for both $(\mathsf{sym},\mathsf{CPA})$ and $(\mathsf{pub},\mathsf{CPA})$ types.*

Proof. The proof follows by simple concatenation: the new encryption/decryption algorithms encrypts/decrypts the message/ciphertext by applying the original encryption/decryption algorithm in a bit by bit manner. Hence, a KDM encryption query in Proj^t_{k,k^a} for the new scheme can be emulated by k^a KDM encryption queries in $\mathsf{Proj}^t_{k,1}$ for the original scheme. □

[8] Recall that for non-regular functions the privacy of the encoding is measured as a function of the input length $kt = \mathrm{poly}(k)$, and so Thm 4.1 holds in this setting as well.

As shown in [8], we can use the standard encrypt-then-MAC transformation to upgrade the security of a scheme that satisfies $(\mathsf{sym}, \mathsf{CPA})$-KDM security into a scheme that satisfies $(\mathsf{sym}, \mathsf{CCA})$-security with respect to the same KDM class. A similar result was proven for the public-key setting by [16] via the Naor-Yung double-encryption paradigm (which relies on the existence of NIZK). Hence, by Proposition 4.6 and Claim 4.7, we have:

Corollary 4.8 (KDM Collapse)**.** *For every polynomials t and p, there exists a $\mathsf{Proj}^t_{k,1}$-KDM secure scheme if and only if there exists p-length dependent KDM secure encryption scheme. This holds unconditionally for every KDM type $\{(\mathsf{sym}, \mathsf{CPA}), (\mathsf{sym}, \mathsf{CCA}), (\mathsf{pub}, \mathsf{CPA})\}$, and for $(\mathsf{pub}, \mathsf{CCA})$ assuming the existence of non-interactive zero-knowledge proof system for* **NP**.

We remark that all the known constructions of affine-KDM secure encryption schemes [12, 4, 13] can be adopted to yield KDM security with respect to single-output projections, (as shown in Appendix A) and so, we get p-length dependent $(\mathsf{pub}, \mathsf{CPA})\}$-KDM (resp., $(\mathsf{sym}, \mathsf{CCA})\}$) based on the DDH, LWE, or QR assumptions (resp., LPN assumption), which can be boosted into $(\mathsf{pub}, \mathsf{CCA})\}$-KDM assuming the existence of NIZK for **NP**.

5 On Full KDM Security

In this section, we study the possibility of constructing a scheme which satisfies KDM security for the class of all functions. In [9] it was shown that such a scheme can be constructed based on the existence of cyclic-secure fully homomorphic encryption (FHH) [20]. We show that a similar assumption is inherently required for full KDM security which is also *simulatable*.

A public-key encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is simulatable fully-KDM secure (with arity $n = 1$) if there exists a polynomial-time simulator S such that for every circuit f of size $\mathrm{poly}(k)$, the ensemble $(\mathsf{sk}, S(\mathsf{pk}, f))$ is indistinguishable from $(\mathsf{sk}, \mathsf{Enc}_{\mathsf{pk}}(f(\mathsf{sk})))$, where the ensembles are indexed by f and $(\mathsf{sk}, \mathsf{pk}) \stackrel{R}{\leftarrow} \mathsf{Gen}(1^k)$.

An FHH allows to translate encryptions of a message M into an encryption of a related message $h(M)$ for any polynomial-size circuit h. More formally, we say that a public-key encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is *fully homomorphic* if there exists an efficient algorithm Eval such that for every circuit h of size $\mathrm{poly}(k)$ and message $M \in \{0,1\}^{\mathrm{poly}(k)}$, the ensemble $(\mathsf{sk}, \mathsf{Eval}(\mathsf{pk}, h, \mathsf{Enc}_{\mathsf{pk}}(M)))$ is computationally indistinguishable from the ensemble $(\mathsf{sk}, \mathsf{Enc}_{\mathsf{pk}}(h(M)))$, where the ensembles are indexed by the security parameter k, the function h and the message M, and where $(\mathsf{sk}, \mathsf{pk}) \stackrel{R}{\leftarrow} \mathsf{Gen}(1^k)$.

In [9], it was shown that any circular-secure fully-homomorphic simulatable encryption scheme is simulatable fully-KDM secure. We show that the other direction holds as well, and so the two notions are equivalent.

Proposition 5.1. *Any simulatable fully-KDM secure encryption scheme of type $(\mathsf{pub}, \mathsf{CPA})$ is also fully-homomorphic circular-secure.*

Proof. Given a simulatable fully-KDM secure encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with simulator S, we define $\mathsf{Eval}(\mathsf{pk}, h, C)$ by invoking S on the pair $(\mathsf{pk}, f_{h,C})$ where $f_{h,C}$ is the mapping $\mathsf{sk} \mapsto h(\mathsf{Dec}_{\mathsf{sk}}(C))$. Note that the circuit size of $f_{h,C}$ is polynomial in the circuit size of h (since Dec is

efficient). Also, by definition, we have for every M and h,

$$\begin{aligned}(\mathsf{sk}, \mathsf{Eval}(\mathsf{pk}, h, \mathsf{Enc}_{\mathsf{pk}}(M))) &\equiv (\mathsf{sk}, S(\mathsf{pk}, f_{h,\mathsf{Enc}_{\mathsf{pk}}(M)})) \\ &\stackrel{c}{\equiv} (\mathsf{sk}, \mathsf{Enc}_{\mathsf{pk}}(h(\mathsf{Dec}_{\mathsf{sk}}(\mathsf{Enc}_{\mathsf{pk}}(M))))) \\ &\equiv (\mathsf{sk}, \mathsf{Enc}_{\mathsf{pk}}(h(M))),\end{aligned}$$

and the proposition follows. □

Next, we show that if one removes the simultability requirement then any encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ which provides KDM security with respect to a function which is slightly stronger than its decryption algorithm Dec, is actually fully-KDM secure. This is done by observing that Gentry's "bootstrapping technique" can be adopted to the KDM setting.

Proposition 5.2. *Suppose that* $\mathsf{PKC} = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is type* $T \in \{(\mathsf{pub}, \mathsf{CPA}), (\mathsf{sym}, \mathsf{CPA})\}$ *KDM secure with respect to single-output projections and in addition with respect to the function family* $\mathcal{F}_k = \{f_{C_1,C_2} : \mathsf{sk} \mapsto \mathrm{NAND}(\mathsf{Dec}_{\mathsf{sk}}(C_1), \mathsf{Dec}_{\mathsf{sk}}(C_2))\}_{C_1,C_2 \in \{0,1\}^{p(k)}}$, *where* $p(k)$ *is the length of an encryption of one-bit message under secret-key of length* k. *Then,* PKC *is fully KDM secure of the same type* T.

Sketch. Since we restrict our attention to the CPA setting, it suffices to prove full KDM security with respect to all circuits of single output. We show how to convert an attacker which sends arbitrary KDM queries into one which uses only queries from $\mathcal{F}_k$. Let h be a circuit of size t, which is wlog composed of NAND gates, and let h_i denote the function computed by the i-th gate of h, where gates are ordered under some topological ordering. We translate a KDM query for h into t KDM calls to $\mathcal{F}_k$ by traversing the circuit from bottom to top in a gate by gate manner. At the i-th query we will have a ciphertext C_i such that, if the oracle is in the real mode $C_i = \mathsf{Enc}_{\mathsf{pk}}(h_i(\mathsf{sk}))$ and if it is in the fake mode $C_i = \mathsf{Enc}_{\mathsf{pk}}(0)$. This can be achieve directly for the input gates by making a single KDM query with a single-bit wise operation. To do this, for an internal gate h_ℓ whose input wires are connected to h_i and h_j for some $i, j < \ell$, we use a KDM query to f_{C_i,C_j}. It is not hard to see that the invariant holds, and therefore the claim follows. □

References

[1] Abadi and Rogaway. Reconciling two views of cryptography (the computational soundness of formal encryption). *JCRYPTOL: Journal of Cryptology*, 15, 2002.

[2] T. Acar, M. Belenkiy, M. Bellare, and D. Cash. Cryptographic agility and its relation to circular encryption. In *Proc. of EUROCRYPT 2010*, pages 403–422, 2010.

[3] P. Adão, G. Bana, J. Herzog, and A. Scedrov. Soundness and completeness of formal encryption: The cases of key cycles and partial information leakage. *Journal of Computer Security*, 17(5):737–797, 2009.

[4] B. Applebaum, D. Cash, C. Peikert, and A. Sahai. Fast cryptographic primitives and circular-secure encryption based on hard learning problems. In *Advances in Cryptology - CRYPTO 2009*, 2009.

[5] B. Applebaum, Y. Ishai, and E. Kushilevitz. Computationally private randomizing polynomials and their applications. *Computional Complexity*, 15(2):115–162, 2006. Preliminary version in Proc. 20th CCC, 2005.

[6] B. Applebaum, Y. Ishai, and E. Kushilevitz. Cryptography in NC^0. *SIAM J. Comput.*, 36(4):845–888, 2006. Preliminary version in Proc. 45th FOCS, 2004.

[7] M. Backes, M. Dürmuth, and D. Unruh. Oaep is secure under key dependent messages. In *ASIACRYPT '08*, 2008. To appear.

[8] M. Backes, B. Pfitzmann, and A. Scedrov. Key-dependent message security under active attacks - brsim/uc-soundness of symbolic encryption with key cycles. In *Proceedings of 20th IEEE Computer Security Foundation Symposium (CSF)*, June 2007. Preprint on IACR ePrint 2005/421.

[9] B. Barak, I. Haitner, D. Hofheinz, and Y. Ishai. Bounded key-dependent message security. In *Advances in Cryptology - EUROCRYPT 2010*, pages 423–444, 2010.

[10] A. Beimel and A. Gál. On arithmetic branching programs. *J. Comput. Syst. Sci*, 59(2):195–220, 1999.

[11] J. Black, P. Rogaway, and T. Shrimpton. Encryption-scheme security in the presence of key-dependent messages. In *SAC '02*, pages 62–75, 2002.

[12] D. Boneh, S. Halevi, M. Hamburg, and R. Ostrovsky. Circular-secure encryption from decision diffie-hellman. In *CRYPTO '08*, pages 108–125, 2008.

[13] Z. Brakerski and S. Goldwasser. Circular and leakage resilient public-key encryption under subgroup indistinguishability (or: Quadratic residuosity strikes back). In *Advances in Cryptology - CRYPT0 2010*, 2010.

[14] Z. Brakerski, S. Goldwasser, and Y. Kalai. Circular-secure encryption beyond affine functions. Cryptology ePrint Archive, Report 2009/485, 2009.

[15] Camenisch and Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In *EUROCRYPT '01*, 2001.

[16] J. Camenisch, N. Chandran, and V. Shoup. A public key encryption scheme secure against key dependent chosen plaintext and adaptive chosen ciphertext attacks. In *Proc. of EUROCRYPT 2009*, pages 351–368, 2009.

[17] D. Dolev, C. Dwork, and M. Naor. Non-malleable cryptography (extended abstract). In *STOC*, 1991.

[18] S. Even, O. Goldreich, and A. Lempel. A randomized protocol for signing contracts. *CACM: Communications of the ACM*, 28, 1985.

[19] Freedman, Ishai, Pinkas, and Reingold. Keyword search and oblivious pseudorandom functions. In *Theory of Cryptography Conference (TCC), LNCS*, volume 2, 2005.

[20] C. Gentry. Fully homomorphic encryption using ideal lattices. In *Proc. 41st STOC*, pages 169–178, 2009.

[21] Gertner, Kannan, Malkin, Reingold, and Viswanathan. The relationship between public key encryption and oblivious transfer. In *Proc. of 41st FOCS*, 2000.

[22] O. Goldreich. *Foundations of Cryptography: Basic Tools*. Cambridge University Press, 2001.

[23] O. Goldreich. *Foundations of Cryptography: Basic Applications*. Cambridge University Press, 2004.

[24] S. Goldwasser and S. Micali. Probabilistic encryption. *JCSS*, 28(2):270–299, 1984. Preliminary version in Proc. STOC, 1982.

[25] I. Haitner and T. Holenstein. On the (im)possibility of key dependent encryption. Cryptology ePrint Archive, Report 2008/164, 2008.

[26] S. Halevi and H. Krawczyk. Security under key-dependent inputs. In *CCS '07*, pages 466–475, 2007.

[27] D. Hofheinz and D. Unruh. Towards key-dependent message security in the standard model. In *EUROCRYPT '08*, pages 108–126, 2008.

[28] Impagliazzo and Rudich. Limits on the provable consequences of one-way permutations. In *Advances in Cryptology: Proc. of CRYPTO '88*, 1988.

[29] Y. Ishai and E. Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *Proc. 41st FOCS*, pages 294–304, 2000.

[30] Y. Ishai and E. Kushilevitz. Perfect constant-round secure computation via perfect randomizing polynomials. In *Proc. 29th ICALP*, pages 244–256, 2002.

[31] M. Naor and M. Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *Proc. of the 22nd STOC*, pages 427–437, 1990.

[32] M. Rabin. Digitalized signatures and public key functions as intractable as factoring. Technical Report 212, LCS, MIT, 1979.

[33] Rackoff and Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In *Advances in Cryptology - CRYPTO 1991*, 1991.

[34] M. van Dijk, C. Gentry, S. Halevi, and V. Vaikuntanathan. Fully homomorphic encryption over the integers. In *Proc. of EUROCRYPT 2010*, pages 24–43, 2010.

[35] A. C. Yao. How to generate and exchange secrets. In *Proc. 27th FOCS*, pages 162–167, 1986.

A From affine functions to projections

Converting affine-security to security under single-output projections is immediate if the affine functions are defined over the binary field $\mathbb{F}_2$ (as in the LPN based scheme of [16]), but can also be established in more general cases, which capture all known schemes, as follows.

Suppose that we have KDM security for affine functions over a ring $\mathcal{R}$ (which may, in general, be a ring family whose size depends on the security parameter). Namely, the scheme encrypts ring elements, the secret-key consists of n ring elements $\mathsf{sk} = (\mathsf{sk}_i)_{i\in[n]}$, and KDM security holds with respect to affine functions: $f_{a,b} : \mathsf{sk} \mapsto (\sum_i a_i \cdot \mathsf{sk}_i) + b$. Let $\langle \mathsf{sk} \rangle = (\langle \mathsf{sk} \rangle_1, \ldots, \langle \mathsf{sk} \rangle_k)$ denote the representation of the secret key sk as a k-bit string. We will show that affine functions can encode projections and thus, by Thm. 4.1, get a new scheme with bit-operation KDM security. This is possible as long as the bit representation is "meaningful" in terms of the group $\mathcal{R}$. Formally, we distinguish between the following two cases.

In the first case, each key element sk_i is represented by a single bit $\langle \mathsf{sk} \rangle_i$. That is, there exists a list of non-zero public elements $g = (g_1, \ldots, g_\ell)$ such that $\mathsf{sk}_i = g_i \cdot \langle \mathsf{sk} \rangle_i$ (where, for the sake of ring arithmetics, we think of a bit β as either the zero element or the one element of the ring). This is the case, for example, in the schemes of [12, 13]. Let us assume that $\langle \mathsf{sk} \rangle$ is used as the bit-representation of the key. Then we can use the RE approach to amplify affine security (over $\mathcal{R}$) into security against projections by showing that the former encodes the latter. Formally, every single-output bit operation function $f_{i,\sigma}(\langle \mathsf{sk} \rangle) = \langle \mathsf{sk} \rangle_i \oplus \sigma$ is encoded by $\hat{f}_{i,\sigma}(\langle \mathsf{sk} \rangle; r) = (\langle \mathsf{sk} \rangle_i - \sigma) \cdot g_i \cdot r$ where r is a randomly chosen non-zero element of $\mathcal{R}$. This encoding enjoys perfect correctness via a universal decoder (a zero element is decoded to 0 and any other element is decoded to 1) and perfect privacy via a universal simulator (given an output β of f simulate the corresponding output of $\hat{f}$ by multiplying it with a random non-zero element). Moreover, when the randomness is fixed we get a linear function over $\mathcal{R}$. Hence, by Thm. 4.1, the security of the scheme can be amplified to hold with respect to single-output projections.

We proceed with the second case. Let us assume that the mapping from sk to each bit of the representation $\langle \mathsf{sk} \rangle$ can be computed by a polynomial-size arithmetic branching program (ABP) (see [10, 19]) over $\mathcal{R}$. (This is possible in a trivial way whenever the ring is of polynomial size, as in the LWE-based scheme of [4].) Then, the mappings $f_{i,0} : \mathsf{sk} \mapsto \langle \mathsf{sk} \rangle_i$ and $f_{i,1} : \mathsf{sk} \mapsto 1 - \langle \mathsf{sk} \rangle_i$ can also be computed by a polynomial-size ABP. Hence, by [19], there exists a perfect (universal) RE $\hat{f}_{i,\sigma}(\mathsf{sk}; r)$ such that for every fixed choice of r, $\hat{f}_{r,i,\sigma}(\mathsf{sk}) = \hat{f}_{i,\sigma}(\mathsf{sk}; r)$ is an affine function over $\mathcal{R}$. Hence, by Thm. 4.1, the security of the scheme can be amplified to hold with respect to single-output projections.