Secure Hybrid Encryption from Weakened Key Encapsulation

Dennis Hofheinz and Eike Kiltz*

Cryptology and Information Security Research Theme
CWI Amsterdam
The Netherlands
{hofheinz,kiltz}@cwi.nl

Abstract. We put forward a new paradigm for building hybrid encryption schemes from *constrained chosen-ciphertext secure* (CCCA) key-encapsulation mechanisms (KEMs) plus authenticated symmetric encryption. Constrained chosen-ciphertext security is a new security notion for KEMs that we propose. It has less demanding security requirements than standard CCCA security (since it requires the adversary to have a certain plaintext-knowledge when making a decapsulation query) yet we can prove that it is CCCA sufficient for secure hybrid encryption.
Our notion is not only useful to express the Kurosawa-Desmedt public-key encryption scheme and its generalizations to hash-proof systems in an abstract KEM/DEM security framework. It also has a very constructive appeal, which we demonstrate with a new encryption scheme whose security relies on a class of intractability assumptions that we show (in the generic group model) strictly weaker than the Decision Diffie-Hellman (DDH) assumption. This appears to be the first practical public-key encryption scheme in the literature from an algebraic assumption strictly weaker than DDH.

1 Introduction

One of the main fields of interest in cryptography is the design and analysis of encryption schemes in the public-key setting (PKE schemes) that are secure against a very strong type of attacks — indistinguishability against chosen-ciphertext attacks (IND-CCA) [24]. In this work, we are interested in *practical schemes* with proofs of security under *reasonable security assumptions* (without relying on heuristics such as the random oracle model) and in *general methods* for constructing such schemes.

The first practical IND-CCA secure PKE scheme without random oracles was proposed in a seminal paper by Cramer and Shoup [11, 13]. Their construction was later generalized to hash proof systems [12]. In [30, 13] Cramer and Shoup also give a hybrid variant that encrypts messages of arbitrary length. The idea

* Supported by the research program Sentinels (http://www.sentinels.nl). Sentinels is being financed by Technology Foundation STW, the Netherlands Organization for Scientific Research (NWO), and the Dutch Ministry of Economic Affairs.

is to conceptually separate the key-encapsulation (KEM) part from the symmetric (DEM) part. Generally, this hybrid approach greatly improved practicality of encryption schemes. A folklore composition theorem (formalized in [13]) shows that if both KEM and DEM are CCA-secure then the hybrid encryption is CCA-secure. Common wisdom was that this sufficient condition was also necessary. However, at CRYPTO 2004, Kurosawa and Desmedt challenged this common wisdom by presenting a hybrid encryption scheme that demonstrates that a weaker security condition on the KEM may suffice for full CCA-secure hybrid encryption. Compared to the original Cramer-Shoup scheme, the scheme by Kurosawa and Desmedt improved efficiency and ciphertext expansion by replacing some of its algebraic components with *information theoretically* secure symmetric primitives. More recently, the KEM part of their scheme was indeed shown to be not CCA secure [15].

One natural open problem from [21] is if there exists a weaker yet natural security condition on the KEM such that, in combination with sufficiently strong symmetric encryption, chosen-ciphertext secure hybrid encryption can be guaranteed.

Extending the work of Cramer and Shoup [12], it was demonstrated in [21, 2, 14] that a variant of hash-proof systems (HPS) can be combined with symmetric encryption and a message authentication code (MAC) to obtain hybrid encryption. If the hash-proof system is $universal_2$, then the encryption scheme is chosen-ciphertext secure. However, the Kurosawa-Desmedt hybrid scheme could not be rigorously explained in this general HPS framework since the underlying hash-proof system is not universal_2. (Roughly, this is since universal_2 is a *statistical* property whereas the Kurosawa-Desmedt system contains a *computational* component, namely a target collision resistant (TCR) hash function.) In [21] (and [12]) only less efficient "hash-free variants" of their schemes could be explained through hash proof systems; security of all efficient TCR-based schemes had to be proved separately.

Surprisingly, almost all *practical* standard-model encryption schemes [11, 13, 21, 2, 10, 9, 19, 20] are based on the difficulty of Decision Diffie-Hellman (DDH) or stronger assumptions. This is contrasted by the existence of many natural groups in which the DDH assumption is known to be wrong; examples include pairing-groups and certain non prime-order groups like $\mathbb{Z}_p^*$. This often overlooked fact may turn into a serious problem in case DDH turns out to be wrong in all cryptographically interesting groups. In particular, [16] give evidence that groups with easy DDH problem, but hard computational Diffie-Hellman problem exist. [16] interpret this as an argument to rely on weaker assumptions than DDH.

1.1 Our contributions

A NEW KEM/DEM COMPOSITION THEOREM. We put forward the security notion of *indistinguishability against constrained chosen-ciphertext attacks* (IND-CCCA) for KEMs which is stronger than IND-CPA (CPA stands for chosen-plaintext attacks) yet strictly weaker than IND-CCA. Intuitively, CCCA is separated from CCA security by only allowing an adversary to make a decapsulation

query if it has sufficient "implicit knowledge" about the plaintext key to be decapsulated (hence the name "constrained chosen-ciphertext security").[1]

As our main technical contribution we formalize the above notion and prove a composition theorem that shows that *any* IND-CCCA secure KEM combined with *any* authenticated (symmetric) encryption scheme yields IND-CCA secure hybrid encryption. This gives a positive answer to the open question from [21] mentioned before. Authenticated encryption is a quite general symmetric primitive and examples include "encrypt-then-mac" schemes (based on computationally secure primitives), and also more efficient single-pass schemes (see, e.g., [25]).

Constrained chosen-ciphertext secure KEMs formalize a new design paradigm for efficient hybrid encryption. To guarantee chosen-ciphertext security for hybrid encryption schemes it is sufficient to verify a natural security condition on the key encapsulation part. We assess the constructive appeal of this framework by demonstrating that the original Kurosawa-Desmedt scheme [21], along with its variants [2, 23] and all hash-proof systems based schemes [12, 21], can be thoroughly explained through it. We furthermore present a new IND-CCCA secure KEM from the DDH assumption and show how to build a class of practical KEMs from progressively weaker assumptions than DDH.

Constrained chosen-ciphertext secure KEM from DDH. We propose a new KEM which is IND-CCCA secure under the DDH assumption. Although it relies on different proof techniques (it is not based on hash proof systems), syntactically it is reminiscent to the one by Kurosawa and Desmedt and can in fact be viewed as its *dual* (in the sense that certain parts from the ciphertext and the symmetric key are swapped in our scheme).

Constrained chosen-ciphertext secure KEM from n-Linear. Building on [8, 18] we introduce a new class of purely algebraic intractability assumptions, the n-Linear assumptions, where $n \geq 1$ is a parameter. They are such that the DDH assumption equals the 1-Linear assumption, the Linear assumption [8] equals the 2-Linear assumption, and the n-Linear assumptions become *strictly weaker* as the parameter n grows. More precisely, 1-Linear = DDH, and n-Linear implies $n+1$-Linear, but (in the generic group model [29]) $n+1$-Linear is still hard relative to an n-Linear oracle. In fact, for $n \geq 2$ the n-Linear assumption does not seem to be invalid in any obvious sense even in the groups from [16], in which the DDH problem is easy, and the computational Diffie-Hellman problem is supposedly hard. We generalize the KD scheme and its dual to a class of parametrized KEMs and prove their IND-CCCA security assuming n-Linear. These appear to be the first practical encryption schemes in the literature from a purely algebraic assumption which is strictly weaker than DDH.

Computational Hash-Proof Systems. We propose a purely computational variant of hash-proof systems. Generalizing [12, 21], we prove that computa-

[1] This is reminiscent to the notion of "plaintext awareness" for public-key encryption [5] where it is infeasible for an adversary to come up with a valid ciphertext without being aware of the corresponding plaintext. Our definition is weaker in the sense that it only requires the adversary to have *implicit knowledge* on the plaintext.

tional hash-proof systems directly imply IND-CCCA secure KEMs. Hence, in combination with authenticated encryption, they yield efficient IND-CCA secure hybrid encryption. The Kurosawa-Desmedt scheme fits this framework, i.e. the underlying HPS is computational. This gives the first full explanation of the Kurosawa-Desmedt scheme in terms of HPS. As a generalization we provide computational hash-proof systems from the n-Linear assumptions hence explaining IND-CCCA security of our class of KEMs from the n-Linear assumptions.

1.2 Discussion and related work

In [1] (which is the full version of [2]), Abe et al. address the question from [21] about the existence of a natural weaker security condition for KEMs. They propose the notion of *LCCA secure KEMs with respect to the predicate* $\mathcal{P}^{\mathrm{mac}}$ and prove it sufficient to obtain, in combination with a MAC, IND-CCA secure tag-KEMs (and hence IND-CCA secure hybrid encryption). Though syntactically similar to ours, their notion *mingles* security of the KEM with the MAC part of the symmetric encryption scheme. The conceptual difference in our notion is that we give a general security definition for KEMs that is *completely independent* of any particular symmetric primitive. We think that this is more natural and more closely follows the spirit of the KEM/DEM approach [13], where (for good reason) KEM and DEM are viewed as independent components.

Independent from this work Shacham [28] also proposes a family of hybrid encryption schemes from the n-Linear assumptions. His schemes can be viewed as a (slightly less efficient) Cramer-Shoup variant of our schemes from Section 4.2.

The 2-Linear assumption was introduced by Boneh, Boyen, and Shacham [8] and was later used in gap-groups to build an IND-CCA secure KEM [19]. For $n > 2$, Kiltz [18] introduced the class of *gap n-Linear* assumptions and (generalizing [19]) built a class of IND-CCA secure KEMs from it. Compared to n-Linear, in the latter gap-assumptions an adversary gets access to a DDH oracle which makes (for example) the gap 2-Linear assumption incomparable to DDH. In contrast, our motivation is to build schemes from an assumption weaker than DDH.

2 Hybrid encryption from constrained CCA secure KEMs

2.1 Key Encapsulation Mechanisms

A *key-encapsulation mechanism* $\mathcal{KEM} = (\mathsf{KEM.Kg}, \mathsf{KEM.Enc}, \mathsf{KEM.Dec})$ with key-space $\mathcal{K}(k)$ consists of three polynomial-time algorithms (PTAs). Via $(pk, sk) \xleftarrow{\$} \mathsf{KEM.Kg}(1^k)$ the randomized key-generation algorithm produces public/secret keys for security parameter $k \in \mathbb{N}$; via $(K, C) \xleftarrow{\$} \mathsf{KEM.Enc}(pk)$ the randomized encapsulation algorithm creates an uniformly distributed symmetric key $K \in \mathcal{K}(k)$ together with a ciphertext C; via $K \leftarrow \mathsf{KEM.Dec}(sk, C)$ the possessor of secret key sk decrypts ciphertext C to get back a key K which is an element in $\mathcal{K}$ or a special rejection symbol $\perp$. For consistency, we require that for all

$k \in \mathbb{N}$, and all $(K, C) \xleftarrow{\$} \mathsf{KEM.Enc}(pk)$ we have $\Pr[\,\mathsf{KEM.Dec}(sk, C) = K\,] = 1$, where the probability is taken over the choice of $(pk, sk) \xleftarrow{\$} \mathsf{KEM.Kg}(1^k)$, and the coins of all the algorithms in the expression above. Here we only consider only KEMs that produce perfectly uniformly distributed keys (i.e., we require that for all public keys pk that can be output by $\mathsf{KEM.Kg}$, the first component of $\mathsf{KEM.Enc}(pk)$ has uniform distribution).[2]

Constrained Chosen-Ciphertext Security. The common requirement for a KEM is indistinguishability against chosen-ciphertext attacks (IND-CCA) [13] where an adversary is allowed to adaptively query a decapsulation oracle with ciphertexts to obtain the corresponding session key. We relax this notion to indistinguishability against *constrained chosen-ciphertext attacks* (IND-CCCA). Intuitively, we only allow the adversary to make a decapsulation query if it already has some "a priori knowledge" about the decapsulated key. This partial knowledge about the key is modeled implicitly by letting the adversary additionally provide an efficiently computable Boolean predicate $\mathrm{pred} : \mathcal{K} \to \{0,1\}$. If $\mathrm{pred}(K) = 1$ then the decapsulated key K is returned, and $\perp$ otherwise. The amount of uncertainty the adversary has about the session key (denoted as *plaintext uncertainty* $\mathrm{uncert}_{\mathcal{A}}$) is measured by the fraction of keys the predicate evaluates to 1. We require this fraction to be negligible for every query, i.e. the adversary has to have a high a priori knowledge about the decapsulated key when making a decapsulation query. More formally, for an adversary $\mathcal{A}$ we define the advantage function

$$\mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},\mathcal{A}}(k) \;=\; \left|\Pr[\mathbf{Exp}^{\text{ccca-1}}_{\mathcal{KEM},\mathcal{A}}(k) = 1] - \Pr[\mathbf{Exp}^{\text{ccca-0}}_{\mathcal{KEM},\mathcal{A}}(k) = 1]\right|$$

where, for $b \in \{0,1\}$, $\mathbf{Exp}^{\text{ccca-b}}_{\mathcal{KEM},\mathcal{A}}$ is defined by the following experiment.

Experiment $\mathbf{Exp}^{\text{ccca-b}}_{\mathcal{KEM},\mathcal{A}}(k)$
$(pk, sk) \xleftarrow{\$} \mathsf{KEM.Kg}(1^k)$
$K_0^* \xleftarrow{\$} \mathcal{K}(k)\,;\; (K_1^*, C^*) \xleftarrow{\$} \mathsf{KEM.Enc}(pk)$
$b' \xleftarrow{\$} \mathcal{A}^{\textsc{Dec}(\cdot,\cdot)}(pk, K_b^*, C^*)$
Return b'

$\textsc{Dec}(\mathrm{pred}_i, C_i)$
$K \leftarrow \mathsf{KEM.Dec}(sk, C_i)$
If $K = \perp$ or $\mathrm{pred}_i(K) = 0$ then $\perp$
Else return $K \in \mathcal{K}$

with the restriction that $\mathcal{A}$ is only allowed to query $\textsc{Dec}(\mathrm{pred}_i, C_i)$ on predicates pred_i that are provided as PTA[3] and on ciphertexts C_i different from the challenge ciphertext C^*.

For an adversary $\mathcal{A}$, let $t_{\mathcal{A}}$ denote the number of computational steps $\mathcal{A}$ runs (that includes the maximal time to *evaluate* each pred_i once), and let $Q_{\mathcal{A}}$ be the number of decapsulation queries $\mathcal{A}$ makes to its decapsulation oracle.

[2] This requirement is met by all popular KEMs and makes our reduction in Theorem 1 tighter. However, we can show Theorem 1 also without this assumption, and derive that the keys are computationally close to uniform from our upcoming KEM security assumption. This comes at the price of a less tight security reduction in Theorem 1.

[3] Technically, we charge the time required to evaluate each pred_i to $\mathcal{A}$'s runtime and require that $\mathcal{A}$ be polynomial-time.

For simplicity and without losing on generality, we consider only adversaries for which $t_{\mathcal{A}}$ and $Q_{\mathcal{A}}$ are independent of the environment that $\mathcal{A}$ runs in. To adversary $\mathcal{A}$ in the above experiment we also associate $\mathcal{A}$'s (implicit) plaintext uncertainty $\mathrm{uncert}_{\mathcal{A}}(k)$ when making decapsulation queries, measured by

$$\mathrm{uncert}_{\mathcal{A}}(k) \;=\; \frac{1}{Q} \sum_{1 \le i \le Q} \Pr_{K \in \mathcal{K}}[\mathrm{pred}_i(K) = 1] \;,$$

where $\mathrm{pred}_i : \mathbb{G} \to \{0,1\}$ is the predicate $\mathcal{A}$ submits in the ith decapsulation query. Let, for integers k, t, Q and $0 \le \mu \le 1$,

$$\mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},t,Q,\mu}(k) \;=\; \max_{\mathcal{A}} \mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},\mathcal{A}}(k),$$

where the maximum is over all $\mathcal{A}$ with $t_{\mathcal{A}} \le t$, $Q_{\mathcal{A}} \le Q$, and $\mathrm{uncert}_{\mathcal{A}}(k) \le \mu$.

A key encapsulation mechanism $\mathcal{KEM}$ is said to be *indistinguishable against constrained chosen ciphertext attacks* (IND-CCCA) if for all PTA adversaries $\mathcal{A}$ with negligible $\mathrm{uncert}_{\mathcal{A}}(k)$ (in any environment), the advantage $\mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},\mathcal{A}}(k)$ is a negligible function in k.

It is worth pointing out that by making different restrictions on $\mathrm{uncert}(k)$ our notion of CCCA security leads to an interesting continuum between CPA and CCA security. With the restriction $\mathrm{uncert}(k) = 0$ then CCCA = CPA; with the trivial restriction $\mathrm{uncert}(k) \le 1$ (which makes is possible to always use the constant predicate $\mathrm{pred}(\cdot) := 1$) then CCCA = CCA. Here, we require a negligible $\mathrm{uncert}(k)$, which syntactically makes IND-CCCA more similar to IND-CPA than to IND-CCA security. Yet, since it in principle allows decryption queries, IND-CCCA is substantially stronger than IND-CPA, and — as we will show — is a good base for hybrid IND-CCA security.

2.2 Authenticated Encryption

An authenticated symmetric encryption (AE) scheme $\mathcal{AE} = (\mathsf{AE.Enc}, \mathsf{AE.Dec})$ is specified by its encryption algorithm $\mathsf{AE.Enc}$ (encrypting $M \in \mathit{MsgSp}(k)$ with keys $K \in \mathcal{K}(k)$) and decryption algorithm $\mathsf{AE.Dec}$ (returning $M \in \mathit{MsgSp}(k)$ or $\perp$). Here we restrict ourselves to deterministic PTAs $\mathsf{AE.Enc}$ and $\mathsf{AE.Dec}$. The AE scheme needs to provide privacy (indistinguishability against one-time attacks) and authenticity (ciphertext authenticity against one-time attacks). This is *simultaneously* captured (similar to the more-time attack case [26]) by defining the ae-ot-advantage of an adversary $\mathcal{B}_{ae}$ as $\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k) =$

$$2 \left| \Pr[K \xleftarrow{\$} \mathcal{K}(k) \,;\, b \xleftarrow{\$} \{0,1\} \,;\, b' \xleftarrow{\$} \mathcal{B}_{ae}^{\mathrm{LoR}_b(\cdot,\cdot),\mathrm{DoR}_b(\cdot)}(1^k) \;:\; b = b'] - 1 \right| \;.$$

Here, $\mathrm{LoR}_b(M_0, M_1)$ returns $\psi \leftarrow \mathsf{AE.Enc}(K, M_b)$, and $\mathcal{B}_{ae}$ is allowed only one query to this left-or-right encryption oracle (one-time attack), with a pair of equal-length messages. Furthermore, the decrypt-or-reject oracle $\mathrm{DoR}_1(\psi)$ returns $M \leftarrow \mathsf{AE.Dec}(K, \psi)$ and $\mathrm{DoR}_0(\psi)$ always returns $\perp$ (reject), $\mathcal{B}_{ae}$ is allowed

only one query to this decrypt-or-reject oracle which must be different from the output of the left-or-right oracle.

We say that $\mathcal{AE}$ is a *one-time secure authenticated encryption scheme* (AE-OT secure) if the advantage function $\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k)$ is negligible for all PTA $\mathcal{B}_{ae}$. Again, for integers k, t, $\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},t}(k) = \max_{\mathcal{B}_{ae}} \mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k)$, where the maximum is over all $\mathcal{B}_{ae}$ that fulfill $t_{\mathcal{B}_{ae}} \leq t$.

2.3 Hybrid Encryption

Let $\mathcal{KEM} = (\mathsf{KEM.Kg}, \mathsf{KEM.Enc}, \mathsf{KEM.Dec})$ be a KEM and let $\mathcal{AE} = (\mathsf{AE.Enc}, \mathsf{AE.Dec})$ be an authenticated encryption scheme. We assume that the two schemes are compatible in the sense that for all security parameters k, we have that the KEM's and the AE's key-space are equal. Then we can consider a hybrid public key encryption scheme (whose syntax and security definition is standard and can be looked up in the full version) that encrypts arbitrary messages $M \in \mathit{MsgSp}$. The construction of $\mathcal{PKE} = (\mathsf{PKE.kg}, \mathsf{PKE.Enc}, \mathsf{PKE.Dec})$ is as follows.

$\mathsf{PKE.kg}(1^k)$	$\mathsf{PKE.Enc}(pk, M)$	$\mathsf{PKE.Dec}(sk, C_{pke} = (C, \psi))$
$(pk, sk) \xleftarrow{\$} \mathsf{KEM.Kg}(1^k)$	$(K, C) \xleftarrow{\$} \mathsf{KEM.Enc}(pk)$	$K \leftarrow \mathsf{KEM.Dec}(sk, C)$
Return (pk, sk)	$\psi \leftarrow \mathsf{AE.Enc}(K, M)$	$M \leftarrow \mathsf{AE.Dec}(K, \psi)$
	Return $C_{pke} = (C, \psi)$	Return M or $\perp$

Here $\mathsf{PKE.Dec}$ returns $\perp$ if either $\mathsf{KEM.Dec}$ or $\mathsf{AE.Dec}$ returns $\perp$.

Theorem 1. *Assume $\mathcal{KEM}$ is secure in the sense of* IND-CCCA *and $\mathcal{AE}$ is secure in the sense of* AE-OT. *Then $\mathcal{PKE}$ is secure in the sense of* IND-CCA. *In particular,*

$$\mathbf{Adv}^{\mathrm{cca}}_{\mathcal{PKE},t,Q}(k) \leq \mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},t,Q,Q\cdot\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},t}(k)}(k) + (Q+1)\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},t}(k) + \frac{Q}{|\mathcal{K}|} .$$

Proof. Let $\mathcal{A}$ be an adversary on the IND-CCA security of the hybrid scheme. We will consider a sequence of games, Game 1, Game 2, ..., each game involving $\mathcal{A}$. Let X_i be the event that in Game i, it holds that $b = b'$, i.e., that the adversary succeeds. We will make use of the following simple "Difference Lemma" [13].

Lemma 1. *Let X_1,X_2, B be events, and suppose that $X_1 \wedge \neg B \Leftrightarrow X_2 \wedge \neg B$. Then $|\Pr[X_1] - \Pr[X_2]| \leq \Pr[B]$.*

Game 1. The original PKE IND-CCA game, i.e. we have

$$|\Pr[X_1] - 1/2| = \mathbf{Adv}^{\mathrm{cca}}_{\mathcal{PKE},\mathcal{A}}(k) .$$

Game 2. Let $C^*_{pke} = (C^*, \psi^*)$ be the challenge ciphertext in the PKE IND-CCA game. In this game the decryption oracle in the first phase rejects all ciphertexts of the form $C_{pke} = (C^*, *)$. The view of adversary $\mathcal{A}$ is identical in Games 1 and 2 until a decryption query $(C^*, *)$ is made in the first phase of the IND-CCA experiment (so *before* $\mathcal{A}$ gets to see C^*).

Since the key K encapsulated in C^* is uniformly distributed and independent of $\mathcal{A}$'s view in the first phase, we have

$$|\Pr[X_2] - \Pr[X_1]| \leq \frac{Q}{|\mathcal{K}|} .$$

Note that each ciphertext uniquely determines a key.

Game 3. Replace the symmetric key K^* used to create the PKE challenge ciphertext with a random key K^*, uniformly independently chosen from $\mathcal{K}$. The proof of the following lemma is postponed until later.

Lemma 2. $|\Pr[X_3] - \Pr[X_2]| \leq \mathbf{Adv}^{\text{ccca}}_{\mathcal{KEM},t,Q,Q\cdot\mathbf{Adv}^{ae\text{-ot}}_{\mathcal{AE},t}(k)}(k)$.

Game 4. Reject all ciphertexts C_{pke} of the form $(C^*, *)$. Since ψ^* was generated using a random key $K^* \in \mathcal{K}$ that only leaks to $\mathcal{A}$ through ψ^*, authenticity of $\mathcal{AE}$ implies

$$|\Pr[X_4] - \Pr[X_3]| \;\leq\; Q_{\mathcal{A}} \cdot \mathbf{Adv}^{ae\text{-ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k)$$

for a suitable adversary $\mathcal{B}_{ae}$ that simulates Game 3, using the LoR_b with two identical messages to obtain the AE part of the challenge ciphertext. $\mathcal{B}_{ae}$ simply uniformly picks one AE part of a decryption query of the form (C^*, ψ) to submit to the decrypt-or-reject oracle $\text{DoR}_1(\cdot)$.
Finally, Game 4 models one-time security of the AE scheme, and we have

$$|\Pr[X_4] - 1/2| \leq \mathbf{Adv}^{ae\text{-ot}}_{\mathcal{AE},t}(k) .$$

Collecting the probabilities proves the theorem. It leaves to prove Lemma 2.

Proof (Lemma 2). We show that there exists an adversary $\mathcal{B}_{kem}$ against the IND-CCCA security of $\mathcal{KEM}$ with $t_{\mathcal{B}_{kem}} = t_{\mathcal{A}}$, $Q_{\mathcal{B}_{kem}} = Q_{\mathcal{A}}$, and an adversary $\mathcal{B}_{ae}$ against $\mathcal{AE}$ with $t_{\mathcal{B}_{ae}} = t_{\mathcal{A}}$, such that

$$\text{uncert}_{\mathcal{B}_{kem}}(k) \leq Q_{\mathcal{A}} \cdot \mathbf{Adv}^{ae\text{-ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k) \tag{1}$$

$$\Pr[X_2] = \Pr[\mathbf{Exp}^{\text{ccca-1}}_{\mathcal{KEM},\mathcal{B}_{kem}}(k) = 1] \tag{2}$$

$$\Pr[X_3] = \Pr[\mathbf{Exp}^{\text{ccca-0}}_{\mathcal{KEM},\mathcal{B}_{kem}}(k) = 1] . \tag{3}$$

The adversary $\mathcal{B}_{kem}$ against the CCCA security of $\mathcal{KEM}$ is defined as follows. $\mathcal{B}_{kem}$ inputs (pk, K^*_b, C^*) for an unknown bit b. First, $\mathcal{B}_{kem}$ runs $\mathcal{A}_1$ on input pk. For the ith decryption query (C_i, ψ_i) made by adversary $\mathcal{A}$, adversary $\mathcal{B}_{kem}$ defines the function $\text{pred}_i : \mathcal{K} \to \{0, 1\}$ as

$$\text{pred}_i(K) := \begin{cases} 0 & : \quad \text{if } \mathsf{AE.Dec}(K, \psi_i) \text{ returns } \perp \\ 1 & : \quad \text{otherwise} \end{cases}$$

Note that the symmetric ciphertext ψ_i is hard-coded into $\text{pred}_i(\cdot)$ which is clearly efficiently computable. $\mathcal{B}_{kem}$ queries (pred_i, C_i) to its own oracle $\text{DEC}(\cdot,\cdot)$ and receives the following answer. If $\mathsf{KEM.Dec}(sk, C_i)$ returns a key $K_i \in \mathcal{K}$ with

$\mathsf{AE.Dec}(K_i, \psi_i)$ returns $\perp$ then $\textsc{Dec}(\text{pred}_i, C_i)$ returns the key K_i. Otherwise (if $\mathsf{KEM.Dec}(sk, C_i)$ returns $\perp$ or if $\mathsf{AE.Dec}(K_i, \psi_i)$ returns $\perp$), $\textsc{Dec}(\text{pred}_i, C_i)$ returns $\perp$. Note that by the syntax of $\mathcal{AE}$ this perfectly simulates $\mathcal{A}$'s decryption queries.

For $\mathcal{A}$'s encryption challenge for two messages M_0, M_1, $\mathcal{B}_{kem}$ uses its own input (K_b^*, C^*) together with a random bit δ to create a challenge ciphertext $C_{pke}^* = (C^*, \psi^* \leftarrow \mathsf{AE.Enc}(K^*, M_\delta))$ of message M_δ. Adversary $\mathcal{B}_{kem}$ runs $\mathcal{A}_2(C_{pke}^*, St_1)$ and inputs a guess bit δ' for δ. Finally, $\mathcal{B}_{kem}$ concludes its game with outputting $b' = 1$ if $\delta = \delta'$ and $b' = 0$, otherwise. This completes the description of $\mathcal{B}_{kem}$.

Adversary $\mathcal{B}_{kem}$ always perfectly simulates $\mathcal{A}$'s decapsulation queries. In case $b = 1$, $\mathcal{B}_{kem}$ uses the real key K_1^* for $\mathcal{A}$'s simulation which implies Equation (2). In case $b = 0$, $\mathcal{B}_{kem}$ uses a random key K_0^* for $\mathcal{A}$'s simulation which implies Equation (3). The complexity bounds for $\mathcal{B}_{kem}$ are clear from the construction, and it is left to show that $\text{uncert}_{\mathcal{B}_{kem}}(k) \leq Q \cdot \mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k)$ for a suitable $\mathcal{B}_{ae}$.

To this end we build an adversary $\mathcal{B}_{ae}$ against the AE security of $\mathcal{AE}$ as follows. $\mathcal{B}_{ae}$ inputs 1^k and, using its own pair of KEM keys $(pk, sk) \xleftarrow{\$} \mathsf{KEM.Kg}(1^k)$, emulates the same simulation for $\mathcal{A}$ as $\mathcal{B}_{kem}$ did above (using sk to answer its own $\textsc{Dec}(\cdot,\cdot)$ queries). It additionally picks a random index $j^* \in \{1, \ldots, Q\}$. On $\mathcal{A}$'s j^* decryption query (C_{j^*}, ψ_{j^*}), $\mathcal{B}_{ae}$ submits ψ_{j^*} to its own decryption-or-reject oracle $\textsc{DoR}_b(\cdot)$, and outputs $b' = 0$ iff $\textsc{DoR}_b(\cdot)$ rejects with $\perp$.

Now $\mathcal{B}_{ae}$ will always output $b' = 0$ if $b = 0$ by definition of $\textsc{DoR}_0$. In case $b = 1$, $\mathcal{B}_{ae}$ will output $b' = 1$ iff the ciphertext ψ_{j^*} is valid in the sense $\mathsf{AE.Dec}(K', \psi_{j^*}) \neq \perp$ for an independent, uniformly (by the AE experiment) chosen key K'. So adversary $\mathcal{B}_{ae}$'s advantage is as follows.

$$\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k) = \Pr[K' \xleftarrow{\$} \mathcal{K} : \mathsf{AE.Dec}(K', \psi_{j^*}) \neq \perp]$$

The above equals $\Pr[K' \xleftarrow{\$} \mathcal{K} : \text{pred}_{j^*}(K') = 1]$, where $\text{pred}_{j^*}(\cdot) = \mathsf{AE.Dec}(\cdot, \psi_{j^*})$ is the predicate $\mathcal{B}_{kem}$ submits to oracle $\textsc{Dec}$ as the j^*th query. For a uniformly chosen $j^* \in \{1, \ldots, Q\}$, the above equals $\text{uncert}_{\mathcal{B}_{ae}}(k)$. Consequently, $\mathbf{Adv}^{ae\text{-}\mathrm{ot}}_{\mathcal{AE},\mathcal{B}_{ae}}(k) \geq \frac{1}{Q} \cdot \text{uncert}_{\mathcal{B}_{ae}}(k)$. □

3 Efficient Key Encapsulation from DDH

3.1 Building blocks

We describe the building blocks used and assumptions made about them.

GROUP SCHEMES. A group scheme $\mathcal{GS}$ [13] specifies a sequence $(\mathcal{GR}_k)_{k\in\mathbb{N}}$ of group descriptions. For every value of a security parameter $k \in \mathbb{N}$, $\mathcal{GR}_k$ specifies the four tuple $\mathcal{GR}_k = (\hat{\mathbb{G}}_k, \mathbb{G}_k, p_k, g_k)$ (for notational convenience we sometimes drop the index k). $\mathcal{GR}_k = (\hat{\mathbb{G}}, \mathbb{G}, p, g)$ specifies a finite abelian group $\hat{\mathbb{G}}$, along with a prime-order subgroup $\mathbb{G}$, a generator g of $\mathbb{G}$, and the order p of $\mathbb{G}$. We denote the identity element of $\mathbb{G}$ as $1_{\mathbb{G}} \in \mathbb{G}$. We assume the existence of an

efficient sampling algorithm $x \xleftarrow{\$} \mathbb{G}$ and an efficient membership algorithm that test if a given element $x \in \hat{\mathbb{G}}$ is contained in the subgroup $\mathbb{G}$.

We further assume the DDH problem is hard in $\mathcal{GS}$, captured by defining the ddh-advantage of an adversary $\mathcal{B}_{\mathrm{ddh}}$ as

$$\mathbf{Adv}^{\mathrm{ddh}}_{\mathcal{GS},\mathcal{B}_{\mathrm{ddh}}}(k) = |\Pr[\mathcal{B}_{\mathrm{ddh}}(g,h,g^a,h^a)=1] - \Pr[\mathcal{B}_{\mathrm{ddh}}(g,h,g^a,K)=1]|\,,$$

where $g,\ h, K \xleftarrow{\$} \mathbb{G}$ and $a \leftarrow \mathbb{Z}_p^*$.

Authenticated Encryption. We need an abstract notion of algebraic authenticated encryption where the keyspace consists of $\mathbb{G}$, secure in the sense of OT-AE. In the full version we recall (following the encrypt-then-mac approach [4, 13]) how to build such algebraic AE satisfying all required functionality and security from the following basic primitives:

- A (computationally secure) one-time symmetric encryption scheme with binary k-bit keys (such as AES or padding with a PRNG)
- A (computationally secure) MAC (existentially unforgeable) with k-bit keys
- A (computationally secure) key-derivation function (pseudorandom).

We remark that for our purposes it is also possible to use a more efficient single-pass authenticated encryption scheme (see, e.g., [25]). In both cases the ciphertext expansion (i.e., ciphertext size minus plaintext size) of the AE scheme is only k (security parameter) bits which is optimal with respect to our security notion.

Target Collision Resistant Hashing. $\mathcal{TCR} = (\mathsf{TCR}_k)_{k\in\mathbb{N}}$ is a family of keyed hash functions $\mathsf{TCR}^s_k : \mathbb{G} \to \mathbb{Z}_p$ for each k-bit key s. It is assumed to be target collision resistant (TCR) [13], which is captured by defining the tcr-advantage of an adversary $\mathcal{B}_{\mathrm{tcr}}$ as $\mathbf{Adv}^{\mathrm{tcr}}_{\mathsf{TCR},\mathcal{B}_{\mathrm{tcr}}}(k) =$

$$\Pr[\mathsf{TCR}^s(c^*) = \mathsf{TCR}^s(c) \wedge c \neq c^* \ :\ s \xleftarrow{\$} \{0,1\}^k\ ;\ c^* \xleftarrow{\$} \mathbb{G}\ ;\ c \xleftarrow{\$} \mathcal{B}_{\mathrm{tcr}}(s,c^*)].$$

Note TCR is a weaker requirement than collision-resistance, so that, in particular, any practical collision-resistant function can be used. Also note that our notion of TCR is related to the stronger notion of universal one-way hashing [22], where in the security experiment of the latter the target value c^* is chosen by the adversary (but before seeing the hash key s).

Commonly [13, 21] this function is implemented using a dedicated cryptographic hash function like MD5 or SHA, which we assume to be target collision resistant. Since $|\mathbb{G}| = |\mathbb{Z}_p| = p$ we can alternatively also use a fixed (non-keyed) bijective encoding function $\mathsf{INJ} : \mathbb{G} \to \mathbb{Z}_p$. In that case we have a perfectly collision resistant hash function, i.e. $\mathbf{Adv}^{\mathrm{tcr}}_{\mathsf{INJ},\mathcal{B}_{\mathrm{tcr}}}(k) = 0$. In the full version, we show how to build such bijective encodings for a number of concrete group schemes.

3.2 The key-encapsulation mechanism

Let $\mathcal{GS}$ be a group scheme where $\mathcal{GR}_k$ specifies $(\hat{\mathbb{G}}, \mathbb{G}, g, p)$ and let $\mathsf{TCR} : \mathbb{G} \to \mathbb{Z}_p$ be a target collision resistant hash function (for simplicity we assume TCR to be non-keyed). We build a key encapsulation mechanism $\mathcal{KEM} = (\mathsf{KEM.kg}, \mathsf{KEM.Enc}, \mathsf{KEM.Dec})$ with $\mathcal{K} = \mathbb{G}$ as follows.

$\mathsf{KEM.Kg}(1^k)$	$\mathsf{KEM.Enc}(pk)$	$\mathsf{KEM.Dec}(sk, C)$
$x, y, \omega \xleftarrow{\$} \mathbb{Z}_p^*$	$r \xleftarrow{\$} \mathbb{Z}_p^*$; $c \leftarrow g^r$	Parse C as $(c, \pi) \in \hat{\mathbb{G}}^2$
$u \leftarrow g^x$; $v \leftarrow g^y$; $h \leftarrow g^\omega$	$t \leftarrow \mathsf{TCR}(c)$; $\pi \leftarrow (u^t v)^r$	if $c \notin \mathbb{G}$ return $\perp$
$pk \leftarrow (u, v, h) \in \mathbb{G}^3$	$C \leftarrow (c, \pi) \in \mathbb{G}^2$	$t \leftarrow \mathsf{TCR}(c)$
$sk \leftarrow (x, y, \omega) \in (\mathbb{Z}_p)^3$	$K \leftarrow h^r \in \mathbb{G}$	if $c^{xt+y} \neq \pi$ return $\perp$
Return (sk, pk)	Return (C, K)	Return $K \leftarrow c^\omega$

We stress that decryption never explicitly checks if $\pi \in \mathbb{G}$; this check happens implicitly when $c \in \mathbb{G}$ and $c^{xt+y} = \pi$ is checked. A correctly generated ciphertext has the form $C = (c, \pi) \in \mathbb{G} \times \mathbb{G}$, where $c = g^r$ and $\pi = (u^t v)^r = (g^{xt+y})^r = c^{xt+y}$. Hence decapsulation will not reject and compute the key $K = c^\omega = h^r$, as in encapsulation.

Encryption takes four standard exponentiations plus one application of TCR, where the generation of π can also be carried out as one single multi-exponentiation [6]. Decryption takes two exponentiations plus one application of TCR, where the two exponentiations can also be viewed as one sequential exponentiation [6] (which is as efficient as a multi-exponentiation) to simultaneously compute c^{xt+y} and c^ω. The proof of the the following theorem is given in the full version.

Theorem 2. *Let $\mathcal{GS}$ be a group scheme where the DDH problem is hard and assume $\mathcal{TCR}$ is target collision resistant. Then $\mathcal{KEM}$ is secure in the sense of* IND-CCCA. *In particular,*

$$\mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},t,Q,\mathrm{uncert}(k)}(k) \leq \mathbf{Adv}^{\mathrm{ddh}}_{\mathcal{GS},t}(k) + \mathbf{Adv}^{\mathrm{tcr}}_{\mathcal{TCR},t}(k) + \mathrm{uncert}(k) + \frac{Q}{p} \,.$$

3.3 Comparison with Cramer-Shoup and Kurosawa-Desmedt

The following table summarizes the key-encapsulation part of the Cramer-Shoup encryption scheme [13], the Kurosawa-Desmedt scheme [21], and ours.

Scheme	Ciphertext	Encapsulated Key
Cramer-Shoup	$g^r, \hat{g}^r, (u^t v)^r$	h^r
Kurosawa-Desmedt	$g^r, \hat{g}^r$	$(u^t v)^r$
Ours	$g^r, (u^t v)^r$	h^r

Here $\hat{g}$ is another element from the public-key. Compared to the Cramer-Shoup scheme, the Kurosawa-Desmedt scheme leaves out the value h^r and defines $(u^t v)^r$ as the session key. Our results shows that it is also possible to leave out the element $\hat{g}^r$ from the ciphertext and that $\pi = (u^t v)^r$ is sufficient to authenticate $c = g^r$. Hence, our scheme can be viewed as the *dual* of (the KEM part of) the Kurosawa-Desmedt scheme [21].

From a technical point of view, our scheme mixes Cramer-Shoup like techniques [12] to obtain a form of "plaintext awareness" for inconsistent ciphertexts with an "algebraic trick" from the Boneh-Boyen identity-based encryption

scheme [7] to decrypt consistent ciphertexts. Compared to Cramer-Shoup based proofs [11, 13, 21, 2] the most important technical difference, caused by the mentioned ability to decrypt consistent ciphertexts without knowing the full secret key, is that during our simulation the challenge ciphertexts is never made inconsistent. Intuitively this is the reason why we manage to maintain a consistent simulation using less redundancy in the secret key. This demonstrates that IND-CCCA security can be obtained with constructions that differ from hash proof systems.

On the other hand, the security proofs of all known schemes based on IBE techniques [10, 9, 19, 20, 18] inherently rely on some sort of external consistency check for the ciphertexts. This can be seen as the main reason why security of the IBE-based PKE schemes could only be proved in pairing groups (or relative to a gap-assumption), where the pairing was necessary for helping the proof identifying inconsistent ciphertexts. In our setting, the consistency check is done implicitly, using information-theoretic arguments borrowed from hash proof systems.

3.4 Efficiency

We compare our new DDH-based scheme's efficiency with the one of Kurosawa and Desmedt (in its more efficient "explicit-rejection" variant from [23]). Most importantly, the number of exponentiations for encryption and decryption are equal in both schemes. Although our security result is much more general (our KEM can be combined with any authenticated encryption scheme) this is not an exclusive advantage of our scheme. In fact we can derive the same result for the KD scheme from a more general theorem that we will prove in Section 5. (A similar result about combining the Kurosawa-Desmedt scheme with authenticated encryption was already obtained in [3] in the context of statefull encryption.)

However, there is one crucial difference in case one needs a scheme that is provably secure *solely* on the DDH assumption. Note that security (of the KD scheme and ours) relies on the DDH assumption *and* the assumption that $\mathcal{TCR}$ is target collision resistant. So as long as one does not want to sacrifice *provable* security by implementing the TCR function with a dedicated hash function like SHA-x or MD5 (what potentially renders the whole scheme insecure given the recent progress in attacking certain hash functions), one must either resort to inefficient generic constructions of TCR functions [22, 27], or one can use the "hash-free technique" described in [13]. With this latter technique, one can get rid of the TCR function completely; however, this comes at the cost of additional elements in the public and the secret key, and additional exponentiations during encryption. This overhead is linear in the number of elements that would have been hashed with the TCR. In the Kurosawa-Desmedt scheme, TCR acts on two group elements whereas in our scheme only on one. Hence the hash-free variant of our scheme is more efficient.

More importantly, since in our scheme a TCR is employed which maps *one* group element to integers modulo the group-order this can also be a bijection. In many concrete groups, e.g., when using the subgroup of quadratic residues

modulo a safe prime or certain elliptic curves, this bijection can be trivially implemented at zero cost [13, 9], without any additional computational assumption, and without sacrificing provable security. See the full version for more details. In terms of efficiency we view this as the main benefit of our scheme.

4 Key Encapsulation from n-Linear

4.1 Linear Assumptions

Let $n = n(k)$ be a polynomial in k. Generalizing [8, 18] we introduce the class of n-Linear assumptions which can be seen as a natural generalization of the DDH assumption and the Linear assumption.

Let $\mathcal{GS}$ be a group scheme. We define the n-lin-advantage of an adversary $\mathcal{B}_{n\text{-lin}}$ as

$$\mathbf{Adv}^{n\text{-lin}}_{\mathcal{GS},\mathcal{B}_{n\text{-lin}}}(k) = \big|\Pr[\mathcal{B}_{n\text{-lin}}(g_1,\ldots,g_n,g_1^{r_1},\ldots,g_n^{r_n},h,h^{r_1+\ldots+r_n})=1] - \Pr[\mathcal{B}_{n\text{-lin}}(g_1,\ldots,g_n,g_1^{r_1},\ldots,g_n^{r_n},h,K)=1]\big|,$$

where $g_1,\ldots,g_n,\ h,K \xleftarrow{\$} \mathbb{G}$ and all $r_i \leftarrow \mathbb{Z}_p^*$. We say that the *$n$-Linear Decisional Diffie-Hellman (n-Linear) assumption relative to group scheme $\mathcal{GS}$* holds if $\mathbf{Adv}^{n\text{-lin}}_{\mathcal{GS},\mathcal{B}_{n\text{-lin}}}$ is a negligible function in k for all polynomial-time adversaries $\mathcal{B}_{n\text{-lin}}$.

The n-Linear assumptions form a strict hierarchy of security assumptions with 1-Linear = DDH, 2-Linear=Linear [8] and, the larger the n, the weaker the n-Linear assumption. More precisely, for any $n \geq 1$ we have that n-Linear implies $n{+}1$-Linear. On the other hand (extending the case of $n = 1$ [8]) we can show that in the generic group model [29], the $n{+}1$-Linear assumption holds, even relative to an n-Linear oracle.

Lemma 3. DDH = 1-*Linear* $\not\Rightarrow$ 2-*Linear* $\not\Rightarrow$ 3-*Linear* $\not\Rightarrow \ldots$

4.2 The key-encapsulation mechanism

Let $\mathcal{GS}$ be a group scheme where $\mathcal{GR}_k$ specifies $(\hat{\mathbb{G}},\mathbb{G},g,p)$ and let $\mathsf{TCR}:\mathbb{G}^{n+1}\to\mathbb{Z}_p$ be a target collision resistant hash function. Generalizing the Kurosawa-Desmedt KEM, for a parameter $n = n(k) \geq 1$, we build $\mathcal{KEM} = (\mathsf{KEM.Kg}, \mathsf{KEM.Enc}, \mathsf{KEM.Dec})$ as follows.

Key generation $\mathsf{KEM.Kg}(1^k)$ generates random group elements $g_1,\ldots,g_n,h \in \mathbb{G}$. Furthermore, it defines $u_j = g_j^{x_j}h^z$ and $v_j = g_j^{y_j}h^{z'}$ for random $z,z' \in \mathbb{Z}_p$ and $x_j,y_j \in \mathbb{Z}_p$ $(j \in \{1,\ldots,n\})$. The public key contains the elements h, $(g_j,u_j)_{1\leq i\leq n}$, and the secret key contains all corresponding indices.

$\mathsf{KEM.Enc}(pk)$	$\mathsf{KEM.Dec}(sk,C)$
$\forall j \in \{1,\ldots,n\}$: $r_j \xleftarrow{\$} \mathbb{Z}_p^*$; $c_j \leftarrow g_j^{r_j}$	$\forall j \in \{1,\ldots,n\}$: check if $c_j \in \mathbb{G}$
$d \leftarrow h^{r_1+\ldots+r_n}$; $t \leftarrow \mathsf{TCR}(c_1,\ldots,c_n,d)$	Check if $d \in \mathbb{G}$
$C \leftarrow (c_1,\ldots,c_n,d)$; $K = \prod_{i=1}^n (u_i^t v_i)^{r_i}$	$t \leftarrow \mathsf{TCR}(c_1,\ldots,c_n,d)$
Return (C,K)	Return $K \leftarrow d^{zt+z'} \cdot \prod_{j=1}^n c_j^{x_j t+y_j}$

Ciphertexts contain $n+1$ group elements, public/secret keys $2n+1$ elements. The scheme instantiated with $n = 1$ precisely reproduces the KEM part of the Kurosawa-Desmedt encryption scheme [21]. Security of the schemes can be explained using the more general framework of computational hash-proof systems. This will be done in Section 5.

Theorem 3. *Let $\mathcal{GS}$ be a group scheme where the n-Linear problem is hard, assume $\mathcal{TCR}$ is target collision resistant. Then $\mathcal{KEM}$ is secure in the sense of* IND-CCCA.

We remark that it is also possible to give the scheme in its explicit-rejection variant [13]. Furthermore, in the full version we also provide a class of alternative schemes generalizing our dual KD scheme from Section 3 to the n-Linear assumption.

5 Key encapsulation from Hash Proof Systems

In [12], Cramer and Shoup showed that their original scheme in [13] was a special instance of a generic framework based on hash proof systems (HPS). Following [12] we recall the basic ideas of hash proof systems and show (generalizing [21]) how to build IND-CCCA secure key encapsulation based on a computational variant of hash proof systems. Here we use a slightly different notation for HPS that better reflects our primary application of hash-proof systems to key-encapsulation mechanisms.

5.1 Hash proof systems

Let $\mathcal{C}, \mathcal{K}$ be sets and $\mathcal{V} \subset \mathcal{C}$ a language. Let $\mathsf{D}_{sk} : \mathcal{C} \to \mathcal{K}$ be a hash function indexed with $sk \in \mathcal{S}$, where $\mathcal{S}$ is a set. A hash function D_{sk} is *projective* if there exists a projection $\mu : \mathcal{S} \to \mathcal{P}$ such that $\mu(sk) \in \mathcal{P}$ defines the action of D_{sk} over the subset $\mathcal{V}$. That is, for every $C \in \mathcal{V}$, the value $K = \mathsf{D}_{sk}(C)$ is uniquely determined by $\mu(sk)$ and C. In contrast, nothing is guaranteed for $C \in \mathcal{C} \setminus \mathcal{V}$, and it may not be possible to compute $\mathsf{D}_{sk}(C)$ from $\mu(sk)$ and C. A *strongly universal*$_2$ projective hash function has the additional property that for $C \in \mathcal{C} \setminus \mathcal{V}$, the projection key $\mu(sk)$ actually says nothing about the value of $K = \mathsf{D}_{sk}(C)$, even given an instance (C^*, K^*) such that $C^* \in \mathcal{C} \setminus \mathcal{V}$ and $K^* = \mathsf{D}_{sk}(C)$. More precisely, for all $pk \in \mathcal{P}$, C, all $C^* \in \mathcal{C} \setminus \mathcal{V}$ with $C \neq C^*$, all $K, K^* \in \mathcal{K}$,

$$\Pr_{\substack{sk \in \mathcal{S} \\ \mathsf{D}_{sk}(C^*) = K^* \\ \mu(sk) = pk}} [\mathsf{D}_{sk}(C) = K] = 1/|\mathcal{K}|. \tag{4}$$

A hash proof system $\mathcal{HPS} = (\mathsf{HPS.param}, \mathsf{HPS.pub}, \mathsf{HPS.priv})$ consists of three algorithms where the randomized algorithm $\mathsf{HPS.param}(1^k)$ generates instances of $params = (group, \mathcal{C}, \mathcal{V}, \mathcal{P}, \mathcal{S}, \mathsf{D}_{(\cdot)} : \mathcal{C} \to \mathcal{K}, \mu : \mathcal{S} \to \mathcal{P})$, where $group$ may contain some additional structural parameters. The deterministic public evaluation algorithm $\mathsf{HPS.pub}$ inputs the projection key $pk = \mu(sk)$, $C \in \mathcal{V}$ and

a witness w of the fact that $C \in \mathcal{V}$ and returns $K = \mathsf{D}_{sk}(C)$. The deterministic private evaluation algorithm inputs $sk \in \mathcal{S}$ and returns $\mathsf{D}_{sk}(C)$, without knowing a witness. We further assume there are efficient algorithms given for sampling $sk \in \mathcal{S}$ and sampling $C \in \mathcal{V}$ uniformly together with a witness w.

As computational problem we require that the *subset membership problem* is hard in $\mathcal{HPS}$ which means that the two elements C and C' are computationally indistinguishable, for random $C \in \mathcal{V}$ and random $C' \in \mathcal{C} \setminus \mathcal{V}$. This is captured by defining the advantage function $\mathbf{Adv}^{\mathrm{sm}}_{\mathcal{HPS},\mathcal{A}}(k)$ of an adversary $\mathcal{A}$ as

$$\mathbf{Adv}^{\mathrm{sm}}_{\mathcal{HPS},\mathcal{A}}(k) := \big|\Pr[C_1 \xleftarrow{\$} \mathcal{C}\,;\, b' \xleftarrow{\$} \mathcal{A}(\mathcal{C},\mathcal{V},C_1) \,:\, b'=1\,] - \Pr[C_0 \xleftarrow{\$} \mathcal{C}\setminus\mathcal{V}\,;\, b' \xleftarrow{\$} \mathcal{A}(\mathcal{C},\mathcal{V},C_0) :\, b'=1\,]\big| .$$

5.2 Key encapsulation from HPS

Using the above notion of a hash proof system, Kurosawa and Desmedt [21] proposed a hybrid encryption scheme which improved the schemes from [12]. The key-encapsulation part of it is as follows. The system parameters of the scheme consist of $params \xleftarrow{\$} \mathsf{HPS.param}(1^k)$.

$\mathsf{KEM.Kg}(k)$**.** Choose random $sk \xleftarrow{\$} \mathcal{S}$ and define $pk = \mu(sk) \in \mathcal{P}$. Return (pk, sk).

$\mathsf{KEM.Enc}(pk)$**.** Pick $C \xleftarrow{\$} \mathcal{V}$ together with its witness ω that $C \in \mathcal{V}$. The session key $K = \mathsf{D}_{sk}(C) \in \mathcal{K}$ is computed as $K \xleftarrow{\$} \mathsf{HPS.pub}(pk, C, \omega)$. Return (K, C).

$\mathsf{KEM.Dec}(sk, C)$**.** Reconstruct the key $K = \mathsf{D}_{sk}(C)$ as $K \leftarrow \mathsf{HPS.priv}(sk, C)$ and return K.

We can prove the following theorem that is a slight generalization of [21].

Theorem 4. *If $\mathcal{HPS}$ is strongly universal$_2$ and the subset membership problem is hard in $\mathcal{HPS}$ then $\mathcal{KEM}$ is secure in the sense of* IND-CCCA*.*

Unfortunately, the original KEM part of the Kurosawa Desmedt DDH-based hybrid encryption scheme [21] cannot be explained using this framework and hence needed a separate proof of security. This is since the underlying DDH-based hash proof system involves a target collision resistant hash function TCR which is a "computational primitive" whereas the strongly universal$_2$ property from Equation (4) is a *statistical property* which is in particularly not fulfilled by the DDH-based HPS from [12] used in [21]. In fact, the most efficient HPS-based schemes that are known involve computation of a TCR function and hence all need a separate proof of security. We note that this problem is inherited from the original HPS approach [13].

We overcome this problem we defining the weaker notion of *computational hash proof systems*.

5.3 Computational hash proof systems

We now define a weaker computational variant of strongly universal$_2$ hashing. For an adversary $\mathcal{B}$ we define the advantage function $\mathbf{Adv}^{\mathrm{cu}_2}_{\mathcal{HPS},\mathcal{B}}(k) = |\Pr[\mathbf{Exp}^{\mathrm{cu}_2\text{-}1}_{\mathcal{HPS},\mathcal{B}}(k)=1] - \Pr[\mathbf{Exp}^{\mathrm{cu}_2\text{-}0}_{\mathcal{HPS},\mathcal{B}}(k)=1]|$ where, for $b \in \{0,1\}$, $\mathbf{Exp}^{\mathrm{cu}_2\text{-}\mathrm{b}}_{\mathcal{HPS},\mathcal{B}}$ is defined by the following experiment.

Experiment $\mathbf{Exp}^{\mathrm{cu}_2\text{-}\mathrm{b}}_{\mathcal{HPS},\mathcal{B}}(k)$

$params \xleftarrow{\$} \mathsf{HPS.param}(1^k)$; $sk \xleftarrow{\$} \mathcal{S}$; $pk \leftarrow \mu(sk)$
$C^* \xleftarrow{\$} \mathcal{C} \setminus \mathcal{V}$; $K^* \leftarrow \mathsf{D}_{sk}(C^*)$; $(C, St) \xleftarrow{\$} \mathcal{B}_1^{\textsc{EvalD}(\cdot)}(pk, C^*, K^*)$
$K_0 \xleftarrow{\$} \mathcal{K}$; $K_1 \leftarrow \mathsf{D}_{sk}(C)$; $b' \xleftarrow{\$} \mathcal{B}_2(St, K_b)$
Return b'

where the evaluation oracle $\textsc{EvalD}(C)$ returns $K = \mathsf{D}_{sk}(C)$ if $C \in \mathcal{V}$ and $\perp$, otherwise. We also restrict to adversaries that only return ciphertexts $C \neq C^*$ and that ensure $C \in \mathcal{C} \setminus \mathcal{V}$. This is without losing generality, since $\mathcal{B}_1$ can check $C \in \mathcal{V}$ with its oracle $\textsc{EvalD}$. A hash proof system $\mathcal{HPS}$ is said to be *computationally universal*$_2$ (CU_2) if for all polynomial-time adversaries $\mathcal{B}$ that satisfy these requirements, the advantage function $\mathbf{Adv}^{\mathrm{cu}_2}_{\mathcal{HPS},\mathcal{B}}(k)$ is a negligible function in k.

The following theorem strengthens Theorem 4. A proof will be given in the full version.

Theorem 5. *If $\mathcal{HPS}$ is computationally universal$_2$ and the subset membership problem is hard then $\mathcal{KEM}$ from Section 5.2 is* IND-CCCA *secure. In particular,*

$$\mathbf{Adv}^{\mathrm{ccca}}_{\mathcal{KEM},t,Q,\mathrm{uncert}_{\mathcal{A}}(k)}(k) \leq \mathbf{Adv}^{\mathrm{sm}}_{\mathcal{HPS},t}(k) + (Q+1)\cdot(\mathrm{uncert}_{\mathcal{A}}(k) + \mathbf{Adv}^{\mathrm{cu}_2}_{\mathcal{HPS},t}(k)) \ .$$

5.4 A computational HPS from n-Linear

Let $\mathcal{GS}$ be a group scheme where $\mathcal{GR}_k$ specifies $(\hat{\mathbb{G}}, \mathbb{G}, g, p)$. Let $group = (\mathcal{GR}, g_1, \ldots, g_n, h)$, where $g_1, \ldots, g_n, h$ are independent generators of $\mathbb{G}$. Define $\mathcal{C} = \mathbb{G}^{n+1}$ and $\mathcal{V} = \{(g_1^{r_1}, \ldots, g_n^{r_n}, h^{r_1+\ldots+r_n}) \subset \mathbb{G}^{n+1} \ : \ r_1, \ldots, r_n \in \mathbb{Z}_p\}$ The values $(r_1, \ldots, r_n) \in \mathbb{Z}_p^n$ are a witness of $C \in \mathcal{V}$. Let $\mathsf{TCR} : \mathbb{G}^{n+1} \to \mathbb{Z}_p$ be a target collision resistant hash function. Let $\mathcal{S} = \mathbb{Z}_p^{2n+2}$, $\mathcal{P} = \mathbb{G}^{2n}$, and $\mathcal{K} = \mathbb{G}$. For $sk = (x_1, y_1, \ldots, x_n, y_n, z, z') \in \mathbb{Z}^{2n+2}$, define $\mu(sk) = (u_1, \ldots, u_n, v_1, \ldots, v_n)$, where, for $1 \leq i \leq n$, $u_i = g_i^{x_i} h^z$ and $v_i = g_i^{y_i} h^{z'}$. This defines the output of $\mathsf{HPS.param}(1^k)$. For $C = (c_1, \ldots, c_n, d) \in \mathcal{C}$ define

$$\mathsf{D}_{sk}(C) := d^{zt+z'} \cdot \prod_{i=1}^{n} c_i^{x_i t + y_i}, \text{ where } t = \mathsf{TCR}(c_1, \ldots, c_n) \ . \tag{5}$$

This defines $\mathsf{HPS.priv}(sk, C)$. Given $pk = \mu(sk)$, $C \in \mathcal{V}$ and a witness $w = (r_1, \ldots, r_n) \in (\mathbb{Z}_p)^n$ such that $C = (c_1, \ldots, c_n, d) = (g_1^{r_1}, \ldots, g_n^{r_n}, h^{r_1+\ldots+r_n})$ public evaluation $\mathsf{HPS.pub}(pk, C, w)$ computes $K = \mathsf{D}_{sk}(C)$ as

$$K = \prod_{i=1}^{n} (u_i^t v_i)^{r_i} \ .$$

Correctness follows by Equation (5) and the definition of μ. This completes the description of $\mathcal{HPS}$. Clearly, under the n-Linear assumption, the subset membership problem is hard in $\mathcal{HPS}$.

Obviously, the above defined HPS is not strongly universal$_2$ in the sense of Equation (4). But it is still computationally universal$_2$.

Lemma 4. *The n-Linear based HPS is computationally universal$_2$.*

Together with Theorem 5 this proves Theorem 3. For the case $n = 1$ this also gives an alternative security proof for the Kurosawa-Desmedt scheme [21].

Proof. Consider an adversary $\mathcal{B}$ in the CU_2 experiment such that $\mathcal{B}_1$ outputs a ciphertext $C \in \mathcal{C} \setminus \mathcal{V}$ and let $K \leftarrow \mathsf{D}_{sk}(C)$. Let COL be the event that $C \neq C^*$ but $\mathsf{TCR}(C) = \mathsf{TCR}(C^*)$. We claim that for the following adversary $\mathcal{B}_{\mathrm{tcr}}$ we have $\mathbf{Adv}^{\mathrm{tcr}}_{\mathsf{TCR},\mathcal{B}_{\mathrm{tcr}}}(k) = \Pr[\mathrm{COL}]$. Adversary $\mathcal{B}_{\mathrm{tcr}}$ inputs (s, C^*) and generates a random instance of *params* with known indices α_i such that $h = g^{\alpha_i}$. Furthermore, $\mathcal{B}_{\mathrm{tcr}}$ picks a random $sk \in \mathcal{S}$ and runs $\mathcal{B}_1$ on $pk = \mu(sk)$, a random $C^* \in \mathcal{C} \setminus \mathcal{V}$, and $K^* = \mathsf{D}_{sk}(C^*)$. To answer a query to the evaluation oracle $\textsc{Eval}\mathsf{D}(\cdot)$, $\mathcal{B}_{\mathrm{tcr}}$ fist verifies $C = (c_1, \ldots, c_n, d) \in \mathcal{V}$ by checking if $\prod c_i^{\alpha_i} = d$. If not, return $\perp$. Otherwise it returns $K = \mathsf{D}_{sk}(C)$. If for a decapsulation query C event COL happens, $\mathcal{B}_{\mathrm{tcr}}$ returns C to its TCR experiment and terminates.

Now we claim that conditioned under $\neg$COL, the key $K = \mathsf{D}_{sk}(C)$ is a uniform element in $\mathcal{K}$ independent of the adversary's view. This implies that not even a *computationally unbounded* $\mathcal{B}_2$ could succeed in the second stage. Hence, $\mathbf{Adv}^{\mathrm{cu}_2}_{\mathcal{HPS},\mathcal{B}}(k) \leq \mathbf{Adv}^{\mathrm{tcr}}_{\mathsf{TCR},\mathcal{B}_{\mathrm{tcr}}}(k)$, which proves the lemma.

Let $\log(\cdot) = \log_g(\cdot)$. Consider the view of $\mathcal{B}_2$ consisting of the random variables (pk, C^*, K^*, C), where $sk = (x_1, y_1, \ldots, x_n, y_n, z, z') \xleftarrow{\$} \mathbb{Z}^{2n+2}$, $pk = \mu(sk) = (u_1, \ldots, u_n, v_1, \ldots, v_n)$, $C^* = (c_1^*, \ldots, c_n^*, d^*) = (g_1^{r_1^*}, \ldots, g_n^{r_n^*}, h^{r^*})$ with $\sum r_i^* \neq r^*$ since $C^* \in \mathcal{C} \setminus \mathcal{V}$, $K^* = \mathsf{D}_{sk}(C^*)$, and $C = (c_1, \ldots, c_n, d) = (g_1^{r_1}, \ldots, g_n^{r_n}, h^r)$ ($\sum r_i \neq r$ since $C \in \mathcal{C} \setminus \mathcal{V}$). From the system parameters $g_1, \ldots, g_n, h$, adversary $\mathcal{B}_2$ learns $\omega = \log h$, $\omega_i = \log g_i$, and from pk

$$\text{for } 1 \leq i \leq n : \log u_i = \omega_i x_i + \omega z, \quad \log v_i = \omega_i y_i + \omega z' \,. \tag{6}$$

From C^* the adversary learns $r_i^* = \log_{g_i} c_i^*$, $r^* = \log_h d^*$, and from K^* (by Equation (5)) the value

$$\log K^* = \sum \omega_i r_i^* (x_i t^* + y_i) + \omega(z t^* + z') \,, \tag{7}$$

and $t^* = \mathsf{TCR}(c_1^*, \ldots, c_n^*, d^*)$. Furthermore, from C, $\mathcal{B}_2$ learns $r_i = \log_{g_i} c_i$ and $r = \log_h d$. Let $K = \mathsf{D}_{sk}(C)$. Our claim is that

$$\log K = \sum \omega_i r_i (x_i t + y_i) + \omega(z t + z') \,, \tag{8}$$

with $t = \mathsf{TCR}(C) \neq t^*$, is a uniform and independent element in $\mathbb{Z}_p$. Consider the set of linear equations over the hidden values $x_1, \ldots, x_n, y_1, \ldots, y_n, z, z'$ defined

by Equations (6), (7), and (8), defined by the matrix $M \in \mathbb{Z}_p^{n+2 \times n+2}$,

$$M = \begin{array}{c} \begin{array}{cccccccc} x_1 & \dots & x_n & y_1 & \dots & y_n & z & z' \end{array} \\ \left(\begin{array}{cccccccc} \omega_1 & & & & & & \omega & \\ & \ddots & & & 0 & & \vdots & 0 \\ & & \omega_n & & & & \omega & \\ & & & \omega_1 & & & & \omega \\ & 0 & & & \ddots & & 0 & \vdots \\ & & & & & \omega_n & & \omega \\ \omega_1 r_1^* t^* & \cdots & \omega_n r_n^* t^* & \omega_1 r_1^* & \dots & \omega_n r_n^* & \omega t^* r^* & \omega r^* \\ \omega_1 r_1 t & \cdots & \omega_n r_n t & \omega_1 r_1 & \dots & \omega_n r_n & \omega t r & \omega r \end{array} \right) \end{array}$$

Since $\det(M) = \omega^2 \prod \omega_i (t - t^*)(\sum_{i=1}^n r_i - r)(\sum_{i=1}^n r_i^* - r^*) \neq 0$, Equation (8) is linearly independent of (6) and (7).

We note that (generalizing [12]) we can also give a computationally universal$_2$ hash-proof system based on Paillier's decision composite residue (DCR) assumption.

References

1. Masayuki Abe, Rosario Gennaro, and Kaoru Kurosawa. Tag-KEM/DEM: A new framework for hybrid encryption. Cryptology ePrint Archive, Report 2005/027, 2005. http://eprint.iacr.org/.
2. Masayuki Abe, Rosario Gennaro, Kaoru Kurosawa, and Victor Shoup. Tag-KEM/DEM: A new framework for hybrid encryption and a new analysis of Kurosawa-Desmedt KEM. In EUROCRYPT 2005, volume 3494 of LNCS, pages 128–146. Springer-Verlag, 2005.
3. Mihir Bellare, Tadayoshi Kohno, and Victor Shoup. Stateful public-key cryptosystems: How to encrypt with one 160-bit exponentiation. In ACM CCS 2006, pages 380–389. ACM Press, 2006.
4. Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In ASIACRYPT 2000, volume 1976 of LNCS, pages 531–545. Springer-Verlag, 2000.
5. Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In ACM CCS 1993, pages 62–73. ACM Press, 1993.
6. D. J. Bernstein. Pippenger's exponentiation algorithm. Available from http://cr.yp.to/papers.html#pippenger, 2001.
7. Dan Boneh and Xavier Boyen. Efficient selective-ID secure identity based encryption without random oracles. In EUROCRYPT 2004, volume 3027 of LNCS, pages 223–238. Springer-Verlag, 2004.
8. Dan Boneh, Xavier Boyen, and Hovav Shacham. Short group signatures. In CRYPTO 2004, volume 3152 of LNCS, pages 41–55. Springer-Verlag, 2004.
9. Xavier Boyen, Qixiang Mei, and Brent Waters. Direct chosen ciphertext security from identity-based techniques. In ACM CCS 2005, pages 320–329. ACM Press, 2005.
10. Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from identity-based encryption. In EUROCRYPT 2004, volume 3027 of LNCS, pages 207–222. Springer-Verlag, 2004.

11. Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In CRYPTO 1998, volume 1462 of LNCS, pages 13–25. Springer-Verlag, 1998.
12. Ronald Cramer and Victor Shoup. Universal hash proofs and a paradigm for adaptive chosen ciphertext secure public-key encryption. In EUROCRYPT 2002, volume 2332 of LNCS, pages 45–64. Springer-Verlag, 2002.
13. Ronald Cramer and Victor Shoup. Design and analysis of practical public-key encryption schemes secure against adaptive chosen ciphertext attack. *SIAM Journal on Computing*, 33(1):167–226, 2003.
14. Rosario Gennaro and Victor Shoup. A note on an encryption scheme of Kurosawa and Desmedt. Cryptology ePrint Archive, Report 2004/194, 2004.
15. D. Hofheinz, J. Herranz, and E. Kiltz. The Kurosawa-Desmedt key encapsulation is not chosen-ciphertext secure. Cryptology ePrint Archive, Report 2006/207, 2006.
16. Antoine Joux and Kim Nguyen. Separating decision Diffie-Hellman from computational Diffie-Hellman in cryptographic groups. *Journal of Cryptology*, 16(4):239–247, September 2003.
17. Jonathan Katz and Moti Yung. Unforgeable encryption and chosen ciphertext secure modes of operation. In FSE 2000, volume 1978 of LNCS, pages 284–299. Springer-Verlag, 2000.
18. E. Kiltz. Chosen-ciphertext secure key-encapsulation based on Gap Hashed Diffie-Hellman. In PKC 2007, volume 4450 of LNCS, pages 282 – 297, 2007.
19. Eike Kiltz. Chosen-ciphertext security from tag-based encryption. In TCC 2006, volume 3876 of LNCS, pages 581–600. Springer-Verlag, 2006.
20. Eike Kiltz. On the limitations of the spread of an IBE-to-PKE transformation. In PKC 2006, volume 3958 of LNCS, pages 274–289. Springer-Verlag, 2006.
21. Kaoru Kurosawa and Yvo Desmedt. A new paradigm of hybrid encryption scheme. In CRYPTO 2004, volume 3152 of LNCS, pages 426–442. Springer-Verlag, 2004.
22. Moni Naor and Moti Yung. Universal one-way hash functions and their cryptographic applications. In 21st ACM STOC, pages 33–43. ACM Press, 1989.
23. Le Trieu Phong and Wakaha Ogata. On a variation of Kurosawa-Desmedt encryption scheme. Cryptology ePrint Archive, Report 2006/031, 2006.
24. Charles Rackoff and Daniel R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In CRYPTO 1991, volume 576 of LNCS, pages 433–444. Springer-Verlag, 1992.
25. Phillip Rogaway, Mihir Bellare, John Black, and Ted Krovetz. OCB: A block-cipher mode of operation for efficient authenticated encryption. In ACM CCS 2001, pages 196–205. ACM Press, 2001.
26. Phillip Rogaway and Thomas Shrimpton. A provable-security treatment of the key-wrap problem. In EUROCRYPT 2006, volume 4004 of LNCS, pages 373–390. Springer-Verlag, 2006.
27. John Rompel. One-way functions are necessary and sufficient for secure signatures. In 22nd ACM STOC, pages 387–394. ACM Press, 1990.
28. Hovav Shacham. A Cramer-Shoup encryption scheme from the Linear Assumption and from progressively weaker linear variants. Cryptology ePrint Archive, Report 2007/074, 2007.
29. Victor Shoup. Lower bounds for discrete logarithms and related problems. In EUROCRYPT 1997, volume 1233 of LNCS, pages 256–266. Springer-Verlag, 1997.
30. Victor Shoup. Using hash functions as a hedge against chosen ciphertext attack. In EUROCRYPT 2000, volume 1807 of LNCS, Springer-Verlag, 2000.