Efficient Chosen Ciphertext Secure Public Key Encryption under the Computational Diffie-Hellman Assumption

Goichiro Hanaoka[1] and Kaoru Kurosawa[2]

[1] RCIS, AIST
[2] Ibaraki University

Abstract. Recently Cash, Kiltz, and Shoup [13] showed a variant of the Cramer-Shoup (CS) scheme [14] whose chosen-ciphertext (CCA) security relies on the *computational Diffie-Hellman* (CDH) assumption. The cost for this high security is that the size of ciphertexts is much longer than the CS scheme (which is based on the decisional Diffie-Hellman assumption). In this paper, we show how to achieve CCA-security under the CDH assumption without increasing the size of ciphertexts. We also show a more efficient scheme under the *hashed Diffie-Hellman* assumption.
Both of our schemes are based on a certain broadcast encryption (BE) scheme while the Cash-Kiltz-Shoup scheme is based on the Twin DH problem. Of independent interest, we also show a generic method of constructing CCA-secure PKE schemes from BE schemes.

1 Introduction

1.1 Background

Chosen-ciphertext security (CCA-security, for short) [35, 16] is considered as a standard notion of security for public key encryption (PKE) in practice. Furthermore, this security also implies universally composable security [11]. So far, many CCA-secure PKE schemes have been proposed, both theoretical ones [31, 16, 36] and practical ones [14, 38, 12, 26, 10, 1, 25, 22], and their security are proven under existence of enhanced trapdoor permutations (for theoretical schemes) or under various number theoretic assumptions (for practical schemes). Theoretical schemes pursue weaker assumptions and practical schemes pursue efficiency.

One of the most important research topics in this field is to design CCA-secure PKE schemes with weaker assumptions and better efficiency. Cramer and Shoup showed the first practical PKE scheme under the decisional Diffie-Hellman (DDH) assumption. Kurosawa and Desmedt showed a more efficient scheme under the DDH assumption [26].

However, there has been no (even theoretical) CCA-secure PKE scheme under the *computational Diffie-Hellman* (CDH) assumption except for a recent work by Cash, Kiltz, and Shoup [13].[3]

[3] We started our work independently of [13]. In fact, the authors of [13] kindly cited an earlier version of our paper as an independent work.

1.2 Our Contribution

In this paper, we present a practical CCA-secure PKE scheme under the CDH assumption such that the size of a ciphertext is much smaller than that of the Cash-Kiltz-Shoup (CKS) scheme. Indeed, the ciphertext length of our scheme is the same as that of the Cramer-Shoup (CS) scheme (which is based on the DDH assumption). Specifically, ciphertext overhead of our CDH-based scheme is only three group elements for arbitrary plaintext length, while that of the CKS scheme is $k/\log k + 2$ group elements where k is the security parameter.

We also present a more efficient CCA-secure PKE scheme under the hashed Diffie-Hellman (HDH) assumption. This scheme is as efficient as the Kurosawa-Desmedt (KD) scheme [26] in terms of both computational costs and data sizes while the HDH assumption is weaker than the DDH assumption.[4]

Both of our schemes are based on the Naor-Pinkas broadcast encryption (BE) scheme while the CKS scheme is based on the Twin DH problem. Of independent interest, we show a generic method of transforming any selectively chosen-plaintext (CPA) secure *verifiable* BE scheme into a CCA-secure key encapsulation mechanism (KEM) with almost no cost, where we say that a BE scheme is verifiable if any receiver can tell whether all receivers decrypt a given ciphertext to the identical result or not.

Further, we show that almost all existing methods for achieving CCA-security, e.g. [16, 14, 12], can be explained by using verifiable BE schemes. It is also possible to construct a new PKE scheme based on this paradigm, for example, from the Boneh-Gentry-Waters (BGW) BE scheme [6]. Moreover, we can generically convert any CPA-secure verifiable BE into a CCA-secure BE with almost no cost. Our results imply that verifiable BE is a powerful tool to obtain CCA-security.

1.3 Related Works

Under Stronger Assumptions than CDH. After the KD scheme, several CCA-secure encryption schemes were constructed under stronger assumptions than the CDH assumption. The scheme of Boyen, Mei, and Waters [10] is based on the *bilinear Diffie-Hellman* (BDH) assumption. The scheme of Kiltz [25] is based on the *gap hashed Diffie-Hellman* (GHDH) assumption. The scheme of Hofheinz and Kiltz [22] is based on the *n-linear* DDH assumption.

KEM/DEM Framework. The KEM/DEM framework was formalized by Shoup [38] for the design of hybrid encryption schemes, and the CS hybrid encryption scheme was constructed. However, the KD scheme does not fit into this framework. To explain the KD scheme in a general framework, Abe, Gennaro, Kurosawa, and Shoup [1] established the Tag-KEM/DEM framework. Hofheinz and Kiltz [22] introduced the notion of Constrained CCA (CCCA) security of KEM.

[4] After an earlier version of this paper [21], in the latest full-version of [13], Cash, Kiltz, and Shoup pointed out that the Hofheinz-Kiltz scheme in [22] can be also proved to be secure under the HDH assumption.

How to Achieve CCA Security. Naor and Yung showed that a non-adaptively CCA-secure encryption scheme can be constructed from any semantically secure encryption [19] and *non-interactive zero knowledge* (NIZK) proof [4]. Dolev, Dwork, and Naor [16] and Sahai [36] improved this idea and presented adaptively CCA-secure constructions. However, it is not known if an NIZK proof can be constructed from any semantically secure encryption scheme. (A partial answer to this question is given in [32].)

Canetti, Halevi, and Katz [12] proposed another generic method such that a CCA-secure PKE scheme can be obtained from a selectively secure identity-based encryption (IBE) scheme [37, 5]. Boneh and Katz [7] improved its efficiency. Kiltz [24] discussed a more relaxed condition for achieving CCA-security.

Broadcast Encryption. In the model of broadcast encryption (BE) schemes, there are multiple receivers. The sender broadcasts a ciphertext such that only privileged receivers can decrypt. Fiat and Naor [17] proposed the first non-trivial construction of BE. Naor, Naor, and Lotspiech [29] presented a significantly more efficient scheme. Naor and Pinkas [30] proposed a public key BE scheme by using ElGamal-like construction, and Dodis and Fazio [15] improved it to be secure against adaptive adversaries as well as chosen-ciphertext adversaries. Boneh, Gentry, and Waters [6] proposed the first fully collusion resistant (public key) BE scheme whose ciphertext and user decryption keys are of constant size.

1.4 Organization

Definitions are given in Sec. 2. Our main idea is described in Sec. 3. The proposed scheme under the CDH assumption is shown in Sec. 4. A more efficient scheme under the HDH assumption is presented in Sec. 5. A comparison with other PKE schemes is given in Sec. 6. Finally, we show a generic method to construct CCA-secure PKE schemes from verifiable BE in Sec. 7.

2 Definitions

2.1 Key Encapsulation Mechanisms

It is well-known that by combining a CCA-secure KEM and a CCA-secure data encryption mechanism (DEM), a CCA-secure PKE scheme is generically obtained [38], and furthermore, there exist some other flexible methods for hybrid encryption as well [1, 22]. It is also known that a CCA-secure DEM can be generically constructed from any pseudorandom functions without redundancy [27, 33].

A KEM consists of the following three algorithms: $\mathbf{Setup}(1^k)$ takes as input the security parameter 1^k and outputs a decryption key dk and a public key PK. $\mathbf{Encrypt}(PK)$ takes as input a public key PK and outputs a pair (ψ, K) where ψ is a ciphertext and $K \in \mathcal{K}$ is a data encryption key. $\mathbf{Decrypt}(dk, \psi, PK)$ takes as input the decryption key dk, a ciphertext ψ, and the public key PK, and outputs

$K \in \mathcal{K}$ which will be used for decrypting the DEM part of hybrid encryption. We require that if $(dk, PK) \stackrel{R}{\leftarrow} \mathbf{Setup}(1^k)$ and $(\psi, K) \stackrel{R}{\leftarrow} \mathbf{Encrypt}(PK)$ then $\mathbf{Decrypt}(dk, \psi, PK) = K$.

CCA-security of a KEM is defined using the following game between an attack algorithm A and a challenger. Both the challenger and A are given 1^k as input.

Setup. The challenger runs $\mathbf{Setup}(1^k)$ to obtain a decryption key dk and a public key PK. The challenger also runs algorithm **Encrypt** to obtain $(\psi^\star, K^\star) \stackrel{R}{\leftarrow} \mathbf{Encrypt}(PK)$ where $K^\star \in \mathcal{K}$. Next, the challenger picks a random $b \in \{0, 1\}$. It sets $K_0 = K^\star$ and picks a random $K_1 \in \mathcal{K}$. It then gives the public key PK and the challenge ciphertext $(\psi^\star, K_b)$ to algorithm A.

Query. Algorithm A adaptively issues decryption queries $\psi_1, ..., \psi_{q_D}$. For query $\psi_i (\neq \psi^\star)$, the challenger responds with $\mathbf{Decrypt}(dk, \psi_i, PK)$.

Guess. Algorithm A outputs its guess $b' \in \{0, 1\}$ for b and wins the game if $b = b'$.

Let $\mathsf{AdvKEM}_\mathsf{A}$ denote the probability that A wins the game.

Definition 1 We say that a KEM is (τ, ϵ, q_D) *CCA-secure* if for all τ-time algorithms A who make a total of q_D decryption queries, we have that $|\mathsf{AdvKEM}_\mathsf{A} - 1/2| < \epsilon$.

2.2 Number Theoretic Assumptions

The CDH, HDH, and DDH Assumptions. Let $\mathbb{G}$ be a multiplicative group with prime order p. Then, the CDH problem on $\mathbb{G}$ is stated as follows. Let A be an algorithm, and we say that A has advantage ϵ in solving the CDH problem on $\mathbb{G}$ if $\Pr[\mathsf{A}(g, g^\alpha, g^\beta) = g^{\alpha\beta}] \geq \epsilon$, where the probability is over the random choice of generator g in $\mathbb{G}$, the random choice of α and β in $\mathbb{Z}_p$, and the random bits consumed by A.

Definition 2 *We say that the (τ, ϵ)-CDH assumption holds in $\mathbb{G}$ if no τ-time algorithm has advantage ϵ in solving the CDH problem on $\mathbb{G}$.*

The *hashed Diffie-Hellman* (HDH) problem on $\mathbb{G}$ and function $h : \mathbb{G} \to \mathcal{D}$ is stated as follows. Let A be an algorithm, and we say that A has advantage ϵ in solving the HDH problem on $\mathbb{G}$ and h if

$$1/2 \cdot |\Pr[\mathsf{A}(g, g^\alpha, g^\beta, h(g^{\alpha\beta})) = 0] - \Pr[\mathsf{A}(g, g^\alpha, g^\beta, T) = 0]| \geq \epsilon,$$

where the probability is over the random choice of generator g in $\mathbb{G}$, the random choice of α and β in $\mathbb{Z}_p$, the random choice of $T \in \mathcal{D}$, and the random bits consumed by A.

Definition 3 *We say that the (τ, ϵ)-HDH assumption holds in $\mathbb{G}$ and h if no τ-time algorithm has advantage ϵ in solving the HDH problem on $\mathbb{G}$ and h. Especially, we say that the (τ, ϵ)-DDH assumption holds in $\mathbb{G}$ if (τ, ϵ)-HDH assumption holds in $\mathbb{G}$ and h, where h is the identity function.*

Important Implications. It is important to note that the HDH assumption is strictly weaker than the DDH assumption for appropriately chosen h. If h is a *key derivation function* [38], then the DDH assumption immediately implies the HDH assumption (but not vice versa). Furthermore, if h is a hardcore bit for the Diffie-Hellman key [18, 9, 8, 23], then the CDH assumption is equivalent to the HDH assumption. Obviously, the CDH assumption is weaker than both the HDH and DDH assumptions.

Hardcore Bits for the Diffie-Hellman Key. Let A be a τ-time algorithm which has advantage ϵ in solving the HDH problem on $\mathbb{G}$ and $h : \mathbb{G} \rightarrow \{0,1\}$.

Definition 4 *We say that function $h : \mathbb{G} \rightarrow \{0,1\}$ is a (p_1, p_2) hardcore bit function in $\mathbb{G}$ if there exists a $p_1(\tau)$-time algorithm B which for any given A, can solve the CDH problem with advantage $p_2(\epsilon)$ for some polynomials p_1 and p_2.*

2.3 Public Key Broadcast Encryption Schemes

Model. Here, we review definitions for public key BE schemes. For simplicity, we define encryption schemes as key encapsulation mechanisms, and borrow the same notations as [6] with some slight modifications. A BE scheme consists of the following three algorithms: $\mathbf{Setup}(1^k, n, t)$ takes as input the security parameter 1^k, the number of receivers n, and the maximum number of revoked users t ($t < n$). It outputs n decryption keys $d_1, ..., d_n$ and a public key PK. $\mathbf{Encrypt}(\mathcal{S}, PK)$ takes as input a subset $\mathcal{S} \subseteq \{1, ..., n\}$ with $|\mathcal{S}| \geq n - t$, and a public key PK. It outputs a pair (ψ, K) where ψ is called the header and $K \in \mathcal{K}$ is a message encryption key. Let M be a message to be broadcast to the set $\mathcal{S}$ and let C_M be the encryption of M under the symmetric key K. The broadcast to users in $\mathcal{S}$ consists of $(\mathcal{S}, \psi, C_M)$. The pair $(\mathcal{S}, \psi)$ is often called the full header and C_M is often called the broadcast body. $\mathbf{Decrypt}(\mathcal{S}, i, d_i, \psi, PK)$ takes as input a subset $\mathcal{S} \subseteq \{1, ..., n\}$, a user index $i \in \{1, ..., n\}$ and the decryption key d_i for user i, a header ψ, and the public key PK. If $i \in \mathcal{S}$ and $|\mathcal{S}| \geq n - t$, then the algorithm outputs the message encryption key $K \in \mathcal{K}$. The key K can then be used to decrypt the broadcast body C_M and obtain the message body M.

As usual, we require that the scheme be correct, namely that for all $\mathcal{S} \subseteq \{1, ..., n\}$ and all $i \in \mathcal{S}$, if $((d_1, ..., d_n), PK) \stackrel{R}{\leftarrow} \mathbf{Setup}(1^k, n, t)$ and $(\psi, K) \stackrel{R}{\leftarrow} \mathbf{Encrypt}(\mathcal{S}, PK)$ then $\mathbf{Decrypt}(\mathcal{S}, i, d_i, \psi, PK) = K$.

CCA Security. We define CCA-security of a BE scheme against a static adversary. Security is defined using the following game between an attack algorithm A and a challenger. Both the challenger and A are given 1^k, n and t, the total number of potential users and the maximum number of revoked users, respectively, as inputs.

Init. Algorithm A begins by outputting a set $\mathcal{S}^\star \subseteq \{1, ..., n\}$ of receivers that A wants to attack, where $|\mathcal{S}^\star| \geq n - t$.

Setup. The challenger runs $\mathbf{Setup}(1^k, n, t)$ to obtain decryption keys $d_1, ..., d_n$ and a public key PK. The challenger also runs algorithm **Encrypt** to obtain $(\psi^\star, K^\star) \stackrel{R}{\leftarrow} \mathbf{Encrypt}(\mathcal{S}^\star, PK)$ where $K^\star \in \mathcal{K}$. Next, the challenger picks a random $b \in \{0, 1\}$. It sets $K_0 = K^\star$ and picks a random $K_1 \in \mathcal{K}$. It then gives $(\psi^\star, K_b)$ to algorithm A.

Query. Algorithm A adaptively issues decryption queries $q_1, ..., q_D$ where a decryption query consists of the triple $(u, \mathcal{S}, \psi)$ where $\psi \neq \psi^\star$, $\mathcal{S} \subseteq \mathcal{S}^\star$ and $u \in \mathcal{S}$. The challenger responds with K (or $\perp$) $= \mathbf{Decrypt}(\mathcal{S}, u, d_u, \psi, PK)$.

Guess. Algorithm A outputs its guess b' for b and wins the game if $b = b'$.

Let $\mathsf{AdvBr}_{\mathsf{A},n,t}$ denote the probability that A wins the game when the challenger is given n and t.

Definition 5 We say that a broadcast encryption scheme is $(\tau, \epsilon, n, t, q_D)$ *CCA-secure* if for all τ-time algorithms A who make a total of q_D decryption queries, we have that $|\mathsf{AdvBr}_{\mathsf{A},n,t} - 1/2| < \epsilon$. Especially, we say that a broadcast encryption scheme is (τ, ϵ, n, t) *semantically secure* if it is $(\tau, \epsilon, n, t, 0)$ CCA-secure.

Verifiability. For achieving CCA-security, we need an important property for underlying BE, which we call *verifiability*. Roughly speaking, we say that a BE scheme has verifiability if a valid receiver of a broadcasted message can verify if his decryption result is the same as that for any other receiver. We can define two flavors of verifiability: *public* verifiability and *private* verifiability. Their difference is that in a publicly verifiable BE scheme, a receiver can verify equality of keys without using his decryption key, and on the other hand, it is necessary in a privately verifiable scheme.

For public verifiability, we define adversary A's advantage $\mathsf{AdvVfy}_{\mathsf{A},n,t}$ as

$$\begin{aligned}&\mathsf{AdvVfy}_{\mathsf{A},n,t}\\ &= \Pr[\exists i, j \in \mathcal{S}^\star,\ \mathbf{Decrypt}(\mathcal{S}^\star, i, d_i, \psi^\star, PK) \neq \mathbf{Decrypt}(\mathcal{S}^\star, j, d_j, \psi^\star, PK)|\\ &\qquad ((d_1, ..., d_n), PK) \stackrel{R}{\leftarrow} \mathbf{Setup}(1^k, n, t);\ (\mathcal{S}^\star, \psi^\star) \stackrel{R}{\leftarrow} \mathsf{A}((d_1, ..., d_n), PK)].\end{aligned}$$

Definition 6 We say that a broadcast encryption scheme is (τ, ϵ, n, t) *publicly verifiable* if for all τ-time algorithms A, we have that $\mathsf{AdvVfy}_{\mathsf{A},n,t} < \epsilon$.

We can also define private verifiability in a similar manner, and its formal definition is given in the full version of this paper [21].

2.4 Other Cryptographic Tools

Target Collision Resistant Hash Functions. Let $\mathsf{TCR} : \mathcal{X} \to \mathcal{Y}$ be a hash function (we individually define the range and domain of TCR for each scheme), A be an algorithm, and A's advantage $\mathsf{AdvTCR}_{\mathsf{A}}$ be $\mathsf{AdvTCR}_{\mathsf{A}} = \Pr[\mathsf{TCR}(x') = \mathsf{TCR}(x) \in \mathcal{Y} \wedge x' \neq x|\ x \stackrel{R}{\leftarrow} \mathcal{X};\ x' \stackrel{R}{\leftarrow} \mathsf{A}(x)]$.

Definition 7 We say that TCR is a (τ, ϵ) *target collision resistant hash function* if for all τ-time algorithms A, we have that $\mathsf{AdvTCR}_{\mathsf{A}} < \epsilon$.

One-Time Signatures. A signature scheme consists of the following three algorithms: $\mathbf{Gen}(1^k)$ takes as input the security parameter 1^k, and outputs a verification key vk and a signing key sk. $\mathbf{Sign}(sk, m)$ takes as input a signing key sk and a message m, and outputs a signature σ. $\mathbf{Verify}(vk, m, \sigma)$ takes as input a verification key vk, a message m, and a signature σ, and outputs a bit $b \in \{0, 1\}$. We require that for all sk, all m in the message space, and all σ output by $\mathbf{Sign}(sk, m)$, we have $\mathbf{Verify}(vk, m, \sigma) = 1$.

Security is defined using the following game between an attack algorithm A and a challenger. Both the challenger and A are given 1^k as input.

Setup. The challenger runs $\mathbf{Gen}(1^k)$ to obtain vk and sk. It gives A the verification key vk.

Query. Algorithm A may issue at most one query m. The challenger responds with $\sigma \stackrel{R}{\leftarrow} \mathbf{Sign}(sk, m)$.

Forge. Algorithm A outputs $(m^\star, \sigma^\star)$ such that $(m^\star, \sigma^\star) \neq (m, \sigma)$.

Let $\mathsf{AdvOTS}_\mathsf{A}$ denote the probability that $\mathbf{Verify}(vk, m^\star, \sigma^\star) = 1$.

Definition 8 We say that a signature scheme is (τ, ϵ) *strongly unforgeable* if for all τ-time algorithms A, we have that $\mathsf{AdvOTS}_\mathsf{A} < \epsilon$.

3 Toward Efficient CCA-Secure Scheme under CDH

The Naor-Pinkas BE scheme [30] is one-way under the CDH assumption. In this section, we construct a verifiable BE scheme from the Naor-Pinkas BE scheme, where we say that a BE scheme is verifiable if any receiver can tell whether all receivers decrypt a given ciphertext to the identical result or not. The main difficulty in this paper is how to add verifiability to the Naor-Pinkas scheme.

Our CCA-secure PKE scheme under the CDH assumption is obtained from this variant of the Naor-Pinkas BE scheme. See Sec. 7 for details on this observation.

3.1 The Naor-Pinkas Broadcast Encryption Scheme

The Naor-Pinkas scheme [30], which was constructed based on [2], is as follows. Let $\mathbb{G}$ be a multiplicative group with prime order p, and $g \in \mathbb{G}$ be a generator. Suppose that there are at most t potential revoked users.

In the setup phase, the center chooses a polynomial $f(x) = \sum_{0 \leq i \leq t} a_i x^i$ over $GF(p)$ randomly, and computes $y_i = g^{a_i}$ for $0 \leq i \leq t$. The public key is $PK = (\mathbb{G}, g, y_0, ..., y_t)$. The center keeps $f(x)$ as the master key, and gives $d_i = f(i)$ to each user $i = 1, ..., p-1$ as his decryption key.

To revoke users $i_1, ..., i_t \in Z_p$, the sender generates a ciphertext $\psi = (g^r, (g^{f(i_1)})^r, ..., (g^{f(i_t)})^r)$ and a key $K = y_0^r$ where $r \stackrel{R}{\leftarrow} \mathbb{Z}_p$. Notice that $g^{f(i)}$ can be computed as $\prod_{0 \leq j \leq t} y_j^{i^j}$ for any $i \in \{1, ..., p-1\}$. On receiving $\psi = (C_0, ..., C_t)$, user $u \notin \{i_1, ..., i_t\}$ computes $C_u = C_0^{d_u}$ and recovers the key

as $K = C_u^{\lambda(u)} \prod_{1 \le j \le t} C_j^{\lambda(i_j)}$ where $\lambda(x)$ is the Lagrange coefficient such that $\lambda(x) = \prod_{i' \in \{i, i_1, ..., i_t\} \setminus \{x\}} i' \cdot (i' - x)^{-1}$ over $\mathbb{Z}_p$.

3.2 Verifiability

As we mentioned, the main difficulty in this paper is how to add verifiability to the Naor-Pinkas scheme. Here we give a solution. Consider a modification of the Naor-Pinkas scheme such that user i is given $(f(i), f(rnd), rnd)$ as his decryption key, where $rnd \xleftarrow{R} \mathbb{Z}_p$. We note that a legitimate user i can decrypt a ciphertext in two different ways according to two different keys, i.e. $f(i)$ and $f(rnd)$. If these decryption results are not identical, then the user can detect that the ciphertext is in an invalid form. Notice that since rnd is random and not known to other users, it is difficult to generate an invalid ciphertext whose decryption results under $f(i)$ and $f(rnd)$ are identical.

Unfortunately, the above idea is faulty. Namely, even if user i is revoked and $f(i)$ does not work for decryption, he still has $f(rnd)$ and can decrypt a ciphertext by using it. Hence, the modified scheme is not secure any more. Therefore, we further modify the Naor-Pinkas scheme as follows: For at most t revoked users, in the setup phase, a polynomial $f(x) = \sum_{0 \le i \le 2t+1} a_i x^i$ is generated in the same manner as the original Naor-Pinkas scheme except that its degree is changed to be $2t + 1$. The public key is $PK = (\mathbb{G}, g, y_0, ..., y_{2t+1})$. We assume that a user i has two unique identities $\mathbf{i}$ and i, where we denote $i = (\mathbf{i}, \mathsf{i}) \in \{1, ..., p-1\}^2$. The center keeps $f(x)$ as the master key, and for user $i = (\mathbf{i}, \mathsf{i}) \in \{1, ..., p-1\}^2$ he publishes $d_i = (f(\mathbf{i}), f(\mathsf{i}), f(rnd), rnd)$ as i's decryption key, where $rnd \xleftarrow{R} \mathbb{Z}_p$. Assuming that users $i_1 (= (\mathbf{i}_1, \mathsf{i}_1)), ..., i_t (= (\mathbf{i}_t, \mathsf{i}_t))$ are revoked, the sender generates $\psi = (g^r, (g^{f(\mathbf{i}_1)})^r, ..., (g^{f(\mathbf{i}_t)})^r, (g^{f(\mathsf{i}_1)})^r, ..., (g^{f(\mathsf{i}_t)})^r)$ and $K = y_0^r$ where $r \xleftarrow{R} \mathbb{Z}_p$.

On receiving $\psi = (C_0, ..., C_{2t})$, a user $i = (\mathbf{i}, \mathsf{i}) (\notin \{i_1, ..., i_t\})$ computes $C_{\mathbf{i}} = C_0^{f(\mathbf{i})}$, $C_{\mathsf{i}} = C_0^{f(\mathsf{i})}$, and $C_{rnd} = C_0^{f(rnd)}$. We notice that ψ can be decrypted by using any two of $C_{\mathbf{i}}$, C_{i}, and C_{rnd} with the Lagrange interpolation (for example, by using $(C_{\mathbf{i}}, C_{\mathsf{i}})$, the session key is recovered as $K = C_{\mathbf{i}}^{\lambda(\mathbf{i})} C_{\mathsf{i}}^{\lambda(\mathsf{i})} \prod_{1 \le j \le t} (C_j^{\lambda(\mathbf{i}_j)} C_{j+t}^{\lambda(\mathsf{i}_j)})$ where $\lambda(x)$ is the Lagrange coefficient such that $\lambda(x) = \prod_{i' \in \{\mathbf{i}, \mathsf{i}, \mathbf{i}_1, \mathsf{i}_1, ..., \mathbf{i}_t, \mathsf{i}_t\} \setminus \{x\}} i' \cdot (i' - x)^{-1}$ over $\mathbb{Z}_p$). Then, user i carries out decryption in three different ways according to the three different choices of $(C_{\mathbf{i}}, C_{\mathsf{i}})$, $(C_{\mathbf{i}}, C_{rnd})$, and $(C_{\mathsf{i}}, C_{rnd})$. Then, user i can be convinced of the equality of decryption results for all legitimate subscribers if i's three decryption results are identical. Furthermore, when i is revoked, he cannot decrypt a ciphertext at all even though he still has $f(rnd)$. Now, we obtain a new verifiable BE scheme from Naor-Pinkas BE, and are ready to convert it into a CCA-secure PKE scheme.

4 Efficient CCA-Secure KEM from CDH

In this section, we show an efficient CCA-secure KEM under the CDH assumption such that the size of ciphertexts is the same as that of the CS scheme. Our

KEM is obtained from a verifiable BE scheme which was shown in Sec. 3. Let $\mathbb{G}$ be a multiplicative group with prime order p, and $g \in \mathbb{G}$ be a generator. Then, the construction of the scheme is as follows:

Setup(1^k)**:** Generate a random polynomial $f(x) = a_0 + a_1 x + \cdots + a_{k+2}x^{k+2}$ over $\mathbb{Z}_p$, and compute $y_i = g^{a_i}$ for $0 \leq i \leq k+2$. The decryption key is $f(x)$, and the public key is $PK = (\mathbb{G}, g, y_0, y_1, ..., y_{k+2}, \mathsf{TCR}_0, \mathsf{TCR}_1, h)$, where $\mathsf{TCR}_b : \mathbb{G} \to \mathcal{S}_b$ $(b = 0, 1)$ are target collision resistant hash functions such that $\mathcal{S}_0 \cup \mathcal{S}_1 \subseteq \mathbb{Z}_p^*$, $\mathcal{S}_0 \cap \mathcal{S}_1 = \emptyset$, and $h : \mathbb{G} \to \{0,1\}$ is a hardcore bit function for the Diffie-Hellman key in $\mathbb{G}$.[5]

Encrypt(PK)**:** Pick a random $r \xleftarrow{R} \mathbb{Z}_p$, and compute

$$\psi = (g^r, g^{r \cdot f(\mathbf{i})}, g^{r \cdot f(\mathsf{i})}), \qquad K = (h(y_0^r)||h(y_1^r)||...||h(y_{k-1}^r))$$

where $\mathbf{i} = \mathsf{TCR}_0(g^r)$ and $\mathsf{i} = \mathsf{TCR}_1(g^r)$. The final output is (ψ, K). (Notice that one can easily compute $g^{f(x)}$ as $g^{f(x)} = \prod_{0 \leq i \leq k+2} y_i^{x^i}$.)

Decrypt(dk, ψ, PK)**:** For a ciphertext $\psi = (C_0, C_1, C_2)$, check whether $(C_1, C_2) \stackrel{?}{=} (C_0^{f(\mathbf{i})}, C_0^{f(\mathsf{i})})$, where $\mathbf{i} = \mathsf{TCR}_0(C_0)$ and $\mathsf{i} = \mathsf{TCR}_1(C_0)$. If not, output $\perp$. Otherwise, output $K = (h(C_0^{a_0})||h(C_0^{a_1})||...||h(C_0^{a_{k-1}}))$.

Theorem 1 *Let $\mathbb{G}$ be a multiplicative group with prime order p, TCR_0 and TCR_1 be (τ, ϵ_{tcr}) target collision resistant hash functions, and h be a (p_1, p_2) hardcore bit function for the Diffie-Hellman key in $\mathbb{G}$. Then, the above scheme is $(p_1^{-1}(\tau) - o(p_1^{-1}(\tau)), k \cdot p_2^{-1}(\epsilon_{cdh}) + 2\epsilon_{tcr} + q_D(2k/(p-3) + 1/(p-k-2)), q_D)$ CCA-secure under the (τ, ϵ_{cdh}) CDH assumption on $\mathbb{G}$.*

Proof. Assume that for challenge ciphertext $(g^\beta, g^{\beta \cdot f(\mathbf{i}^\star)}, g^{\beta \cdot f(\mathsf{i}^\star)})$ such that $\mathbf{i}^\star = \mathsf{TCR}_0(g^\beta)$ and $\mathsf{i}^\star = \mathsf{TCR}_1(g^\beta)$, there exists an adversary A' which distinguishes $(h(y_0^\beta)||h(y_1^\beta)||...||h(y_{k-1}^\beta))$ from a random k-bit string. Then, by a standard hybrid argument, there also exists another adversary A which for some j such that $0 \leq j \leq k-1$ distinguishes

$$(h(y_0^\beta)||h(y_1^\beta)||...||h(y_j^\beta)||random_{k-j-1})$$

from

$$(h(y_0^\beta)||h(y_1^\beta)||...||h(y_{j-1}^\beta)||random_{k-j})$$

where $random_\ell$ denotes an ℓ-bit random string.

Now, assume we are given such an adversary A which distinguishes these two values with running time τ, advantage ϵ, and q_D decryption queries. We use A to construct another adversary B which for given (g, g^α, g^β) distinguishes $h(g^{\alpha\beta})$ from a random bit. Define adversary B as follows:

[5] h is a random string R if it is the Goldreich-Levin (GL) bit [18], where the size of R is equal to that of a group element. See also Appendix of [9] for the GL bit of the Diffie-Hellman keys.

1. For given $(g, g^{\alpha}, g^{\beta})$, B picks target collision resistant hash functions TCR_0 and TCR_1, and computes $\mathbf{i}^{\star} = \mathsf{TCR}_0(g^{\beta})$ and $\mathsf{i}^{\star} = \mathsf{TCR}_1(g^{\beta})$.
2. B sets $y_j = g^{\alpha}$, and picks distinct randoms $rnd_j, ..., rnd_{k-1}$ from $\mathbb{Z}_p^* \backslash \{\mathbf{i}^{\star}, \mathsf{i}^{\star}\}$. B also picks randoms $u_{\mathbf{i}^{\star}}$, $u_{\mathsf{i}^{\star}}$, $a_0, ..., a_{j-1}$, and $u_j, ..., u_{k-1}$ from $\mathbb{Z}_p$.
3. B calculates $y_l = g^{a_l}$ for $0 \leq l \leq j-1$.
4. Let $f(x) = \sum_{i=0}^{k+2} a_i x^i$ be a polynomial over $\mathbb{Z}_p$ such that $a_j = \alpha$, $f(\mathbf{i}^{\star}) = u_{\mathbf{i}^{\star}}$, $f(\mathsf{i}^{\star}) = u_{\mathsf{i}^{\star}}$, and $f(rnd_j) = u_j, ..., f(rnd_{k-1}) = u_{k-1}$. Then, by using the Lagrange interpolation, B calculates $y_{j+1}, ..., y_{k+2}$ such that $g^{f(x)} = \prod_{0 \leq j \leq k+2} y_j^{x^j}$. Notice that $y_l = g^{a_l}$ holds for $0 \leq l \leq k+2$.
5. B inputs public key $PK = (\mathbb{G}, g, y_0, y_1, ..., y_{k+2}, \mathsf{TCR}_0, \mathsf{TCR}_1, h)$ and challenge ciphertext $\psi^{\star} = (g^{\beta}, (g^{\beta})^{u_{\mathbf{i}^{\star}}}, (g^{\beta})^{u_{\mathsf{i}^{\star}}})$ and

$$K^{\star} = (h((g^{\beta})^{a_0}) || h((g^{\beta})^{a_1}) || ... || h((g^{\beta})^{a_{j-1}}) || \gamma || random_{k-j-1})$$

   to A where γ is $h(g^{\alpha\beta})$ or a random bit.
6. When A makes decryption query $\psi = (C_0, C_1, C_2)$, B proceeds as follows:
   (a) If $C_0 = g^{\beta}$, then B responds $\perp$.
   (b) If $C_0 \neq g^{\beta}$ and $\mathsf{TCR}_b(C_0) \in \{\mathbf{i}^{\star}, \mathsf{i}^{\star}, rnd_j, ..., rnd_{k-2}, rnd_{k-1}\}$ for $b = 0$ or 1, then B aborts and outputs a random bit.
   (c) If $C_0 \neq g^{\beta}$ and $\mathsf{TCR}_b(C_0) \notin \{\mathbf{i}^{\star}, \mathsf{i}^{\star}, rnd_j, ..., rnd_{k-2}, rnd_{k-1}\}$ for both $b = 0$ and 1, B computes $C_0^{u_{\mathbf{i}^{\star}}}$, $C_0^{u_{\mathsf{i}^{\star}}}, C_0^{u_j}, ..., C_0^{u_{k-2}}$, and $C_0^{u_{k-1}}$. Let $\mathsf{TCR}_0(C_0) = \mathbf{i}$ and $\mathsf{TCR}_1(C_0) = \mathsf{i}$, and f_1, f_2, and f_3 be polynomials over $\mathbb{Z}_p$ with degree $k+2$ whose coefficient for x^l term is a_l for $0 \leq l \leq j-1$, such that

$$\begin{aligned}
&(f_1(\mathbf{i}), f_1(\mathsf{i}), f_1(\mathbf{i}^{\star}), f_1(\mathsf{i}^{\star}), f_1(rnd_{j+1}), ..., f_1(rnd_{k-1})) \\
&\qquad = (\log_{C_0} C_1, \log_{C_0} C_2, u_{\mathbf{i}^{\star}}, u_{\mathsf{i}^{\star}}, u_{j+1}, ..., u_{k-1}) \\
&(f_2(\mathbf{i}), f_2(\mathsf{i}), f_2(\mathbf{i}^{\star}), f_2(rnd_j), ..., f_2(rnd_{k-1})) \\
&\qquad = (\log_{C_0} C_1, \log_{C_0} C_2, u_{\mathbf{i}^{\star}}, u_j, ..., u_{k-1}) \\
&(f_3(\mathbf{i}), f_3(\mathsf{i}), f_3(\mathsf{i}^{\star}), f_3(rnd_j), ..., f_3(rnd_{k-1})) \\
&\qquad = (\log_{C_0} C_1, \log_{C_0} C_2, u_{\mathsf{i}^{\star}}, u_j, ..., u_{k-1}).
\end{aligned}$$

   Then, B calculates ${C_0}^{a_{1,l}}, {C_0}^{a_{2,l}}, {C_0}^{a_{3,l}}$ by using the Lagrange interpolation where $a_{1,l}$, $a_{2,l}$, and $a_{3,l}$ denote the coefficients of x^l term of f_1, f_2, and f_3 for $j \leq l \leq k-1$, respectively, and responds

$$K = (h(C_0^{a_0}) || ... || h(C_0^{a_{j-1}}) || h(C_0^{a_{1,j}}) || ... || h(C_0^{a_{1,k-1}}))$$

   if ${C_0}^{a_{1,j}} = {C_0}^{a_{2,j}} = {C_0}^{a_{3,j}}$, or "$\perp$" otherwise.
7. Finally, A outputs a bit b as his guess, and B outputs the same bit b as his own guess for $h(g^{\alpha\beta})$.

Let Win denote the event that A's guess is correct in the real world, Abort denote the event that A submits a ciphertext $\psi = (C_0, C_1, C_2)$ such that $C_0 \neq g^{\beta}$ and $\mathsf{TCR}_b(C_0) \in \{\mathbf{i}^{\star}, \mathsf{i}^{\star}, rnd_j, ..., rnd_{k-2}, rnd_{k-1}\}$ for $b = 0$ or 1, and $\mathsf{Invalid}$ denote the event that A submits a ciphertext $\psi = (C_0, C_1, C_2)$ such that B does not abort, ${C_0}^{a_{1,j}} = {C_0}^{a_{2,j}} = {C_0}^{a_{3,j}}$, but $(C_1, C_2) \neq (C_0^{f(\mathbf{i})}, C_0^{f(\mathsf{i})})$.

Then, B's advantage for guessing $h(g^{\alpha\beta})$ is estimated as follows:

$$\begin{aligned}\frac{1}{2}\cdot|\Pr[\mathsf{B}(g,g^{\alpha},g^{\beta},h(g^{\alpha\beta}))=0]-\Pr[\mathsf{B}(g,g^{\alpha},g^{\beta},T)=0]| \\ \geq|\Pr[\mathsf{Win}|\overline{\mathsf{Abort}}\wedge\overline{\mathsf{Invalid}}]\Pr[\overline{\mathsf{Abort}}\wedge\overline{\mathsf{Invalid}}]-\frac{1}{2}| \\ \geq|\Pr[\mathsf{Win}]-\Pr[\mathsf{Abort}]-\Pr[\mathsf{Invalid}]-\frac{1}{2}|.\end{aligned}$$

Now, we prove following lemmas.

Lemma 1 $\Pr[\mathsf{Abort}]\leq 2\epsilon_{tcr}+\frac{2q_Dk}{p-3}$.

Proof. Assume we are given an adversary A with $\Pr[\mathsf{Abort}]=p_A$. Then, we can construct another adversary B' which for given $C\xleftarrow{R}\mathbb{G}$, finds $C'(\neq C)\in\mathbb{G}$ such that $\mathsf{TCR}_b(C')=\mathsf{TCR}_b(C)$ for $b=0$ or 1 as follows: For given C, B' generates decryption key $f(x)$ and public key $PK=(\mathbb{G},g,y_0,y_1,...,y_{k+2},\mathsf{TCR}_0,\mathsf{TCR}_1,h)$, and computes challenge ciphertext $\psi^{\star}=(C,C^{u_{\mathbf{i}^{\star}}},C^{u_{\mathsf{i}^{\star}}})$, where $u_{\mathbf{i}^{\star}}=f(\mathbf{i}^{\star})$, $u_{\mathsf{i}^{\star}}=f(\mathsf{i}^{\star})$, $\mathbf{i}^{\star}=\mathsf{TCR}_0(C)$, and $\mathsf{i}^{\star}=\mathsf{TCR}_1(C)$. B' also picks distinct randoms $rnd_j,...,rnd_{k-1}$ from $\mathbb{Z}_p^*\backslash\{\mathbf{i}^{\star},\mathsf{i}^{\star}\}$, and gives PK and $(\psi^{\star},K^{\star})$ to A, where $K^{\star}$ is a correct key under $f(x)$ or a random element of $\mathbb{G}$ with probability $1/2$.

Since $rnd_j,...,rnd_{k-1}$ are information-theoretically hidden to A, for a query $\psi=(C_0,C_1,C_2)$, $\mathsf{TCR}_0(C_0)$ or $\mathsf{TCR}_1(C_0)\in\{rnd_j,...,rnd_{k-2},rnd_{k-1}\}$ happens with probability at most $2(k-j)/(p-3)$. Therefore, the probability that A submits a ciphertext $\psi=(C_0,C_1,C_2)$ $(C_0\neq C)$ such that $\mathsf{TCR}_0(C_0)=\mathbf{i}^{\star}$ or $\mathsf{TCR}_1(C_0)=\mathsf{i}^{\star}$ is at least $p_A-2q_D(k-j)/(p-3)$. B' outputs such C_0 as C'.

By using B' as it is, we immediately have an algorithm B'' which for given $C\xleftarrow{R}\mathbb{G}$, finds $C''(\neq C)\in\mathbb{G}$ such that $\mathsf{TCR}_0(C'')=\mathsf{TCR}_0(C)$ with probability at least $p_A-2q_D(k-j)/(p-3)-p_1$, where p_1 is the probability that B' outputs C' such that $\mathsf{TCR}_1(C')=\mathsf{TCR}_1(C)$. Since $p_1\leq\epsilon_{tcr}$, B'''s advantage is at least $p_A-2q_D(k-j)/(p-3)-\epsilon_{tcr}$. Hence, $\epsilon_{tcr}\geq p_A-2q_D(k-j)/(p-3)-\epsilon_{tcr}$, and therefore, we have $2\epsilon_{tcr}+2q_D(k-j)/(p-3)\geq p_A$. □

Lemma 2 $\Pr[\mathsf{Invalid}]\leq\frac{q_D}{p-k-2}$.

Proof. Let $f_0(x)=\sum_{0\leq l\leq j-1}a_lx^l$, and $f_1'(x),f_2'(x)$, and $f_3'(x)$ be polynomials such that $f_l(x)=f_0(x)+x^j\cdot f_l'(x)$ for $l=1,2,3$. Let $f'(x)$ be a polynomial such that $f(x)=f_0(x)+x^j\cdot f'(x)$. Suppose $\psi=(C_0,C_1,C_2)$ is a ciphertext such that B does not abort, ${C_0}^{f_1'(0)}={C_0}^{f_2'(0)}={C_0}^{f_3'(0)}$, but $(C_1,C_2)\neq(C_0^{f(\mathbf{i})},C_0^{f(\mathsf{i})})$. Then, we notice that f_1' and f_2' which are polynomials with degree $k-j+2$ have $k-j+3$ intersections, and consequently they have to be identical. Similarly, we have that $f_1'=f_2'=f_3'$. This implies that for $[\mathsf{Invalid}=\text{true}]$, A has to choose C_1 and C_2 (without knowing $rnd_j,...,rnd_{k-1}$) such that f_1' (with degree $k-j+2$) satisfies

1. $(f_1'(\mathbf{i}),f_1'(\mathsf{i}),f_1'(\mathbf{i}^{\star}),f_1'(\mathsf{i}^{\star}),f_1'(rnd_j),...,f_1'(rnd_{k-1}))$
$=((\log_{C_0}C_1-f_0(\mathbf{i}))\cdot\mathbf{i}^{-j},(\log_{C_0}C_2-f_0(\mathsf{i}))\cdot\mathsf{i}^{-j},f'(\mathbf{i}^{\star}),f'(\mathsf{i}^{\star}),$
$f'(rnd_j),...,f'(rnd_{k-1})),$

2. $f'_1 \neq f'$.

Since f'_1 and f' have at most $k-j+2$ intersections and $k-j+1$ of them are $(\mathbf{i}^\star, f'(\mathbf{i}^\star))$, $(\mathsf{i}^\star, f'(\mathsf{i}^\star))$, $(rnd_{j+1}, f'(rnd_{j+1})), ..., (rnd_{k-1}, f'(rnd_{k-1}))$, there is only one remained intersection which must be $(rnd_j, f'(rnd_j))$. Therefore, [Invalid = true] happens only when A correctly guesses the value of rnd_j (even if A is given $rnd_{j+1}, ..., rnd_{k-1}$). Hence, for any invalid query ψ, the probability that B does not respond "$\perp$" is at most $1/(p-k+j-2)(\leq 1/(p-k-2))$. □

A's advantage is estimated as at least $1/k$ times A′'s advantage due to the hybrid argument. □

5 Efficient CCCA-Secure KEM from HDH

In this section, based on the strategy in Sec. 3, we propose another KEM which is CCCA-secure [22] under the HDH assumption. This scheme is as efficient as the KD scheme [26] with a weaker assumption. As shown in [22], a CCA-secure PKE scheme can be constructed by combining any CCCA-secure KEM and authenticated symmetric key encryption [3] as a DEM. Let $\mathbb{G}$ be a multiplicative group with prime order p, and $g \in \mathbb{G}$ be a generator. Then, the construction of our CCCA-secure KEM is as follows:

Setup(1^k)**:** Generate a random polynomial $f(x) = a_0 + a_1 x + a_2 x^2$ over $\mathbb{Z}_p$, and compute $y_j = g^{a_j}$ for $0 \leq j \leq 2$. The decryption key is $f(x)$, and the public key is $PK = (\mathbb{G}, g, y_0, y_1, y_2, \mathsf{TCR}, h)$, where $\mathsf{TCR} : \mathbb{G} \to \mathbb{Z}_p^*$ is a target collision resistant hash function and $h : \mathbb{G} \to \{0,1\}^\nu$ is a hash function.

Encrypt(PK)**:** Pick a random $r \xleftarrow{R} \mathbb{Z}_p$, and compute $\psi = (g^r, g^{r \cdot f(i)})$ and $K = h(y_0^r)$, where $i = \mathsf{TCR}(g^r)$. The final output is (ψ, K). (Notice that one can easily compute $g^{f(x)}$ as $g^{f(x)} = \prod_{0 \leq j \leq 2} y_j^{x^j}$.)

Decrypt(dk, ψ, PK)**:** For a ciphertext $\psi = (C_0, C_1)$, check whether $C_1 \stackrel{?}{=} C_0^{f(i)}$, where $i = \mathsf{TCR}(C_0)$. If not, output $\perp$. Otherwise, output $K = h(C_0^{a_0})$.

The above scheme can be proved to be CCCA-secure, and its security is formally addressed in the full version of this paper [21].

6 Comparison

Table 1 shows a comparison of our schemes with other CCA-secure schemes, i.e. Cramer-Shoup (CS) [14, 38], Kurosawa-Desmedt (KD) [26], Boyen-Mei-Waters (BMW) [10], Kiltz [25], Cash-Kiltz-Shoup (CKS) [13], and Hofheinz-Kiltz (HK) [22]. In the comparison, we utilize a redundancy-free CCA-secure DEM [20, 33] for constructing a CCA-secure hybrid encryption scheme from a CCA-secure KEM.

As seen in Table 1, our proposed scheme in Sec. 4 yields both provable security under the CDH assumption and short ciphertext length which is comparable

Table 1. Efficiency comparison for CCA-secure PKE schemes. Some figures are borrowed from [10, 25]. For efficiency, we count the number of pairings, multi(or sequential)-exponentiations [34], regular-exponentiations, and other group operations ("ops" denotes group operations) used for encryption and decryption. All symmetric operations (such as hash function/MAC/KDF) are ignored. Ciphertext overhead represents the difference between ciphertext and plaintext length, and $|g|$ and $|mac|$ are the length of a group element and an authentication tag, respectively. In the table, we let $k' = k/\log k$ where k is the security parameter, i.e. DEM-key length.

	Security Assumption	Ciphertext Overhead	Encryption	Decryption
			#pairings + #[multi,regular]-exp (+ #ops)	
CS [14]	DDH	$3\|g\|$	$0 + [1, 3]$	$0 + [1, 1]$
KD [26]	DDH	$2\|g\| + \|mac\|$	$0 + [1, 2]$	$0 + [1, 0]$
BMW [10]	BDH	$2\|g\|$	$0 + [1, 2]$	$1 + [0, 1]$
Kiltz [25]	GHDH	$2\|g\|$	$0 + [1, 2]$	$0 + [1, 0]$
CKS [13]	CDH	$(k' + 2)\|g\|$	$0 + [k' + 1, k' + 1]$	$0 + [1^{\ddagger}, 0]$
	HDH	$3\|g\|$	$0 + [2, 2]$	$0 + [1, 0]$
HK† [22]	HDH	$2\|g\| + \|mac\|$	$0 + [1, 2]$	$0 + [1, 0]$
Ours §4	CDH	$3\|g\|$	$0 + [2^{\ddagger}, k' + 1]$	$0 + [1^{\ddagger}, 0]$
Ours §5	HDH	$2\|g\| + \|mac\|$	$0 + [1, 2]$	$0 + [1, 0]$
Ours §7.3	2ℓ-BDHE	$2\|g\|$	$0 + [0, 3] + \ell$	$3 + [0, 0] + \ell$

† A slight modification by [13] is applied.

‡ Relatively more expensive computation is needed for one exponentiation.

to other practical schemes. Comparing with the CDH-based CKS scheme, our scheme in Sec. 4 is more efficient, and especially, the ciphertext overhead of our scheme, i.e. three group elements, is much shorter than that of the CKS scheme, i.e. $k/\log k + 2$ group elements, since $k/\log k \simeq 18$ for 128-bit security. In the comparison, we assume that $\log k$ hardcore bits can be extracted from a single DH key [18]. Furthermore, the ciphertext overhead of our scheme is the same as that of the CS scheme. Our scheme in Sec. 5 is as efficient as the KD scheme with a weaker underlying assumption. The Hofheinz-Kiltz scheme [22] (with a modification by [13]) has almost the same property as ours. (See also the footnote in Sec. 1.2.)

7 CCA-Security from BE with Verifiability

In this section, we observe that it is possible to construct a CCA-secure PKE scheme from an arbitrary verifiable BE scheme, and that security of many existing CCA-secure PKE schemes can also be explained from this viewpoint. This observation implies that one of promising approaches for achieving CCA-security is to concentrate on designing verifiable BE schemes. In fact, constructions of our proposed schemes are based on this approach.

7.1 The Generic Conversion

Given a verifiable BE scheme $\Pi' = (\mathbf{Setup}', \mathbf{Encrypt}', \mathbf{Decrypt}')$ which is CPA-secure against selective adversaries, we construct a CCA-secure KEM $\Pi =$

(**Setup**, **Encrypt**, **Decrypt**). In the construction, we use a strong one-time signature scheme $\Sigma = (\mathbf{Gen}, \mathbf{Sign}, \mathbf{Verify})$ in which the verification key generated by $\mathbf{Gen}(1^k)$ has length k. We assume that the maximum number of potential users in Π' is n, and a sender can revoke t users where there exists an injective mapping (or a target collision resistant hash function) $\mathsf{INJ} : \{0,1\}^k \rightarrow \mathcal{P}$ and $\mathcal{P}$ is the set of all subsets $\mathcal{S} \subseteq \{1, ..., n\}$ with $|\mathcal{S}| = n - t$. Notice that for existence of such an injective mapping, it is necessary that ${}_nC_t \geq 2^k$ (for example, $(n, t) = (2k, k)$). The construction of Π is as follows:

Setup(1^k)**:** Choose n and t (which is a possible parameter choice for Π') such that ${}_nC_t \geq 2^k$. Run $\mathbf{Setup}'(1^k, n, t)$ to obtain $(d_1, ..., d_n, PK)$, and pick an injective mapping $\mathsf{INJ} : \{0,1\}^k \rightarrow \mathcal{P}$. The decryption key is $dk = (d_1, ..., d_n)$ and the public key is $\overline{PK} = (PK, \mathsf{INJ})$.

Encrypt$(\overline{PK})$**:** Run $\mathbf{Gen}(1^k)$ to obtain verification key vk and signing key sk (with $|vk| = k$), and compute $\mathcal{S}_{vk} = \mathsf{INJ}(vk)$, $(\psi, K) \leftarrow \mathbf{Encrypt}'(\mathcal{S}_{vk}, PK)$ and $\sigma \leftarrow \mathbf{Sign}(sk, \psi)$. The final output is $((\psi, vk, \sigma), K)$.

Decrypt$(dk, \psi, \overline{PK})$**:** For a ciphertext (ψ, vk, σ), check whether $\mathbf{Verify}(vk, \psi, \sigma) \stackrel{?}{=} 1$. If not, output $\perp$. Otherwise, compute $\mathcal{S}_{vk} = \mathsf{INJ}(vk)$ and output $K \leftarrow \mathbf{Decrypt}'(\mathcal{S}_{vk}, i, d_i, \psi, PK)$ where $i \in \mathcal{S}_{vk}$.

CCA-security of the above construction can be proven in a similar manner to [12]. We give an intuitive explanation for the security. Let A be an algorithm which can break CCA-security of Π. Then, it is possible to construct another algorithm B which can break Π' by using A as follows: B runs $(vk^\star, sk^\star) \leftarrow \mathbf{Gen}(1^k)$, and commits $\mathcal{S}^\star = \mathsf{INJ}(vk^\star)$ as the subset of users which will be attacked. For given public key PK of Π', B passes (PK, INJ) to A as a public key of Π. When A submits decryption query (ψ, vk, σ), B responds to it by simply decrypting the ciphertext with decryption key d_i such that $i \in \mathsf{INJ}(vk) \backslash \mathcal{S}^\star \subseteq \{1, ..., n\}$. We note that there always exists at least one such a decryption key unless $vk = vk^\star$, and $vk \neq vk^\star$ holds with an overwhelming probability if σ is a valid signature. Let $(\psi^\star, K^\star)$ be a challenge ciphertext of Π' from the challenger. Then, B gives $((\psi^\star, vk^\star, \sigma^\star), K^\star)$ to A as a challenge ciphertext of Π where $\sigma^\star \leftarrow \mathsf{Sign}(sk^\star, \psi^\star)$. A formal security proof is given in the full version of this paper [21].

Theorem 2 *If Π' is a $(\tau, \epsilon_{cpa}, n, t)$ semantically secure and $(\tau, \epsilon_{vfy}, n, t)$ publicly verifiable broadcast encryption scheme such that ${}_nC_t \geq 2^k$, and Σ is a (τ, ϵ_{uf}) strongly unforgeable one-time signature scheme, then Π is a $(\tau - o(\tau), \epsilon_{cpa} + \epsilon_{vfy} + \frac{1}{2}\epsilon_{uf}, q_D)$ CCA-secure key encapsulation mechanism.*

A similar result can also be obtained from *privately* verifiable BE schemes.

7.2 Remarks

We notice that the above generic conversion is identical to the Canetti-Halevi-Katz (CHK) paradigm [12] except that the underlying primitive of CHK, i.e. IBE, is replaced with verifiable BE in our construction. Kiltz [24] also showed that IBE is not always necessary for CHK and a weaker primitive which is called

Table 2. Relation among broadcast encryption and public key encryption schemes. The column "(n, t)" denotes a possible and typical parameter setting for each underlying broadcast encryption scheme, and $\mathsf{poly}(k)$ and $\mathsf{exp}(k)$ denote polynomial and exponential functions for the security parameter k, respectively. For verifiability, related cryptographic tools are described, and $\surd$ means that the underlying broadcast encryption has verifiability as it is.

BE Scheme	(n, t)	Verifiability	$\Rightarrow$	PKE Scheme
Trivial BE	$(\mathsf{poly}(k), n/2)$	NIZK		DDN [16]
Naor-Pinkas [30]	$(\mathsf{exp}(k), 1)$	DDH		a variant of CS [14]
		GHDH	$\Rightarrow$	Kiltz [25]
		Sec. 3.2		Ours §4
IBE	$(\mathsf{exp}(k), n-1)$	$\surd$		CHK [12]
BGW [6]	$(\mathsf{poly}(k), n/2)$	$\surd$		Ours §7.3

tag-based encryption (TBE) [28] is sufficient, and demonstrated to construct a concrete TBE scheme without using IBE-related techniques. There are also other CCA-secure schemes whose security can be explained via the TBE framework, e.g. [14, 10, 25]. Our proposed method is a generic construction of TBE from BE with verifiability.

Many existing CCA-secure PKE schemes can be explained via our observation in Sec. 7.1 with different underlying BE schemes, and relations among existing BE and CCA-secure PKE schemes are summarized in Table 2. We give more detailed explanations for this in the full version of this paper [21].

7.3 Another New CCA-Secure KEM from Boneh-Gentry-Waters

Based on the proposed methodology, we can construct yet another new practical CCA-secure KEM from the BGW BE scheme [6]. This can be a further evidence that BE with verifiability is a powerful tool for constructing CCA-secure PKE. The proposed scheme yields tight security reduction to the 2ℓ-BDHE problem [6] for relatively small ℓ, short ciphertexts and short decryption keys. The concrete construction of the scheme is as follows: Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be multiplicative cyclic groups with prime order p, and $e : \mathbb{G}_1 \times \mathbb{G}_1 \to \mathbb{G}_2$ be a bilinear mapping [5]. $\mathbf{Setup}(1^k)$ chooses $\ell \in \mathbb{N}$ such that ${}_{2\ell}C_\ell \geq 2^k$, and picks a random generator $g \in \mathbb{G}_1$ and random $\alpha, \gamma \in \mathbb{Z}_p$. It also generates $g_1, ..., g_{4\ell}, v$, and Z where $g_i = g^{(\alpha^i)}$, $v = g^\gamma$, and $Z = e(g_{2\ell+1}, g)$. The decryption key is $dk = g^{\alpha^{2\ell+1}}$, and the public key is $PK = (g, g_1, ..., g_{2\ell}, g_{2\ell+2}, ..., g_{4\ell}, v, Z, \mathsf{TCR})$, where $\mathsf{TCR} : \mathbb{G}_1 \to \mathcal{P}$ is a target collision resistant hash function and $\mathcal{P} = \{\mathcal{S} | \mathcal{S} \subseteq \{1, ..., 2\ell\}, |\mathcal{S}| = \ell\}$. $\mathbf{Encrypt}(PK)$ picks a random $r \in \mathbb{Z}_p$, sets $K = Z^r \in \mathbb{G}_2$, computes $\mathcal{S} = \mathsf{TCR}(g^r)$, and outputs (ψ, K) where $\psi = (g^r, (v \cdot \prod_{j \in \mathcal{S}} g_{2\ell+1-j})^r) \in \mathbb{G}_1^2$. For ciphertext $\psi = (C_0, C_1)$, $\mathbf{Decrypt}(dk, \psi, PK)$ computes $\mathcal{S} = \mathsf{TCR}(C_0)$, and checks whether $e(g, C_1) \stackrel{?}{=} e(v \cdot \prod_{j \in \mathcal{S}} g_{2\ell+1-j}, C_0)$. It outputs "$\perp$" if it is invalid, or $K = e(dk, C_0)$ otherwise. Security of this scheme can be proven by a straightforward combination of the proofs of Theorem 2 of this paper and Theorem 3.1

of [6]. Unfortunately, this scheme is not very advantageous to other schemes, but it is still comparably efficient to other practical schemes (see Table 1).

7.4 A Generic Construction of CCA-Secure Broadcast Encryption

By using our methodology, it is also generically possible to construct a CCA-secure BE scheme from CPA-secure one with public verifiability. The conversion is fairly simple, and the resulting CCA-secure scheme can be practical. When applying this to the BGW BE scheme, we can have a new CCA-secure BE scheme with verifiability whose computational cost is slightly better than the previous scheme [6]. More detailed explanation is given in the full version of this paper [21].

Acknowledgement

The authors would like to thank Nuttapong Attrapadung, David Cash, Eike Kiltz and Takahiro Matsuda for their helpful comments and suggestions. The authors also would like to thank anonymous reviewers of Asiacrypt'08 for their invaluable comments.

References

1. M. Abe, R. Gennaro, K. Kurosawa, and V. Shoup, "Tag-KEM/DEM: a new framework for hybrid encryption and a new analysis of Kurosawa-Desmedt KEM," Proc. of Eurocrypt'05, pp.128-146, 2005.
2. J. Anzai, N. Matsuzaki, and T. Matsumoto, "A quick group key distribution scheme with "entity revocation"," Proc. of Asiacrypt'99, pp.333-347, 1999.
3. M. Bellare and C. Namprempre, "Authenticated encryption: relations among notions and analysis of the generic composition paradigm," Proc. of Asiacrypt'00, pp.531-545, 2000.
4. M. Blum, P. Feldman, and S. Micali, "Non-interactive zero-knowledge and its applications," Proc. of STOC'88, pp.103-112, 1988.
5. D. Boneh and M.K. Franklin, "Identity-based encryption from the Weil pairing," Proc. of Crypto'01, pp.213-229, 2001.
6. D. Boneh, C. Gentry, and B. Waters, "Collusion resistant broadcast encryption with short ciphertexts and private keys," Proc. of Crypto'05, pp.258-275, 2005.
7. D. Boneh and J. Katz, "Improved efficiency for CCA-secure cryptosystems built using identity-based encryption," Proc. of CT-RSA'05, pp.87-103, 2005.
8. D. Boneh and I. Shparlinski, "On the unpredictability of bits of the elliptic curve Diffie-Hellman scheme," Proc. of Crypto'01, pp.201-212, 2001.
9. D. Boneh and R. Venkatesan, "Hardness of computing the most significant bits of secret keys in Diffie-Hellman and related schemes," Proc. of Crypto'96, pp.129-142, 1996.
10. X. Boyen, Q. Mei, and B. Waters, "Direct chosen ciphertext security from identity-based techniques," Proc. of CCS'05, pp.320-329, 2005.
11. R. Canetti, "Universally composable security: a new paradigm for cryptographic protocols," Proc. of FOCS'01, pp.136-145, 2001.

12. R. Canetti, S. Halevi, and J. Katz, "Chosen-ciphertext security from identity-based encryption," Proc. of Eurocrypt'04, pp.207-222, 2004.
13. D. Cash, E. Kiltz, and V. Shoup, "The twin Diffie-Hellman problem and applications," Proc. of Eurocrypt'08, pp.127-145, 2008. The full version is available from IACR ePrint 2008/067.
14. R. Cramer and V. Shoup, "A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack," Proc. of Crypto'98, pp.13-25, 1998.
15. Y. Dodis and N. Fazio, "Public key trace and revoke scheme secure against adaptive chosen ciphertext attack," Proc. of PKC'03, pp.100-115, 2003.
16. D. Dolev, C. Dwork, and M. Naor, "Non-malleable cryptography," Proc. of STOC'91, pp. 542-552, 1991.
17. A. Fiat and M. Naor, "Broadcast encryption," Proc. of Crypto'93, pp.480-491, 1993.
18. O. Goldreich and L.A. Levin, "A hard-core predicate for all one-way functions," Proc. of STOC'89, pp.25-32, 1989.
19. S. Goldwasser and S. Micali, "Probabilistic encryption," J. Comput. Syst. Sci., 28(2), pp.270-299, 1984.
20. S. Halevi and P. Rogaway, "A tweakable enciphering mode," Proc. of Crypto'03, pp.482-499, 2003.
21. G. Hanaoka and K. Kurosawa, "Efficient chosen ciphertext secure public key encryption under the computational Diffie-Hellman assumption," IACR ePrint 2008/211, 2008. Full version of this paper.
22. D. Hofheinz and E. Kiltz, "Secure hybrid encryption from weakened key encapsulation," Proc. of Crypto'07, pp. 553-571, 2007.
23. E. Kiltz, "A primitive for proving the security of every bit and about universal hash functions & hard core bits," Proc. of FCT'01, pp.388-391, 2001.
24. E. Kiltz, "Chosen-ciphertext security from tag-based encryption," Proc. of TCC'06, pp.581-600, 2006.
25. E. Kiltz, "Chosen-ciphertext secure key-encapsulation based on gap hashed Diffie-Hellman," Proc. of PKC'07, pp.282-297, 2007.
26. K. Kurosawa and Y. Desmedt, "A new paradigm of hybrid encryption scheme," Proc. of Crypto'04, pp.426-442, 2004.
27. M. Luby and C. Rackoff, "How to construct pseudorandom permutations from pseudorandom functions," SIAM J. Comput., 17(2), pp.373-386, 1988.
28. P.D. MacKenzie, M.K. Reiter, and K. Yang, "Alternatives to non-malleability: Definitions, constructions, and applications," Proc. of TCC'04, pp. 171-190, 2004.
29. D. Naor, M. Naor, and J. Lotspiech, "Revocation and tracing schemes for stateless receivers," Proc. of Crypto'01, pp.41-62, 2001.
30. M. Naor and B. Pinkas, "Efficient trace and revoke schemes," Proc. of FC'00, pp. 1-20, 2000.
31. M. Naor and M. Yung, "Public-key cryptosystems provably secure against chosen ciphertext attacks," Proc. of STOC'90, pp.427-437, 1990.
32. R. Pass, a. shelat, and V. Vaikuntanathan, "Construction of a non-malleable encryption scheme from any semantically secure one," Proc. of Crypto'06, pp.271-289, 2006.
33. D.H. Phan and D. Pointcheval, "About the security of ciphers (semantic security and pseudo-random permutations)," Proc. of SAC'04, pp.182-197, 2004.
34. N. Pippenger, "On the evaluation of powers and related problems," Proc. of FOCS'76, pp.258-263, 1976.

35. C. Rackoff and D.R. Simon, "Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack," Proc. of Crypto'91, pp.433-444, 1991.
36. A. Sahai, "Non-malleable non-interactive zero knowledge and adaptive chosen-ciphertext security," Proc. of FOCS'99, pp.543-553, 1999.
37. A. Shamir, "Identity-based cryptosystems and signature schemes," Proc. of Crypto'84, pp.47-53, 1985.
38. V. Shoup, "Using hash functions as a hedge against chosen ciphertext attack," Proc. of Eurocrypt'00, pp.275-288, 2000.