A preliminary version of this paper appears in Topics in Cryptology — CT-RSA '08, Lecture Notes in Computer Science Vol. ??, T. Malkin ed., Springer-Verlag, 2008. This is the full version.

CCA2 Secure IBE: Standard Model Efficiency through Authenticated Symmetric Encryption

Eike Kiltz[1] Yevgeniy Vahlis[2]

[1] Cryptology and Information Security Theme,
CWI Amsterdam, The Netherlands
kiltz@cwi.nl

[2] Computer Science Department,
University of Toronto, Canada
evahlis@cs.toronto.edu

Abstract

We propose two constructions of chosen-ciphertext secure identity-based encryption (IBE) schemes. Our schemes have a security proof in the standard model, yet they offer performance competitive with all known random-oracle based schemes. The efficiency improvement is obtained by combining modifications of the IBE schemes by Waters [41] and Gentry [23] with authenticated symmetric encryption.

Keywords: Chosen-ciphertext security, Identity-Based Encryption, Bilinear Maps.

Contents

1 Introduction

An Identity-Based Encryption (IBE) scheme is a public-key encryption scheme where any string is a valid public key. In particular, email addresses and dates can be public keys. The ability to use identities as public keys avoids the need to distribute public key certificates — which is one of the main technical difficulties when setting up a public-key infrastructure. An efficient construction of an IBE was not given until almost two decades after Shamir posed the initial open question in [38] regarding the existence of such cryptographic primitives. The first efficient IBEs appeared in 2001, given separately by Boneh and Franklin [11, 12], and Sakai et al. [36]. In particular, Boneh and Franklin [11, 12] proposed formal security notions for IBE systems and designed a fully functional secure IBE scheme using bilinear maps. This scheme and the tools developed in its design have been successfully applied in numerous cryptographic settings, transcending by far the identity based cryptography framework.

Despite its only recent invention, IBE is already used extensively in practice. Two companies — Voltage security (`http://voltage.com`) and Identum (`http://www.identum.com/`) —are specialized in identity-based security solutions. This is one of the reasons why IBE is currently in the process of getting standardized — the new IEEE P1363.3 standard for "Identity-Based Cryptographic Techniques using Pairings" is currently in preparation [27]. The schemes that are currently in consideration are the one by Boneh and Franklin [12]; the one by Boneh and Boyen [8, 14]; and the one by Kasahara and Sakai [36, 18].

All the above IBE schemes provide security against *chosen-ciphertext attacks*. In a chosen ciphertext attack [34, 12], the adversary is given access to a decryption oracle that allows to obtain the decryptions of ciphertexts of his choosing. Intuitively, security in this setting means that an adversary obtains (effectively) no information about encrypted messages, provided the corresponding ciphertexts are never submitted to the decryption oracle. Since the dramatic attack by Bleichenbacher [6], the notion of chosen-ciphertext security is commonly agreed as the "right" notion of security for encryption schemes [40]. We stress that, in general, chosen-ciphertext security is a much stronger security requirement than semantic security, where in the latter an attacker is not given access to the decryption oracle.

Random Oracles. The drawback of all the IBE schemes [12, 8, 36, 18] that are currently under submission to the new IEEE P1363.3 standard is that their security can only be guaranteed in the *random oracle* model [4], i.e. in an idealized world where all parties get black-box access to a truly random function. Unfortunately a proof in the random oracle model can only serve as a heuristic argument and, admittedly using quite contrived constructions, has been shown to possibly lead to insecure schemes when the random oracles are implemented in the standard model (see, e.g., [16]). More importantly, there exist results [22] indicating that even certain standardized cryptographic schemes (such as full-domain hash signatures) will always remain in the grey area of schemes having a proof in the random oracle yet are "provably unprovable" in the standard model.

IBE Without Random Oracles. Waters [41] presents the first practical IBE that is chosen-plaintext secure without random oracles. It fits the category of "commutative-blinding" IBE schemes from Boneh and Boyen [8] and its chosen-plaintext security can be reduced to the Bilinear Decisional Diffie-Hellman (BDDH) assumption. Based on Waters scheme several chosen-ciphertext secure IBE schemes were proposed starting with generic constructions [10] whose specific instantiations were later improved [15, 30]. Today's most efficient variant is due to Kiltz and Galindo who successfully applied "direct chosen-ciphertext" techniques from [15, 29] to Waters' IBE scheme. More recently, Gentry [23] proposed yet another practical chosen-ciphertext secure IBE scheme based on the class of "inversion-based" IBE schemes from [8], offering interesting efficiency trade-offs compared to the commutative-blinding schemes [30].

Random Oracles: Theory vs. Practice. The above mentioned drawbacks of the random oracle model readily leads to the question why random-oracle based schemes are sometimes chosen over schemes with a rigorous proof in the standard model. The answer is straight-forward: it is common knowledge that schemes in the random-oracle model are usually much more efficient than schemes in the standard model. As long as the "theoretical problems" from [16, 22] do not lead to an actual break of a non-artificial scheme, using random-oracle schemes seems justifiable in practice. On the other hand it is in the belief of the authors that this general perception about random oracles will change when alternative random-oracle free schemes become available that offer competitive performance.

	Size (bits)		Cost (relative)	
Scheme	Ciphertext	Public Key	Encrypt	Decrypt
Standard model				
Ours: IBE_1 (§4)	422	2376	39	216
Ours: IBE_2 (§5)	1277	2223	110	222
KG [30]	513	2565	40	360
Gentry [23]	2223	3249	146	408
Random Oracle model				
BF [11]	331	171	187	151
BB_1 [8]	502	1386	39	217
KS [18]	331	171	38	152

Figure 1: Efficiency comparison for CCA-secure IBE schemes in the standard/random oracle model for MNT/80-bit security level. Timings are relative to one exponentiation in group $\mathbb{G}$.

1.1 Our contributions

In this paper we demonstrate that there exist identity-based encryption schemes that are provably secure in the standard model, yet their performance is competitive with the best schemes in the random oracle model. We propose two constructions of chosen-ciphertext secure IBE schemes which outperform all such existing standard-model schemes, and have performance comparable to the random-oracle based schemes that were described above.

Scheme I. Our first IBE scheme is based on Waters' semantically secure IBE. Our approach to protecting a ciphertext against chosen ciphertext attacks bears some resemblance to the one used by Cramer and Shoup [20, 21] to obtain chosen ciphertext secure public key encryption. More precisely, we use the more efficient "encrypt-then-mac" or "authenticated symmetric encryption" variant proposed by Kurosawa and Desmedt [31]. More precisely, in our construction decryption of ill-formed ciphertexts (i.e. ciphertexts that could not have been generated by the encryption algorithm) uses randomness which is built into the user private key (and is independent of the master public key). Such ill-Formed ciphertexts can be detected using extra-information that is algebraically encoded into the "identity-carrying" part of the ciphertext (similar to the HIBE construction from [9]). Overall this allows us to obtain a CCA secure IBE scheme by only adding *one exponentiation* to the encryption/decryption algorithm of Waters' scheme, which is secure only against chosen plaintext attacks. We give a standard-model security proof reducing the intractability of the *modified Bilinear Decisional Diffie-Hellman* (mBDDH) problem (a problem closely related to BDDH) to breaking the CCA security of our scheme.

Scheme II. Our second construction is a variant of Gentry's chosen-ciphertext secure IBE scheme. Here our new contribution is to use authenticated symmetric encryption [31, 25] to reduce ciphertext expansion and encryption/decryption cost compared to Gentry's original schemes. We prove chosen-ciphertext security of our scheme with respect to the decisional augmented bilinear Diffie-Hellman exponent (q-ABDHE) assumption [23] in the standard model. We remark that the proof technique is significantly different from the one used for the first scheme.

1.2 Comparison

We carefully review all known chosen-ciphertext secure IBE constructions and make an extensive comparison with our schemes. Our studies also incorporate all relevant practical issues when making a comparison, including the tightness of the security reduction with respect to different assumptions and instantiating the schemes in asymmetric pairing groups. To obtain concrete comparison values we estimate ciphertext expansion and encryption/decryption cost when implemented in different pairing groups using recent (independent) timing data from [14]. This includes pairing groups based on super-singular curves and MNT curves.

The numerical results of our comparison for 80 bits MNT curves are given in Figure 1 (For 80

bits super-singular curves the results are similar. We refer the reader to Figure 6 in Section 6.) The figure shows that our schemes outperform all known IBE schemes in the standard model. Most notably, compared to the standard-model scheme KG from [30] decryption cost and ciphertext expansion is reduced by approximately one third, whereas encryption cost is the same. More importantly, in comparison with the random-oracle based schemes BF from [12], BB_1 from [8, 14], and KS from [36, 18] our schemes offer competitive performance in all parameters, yet are provably secure in the standard model.

1.3 Related work

A special class of authenticated symmetric encryption schemes which is obtained using the "encrypt-then-mac" primitive was recently successfully applied to public-key encryption schemes by Kurosawa and Desmedt [31, 2] who greatly improved efficiency of the original Cramer-Shoup encryption scheme [21]. Their result was generalized to cover arbitrary authenticated encryption schemes [25]. In fact, our second IBE scheme can be seen as the "Kurosawa-Desmedt variant" of the original CCA secure scheme by Gentry. A variant of it was also sketeched in independent work by Boneh, Gentry and Hamburg [7] using their general framework of "hash proof systems". In connection with IBE, authenticated encryption was first used in [37]. This paper is an extended version of an unpublished manuscript [28] by the first author.

2 Preliminaries

2.1 Notation

If x is a string, then $|x|$ denotes its length, while if S is a set then $|S|$ denotes its size. If $k \in \mathbb{N}$ then 1^k denotes the string of k ones. If S is a set then $s \leftarrow_{\mathrm{R}} S$ denotes the operation of picking an element s of S uniformly at random. Unless otherwise indicated, algorithms are randomized and polynomial time. By $z \leftarrow_{\mathrm{R}} \mathsf{A}^{\mathcal{O}_1,\mathcal{O}_2,\cdots}(x, y, \ldots)$ we denote the operation of running algorithm A with inputs $x, y, \ldots$ and access to oracles $\mathcal{O}_1, \mathcal{O}_2, \ldots$, and letting z be the output. An adversary is an algorithm or a tuple of algorithms.

2.2 Identity Based Encryption

An IBE scheme consists of four algorithms: Setup, KeyGen, Enc, and Dec. Setup generates the global public and private keys; KeyGen uses the global private key to generate an individual private key PRI_{id} for a given identity; Enc uses the global public key to encrypt a message to a given identity; and Dec uses the individual private key to decrypt ciphertexts.

The strongest and commonly accepted notion of security for an identity-based key encryption is that of indistinguishability against an adaptive chosen ciphertext attack [12]. This notion, denoted IND-ID-CCA (or simply CCA), is captured by defining the following advantage function for an adversary $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$, and for an IBE scheme IBE:

$$\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(k) \;=\; \left|\Pr[\mathsf{Exp}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(k) = 1] - 1/2\right|$$

where $\mathsf{Exp}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(k)$ is defined by the following experiment.

Experiment $\mathsf{Exp}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(k)$
$(\mathsf{PUB}, \mathsf{PRI}) \leftarrow_{\mathrm{R}} \mathsf{Setup}(1^k)$
$(id^*, m_0, m_1, St) \leftarrow_{\mathrm{R}} \mathsf{A}_1^{\mathsf{KeyGen}(\cdot),\mathsf{Dec}(\cdot,\cdot)}(\mathsf{PUB})$
$b \leftarrow_{\mathrm{R}} \{0,1\}; \quad C^* \leftarrow_{\mathrm{R}} \mathsf{Enc}(\mathsf{PUB}, id^*, m_b)$
$b' \leftarrow_{\mathrm{R}} \mathsf{A}_2^{\mathsf{KeyGen}(\cdot),\mathsf{Dec}(\cdot,\cdot)}(C^*, St)$
If $b = b'$ Return 1 else return 0

The oracle $\mathsf{KeyGen}(\cdot)$ on input id generates a new private key for the identity id and returns it. The oracle $\mathsf{Dec}(\cdot,\cdot)$ on input id and C first generates a new private key for id and then uses it to decrypt C. When A_1 outputs id^* it must not be any of the identities that the adversary queried to the $\mathsf{KeyGen}(\cdot)$ oracle. Furthermore, A_2 is not allowed to query the $\textsc{KeyGen}(\cdot)$ oracle on id^*, and is not allowed to query

the $\mathsf{Dec}(\cdot,\cdot)$ oracle on (id^*, C^*). The variable St represents some internal state information of adversary A and can be any (polynomially bounded) string.

Definition 2.1 An IBE scheme IBE is *secure against chosen-ciphertext attacks* (CCA secure) if for all adversaries A the advantage function $\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(\cdot)$ is negligible.

For a more precise analysis of the tightness of reduction we will sometimes use the following more detailed notation. For integers k, t, q_x, q_d, $\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE},t,q_x,q_d}(k) = \max_{\mathsf{A}} \mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE},\mathsf{A}}(k)$, where the maximum is over all adversaries A that make at most t computational steps, q_x key-derivation, and q_d decryption queries. Here we make the convention to count all decryption queries for $id \neq id^*$ as a key-derivation query.

2.3 Symmetric Encryption

A symmetric encryption scheme $\mathsf{SE} = (\mathsf{E}, \mathsf{D})$ is specified by its encryption algorithm E (encrypting $m \in MsgSp(k)$ with keys $K \in \mathcal{K}(k)$) and decryption algorithm D (returning $m \in MsgSp(k)$ or $\perp$). Here we restrict ourselves to deterministic algorithms E and D.

The most common notion of security for symmetric encryption is that of ciphertext indistinguishability, which requires that all efficient adversaries fail to distinguish between the encryptions of two messages of their choice. Another common security requirement is *ciphertext authenticity*. Ciphertext authenticity requires that no efficient adversary can produce a new valid ciphertext under some key when given one encryption of a message of his choice under the same key. A symmetric encryption scheme which satisfies *both* requirements simultaneously is called secure in the sense of authenticated encryption (AE-OT secure). Note that AE-OT security is a stronger notion than chosen-ciphertext security.

The above requirements are formalized as follows:

Ciphertext Indistinguishability. Let $\mathsf{SE} = (\mathsf{E}, \mathsf{D})$ be a symmetric encryption scheme, and let $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ be an adversary. We define the following experiment:

$$
\begin{array}{l}
\text{Experiment } \mathsf{Exp}^{\mathsf{IND}}_{\mathsf{SE},\mathsf{A}}(k) \\
\quad K \leftarrow_{\mathrm{R}} \mathcal{K}(k) \\
\quad (m_0, m_1, St) \leftarrow_{\mathrm{R}} \mathsf{A}_1(1^k) \\
\quad b \leftarrow_{\mathrm{R}} \{0,1\}\,;\; c^* \leftarrow_{\mathrm{R}} \mathsf{E}_K(m_b) \\
\quad b' \leftarrow_{\mathrm{R}} \mathsf{A}_2(1^k, St, c^*) \\
\quad \text{If } b = b' \text{ Return 1 else return 0}
\end{array}
$$

The advantage of A in breaking the ciphertext indistinguishability security of SE is:

$$\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\mathsf{A}}(k) \stackrel{\text{def}}{=} \left|\Pr[\mathsf{Exp}^{\mathsf{IND}}_{\mathsf{SE},\mathsf{A}}(k) = 1] - 1/2\right|$$

Definition 2.2 The symmetric encryption scheme SE has *indistinguishable ciphertexts* if for every adversary A the advantage $\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},A}(\cdot)$ is negligible.

Ciphertext Authenticity. In this work we are only interested in one-time authenticated schemes. That is, schemes for which no efficient adversary can produce a new valid ciphertext after seeing the encryption of a single message.

Let $\mathsf{SE} = (\mathsf{E}, \mathsf{D})$ be a symmetric encryption scheme, and let $\mathsf{A} = (\mathsf{A}_1, \mathsf{A}_2)$ be an algorithm. We define the following experiment:

$$
\begin{array}{l}
\text{Experiment } \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A}}(k) \\
\quad K \leftarrow_{\mathrm{R}} \mathcal{K}(k) \\
\quad (m, St) \leftarrow_{\mathrm{R}} \mathsf{A}_1(1^k) \\
\quad c \leftarrow \mathsf{E}_K(m) \\
\quad c' \leftarrow_{\mathrm{R}} \mathsf{A}_2(1^k, St, c) \\
\quad \text{If } c' \neq c \text{ and } \mathsf{D}_K(C') \neq \perp \text{ return 1 else return 0}
\end{array}
$$

The advantage of A in breaking the ciphertext integrity of SE is:

$$\mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A}}(k) \stackrel{\text{def}}{=} \Pr[\mathsf{Exp}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A}}(k) = 1]$$

Definition 2.3 The symmetric encryption scheme SE has *ciphertext integrity* if for every adversary A the advantage $\mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A}}(\cdot)$ is negligible.

Authenticated Encryption. A symmetric encryption scheme which is secure according to both Definition 2.2 and Definition 2.3 is secure in the sense of *one time authenticated encryption* (or AE-OT).

Constructions. In our IBE constructions we will require an abstract notion of algebraic symmetric encryption where the key-space $\mathcal{K}$ consists of a cyclic group $\mathbb{G}_T$. It is well know (following the encrypt-then-mac approach [3, 21]) how to build such symmetric encryption schemes 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 [21] that maps elements from $\mathbb{G}_T$ into $2k$-bit strings (such as SHA-1).

We remark that for our purposes it is also possible to use a more efficient single-pass authenticated encryption scheme (see, e.g., [35]). In both cases the the ciphertext expansion (i.e., ciphertext size minus plaintext size) of the AE-OT secure symmetric scheme is only k (security parameter) bits which is optimal with respect to our security notion.

2.4 Target Collision Resistant Hashing

$TCR = (TCR_k)_{k\in\mathbb{N}}$ is a family of keyed hash functions $TCR^s_k : \mathbb{G} \to \mathbb{Z}_p$ for each k-bit key s. It is assumed to be target collision resistant (TCR) [21], which is captured by defining the advantage function $\mathsf{Adv}^{\mathsf{TCR}}_{TCR,\mathsf{B}}(k)$ of an adversary B as

$$\Pr\left[TCR^s(c^*) = TCR^s(c) \wedge c \neq c^* : \begin{array}{l} s \leftarrow_{\mathrm{R}} \{0,1\}^k ;\ c^* \leftarrow_{\mathrm{R}} \mathbb{G} \\ c \leftarrow_{\mathrm{R}} \mathsf{B}(s, c^*) \end{array} \right],$$

Note that TCR is a weaker requirement than collision-resistance, so that, in particular, any practical collision-resistant function can be used. Commonly [21, 31] this function is implemented using a dedicated (non-keyed) cryptographic hash function like MD5 or SHA. To simplify notation we will sometimes drop the superscript key s and simply use TCR.

3 Intractability assumptions

3.1 Bilinear Groups

Our schemes will be parameterized by a *pairing parameter generator*. This is an algorithm $\mathcal{G}$ that on input 1^k returns the description of an multiplicative cyclic group $\mathbb{G}$ of prime order p, where $2^k < p < 2^{k+1}$, the description of a multiplicative cyclic group $\mathbb{G}_T$ of the same order, and a non-degenerate bilinear pairing $\hat{e} : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$. See [12] for a description of the properties of such pairings. We use $\mathbb{G}^*$ to denote $\mathbb{G} \setminus \{1\}$, i.e. the set of all group elements except the neutral element. Throughout the paper we use $\mathbb{PG} = (\mathbb{G}, \mathbb{G}_T, p, \hat{e}, g, g_T)$ as shorthand for the description of bilinear groups, where g is a generator of $\mathbb{G}$ and $g_T = \hat{e}(g, g) \in \mathbb{G}_T$.

3.2 The modified BDDH assumption

Let $\mathbb{PG}$ be the description of pairing groups. The Bilinear Decisional Diffie-Hellman (BDDH) assumption [12] states that the two distributions $(g^x, g^y, g^z, \hat{e}(g,g)^{xyz})$ and $(g^x, g^y, g^z, \hat{e}(g,g)^r)$, for $x, y, z, r \leftarrow_{\mathrm{R}} \mathbb{Z}_p$ are indistinguishable for any adversary. For the modified BDDH assumption we furthermore provide

the adversary with the element $g^{(y^2)}$. More formally we define the advantage function $\mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\mathsf{B}}(k)$ of an adversary B as

$$\left| \begin{array}{l} \Pr[\mathsf{B}(\mathbb{PG}, g^x, g^y, g^{y^2}, g^z, \hat{e}(g,g)^{xyz}) = 1] \\ \quad - \Pr[\mathsf{B}(\mathbb{PG}, g^x, g^y, g^{y^2}, g^z, \hat{e}(g,g)^{r}) = 1] \end{array} \right| ,$$

where $x, y, z, r \leftarrow_{\mathrm{R}} \mathbb{Z}_p$ and $\mathbb{PG} \leftarrow_{\mathrm{R}} \mathcal{G}(1^k)$. We say that the *modified Bilinear Decision Diffie-Hellman (mBDDH) assumption relative to generator* $\mathcal{G}$ holds if $\mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\mathsf{B}}(\cdot)$ is negligible for all adversaries B.

3.3 The truncated q-ABDHE assumption

Let $q = q(k)$ be a polynomial. The q-BDDHI assumption [8] states that the two distributions $(g^x, \ldots, g^{x^q}, \hat{e}(g,g)^{1/x})$ and $(g^x, \ldots, g^{x^q}, \hat{e}(g,g)^r)$, for $x, r \leftarrow_{\mathrm{R}} \mathbb{Z}_p$ are indistinguishable for any adversary. In [23] Gentry proposed the related truncated decisional augmented bilinear Diffie-Hellman exponent (truncated q-ABDHE) assumption which augments the q-BDDHI assumption with additional information to the adversary. We define the advantage function $\mathsf{Adv}^{\text{q-abdhe}}_{\mathcal{G},\mathsf{B}}(k)$ of an adversary B as

$$\left| \begin{array}{l} \Pr[\mathsf{B}(\mathbb{PG}, g^x, \ldots, g^{x^q}, g^z, g^{zx^{q+2}}, \hat{e}(g,g)^{zx^{q+1}}) = 1] \\ \quad - \Pr[\mathsf{B}(\mathbb{PG}, g^x, \ldots, g^{x^q}, g^z, g^{zx^{q+2}}, \hat{e}(g,g)^{r}) = 1] \end{array} \right| ,$$

where $x, z, r \leftarrow_{\mathrm{R}} \mathbb{Z}_p$ and $\mathbb{PG} \leftarrow_{\mathrm{R}} \mathcal{G}(1^k)$. We say that the *truncated q-ABDHE assumption relative to generator* $\mathcal{G}$ holds if $\mathsf{Adv}^{\text{q-abdhe}}_{\mathcal{G},\mathsf{B}}(\cdot)$ is negligible for all B.

3.4 Relations

The next lemma classifies the strength of the modified BDDH assumption we introduced between the well known *standard pairing-based assumptions* BDDH and 2-BDDHI. Here "A $\leq$ B" means that assumption B implies assumption A (in a generic sense), i.e. assumption B is a stronger assumption than A.

Lemma 3.1 BDDH $\leq$ mBDDH $\leq$ 2-BDDHI $\leq \ldots \leq$ q-BDDHI $\leq$ truncated q-ABDHE

The simple proof is postponed until Appendix B.3. We remark that the complexity of q-BDDHI (as well as truncated q-ABDHE) in the in the generic-group model [39] is roughly $\Omega(\sqrt{p/q})$ [8, 23] which matches the recent attack due to Cheon [19].

4 IBE Scheme I

In this section we present our first CCA secure IBE scheme. It is based on the Boneh-Boyen "commutative-blinding" IBE scheme [8] in its full-identity secure variant of Waters [41] which is chosen-plaintext secure. We construct a CCA secure IBE by adding a redundant group element to the ciphertext, and authenticating the two group elements both explicitly, using target collision resistant hash function, and implicitly by using the same randomness to generate both elements.

A similar technique was already used by Cramer and Shoup to obtain chosen-ciphertext secure public-key encryption and later also successfully applied in [15, 29, 30]. All the above works make a distinction between ciphertexts that can be generated by the encryption algorithm (well-formed ciphertexts), and strings that the encryption algorithm would never output (ill-formed ciphertexts) in their security analysis. The first CCA secure IBE that applies this methodology is [30]. The IBE of [30] handles ill-formed ciphertexts by decrypting them to a fresh random value chosen by the decryption algorithm ("implicit rejection"). This approach is sufficient for obtaining CCA security, but is prohibitively expensive as it requires the decryption algorithm to be randomized, and to compute several exponentiations of group elements to handle ill-formed ciphertexts.

We avoid this additional computation by exploiting the fact that in our IBE the decryption of an ill-formed ciphertext depends on the randomness of the private key that was used for the decryption. In other words, we decrypt ill-formed ciphertexts in the same way as we would decrypt well-formed ciphertext, but for a well formed ciphertext the outcome of the decryption is independent of the randomness in

$\mathsf{Setup}(1^k)$	$\mathsf{KeyGen}(\mathsf{PRI}, id)$
$\alpha, u \leftarrow_{\mathrm{R}} \mathbb{G};\ z \leftarrow \hat{e}(g,\alpha);\ H \leftarrow_{\mathrm{R}} \mathsf{HGen}(\mathbb{G}; n)$	$s \leftarrow_{\mathrm{R}} \mathbb{Z}_p$
$\mathsf{PUB} \leftarrow (H, u, z) \in \mathbb{G}^{n+1} \times \mathbb{G} \times \mathbb{G}_T$	$\mathsf{PRI}_{id} \leftarrow (\alpha \cdot H(id)^s, g^{-s}, u^s) \in \mathbb{G}^3$
$\mathsf{PRI} \leftarrow \alpha \in \mathbb{G}$	Return PRI_{id}
Return $(\mathsf{PUB}, \mathsf{PRI})$	
$\mathsf{Enc}(\mathsf{PUB}, id, m)$	$\mathsf{Dec}(\mathsf{PUB}, id, \mathsf{PRI}_{id}, C)$
$r \leftarrow_{\mathrm{R}} \mathbb{Z}_p;\ c_1 \leftarrow g^r$	Parse C as $(c_1, c_2, c_3) \in \mathbb{G} \times \mathbb{G}^* \times \{0,1\}^*$
$t \leftarrow TCR(c_1);\ c_2 \leftarrow (H(id) \cdot u^t)^r$	Parse PRI_{id} as $(d_1, d_2, d_3) \in \mathbb{G}^3$
$K \leftarrow z^r \in \mathbb{G}_T;\ c_3 \leftarrow \mathsf{E}_K(m)$	$t \leftarrow TCR(c_1);\ K \leftarrow \hat{e}(c_1, d_1 \cdot d_3^t) \cdot \hat{e}(c_2, d_2)$
Return ciphertext $C = (c_1, c_2, c_3)$	Return $m \leftarrow \mathsf{D}_K(c_3)$

Figure 2: Our first CCA-secure IBE scheme IBE_1.

the private key. As a result our decryption algorithm is deterministic and significantly faster than [30]. Furthermore, our scheme also has one group element less in the ciphertext than [30]. This is achieved by algebraically integrating the implicit ciphertext consistency check into the part of the ciphertext that carries the information about the recipient's identity.

4.1 The IBE Construction

We assume that $\mathbb{PG} = (\mathbb{G}, \mathbb{G}_T, p, \hat{e}, g, g_T)$ are public system parameters obtained by running the group parameter algorithm $\mathcal{G}(1^k)$ (that may be shared among multiple systems).

We review the hash function $H : \{0,1\}^n \to \mathbb{G}$ used in Waters' identity based encryption schemes [41]. On input of $\mathbb{G}$ and an integer n, the randomized hash key generator $\mathsf{HGen}(\mathbb{G}; n)$ chooses $n+1$ random group elements $h_0, \ldots, h_n \in \mathbb{G}$ and returns $h = (h_0, h_1, \ldots, h_n) \in \mathbb{G}^{n+1}$ as the public description of the hash function. The algebraic hash function $H : \{0,1\}^n \to \mathbb{G}$ is evaluated on a string $id = (id_1, \ldots, id_n) \in \{0,1\}^n$ as the product

$$H(id) = h_0 \prod_{i=1}^{n} h_i^{id_i} \in \mathbb{G}.$$

Let $TCR : \mathbb{G} \to$ be a target collision-resistant hash function and $\mathsf{SE} = (\mathsf{E}, \mathsf{D})$ be a symmetric encryption scheme with key-space $\mathcal{K} = \mathbb{G}_T$. Our IBE scheme IBE_1 with identity space $IDSp = \{0,1\}^n$ is described in Figure 2. Here it is understood that decryption rejects if the ciphertext C does not parse to (c_1, c_2, c_3) with $c_1 \in \mathbb{G}$ and $c_2 \in \mathbb{G}^*$. An IBE scheme with arbitrary identity space $IDSp = \{0,1\}^*$ can be obtained by applying a collision-resistant hash function to the identities. (The choice of $n = 2k$ is due to the birthday paradox.)

We now show correctness of the scheme, i.e. that the symmetric key K computed in the encryption algorithm matches the key K computed in the decryption algorithm.[1] A correctly generated secret key for identity id has the form $\mathsf{PRI}_{id} = (d_1, d_2, d_3) = (\alpha \cdot H(id)^s, g^{-s}, u^s)$ for some $s \in \mathbb{Z}_p$. Therefore the decryption algorithm computes the symmetric key K as

$$\begin{aligned}
K &= \hat{e}(c_1, d_1 \cdot d_3^t) \cdot \hat{e}(c_2, d_2) \\
&= \hat{e}(g^r, \alpha \cdot H(id)^s \cdot (u^s)^t) \cdot \hat{e}((H(id) \cdot u^t)^r, g^{-s}) \\
&= \hat{e}(g^r, \alpha) \cdot \hat{e}(g^r, H(id)^s \cdot (u^s)^t) \cdot \hat{e}((H(id) \cdot u^t)^r, g^{-s}) \\
&= z^r \cdot \hat{e}(g^r, (H(id) \cdot u^t)^s) \cdot \hat{e}((H(id) \cdot u^t)^{-s}, g^r) \\
&= z^r,
\end{aligned}$$

[1] Decryption rejects all ciphertexts with $c_2 = 1 \in \mathbb{G}$. We can assume that encryption does not generate ciphertexts with $c_2 = 1$. In case it does encryption can pick fresh randomness r. Note that this implies that our encryption algorithm runs in expected polynomial time (rather then in strict polytime). However, this has little significance in practice since it is unlikely that for two independent choices of r we would get $c_2 = 1$.

which is the same as the key computed in the encryption algorithm. Now correctness of the scheme is implied by correctness of SE.

4.2 Security

Theorem 4.1 Assume *TCR* is a target collision resistant hash function and (E, D) is a AE-OT-secure symmetric scheme. Under the modified Bilinear Decisional Diffie-Hellman (mBDDH) assumption relative to generator $\mathcal{G}$, the IBE scheme IBE_1 is CCA secure. In particular, for $\varepsilon(k) = \mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE}_1,t,q_x,q_d}(k)$ and $\tilde{\varepsilon}(k) = \mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)$ we have

$$\begin{aligned} \varepsilon(k) &\leq (\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\tilde{t}}(k) + \tilde{\varepsilon}(k)) \cdot 10nq + \mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k) + q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},t}(k) + 2q_d^2/p; \\ t &\geq \tilde{t} - \mathcal{O}(\tilde{\varepsilon}^{-2}(k) \cdot \ln(\tilde{\varepsilon}^{-1}(k)) + q_d + q_x) \end{aligned}$$

The full proof is given in Appendix A.1. We give a brief overview here. Our proof for this system has many similarities with [30] (which in turn is based on [41]). The key difference between the two proofs is the treatment of ill-formed ciphertexts. [30] use the fact that anyone that has the global public key can check whether a ciphertext is well-formed. Then, if the ciphertext is ill-formed the decryption algorithm chooses a random value for K, and uses it to attempt and decrypt the symmetric ciphertext. Thus, the adversary himself could have decrypted any ill-formed ciphertext, and does not gain any information from querying the decryption oracle on such ciphertexts.

Our approach to dealing with ill-formed ciphertexts is different. We do not rely on the ability of anyone who has the global public key to check whether a ciphertext is well-formed. Instead, we make the observation that an ill-formed ciphertext, i.e. a ciphertext of the form $C = (g^r, (H(id) \cdot u^t)^{r'}, c_3)$, where $r \neq r'$, decrypts in the following way:

1. The intermediate key K is computed: $K = z^r \cdot \hat{e}(g, H(id) \cdot u^t)^{(r-r')s}$, where s is the random value that was used to generate the private key.

2. K is used to attempt and decrypt the AE ciphertext.

Now, the adversary makes a polynomial number of decryption queries with ill-formed ciphertexts. We show that the first such query is likely to decrypt as "reject", and each query after the first is likely to decrypt as "reject" given that all previous ill-formed queries decrypted as reject, which completes the proof. The idea is that the value s remains random in the view of the adversary as he makes decryption queries with valid ciphertexts, or ciphertexts that decrypt as "reject". Since s is random, K is also a random element of $\mathbb{G}_T$. Thus, by the authenticity property of the AE encryption, c_3 will be decrypted to "reject" when the random element K is used as the key.

4.3 Extensions

4.3.1 Trading Public Key Size and Security Reduction

As independently discovered in [17, 33], there exists an interesting trade-off between key-size of Waters' hash H and the security reduction of the IBE schemes. The construction modifies Waters hash H as follows: Let the integer $l = l(k)$ be a new parameter of the scheme. In particular, we represent an identity $id \in \{0,1\}^n$ as an n/l-dimensional vector $id = (id_1, \ldots, id_{n/l})$, where each id_i is an l bit string. Waters hash is then redefined to $H : \{0,1\}^n \to \mathbb{G}$, with $H(id) = h_0 \prod_{i=1}^{n/l} h_i^{id_i}$ for random public elements $h_0, h_1, \ldots, h_{n/l} \in \mathbb{G}$. Waters' original hash function is obtained as the special case $l = 1$. It is easy to see that using this modification in our IBE scheme (i) reduces the size of the public key from $n+2$ to $n/l+2$ elements in $\mathbb{G}$, whereas (ii) it adds another multiplicative factor of 2^l to the security reduction of the IBE scheme (Theorem 4.1).

For concreteness we propose the following value for l (our choice will become clear in Section 6). For a scheme implemented in groups offering 80 bits of security we have $n = 2 \cdot 80 = 160$ bits and use 128. This shrinks the public-key size to reasonable $n/l + 2 \approx 10$ elements in $\mathbb{G}$ (plus one element in $\mathbb{G}_T$).

We further remark that in the random-oracle model we can replace Waters' hash $H : \{0,1\}^* \to \mathbb{G}$ with $H(id) = h_0 \cdot h_1^{RO(id)}$, where $RO : \{0,1\}^* \to \mathbb{Z}_p$ is a cryptographic hash function which is modeled as a random oracle [4] in the security analysis.

4.3.2 Hierarchical Identities

Hierarchical identity-based encryption (HIBE) is a generalization of IBE to identities supporting hierarchical structures [26]. In a HIBE, identities are hierarchical and take the form $id = [id_1.id_2.id_3]$. This particular hierarchical identity has depth 3, and is subordinate to $[id_1]$, $[id_1.id_2]$, but not to $[id_1.id_2.id_3']$. Each user in the hierarchy may act as a local key-generation authority for all subordinate hierarchical identities.

By the relation to Waters IBE scheme it is easy to see that our technique can also be used to obtain a chosen-ciphertext secure HIBE. Using a technique from [9] it is furthermore possible to reduce the HIBE ciphertext size to three elements, i.e. it is independent of the hierarchy's depth. To be more precise, the IBE from Section 4.1 is modified to a HIBE supporting maximal d hierarchies as follows. The setup algorithm chooses d different and independent hash functions $H_i \leftarrow_{\mathrm{R}} \mathsf{HGen}(\mathbb{G}; n)$, for $1 \leq i \leq d$. The user secret key for the hierarchical identity $id = [id_1.\cdots.id_\mu]$ of depth $\mu \leq d$ is defined as $\mathsf{PRI}_{id} = (d_1, d_2, d_3, (d_{ij})_{\mu+i \leq j \leq d, 0 \leq i \leq n}) \in \mathbb{G}^{3+(n+1)\cdot(d-\mu-1)}$, where $d_1 = \alpha \cdot (\prod_{j=1}^{\mu} H_i(id^{(j)}))^r$, $d_2 = g^{-r}$, $d_3 = u^r$, and $d_{ij} = ((h_i^{(j)})^r)$. We remark that the latter $(n+1)\cdot(d-\mu-1)$ elements d_{ij} are only needed for hierarchical key delegation (and may be not included in PRI_{id} if such a feature is not wanted). Encryption of m with respect to id computes the two ciphertext elements $c_1 = g^r$ and $c_2 = (u^t \prod_{j=1}^{\mu} H_i(id^{(j)}))^r$ and uses the key $K = z^r$ to compute the symmetric ciphertext (using an AE-OT-secure scheme). Decryption uses $K = \hat{e}(d_1 \cdot d_3^t, c_1) \cdot \hat{e}(d_2, c_2)$ to reconstruct the plaintext from the symmetric ciphertext. Note that this only needs two pairing operations, independent of the depth of the hierarchy d. (In contrast the HIBE from [30] needs $d+1$ pairings.)

Security can be proved with respect to the *d-modified BDDH assumption*, where compared to the mBDDH assumption the adversary gets the values $g^y, g^{y^2}, \ldots, g^{y^{d+1}}$ (instead of just g^y, g^{y^2}). As in [24, 41] the security reduction is exponential in the depth d of the hierarchy, i.e. it introduces, roughly, a multiplicative factor of $(nq)^d$. Hence the scheme can only be considered practical for small hierarchies, say of depth $d = 4$.

4.3.3 Trading ciphertext size for efficiency

A variant of our IBE scheme can be combined with CCA-secure symmetric encryption. CCA-secure symmetric encryption is less demanding than authenticated encryption and, in particular, strong pseudorandom permutations imply CCA-secure symmetric encryption without any redundancy. This has the advantage of more compact ciphertexts while decryption has to perform some algebraic consistency checks and is therefore less efficient.

5 IBE Scheme II

In this section we present our second chosen-ciphertext secure IBE scheme from the q-ABDHE assumption. It is based on the Boneh-Boyen "exponent inversion" IBE scheme [8] in its full-identity secure variant of Gentry [23]. Building on techniques by Cramer and Shoup [21], Gentry also presents a chosen-ciphertext secure variant of his basic chosen-plaintext secure scheme. Our main improvement is to combine it with a strongly secure symmetric encryption scheme to considerably reduce ciphertext size and encryption/decryption cost.

5.1 The IBE construction

Let $\mathbb{PG} = (\mathbb{G}, \mathbb{G}_T, p, \hat{e}, g, g_T = \hat{e}(g,g))$ be a pairing group. Let $TCR : \mathbb{G} \times \mathbb{G} \to \mathbb{Z}_p$ be a target collision-resistant hash function. Let (E, D) be a symmetric cipher. Our IBE scheme $\mathsf{IBE}_2 = (\mathsf{Setup}, \mathsf{KeyGen}, \mathsf{Enc}, \mathsf{Dec})$ with identity space $IDSp = \mathbb{Z}_p$ is depicted in Figure 3.

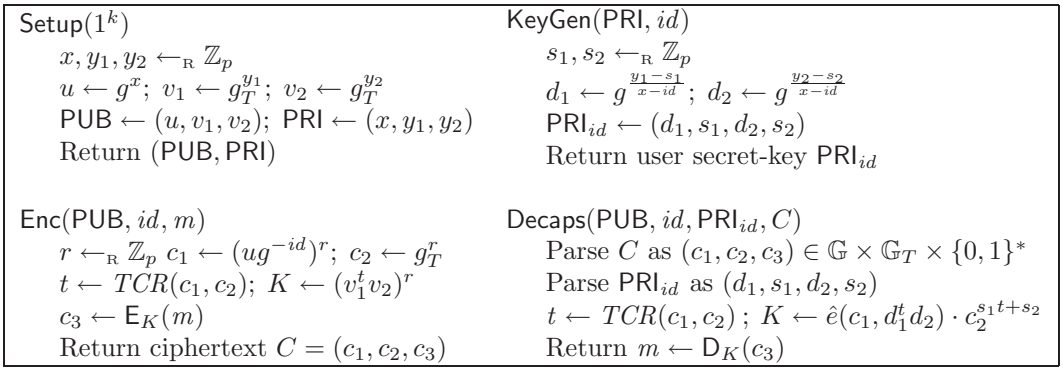

Figure 3: Our CCA-secure IBE scheme IBE_2.

To show correctness consider a ciphertext (c_1, c_2, c_3) generated for identity id that gets decrypted with a valid user secret key $\mathsf{PRI}_{id} = (d_1, d_2, s_1, s_2)$ by computing the symmetric key K as follows

$$\begin{aligned} K &= \hat{e}(c_1, d_1^t d_2) \cdot c_2^{s_1 t + s_2} \\ &= \hat{e}(g^{(x-id)r}, g^{\frac{(y_1 - s_1)t + (y_2 - s_2)}{x - id}}) \cdot \hat{e}(g, g)^{(s_1 t + s_2) r} \\ &= \hat{e}(g^r, g^{y_1 t + y_2}) \\ &= (v_1^t v_2)^r, \end{aligned}$$

as in the encryption algorithm.

5.2 Security

Theorem 5.1 Assume TCR is a target collision resistant hash function and (E, D) is a AE-OT-secure symmetric scheme. Let $q = q_x + 1$, where q_x is the number of key-derivation queries. Under the truncated q-ABDHE assumption relative to generator $\mathcal{G}$, the IBE scheme IBE_2 is IND-CCA secure. In particular, we have

$$\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE}_2, t, q_x, q_d}(k) \leq \mathsf{Adv}^{\text{q-abdhe}}_{\mathcal{G}, t}(k) + \mathsf{Adv}^{\mathsf{TCR}}_{TCR, t}(k) + 2q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE}, t}(k) + \mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE}, t}(k) + \frac{q_d}{p}.$$

The proof of Theorem 5.1 will be given in Appendix A.2. We give some intuition why the scheme is IND-CCA secure. First, the proof of Gentry [23] can be used to show that user secret-key queries, as well as *consistent* decryption queries for the challenge identity id^* are basically useless for an adversary attacking the scheme (unless it can efficiently solve the q-ABDHE problem). However, inconsistent decryption queries with respect to the challenge identity id^* may leak information about the hidden bit b. Here we use a Cramer-Shoup argument. The idea is that the user secret-key $\mathsf{PRI}_{id^*} = (d_1^*, s_1^*, d_2^*, s_2^*)$ used to answer such decryption queries contains some internal randomness $(s_1, s_2) \in \mathbb{Z}_p^2$ that is initially hidden from the adversary's view. During the simulation of the IND-CCA environment the challenge ciphertext will leak (in an information-theoretic sense) one linear equation on the hidden randomness (s_1^*, s_2^*). Decryption queries of inconsistent ciphertexts will use a key K for symmetric decryption that is computed as a linear equation in s_1^*, s_2^* which is linearly independent from the equation the adversary knows. Hence, one single key K is uniformly distributed over $\mathbb{G}_T$. By the ciphertext authenticity property of SE the adversary will not be able to come up with an inconsistent ciphertext (c_1, c_2, c_3) such that $\mathsf{D}_K(c_3)$ does not reject. Consequently, all inconsistent ciphertext will get rejected by the scheme.

5.3 Extensions

Using techniques from [1] it is further possible to prove IBE_2 anonymous in the sense that the ciphertext does not leak any information about the sender's identity. This property has recently proved useful in

the area of public-key encryption with keyword search [1].

We remark that in contrast to the IBE construction from Section 4 it is not possible to trade algebraic consistency checks for a weaker symmetric encryption scheme. In general, the class of inversion-based IBE schemes are less versatile than the commutative-blinding IBE schemes; for example, adding extensions like hierarchical key delegation to inversion-based IBE schemes seems a difficult task.

6 Comparison

In this section we compare our schemes with known CCA-secure IBE schemes from the literature.

6.1 Considered schemes

We consider the following standard-model IBE schemes.

IBE_1**:** Our scheme from Section 4 with the shorter public-parameters. See Section 4.3.1 for details.

IBE_2**:** Our scheme from Section 5.

KG**:** The scheme from Kiltz and Galindo [30].

Gentry**:** The scheme from Gentry [23] (IND-CCA variant).

We furthermore consider the following three IBE schemes that only have a proof in the random-oracle model. All of them are currently in submission for the IEEE1363.3 standardization project [27].

BF**:** The (FullIdent) scheme from Boneh and Franklin [12].

BB_1**:** The scheme from Boneh and Boyen [8] in its "hashed identities" variant [14].

KS**:** The scheme from Kasahara and Sakai [36] as described in [18].

We remark that when assuming the interactive *gap Bilinear Diffie-Hellman* (gap-BDH) assumption efficiency of BF and BB_1 can be further improved [14]. Due to the strong assumption we will not consider those schemes.

6.2 Security reductions

For determining the parameters of the compared schemes, we make the following assumptions, most of the are conservative towards the efficiency of our new schemes. For $k = 80$ bit security we estimate (following Bellare and Rogaway [5]) the number of (random oracle) hash queries as $q_H = 2^{50}$. This seems reasonable since a hash function is in the hand of an adversary and can be attacked offline. Similar to signatures schemes we think that a reasonable estimate for the number of key-derivation queries is $q_x \approx 2^{25}$. This is much smaller than the number of hash queries since key-derivation queries can only be made online, in interaction with the system. In practice it is easy to limit the number of key-derivation queries.

The IBE schemes IBE_1 and KG have two additional integer parameters: n, l. Parameter $n = 2k$ resembles the bit size $n = 160 \approx 2^7$ of the identity space and $l(k)$ defines the tradeoff between public parameters and security-reduction (cf. Section 4.3.1). We choose $l = 18$ to obtain a security loss of $2^{18+7+25} = 2^{50} = q_H$. This explains our choice of $l(k)$: it is chosen such that the security loss of the above schemes matches the one of all random-oracle schemes.

The concrete security reductions are given in Figure 4. For a fair comparison the security reductions of the random-oracle based schemes are given relative to the respective decisional assumption (e.g., BDDH instead of BCDH for BB_1). We note that the two schemes IBE_2 and Gentry have a tight security reduction to a much stronger security assumption. Due to the recent attacks by Cheon [19] it seems reasonable that the q-xxxx assumption are $\sqrt{q}$ times "less secure" than the BDDH assumption. This in particular implies (by Lemma 3.1) that we can treat the mBDDH assumption as "as secure" as the BDDH assumption. To simplify the comparison we make the conservative assumption that all the above schemes with the given parameters have the same security loss with respect to the BDDH assumption.

Scheme	Standard Model?	Assumption	Security reduction	
			Bounds	concrete ($k = 80$)
IBE_1	$\surd$	mBDDH	$2^l n q_x$	2^{50}
IBE_2	$\surd$	q-ABDHE	1	1
KG	$\surd$	BDDH	$2^l n q_x$	2^{50}
Gentry	$\surd$	q-ABDHE	1	1
BF	—	BDDH	$> q_H$	2^{50}
BB_1	—	BDDH	q_H	2^{50}
KS	—	q-BDDHI	q_H^3	$\gg 2^{50}$

Figure 4: Security assumptions and (concrete) reduction factors for IBE schemes.

6.3 Implementation details and curves

In pairing based cryptography efficiency always depends on the chosen curve and how well the scheme can be adapted to it. In particular, in asymmetric pairing groups the bilinear mapping is defined as $\hat{e} : \mathbb{G} \times \hat{\mathbb{G}} \rightarrow \mathbb{G}_T$, where elements in $\hat{\mathbb{G}}$ need much larger representations than in $\mathbb{G}$. We optimize all schemes with respect to short ciphertext sizes. See Figure 5 for concrete values.

We note that whenever the IBE scheme can be proved secure when using a CCA-secure symmetric schemes, we assume it is instantiated with a redundancy-free scheme (cf. Section 2).

Boyen [14] computes estimated relative timings for all atomic asymmetric operations (exponentiations and pairings) and representation sizes for group elements when instantiated in super-singular curves [12] with 80 bits security (symmetric pairing groups) and MNT curves with 80 bits security [32, 13] (asymmetric pairing groups). In Figure 5 (right side) we recall the values from [14]. Here a *ratio of pairings* (sometimes also called multi-pairing) denotes the product of two pairings. For simplicity we decided not to take into account possible savings when performing fixed-based or multi exponentiations.

Table 5 gives concrete values of encryption/decryption operations and ciphertext expansion for the considered IBE schemes. We note that for IBE_1 and KG, computing $H(id)^r = (h_0 \prod_{i=1}^{n/l} h_i^{id_i})^r$ ($id_i \in \{0,1\}^l$) is counted as two exponentiations.

6.4 Results

A comparison with concrete timing values from Boyen [14] is carried out in Figure 1 (Section 1) and Figure 6. Ciphertext overhead represents the difference (in bits) between the ciphertext length and the message length. All timings are given in multiplicative factors relative to one exponentiation in $\mathbb{G}$. As usual, all symmetric operations (cryptographic hash function, symmetric encryption, etc) are ignored. All schemes come with a security proof based on different security assumption, furthermore introducing a different loss of security in the reduction, depending on several system parameters. A high loss in the security reduction reduces the real-world efficiency of the scheme by making it necessary to increase the size of the groups for any given security level. In order not to compare apples with pears, we attempted to pick the parameters (in particular the parameter l for IBE_1 and KG) such that we obtain the *same concrete security reduction* for all schemes.

We conclude that our schemes are the most efficient chosen-ciphertext secure IBE schemes in the standard model. Furthermore its performance and ciphertext expansion seems comparable to the known random-oracle based schemes, in particular to the one by Boneh and Franklin which is intensively used in practice (see, e.g., `http://www.voltage.com`).

Acknowledgements

We thank Charles Rackoff, Ian Blake, and the anonymous CT-RSA reviewers for useful comments. The first author was supported by the research program Sentinels (`http://www.sentinels.nl`). Sentinels

Scheme	IBE_1	IBE_2	KG	Gentry	BF	BB_1	KS	MNT/80	SS/80
Ciphertext expansion								Repr. factor (bits)	
$\mathbb{G}$	2	1	3	1	1	2	1	171	512
$\hat{\mathbb{G}}$								1026	512
$\mathbb{G}_T$		1		2				1026	1024
$\mathbb{Z}_p$/RO					1	1	1	160	160
MAC	1	1						80	80
Public-Key size									
$\mathbb{G}$	10	1	9	1	1	2	1	171	512
$\hat{\mathbb{G}}$								1026	512
$\mathbb{G}_T$	1	2	1	3		1		1026	1024
Encryption cost								Cost factor (relative)	
Exp in $\mathbb{G}$	4	2	5	2	1	3	2	1	1
Hash to $\hat{\mathbb{G}}$					1			36	1
Exp in $\mathbb{G}_T$	1	3	1	4		1	1	36	4
Pairing					1			150	20
Decryption cost									
Exp in $\mathbb{G}$					1	1	2	1	1
Exp in $\hat{\mathbb{G}}$	1	1	5	1				36	1
Exp in $\mathbb{G}_T$		1		2		1		36	4
Pairing		1		2	1		1	150	20
Ratio of pairings	1		1			1		180	24

Figure 5: Efficiency of encryption/decryption operations plus ciphertext expansion (left side). Cost of group operations (normalized to one exponentiation in $\mathbb{G}$) and representation sizes from [14] (right side).

is being financed by Technology Foundation STW, the Netherlands Organization for Scientific Research (NWO), and the Dutch Ministry of Economic Affairs.

References

[1] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi. Searchable encryption revisited: Consistency properties, relation to anonymous IBE, and extensions. In Victor Shoup, editor, *CRYPTO 2005*, volume 3621 of *LNCS*, pages 205–222. Springer-Verlag, Berlin, Germany, August 2005. (Cited on page 10, 11.)

[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 Ronald Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 128–146. Springer-Verlag, Berlin, Germany, May 2005. (Cited on page 3, 22.)

[3] Mihir Bellare and Chanathip Namprempre. Authenticated encryption: Relations among notions and analysis of the generic composition paradigm. In Tatsuaki Okamoto, editor, *ASIACRYPT 2000*, volume 1976 of *LNCS*, pages 531–545. Springer-Verlag, Berlin, Germany, December 2000. (Cited on page 5.)

[4] Mihir Bellare and Phillip Rogaway. Random oracles are practical: A paradigm for designing efficient protocols. In V. Ashby, editor, *ACM CCS 93*, pages 62–73. ACM Press, November 1993. (Cited on page 1, 9.)

[5] Mihir Bellare and Phillip Rogaway. The exact security of digital signatures: How to sign with RSA and Rabin. In Ueli M. Maurer, editor, *EUROCRYPT'96*, volume 1070 of *LNCS*, pages 399–416. Springer-Verlag, Berlin, Germany, May 1996. (Cited on page 11.)

	Size (bits)		Cost (relative)	
Scheme	Ciphertext	Public Key	Encrypt	Decrypt
Standard model				
Ours: IBE_1 (§4)	1104	6144	8	25
Ours: IBE_2 (§5)	1616	2560	14	25
KG [30]	1536	5632	9	29
Gentry [23]	2560	3584	18	49
Random Oracle model				
BF [11]	672	512	22	21
BB_1 [8]	1184	2048	7	29
KS [18]	672	512	6	22

Figure 6: Efficiency comparison for CCA-secure IBE schemes in the standard/random oracle model for SS/80-bit security level. Timings are relative to one exponentiation in group $\mathbb{G}$.

[6] Daniel Bleichenbacher. Chosen ciphertext attacks against protocols based on the RSA encryption standard PKCS #1. In Hugo Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 1–12. Springer-Verlag, Berlin, Germany, August 1998. (Cited on page 1.)

[7] D. Boneh, C. Gentry, and M. Hamburg. Space-efficient identity based encryption without pairings. In *Proceedings of FOCS 2007*, pages 647–657. IEEE, 2007. (Cited on page 3.)

[8] Dan Boneh and Xavier Boyen. Efficient selective-ID secure identity based encryption without random oracles. In Christian Cachin and Jan Camenisch, editors, *EUROCRYPT 2004*, volume 3027 of *LNCS*, pages 223–238. Springer-Verlag, Berlin, Germany, May 2004. (Cited on page 1, 2, 3, 6, 9, 11, 14, 23.)

[9] Dan Boneh, Xavier Boyen, and Eu-Jin Goh. Hierarchical identity based encryption with constant size ciphertext. In Ronald Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 440–456. Springer-Verlag, Berlin, Germany, May 2005. (Cited on page 2, 9.)

[10] Dan Boneh, Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from identity-based encryption. *SIAM Journal on Computing*, 5(36):1301–1328, 2006. (Cited on page 1.)

[11] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the Weil pairing. In Joe Kilian, editor, *CRYPTO 2001*, volume 2139 of *LNCS*, pages 213–229. Springer-Verlag, Berlin, Germany, August 2001. (Cited on page 1, 2, 14, 23.)

[12] Dan Boneh and Matthew K. Franklin. Identity based encryption from the Weil pairing. *SIAM Journal on Computing*, 32(3):586–615, 2003. (Cited on page 1, 3, 5, 11, 12.)

[13] Dan Boneh, Ben Lynn, and Hovav Shacham. Short signatures from the Weil pairing. In Colin Boyd, editor, *ASIACRYPT 2001*, volume 2248 of *LNCS*, pages 514–532. Springer-Verlag, Berlin, Germany, December 2001. (Cited on page 12.)

[14] Xavier Boyen. The BB1 identity-based cryptosystem: A standard for encryption and key encapsulation. Submitted to IEEE 1363.3, aug 2006. `http://grouper.ieee.org/groups/1363/`. (Cited on page 1, 2, 3, 11, 12, 13.)

[15] Xavier Boyen, Qixiang Mei, and Brent Waters. Direct chosen ciphertext security from identity-based techniques. In Vijayalakshmi Atluri, Catherine Meadows, and Ari Juels, editors, *ACM CCS 05*, pages 320–329. ACM Press, November 2005. (Cited on page 1, 6.)

[16] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. In *30th ACM STOC*, pages 209–218. ACM Press, May 1998. (Cited on page 1.)

[17] Sanjit Chatterjee and Palash Sarkar. Trading time for space: Towards an efficient IBE scheme with short(er) public parameters in the standard model. In Dongho Won and Seungjoo Kim, editors, *ICISC 05*, volume 3935 of *LNCS*, pages 424–440. Springer-Verlag, Berlin, Germany, December 2005. (Cited on page 8.)

[18] L. Chen, Z. Cheng, J. Malone-Lee, and N.P. Smart. An efficient ID-KEM based on the Sakai-Kasahara key construction. *IEE Proceedings Information Security*, 152:19–26, 2006. (Cited on page 1, 2, 3, 11, 14.)

[19] Jung Hee Cheon. Security analysis of the strong Diffie-Hellman problem. In Serge Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 1–11. Springer-Verlag, Berlin, Germany, May / June 2006. (Cited on page 6, 11.)

[20] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In Hugo Krawczyk, editor, *CRYPTO'98*, volume 1462 of *LNCS*, pages 13–25. Springer-Verlag, Berlin, Germany, August 1998. (Cited on page 2.)

[21] 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. (Cited on page 2, 3, 5, 9.)

[22] Yevgeniy Dodis, Roberto Oliveira, and Krzysztof Pietrzak. On the generic insecurity of the full domain hash. In Victor Shoup, editor, *CRYPTO 2005*, volume 3621 of *LNCS*, pages 449–466. Springer-Verlag, Berlin, Germany, August 2005. (Cited on page 1.)

[23] Craig Gentry. Practical identity-based encryption without random oracles. In Serge Vaudenay, editor, *EUROCRYPT 2006*, volume 4004 of *LNCS*, pages 445–464. Springer-Verlag, Berlin, Germany, May / June 2006. (Cited on page 1, 2, 6, 9, 10, 11, 14.)

[24] Craig Gentry and Alice Silverberg. Hierarchical ID-based cryptography. In Yuliang Zheng, editor, *ASIACRYPT 2002*, volume 2501 of *LNCS*, pages 548–566. Springer-Verlag, Berlin, Germany, December 2002. (Cited on page 9.)

[25] Dennis Hofheinz and Eike Kiltz. Secure hybrid encryption from weakened key encapsulation. In Alfred Menezes, editor, *CRYPTO 2007*, LNCS, pages 553–571. Springer-Verlag, Berlin, Germany, August 2007. (Cited on page 2, 3, 22.)

[26] Jeremy Horwitz and Ben Lynn. Toward hierarchical identity-based encryption. In Lars R. Knudsen, editor, *EUROCRYPT 2002*, volume 2332 of *LNCS*, pages 466–481. Springer-Verlag, Berlin, Germany, April / May 2002. (Cited on page 9.)

[27] IEEE P1363.3 Committee. IEEE 1363.3 — standard for identity-based cryptographic techniques using pairings. `http://grouper.ieee.org/groups/1363/`, April 2007. (Cited on page 1, 11.)

[28] Eike Kiltz. Chosen-ciphertext secure identity-based encryption in the standard model with short ciphertexts. Cryptology ePrint Archive, Report 2006/122, 2006. `http://eprint.iacr.org/`. (Cited on page 3.)

[29] Eike Kiltz. Chosen-ciphertext security from tag-based encryption. In Shai Halevi and Tal Rabin, editors, *TCC 2006*, volume 3876 of *LNCS*, pages 581–600. Springer-Verlag, Berlin, Germany, March 2006. (Cited on page 1, 6.)

[30] Eike Kiltz and David Galindo. Direct chosen-ciphertext secure identity-based key encapsulation without random oracles. In *ACISP 2006*, volume 4058 of *LNCS*. Springer-Verlag, 2006. (Cited on page 1, 2, 3, 6, 7, 8, 9, 11, 14, 18.)

[31] Kaoru Kurosawa and Yvo Desmedt. A new paradigm of hybrid encryption scheme. In Matthew Franklin, editor, *CRYPTO 2004*, volume 3152 of *LNCS*, pages 426–442. Springer-Verlag, Berlin, Germany, August 2004. (Cited on page 2, 3, 5.)

[32] Atsuko Miyaji, Masaki Nakabayash, and Shunzou Takano. New explicit conditions of elliptic curve traces for FR-reduction. *IEICE Trans. Fundamentals*, E84(5):1234–1243, 2001. (Cited on page 12.)

[33] David Naccache. Secure and *practical* identity-based encryption. Cryptology ePrint Archive, Report 2005/369, 2005. `http://eprint.iacr.org/`. (Cited on page 8.)

[34] Charles Rackoff and Daniel R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In Joan Feigenbaum, editor, *CRYPTO'91*, volume 576 of *LNCS*, pages 433–444. Springer-Verlag, Berlin, Germany, August 1992. (Cited on page 1.)

[35] Phillip Rogaway, Mihir Bellare, John Black, and Ted Krovetz. OCB: A block-cipher mode of operation for efficient authenticated encryption. In *ACM CCS 01*, pages 196–205. ACM Press, November 2001. (Cited on page 5.)

[36] R. Sakai, K. Ohgishi, and M. Kasahara. Cryptosystems based on pairing over elliptic curve (in japanese). In *Proceedings of the Symposium on Cryptography and Information Security — SCIS 2001*, jan 2001. (Cited on page 1, 3, 11.)

[37] Palash Sarkar and Sanjit Chatterjee. Transforming a cpa-secure hibe protocol into a cca-secure hibe protocol without loss of security. Cryptology ePrint Archive, Report 2006/362, 2006. `http://eprint.iacr.org/`. (Cited on page 3.)

[38] Adi Shamir. Identity-based cryptosystems and signature schemes. In G. R. Blakley and David Chaum, editors, *CRYPTO'84*, volume 196 of *LNCS*, pages 47–53. Springer-Verlag, Berlin, Germany, August 1985. (Cited on page 1.)

[39] Victor Shoup. Lower bounds for discrete logarithms and related problems. In Walter Fumy, editor, *EUROCRYPT'97*, volume 1233 of *LNCS*, pages 256–266. Springer-Verlag, Berlin, Germany, May 1997. (Cited on page 6.)

[40] Victor Shoup. Why chosen ciphertext security matters. IBM Research Report RZ 3076, November 1998. (Cited on page 1.)

[41] Brent R. Waters. Efficient identity-based encryption without random oracles. In Ronald Cramer, editor, *EUROCRYPT 2005*, volume 3494 of *LNCS*, pages 114–127. Springer-Verlag, Berlin, Germany, May 2005. (Cited on page 1, 6, 7, 8, 9, 18.)

A Proofs

A.1 Proof of Theorem 4.1

We will start by defining the first game to be the experiment which is described in the definition of CCA security.

Game 1. Game 1 is the experiment of CCA security. At some point during the game the adversary chooses a challenge identity, we will refer to this identity as id^*. In the challenge phase a challenge encryption is generated and is given to the adversary. We will refer to the challenge ciphertext as $C^* = (c_1^*, c_2^*, c_3^*)$. We will also use $*$ to mark intermediate values that were used in generating the challenge ciphertext. Let X_1 be the event that A is successful in Game 1. Our goal is to put an upper bound on $\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE}_1,A,t}(k) \equiv \left|\Pr[X_1] - \frac{1}{2}\right|$. For each of the following experiments we will call X_i the event that the adversary is successful in Game i.

Game 2. We define Game 2 to be the same as Game 1 except that if the adversary makes a decryption query with

$$c_1 \neq c_1^* \text{ and } t = t^*$$

the decryption oracle outputs "reject" (or $\perp$). Furthermore, if the adversary makes a decryption query with $c_1 = c_1^*$ before seeing the challenge, then the experiment is stopped and a random outcome is chosen. Let F_2 be the event that the experiment is stopped, and let F_2' be the event that a ciphertext was rejected in Game 2 by the decryption oracle but would not have been rejected in Game 1. Game 1 and Game 2 proceed identically unless event $F_2 \vee F_2'$ occurs; in particular, the events $X_1 \wedge \neg(F_2 \vee F_2')$ and $X_2 \wedge \neg(F_2 \vee F_2')$ are the same. Therefore we have:

$$|\Pr[X_2] - \Pr[X_1]| \leq (F_2 \vee F_2')$$

It is easy to see that $\Pr[F_2] \leq q_d/p$. Also, we have that $\Pr[F_2'] \leq \mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k)$, otherwise we could break the target collision resistance of TCR with a bigger advantage than $\mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k)$.

Game 3. We call a ciphertext $C = (c_1, c_2, c_3)$ ill-formed with respect to identity id if $\hat{e}(g, c_2) \neq \hat{e}(c_1, H(id) \cdot u^t)$ where $t = H(c_1)$. It is equivalent to saying that C is of the form $(g^r, (H(id) \cdot u^t)^{r'}, c_3)$ and $r \neq r'$. We call a decryption query ill-formed if it contains an ill-formed ciphertext.

Game 3 is the same as Game 2 except that we add the restriction that the adversary is not allowed to ask ill-formed decryption queries. We define a new adversary A_3 that simulates A and whenever A makes an ill-formed decryption query, A_3 responds with *reject*. To see how the probabilities of success of A_3 and A are related we define F_3 to be the event that A makes an ill-formed decryption query which would not have been rejected in Game 2 but is rejected in Game 3. Clearly, if event F_3 does not occur then A_3 simulates A perfectly, therefore we get $|\Pr[X_3] - \Pr[X_2]| \leq \Pr[F_3]$.

Claim A.1 $\Pr[F_3] \leq q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A},t}(k) + q_d^2/p$.

Proof: Let (id, C), where $C = (c_1, c_2, c_3)$, be an ill-formed decryption query. Let's consider the process of decrypting the ciphertext C. The value K is computed using the ciphertext and some private key $\mathsf{PRI}_{id^*} = (\alpha \cdot H(id^*)^s, g^{-s}, u^s)$ of the challenge identity in the following way:

$$\begin{aligned}
K &= \hat{e}(c_1, d_1 \cdot d_3^t) \cdot \hat{e}(c_2, d_2) \\
&= \hat{e}(g^r, \alpha(H(id^*))^s \cdot u^{st}) \cdot \hat{e}((H(id^*) \cdot u^t)^{r'}, g^{-s}) \\
&= \hat{e}(g, \alpha^r (H(id^*) \cdot u^t)^{rs}) \cdot \hat{e}(g, (H(id^*) \cdot u^t)^{-r's}) \\
&= \hat{e}(g, \alpha)^r \cdot \hat{e}(g, (H(id^*) \cdot u^t)^{rs}) \cdot \hat{e}(g, (H(id^*) \cdot u^t)^{-r's}) \\
&= z^r \cdot \hat{e}(g, (H(id^*)) \cdot u^t)^{(r-r')s}
\end{aligned}$$

The value s represents the randomness in the private key PRI_{id^*}. Notice that for well formed ciphertexts the intermediate key K is independent of s (it only depends on z^r). Thus, the adversary did not gain any information about s from any well-formed decapsulation queries that he has made. Now, consider previously failed ill-formed decryption queries. For each such query the adversary learns at best one $\mathbb{Z}_p$ value which is *not* equal to s. Thus, after a polynomial number of failed ill-formed decryption queries, s remains almost uniformly distributed in $\mathbb{Z}_p$ in the adversary's view. Using this, and the fact that $c_2 \neq 1$ we obtain that in the adversary's view K is distributed almost uniformly in $\mathbb{G}_T$. That is, the distance between the uniform distribution on $\mathbb{G}_T$ and the distribution on K is $\leq q/p$ where q is the number of decryption queries that the adversary asks.

Now, we know that the probability of A_3 generating an AE ciphertext c_3 such that it does not decrypt as "reject" when using a random K is $\leq \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A},t}(k)$. Thus, when using the almost random K that is obtained from decrypting as above we get that the probability of the ciphertext not decrypting as "reject" is $\leq \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A},t}(k) + q_d/p$. Since the adversary may attempt to generate such a ciphertext q_d times, the total probability of success is $\leq q_d \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},\mathsf{A},t}(k) + q_d^2/p$. ∎

Game 4. Game 4 is the same as Game 3 except for the following changes:

- To generate u, α, z two random exponents are chosen $a, b \leftarrow \mathbb{Z}_p$, and the following values are assigned: $u \leftarrow g^b$; $\alpha \leftarrow g^{ab}$; $z \leftarrow \hat{e}(g^a, g^b)$.
- Let $q = q_x + q_d$ be the total number of queries (private key and decryption) that the adversary makes. To generate $h_0, h_1, ..., h_n$ compute: $m \leftarrow 4q$; $w \leftarrow_{\mathrm{R}} [0, n]$; $x', x_1, ..., x_n \leftarrow_{\mathrm{R}} [0, q-1]$; $y', y_1, ..., y_n \leftarrow_{\mathrm{R}} [0, m-1]$; $h_0 \leftarrow (g^a)^{p-wm+y'}(g^b)^{-t^*} g^{x'}$; for $1 \leq i \leq n$, $h_i \leftarrow g^{x_i}(g^a)^{y_i}$.
- (Forced abort) For an identity id let

$$x(id) \stackrel{\text{def}}{=} x' + \sum_{i=1}^{n} x_i^{id_i}$$

and

$$y(id) \stackrel{\text{def}}{=} p - km + y' + \sum_{i=1}^{n} y_i^{id_i}$$

We call F_4 the event that one of the following two conditions is true. If F_4 occurs then the experiment is aborted and the outcome is chosen randomly.

1. The adversary asks a private key query for an identity id such that $y(id) \equiv 0 \bmod q$.
2. The adversary chooses a challenge identity id^* such that $y(id^*) \not\equiv 0 \bmod q$.

- (Artificial Abort) Let $view_A$ be the adversary's random tape, and the transcript of its interactions with its oracles in the current run of the experiment of Game 4. Let $Y = (y', y_1, ..., y_n, k)$ where the random variables $y', y_1, ..., y_n, k$ are distributed as described above. Clearly, if we fix $view_A$ and rerun the experiment the random variable Y has the same distribution as for a run of the experiment without a fixed view of the adversary. This is true due to the random "masks" $x_i \in \mathbb{Z}_p$. We define

$$\eta(view_A) \stackrel{\text{def}}{=} \Pr_Y[F_4 | view_A]$$

Let $\lambda = 1/4(n+1)q$, and let $0 < \rho(k) \leq 1$ be a function in k which will be specified later. At the end of the experiment, before comparing the output b' of the adversary to b, $1/2 \cdot \lambda^{-2} \cdot (\rho(k)/r)^{-2} \ln((\lambda\rho)^{-1})$ samples of Y are taken, and an estimate $\eta'(view_A)$ for $\eta(view_A)$ is computed. If $\eta'(view_A) > 1/4q(n+1)$ then the experiment is aborted, that is, a random value b'' is chosen instead of b' and b'' is compared with b.

Note that the values $u, \alpha, z, h_0, h_1, ..., h_n$ as they are computed in Game 4 are distributed identically to the corresponding values in Game 3.

Claim A.2 Let q be the total number of private key and decryption queries made by the adversary. Then, $|\Pr[X_3] - (1/2 + (\Pr[X_4] - 1/2) \cdot 4q(n+1))| \leq \rho(k)$.

Proof: The proof of this claim is somewhat technical, and is omitted here. We refer the reader to [30] for the proof (which is originally due to Waters [41]). ∎

Game 5. Game 5 is the same as Game 4 except that we change the way that private keys are computed, and the way that decryption queries are answered.

1. The private key and decryption oracles generate private keys as follows. Choose a random exponent $s' \leftarrow_{\mathrm{R}} \mathbb{Z}_p$, compute

$$\begin{aligned} d_1 &\leftarrow (g^a)^{s' \cdot y(id)} (g^b)^{-x(id)/y(id) - s' t^*} (g^{b^2})^{t^*/y(id)} g^{s' \cdot x(id)} \\ d_2 &\leftarrow (g^b)^{-1/y(id)} g^{s'} \\ d_3 &\leftarrow (g^{b^2})^{-1/y(id)} (g^b)^{s'} \end{aligned}$$

and output (d_1, d_2, d_3). In the above calculation, a and b are the exponents that were chosen during the setup stage, as described in the definition of Game 4. Notice that we only need to be able to compute private keys for identities for which $y(id) \not\equiv 0 \bmod m$.

2. Answering decryption queries: let (c_1, c_2, c_3) be a decryption query. Compute

$$K \leftarrow \hat{e}(c_2/c_1^{x(id)}, g^a)^{(t-t^*)^{-1}}$$

and return $\mathsf{D}_K(c_3)$.

Following a technical argument it is easy to check that the private keys and decryptions in games 3 and 4 are distributed identically. Thus, the probabilities of success in both games are equal: $\Pr[X_4] = \Pr[X_5]$.

Game 6. Game 6 is the same as Game 5 except that the value K^* is replaced by a random value $\tilde{K} \in \mathbb{G}_T$ which is chosen at the setup stage. Recall that K^* in Game 4 is z^r for some $r \in \mathbb{Z}_p$ which is chosen randomly at the challenge stage. For clarity we will describe the computation of the challenge encryption from scratch:

$b^* \leftarrow_{\mathrm{R}} \{0,1\}$; $r^* \leftarrow_{\mathrm{R}} \mathbb{Z}_p$; $c_1^* \leftarrow g^{r^*}$; $t^* \leftarrow H(c_1^*)$; $c_2^* \leftarrow \left(H(id) \cdot u^{t^*}\right)^{r^*}$; $\tilde{K} \leftarrow_{\mathrm{R}} \mathbb{G}_T$; $c_3^* \leftarrow E_{\tilde{K}}(M_{b^*})$; $C^* \leftarrow (c_1^*, c_2^*, c_3^*)$; Return C^*.

Claim A.3 $|\Pr[X_5] - \Pr[X_6]| \leq \mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)$.

Proof: The idea of the proof is the following. We are given g^a, g^b, g^{b^2}, g^c and Z which is either a random element of $\mathbb{G}_T$ or $\hat{e}(g,g)^{abc}$. We simulate the adversary. Notice that since Game 5 the exponents a and b are not necessary for the simulation except for decrypting ciphertexts in which $(c_1, c_2) = (c_1^*, c_2^*)$. The answers to the private key queries and decryption queries can be computes using g^a, g^b, g^{b^2}. To compute the challenge ciphertext we set $c_1^* = g^c$, and we use Z as the key for the symmetric encryption. Now, if $Z = \hat{e}(g,g)^{abc}$ then we simulate the adversary perfectly in Game 5. If Z is random then we simulate the adversary in Game 6. Thus, if the adversary distinguishes between games 5 and 6, then we distinguish between the two possible values for Z with the same probability. ∎

The proof of the following claim follows directly from the definition of ciphertext indistinguishability for AE:

Claim A.4 $|\Pr[X_6] - 1/2| \leq \mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\tilde{t}}(k)$

Summary. We now summarize the above statements into a bound on the advantage of the adversary in the CCA game:

$$\Pr[X_1] \leq 1/2 + (\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\tilde{t}}(k) + \mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)) \cdot 4q(n+1) + \mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k) + q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},t}(k) + 2q_d^2/p + \rho(k)\ .$$

Now, if we set $\rho(k) \stackrel{\mathrm{def}}{=} (\mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)) \cdot 4q(n+1)$ then we obtain

$$\begin{aligned}\Pr[X_1] &\leq 1/2 + (\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\tilde{t}}(k) + 2\mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)) \cdot 4q(n+1) + \mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k) + q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},t}(k) + 2q_d^2/p \\ &\leq 1/2 + (\mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},\tilde{t}}(k) + \mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)) \cdot 10nq + +\mathsf{Adv}^{\mathsf{TCR}}_{TCR,t}(k) + q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},t}(k) + 2q_d^2/p\end{aligned}$$

The only additional computation that we are required to perform to transform an adversary that breaks the CCA security of the IBE to a distinguisher for the mBDDH problem is the sampling of Y that is described in the "artificial abort" stage of game 4. Thus,

$$t \geq \tilde{t} - \mathcal{O}(\tilde{\varepsilon}^{-2}(k) \cdot \ln(\tilde{\varepsilon}^{-1}(k)) + q_d + q_x)$$

where $\tilde{\varepsilon}(k) = \mathsf{Adv}^{\mathrm{mbddh}}_{\mathcal{G},\tilde{t}}(k)$.

A.2 Proof of Theorem 5.1

We make the following definition. Fix public key PUB and identity id. For a tuple $(c_1, c_2) \in \mathbb{G} \times \mathbb{G}_T$ we consider $r_1 = \log_{ug^{-id}}(c_1)$, $r_2 = \log_{g_T}(c_2)$, where $t = TCR(c_1, c_2)$. We say that (c_1, c_2) is consistent with id if $r_1 = r_2$ and inconsistent otherwise. With the knowledge of $x \in \mathsf{PRI}$ the latter one can be efficiently checked by verifying if $\hat{e}(c_1^{1/(x-id)}, g) = c_2$. We say that a tuple (c_1, c_2) *yields the symmetric key* K *under identity* id, where K is defined as $K = \hat{e}(c_1, d_1^t d_2) \cdot c_2^{s_1 t + s_2}$ and $(d_1, d_2) \leftarrow \mathsf{KeyGen}(\mathsf{PRI}, id)$.

Let A be an adversary on the IND-CCA security of IBE_2. We will consider a sequence of games, Game 1, Game 2, ..., each game involving A. Let X_i be the event that in Game i, it holds that $b = b'$, i.e., that the adversary succeeds.

Game 1. Let Game 1 be the CCA security experiment run with adversary A, i.e., we have

$$\mathsf{Adv}^{\mathsf{CCA}}_{\mathsf{IBE}_2,\mathsf{A}} = |\Pr[X_1] - 1/2| \ .$$

We assume adversary A makes exactly $q_x = q - 1$ key generation queries, all with distinct identities. We further assume that A makes exactly q_d decryption queries (id^*, C_i), all with respect to the challenge identity id^*.

Game 2. We now change the generation of the challenge ciphertext C^* for id^* as follows. The experiment first internally generates a random instance of the user secret key $\mathsf{PRI}_{id^*} = (d_1^*, s_1^*, d_2^*, s_2^*) \leftarrow_{\mathrm{R}} \mathsf{KeyGen}(\mathsf{PRI}, id^*)$. Then it picks a random $r_1 \in \mathbb{Z}_p$ and computes

$$c_1^* = (ug^{-id^*})^{r_1}, \quad c_2^* = \hat{e}(g, g)^{r_1}. \tag{1}$$

The symmetric key K^* is then computed as in decryption as

$$K^* = \hat{e}(c_1^*, {d_1^*}^{t^*} d_2^*) \cdot {c_2^*}^{s_1^* t^* + s_2^*} \tag{2}$$

where $t^* = TCR(c_1^*, c_2^*)$. Finally, c_3^* is computed as $c_3 \leftarrow \mathsf{E}_{K^*}(m_b)$. Since this change is purely conceptional,

$$\Pr[X_2] = \Pr[X_1].$$

Game 3. In this game the experiment stops if the adversary queries the challenge ciphertext in the first phase. Since c_2^* is generated as $c_2^* = \hat{e}(g, g)^{r_1}$, independently from A's view until it sees the challenge ciphertext, we have

$$|\Pr[X_2] - \Pr[X_3]| \leq \frac{q_d}{p} \ .$$

Game 4. For generation of the challenge ciphertext the experiment proceeds as follows. The experiment now generates c_2^* from Equation (1) by picking $r_2 \leftarrow_{\mathrm{R}} \mathbb{Z}_p \setminus \{r_1\}$ and computing $c_2^* = \hat{e}(g, g)^{r_2}$.

Lemma A.5 $|\Pr[X_3] - \Pr[X_4]| \leq \mathsf{Adv}^{\text{q-abdhe}}_{\mathcal{G},t}(k)$.

Proof: We show that there exists an adversary B with $t_{\mathsf{B}} \approx t_{\mathsf{A}}$ such that $\mathsf{Adv}^{\text{q-abdhe}}_{\mathcal{G},\mathsf{B}}(k) = |\Pr[X_2] - \Pr[X_4]|$. Adversary B inputs a truncated q-ABDHE instance

$$(g, g^x, \ldots, g^{x^q}, g^z, g^{zx^{q+2}}, T) \ , \tag{3}$$

and has to distinguish $T = \hat{e}(g^z, g)^{x^{q+1}}$ from a random element in $\mathbb{G}_T$. For key-generation B picks two random degree q polynomial $f_1(X), f_2(X)$ and defines

$$u = g^x, \quad v_1 = \hat{e}(g, g)^{f_1(x)}, \quad v_2 = \hat{e}(g, g)^{f_2(x)},$$

using the values $g, g^x, \ldots, g^{x^q}$ from Equation (3). Note that this does not change the distribution of the public-key $\mathsf{PUB} = (u, v_1, v_2)$. This implicitly defines the secret key values as $y_1 = f_1(x)$ and $y_2 = f_2(x)$. Adversary A_1 is run on PUB.

For a key-derivation query $\mathsf{KeyGen}(\cdot)$ for identity $id_i \in \mathbb{Z}_p$ $(1 \leq i \leq q_x = q-1)$, B defines the two degree $q-1$ polynomials

$$F_{1,id_i}(X) = \frac{f_1(X) - f_1(id_i)}{X - id_i}, \quad F_{2,id_i}(X) = \frac{f_2(X) - f_2(id_i)}{X - id_i}$$

and returns

$$d_{1,id_i} = g^{F_{1,id_i}(x)}, s_{1,id_i} = f_1(id_i), \quad d_{2,id_i} = g^{F_{2,id_i}(x)}, s_{2,id_i} = f_2(id_i) \tag{4}$$

as the user secret key $\mathsf{PRI}_{id_i} = (d_{1,id_i}, d_{2,id_i}, s_{1,id_i}, s_{2,id_i})$ of id_i, which can be computed from the values from Equation (3). This is a correct user secret key since (for $j = 1, 2$), $d_{j,id_i} = g^{F_{j,id_i}(x)} = g^{\frac{y_j - s_{j,id_i}}{x - id_i}}$.

Let id^* be the challenge identity. For a decryption query $\textsc{Dec}(id^*, \cdot)$, adversary B first computes the user secret key for the challenge identity id^* as $\mathsf{PRI}_{id^*} = (d_1^*, d_2^*, s_1^*, s_2^*)$, using Equation (4) and answers all the decryption query as in the IND-CCA experiment. In total B will compute $q_x + 1 \leq q$ user secret keys: q_x for the key-derivation queries and one for the challenge identity id^*. Since $f_1(X)$ and $f_2(X)$ are random degree q polynomials, for all $q_x + 1 \leq q$ key-derivation queries the values $s_{j,id_i} = f_j(id_i)$ are uniform elements in $\mathbb{Z}_p$ and hence the established user secret keys have the correct distribution.

For generation of the challenge ciphertext for $id^* \in \mathbb{Z}_p$, B proceeds as follows. It defines the degree $q+1$ polynomial

$$F^*(X) = \frac{X^{q+2} - id^{*q+2}}{X - id^*} = \sum_{i=0}^{q+1} F_i^* X^i.$$

The challenge ciphertext $C^* = (c_1^*, c_2^*, c_3^*)$ is

$$c_1^* = g^{z x^{q+2}} \cdot (g^z)^{id^{*q+2}}, \quad c_2^* = T^{F_0^*} \cdot \hat{e}(g^z, \prod_{i=1}^{q} (g^{x^i})^{F_i^*}), \quad c_3^* = \mathsf{D}_{K^*}(m_b) \tag{5}$$

where the challenge key K^* is computed from c_1^*, c_2^* as in Equation (2) and b is a random bit. Note that the challenge ciphertext can be entirely computed from B's input values from Equation (3). Adversary B runs A_2 on input (C^*, St), answering all oracle queries as above, and inputing a bit b'. Finally, B outputs 1 if $b = b'$ and 0, otherwise.

We make the following claim that completes the proof of the lemma: if $T = \hat{e}(g^z, g)^{x^{q+1}}$ then A's view is the same as in Game 2. If $T \in \mathbb{G}_T$, then A's view is the same as in Game 4.

To prove the claim we have to consider the distribution of the challenge ciphertext in Games 2 and 4. Note that the element T only leaks through B's simulation in the element c_2^* from the challenge ciphertext. We write c_1^* as

$$c_1^* = g^{z x^{q+2}} \cdot (g^z)^{(id^*)^{q+2}} = g^{z(x - id^*)F^*(x)} = (u g^{-id^*})^{r_1},$$

for $r_1 = z F^*(x)$. If $T = \hat{e}(g^z, g)^{x^{q+1}}$, then

$$c_2^* = T^{F_0^*} \cdot \hat{e}(g^z, \prod_{i=1}^{q} (g^{x_i})^{F_i^*}) = \hat{e}(g^z, g^{F^*(x)}) = \hat{e}(g, g)^{r_1}.$$

is a consistent ciphertext for identity id^*, as in Game 2. On the other hand, if T is a uniform element in $\mathbb{G}_T$ so is c_2^*, as in Game 4. This proves the claim. ∎

Game 5. Let (c_1^*, c_2^*, c_3^*) be the challenge ciphertext for id^* and let $t^* = TCR(C^*)$. In this game the experiment changes the answers to the decryption oracle as follows. If, for a query $\mathsf{Dec}(id^*, C = (c_1, c_2, c_3))$ it holds that $(c_1, c_2) \neq (c_1^*, c_2^*)$ but $TCR(c_1, c_2) = t \neq t^*$ then the experiment aborts. We claim that there exists an adversary $\mathcal{F}$ with $t_{\mathcal{F}} \approx t_{\mathsf{A}}$ such that

$$|\Pr[X_5] - \Pr[X_4]| \leq \mathsf{Adv}^{\mathsf{TCR}}_{TCR,\mathcal{F}}(k).$$

Game 6. Game 6 is like Game 5 with the difference that all decryption queries $\textsc{Dec}(id^*, C = (c_1, c_2, c_3))$ for which (c_1, c_2) is inconsistent with id^* get rejected.

Lemma A.6 $|\Pr[X_6] - \Pr[X_5]| \leq q_d \cdot \mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},t}(k)$.

Proof: Let $(d_1^*, s_1^*, d_2^*, s_2^*)$ be the uniquely defined and fixed user secret-key for id^*. We first claim that in the view of adversary A, one single decryption query (c_1, c_2, c_3) for which (c_1, c_2) is inconsistent with id^* yields a uniform symmetric key $K \in \mathbb{G}_T$. The consequence is as follows. In Game 5 the decryption oracle returns $\perp$ (reject) if $\mathsf{D}_K(c_3) = \perp$. Since K is uniform in $\mathbb{G}_T$, this happens exactly with probability $\Pr_{K' \leftarrow_R \mathbb{G}_T}[\mathsf{D}_{K'}(c_3) = \perp]$ which equals the advantage of a suitable adversary in the ciphertext integrity experiment of the symmetric ciphertext SE. On the other hand, in Game 6 such a query gets always rejected. A standard argument [2, 25] shows that considering all q_d decryption queries one obtains

$$|\Pr[X_6] - \Pr[X_5]| \leq q_d \cdot \mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},t}(k)\ .$$

To prove the above claim, consider the hidden randomness $(s_1^*, s_2^*) \in \mathbb{Z}_p^2$ from generating the user secret key $\mathsf{PRI}_{id^*} = (d_1^*, d_2^*, s_1^*, s_2^*)$ that is used by the experiment when generating the challenge ciphertext. At the time of their generation, s_1^* and s_2^* are two independent random elements, uniformly distributed over $\mathbb{Z}_p$. Consider the symmetric key K^* which is obtained from the inconsistent challenge ciphertext $(c_1^* = (ug^{-id^*})^{r_1^*}, c_2^* = \hat{e}(g,g)^{r_2^*})$ by computing

$$\begin{aligned} K^* &= \hat{e}(c_1^*, {d_1^*}^{t^*} d_2^*) \cdot {c_2^*}^{s_1^* t^* + s_2^*} \\ &= \hat{e}(g,g)^{r_1^*(y_1 t^* + y_2) + (r_2^* - r_1^*)(s_1^* t^* + s_2^*)}. \end{aligned}$$

Now we consider the knowledge A can obtain from the the challenge ciphertext (c_1^*, c_2^*, c_3^*) and the challenge key K^* in an *information-theoretic sense*. A knows r_1^*, r_2^*, t^*, and $k_1^* = \log_{g_T} K_1^* = r_1^*(y_1 t^* + y_2) + (r_2^* - r_1^*)(s_1^* t^* + s_2^*)$. Hence the knowledge A has about the hidden randomness $(s_1^*, s_2^*) \in \mathbb{Z}_p^*$ is a values $l^* \in \mathbb{Z}_p$ such that

$$l^* = s_1^* t^* + s_2^*,$$

which is a point on the 2-dimensional plane.

Now consider the virtual key K that is computed from an ciphertext (c_1, c_2, c_3) of a decapsulation query such that (c_1, c_2) is not consistent with id^*. We claim that in the view of A, K is a uniform element in $\mathbb{G}_T$. Assume $c_1 = (ug^{-id})^{r_1}$ and $c_2 = \hat{e}(g,g)^{r_2}$ with $r_1 \neq r_2$ (since (c_1, c_2) is inconsistent with id^*). The key K is computed as $K = \hat{e}(c_1, {d_1^*}^t d_2^*) \cdot {c_2}^{s_1^* t + s_2^*}$. Again we consider $k = \log_{g_T} K = r_1(y_1 t + y_2) + (r_2 - r_1)(s_1^* t + s_2^*)$, where $r_1 \neq r_2$ and $t = TCR(c_1, c_2) \neq t^*$. Linear algebra shows that as long as $t \neq t^*$, in the view of A,

$$l = s_1^* t + s_2^*$$

is a uniform element in $\mathbb{Z}_p$ and hence the virtual key K is a uniform element in $\mathbb{G}_T$. This proves the claim. ∎

Game 7. The challenge key K^* is replaced with the random challenge key K^* (instead of computing K^* as in Equation (2)). The proof of Lemma A.6 essentially shows that from the adversary's point of view, K^* looks like a uniform element in $\mathbb{G}_T$ and hence

$$\Pr[X_7] = \Pr[X_6].$$

Finally, in Game 7 the adversary A basically carries out a chosen-ciphertext attack on the symmetric cipher since A is still allowed to query ciphertext of the form $(c_1^*, c_2^*, *)$ which are answered using a uniform key $K^* \in \mathbb{G}_T$. Consequently, using the fact that chosen-ciphertext security is implied by AE-OT security we obtain

$$\Pr[X_7] = q_d \cdot \mathsf{Adv}^{\mathsf{CT\text{-}INT}}_{\mathsf{SE},t}(k) + \mathsf{Adv}^{\mathsf{IND}}_{\mathsf{SE},t}(k) \ .$$

Collecting the probabilities proves the theorem.

B Relations between the Assumptions

B.1 The BDDH assumption

Let $\mathbb{PG}$ be the description of bilinear groups and let $g \in \mathbb{G}$ be a random element from group $\mathbb{G}$ of prime order p. Consider the following problem formalized by Boneh and Franklin [11]: Given $(g, g^a, g^b, g^c, W) \in \mathbb{G}^4 \times \mathbb{G}_2$ as input, output yes if $W = \hat{e}(g,g)^{abc}$ and no otherwise. The corresponding BDDH assumption can be formalized the same way as the modified BDDH assumption.

B.2 The q-BDDHI assumptions

Let $\mathbb{PG}$ as above and let $z \in \mathbb{G}$ be a random element from group $\mathbb{G}$. Let $q = q(k)$ be a function polynomial in the security parameter. Associated to q the following problem introduced by Boneh and Boyen [8]: Given $(h, h^a, h^{(a^2)}, \ldots, h^{(a^q)}, W) \in \mathbb{G}^{q+1} \times \mathbb{G}_2$ as input, output yes if $W = \hat{e}(h,h)^{1/a}$ and no otherwise.

B.3 Proof of Lemma 3.1

Proof: The implications BDDH $\leq$ mBDDH and 1-BDDHI $\leq$ 2-BDDHI $\leq$ 3-BDDHI $\leq \ldots$ are easy to show. To prove "modified BDDH assumption $\leq$ 2-BDDHI assumption", assume there exists an adversary A that breaks the modified BDDH assumption. We show that then there exists an adversary B with oracle access to A that breaks the 2-BDDHI assumption. Let (h, h^a, h^{a^2}, W) be an input instance of the 2-BDDHI problem given to B. B's goal is to find out if $W = \hat{e}(h,h)^{1/a}$ or W is random. B picks two random values y_0, z_0 and defines its output bit as $\gamma := \gamma'$, where γ' is input from A as

$$\gamma' \leftarrow \mathsf{A}(h^{a^2}, h^a, h, h^{y_0}, h^{z_0}, W' = W^{y_0 z_0}).$$

We now show correctness. Defining $g := h^{a^2}$, $x = 1/a$, $y = y_0/a^2$, and $z = z_0/a^2$, we have $h^a = g^{1/a} = g^x$ and $h = g^{1/a^2} = g^{x^2}$. Consequently, $(h^{a^2}, h^a, h, h^{y_0}, h^{z_0}) = (g, g^x, g^{x^2}, g^y, g^z)$. If $W = \hat{e}(h,h)^{1/a}$, then

$$W' = W^{y_0 z_0} = \hat{e}(h,h)^{1/a \cdot y_0 \cdot z_0} = \hat{e}(g,g)^{1/a^5 \cdot y_0 z_0} = \hat{e}(g,g)^{1/a \cdot y_0/a^2 \cdot z_0/a^2} = \hat{e}(g,g)^{xyz}.$$

If W is a random element, so is W'. This proves the lemma. ∎