Universally Anonymous IBE based on the Quadratic Residuosity Assumption

Giuseppe Ateniese[1] and Paolo Gasti[2]

[1] Johns Hopkins University
[2] University of Genova

Abstract. We introduce the first universally anonymous, thus key-private, IBE whose security is based on the standard quadratic residuosity assumption. Our scheme is a variant of Cocks IBE (which is not anonymous) and it is efficient and highly parallelizable. We measure the performance of our scheme and compare it to the one of other popular IBE schemes.

1 Introduction

Identity-based encryption was introduced by Shamir in 1984 [23]. He asked whether it was possible to encrypt a message by just using the identity of the intended recipient. Several partial and inefficient solutions were proposed after Shamir's initial challenge but it was only in 2000 that Sakai et al. [21], Boneh and Franklin [8], and Cocks [11] came up with very practical solutions.

The Boneh-Franklin (BF) work has been the most influential of all: It did not just introduce the first practical IBE scheme but, more importantly, it showed how to correctly use pairings on elliptic curves, how to pick the right curves, how to encode and map elements into points, and introduced appropriate assumptions and definitions.

Cocks' scheme is *per se* revolutionary: It is the first IBE that does not use pairings but rather it works in standard RSA groups and its security relies on the standard quadratic residuosity assumption (within the random oracle model). Cocks IBE, however, encrypts the message bit by bit and thus it is considered very bandwidth consuming. On the other end, Cocks [11] observes that his scheme can be used in practice to encrypt short session keys in which case the scheme becomes very attractive. We may add that the importance of relying on such a standard assumption should not be underestimated. In fact, this is what motivated the recent work of Boneh, Gentry, and Hamburg [9] where a new space-efficient IBE scheme is introduced whose security is also based on the quadratic residuosity assumption. Unfortunately, as the authors point out [9], their scheme is not efficient, it is more expensive than Cocks IBE and in fact it is more expensive than all standard IBE and public-key encryption schemes since its complexity is quartic in the security parameter.

However, the scheme of Boneh et al. [9] has an important advantage over the scheme of Cocks: It provides anonymity, i.e., nobody can tell who the intended recipient is by just looking at the ciphertext. Anonymity, or key-privacy, is a very important property that was first studied by Bellare et al. [4]. Recipient anonymity can be used, for example, to thwart traffic analysis, to enable searching on encrypted data [7], or to anonymously broadcast messages [1]. Several IBE schemes provide anonymity, for instance the Boneh-Franklin scheme is anonymous. Other schemes that do not originally provide anonymity can be either properly modified [10] or adapted to work in the XDH setting [22][6][3][2].

At this point, it is natural to ask whether it is possible to enhance Cocks IBE and come up with a variant that provides anonymity and that, unlike Boneh et al.'s scheme [9], is as efficient as the original scheme of Cocks.

The first attempt in this direction has been proposed recently by Di Crescenzo and Saraswat [14]. They provide the first public-key encryption with keyword search (PEKS) that is not based on pairings. Although their scheme is suitable for PEKS, we note that when used as an IBE it becomes quite impractical: It uses four times the amount of bandwidth required by Cocks and it requires each user to store and use a very large number of secret keys (four keys per each bit of the plaintext, e.g., the recipient must store 512 secret keys and use 128 of them for each message in order to decrypt 128 bits). In addition, the security of their scheme

is based on a new assumption they introduce but later we will show that their assumption is equivalent to the standard quadratic residuosity assumption.

Universal anonymity is a new and exciting notion introduced at Asiacrypt 2005 by Hayashi and Tanaka [18]. An encryption scheme is universally anonymous if ciphertexts can be made anonymous by anyone and not just by whoever created the ciphertexts. Specifically, a universally anonymizable public-key encryption scheme consists of a standard public-key encryption scheme and two additional algorithms: one used to anonymize ciphertexts, which takes as input only the public key of the recipient, and the other is used by the recipient to decrypt anonymized ciphertexts.

The following observations are obvious but worth emphasizing: (1) A universally anonymous scheme is also key-private in the sense of Bellare et al. [4]. What makes universally anonymity interesting and unique is that anyone can anonymize ciphertexts using just the public key of the recipient. (2) Key-private schemes can be more expensive than their non-private counterparts. For instance, RSA-OAEP can be made key-private as shown in [4] but the new anonymous variant is more expensive. (3) The concept of universal anonymity makes sense also for schemes that are already key-private. For instance, ElGamal is key-private only by assuming that all keys are generated in the same group and participants share the same public parameters. But in many scenarios this is not the case. In PGP, for instance, parameters for each user are selected in distinct groups. Evidently, ElGamal applied in different algebraic groups is not anonymous anymore as one can test whether a given ciphertext is in a group or not.

Our contributions are:
(1) We enhance Cocks IBE and make it universally anonymous, and thus key-private in the sense of Bellare et al. [4]. Our variant of Cocks IBE can be seen as the most efficient anonymous IBE whose security is based on the quadratic residuosity assumption. The efficiency of our scheme is comparable to that of Cocks IBE. In fact, it is substantially more efficient than the recent scheme of Boneh et al. [9] and the IBE that derives from the PEKS construction by Di Crescenzo et al. [14]. In addition, the ciphertext expansion of our scheme is comparable to that of Cocks IBE.
(2) We implemented our variant and measured its performance. We show that in practice the efficiency of Cocks IBE and the variant we propose in this paper compare favorably even with that of the Boneh-Franklin scheme.
(3) Incidentally, our solutions and techniques can be used to notably simplify and improve the PEKS construction in [14]. In addition, we prove that the new security assumption introduced in [14] is actually equivalent to the standard quadratic residuosity assumption, thus making the resulting PEKS scheme the first one whose security is based solely on such a standard assumption. (More details are in Appendix A.)

2 Preliminaries

In this section, we recall first the IBE scheme proposed by Cocks [11]. Then we show that Cocks IBE is not anonymous due to a test proposed by Galbraith, as reported in [7]. Finally, we show that Galbraith's test is the "best test" possible against the anonymity of Cocks IBE. We assume that N is a large-enough RSA-type modulus. In addition, we will denote with $\mathbb{Z}_N^*[+1]$ ($\mathbb{Z}_N^*[-1]$) the set of elements in $\mathbb{Z}_N^*$ with Jacobi symbol $+1$ (-1, resp.) and with $\mathbb{QR}(N)$ the set of quadratic residues (or squares) in $\mathbb{Z}_N^*$. In our discussion, we will often omit to consider cases where randomly picked elements are in $\mathbb{Z}_N$ but not in $\mathbb{Z}_N^*$ or elements with Jacobi symbol over N equal to 0 since these cases occur only with negligible probability.

The security of Cocks IBE (and our variants) relies on the standard quadratic residuosity assumption which simply states that the two distributions $DQR(n) = \{(c, N) : N \xleftarrow{R} Gen(1^n),\ c \xleftarrow{R} \mathbb{QR}(N)\}$ and $DQRN(n) = \{(c, N) : N \xleftarrow{R} Gen(1^n),\ c \xleftarrow{R} \mathbb{Z}_N^*[+1] \setminus \mathbb{QR}(N)\}$ are computationally indistinguishable, where n is a security parameter and $Gen(\cdot)$ generates n-bit RSA-type moduli.

2.1 Cocks' IBE Scheme

Let $N = pq$ be an RSA-type modulus, where p and q are Blum-Williams primes. In addition, we consider $H : \{0,1\}^* \to \mathbb{Z}_N^*[+1]$ a full-domain hash which will be modeled as a random oracle in the security analysis.

Master Key: The secret key of the trusted authority is (p, q) while its public key is $N = pq$.

Key Generation: Given the identity ID, the authority generates $a = H(ID)$ (thus the Jacobi symbol $\left(\frac{a}{N}\right)$ is $+1$). The secret key for the identity ID is a value r randomly chosen in $\mathbb{Z}_N^*$ such that $r^2 \equiv a \bmod N$ or $r^2 \equiv -a \bmod N$. This value r is stored and returned systematically.

Encryption: To encrypt a bit $b = \{-1, +1\}$ for identity ID, choose uniformly at random two values $t, v \in \mathbb{Z}_N^*$, such that $\left(\frac{t}{N}\right) = \left(\frac{v}{N}\right) = b$, and compute:

$$(c, d) = \left(t + \frac{a}{t} \bmod N, v - \frac{a}{v} \bmod N\right)$$

Decryption: Given a ciphertext (c, d), first set $s = c$ if $r^2 \equiv a \bmod N$ or $s = d$ otherwise. Then, decrypt by computing:

$$\left(\frac{s + 2r}{N}\right) = b$$

Notice that $s + 2r \equiv w(1 + r/w)^2 \bmod N$, thus the Jacobi symbol of $s + 2r$ is equal to that of w, where w is either t or v.

2.2 Galbraith's Test (GT)

As mentioned in the paper by Boneh et al. [7], Galbraith showed that Cocks' scheme is not anonymous. Indeed, let $a \in \mathbb{Z}_N^*[+1]$ be a public key and consider the following set:

$$S_a[N] = \left\{t + \frac{a}{t} \mid t \in \mathbb{Z}_N^*\right\}$$

(One can easily deduce from the proof provided by Cocks [11] that the size of $S_a[N]$ is $1/4$ the size of $\mathbb{Z}_N^*$.) Given two random public keys $a, b \in \mathbb{Z}_N^*[+1]$, Galbraith's test (which we will denote with "$GT(\cdot)$") allows us to distinguish the uniform distribution on the set $S_a[N]$ from the uniform distribution on the set $S_b[N]$. Given $c \in \mathbb{Z}_N^*$, the test over the public key a is defined as the Jacobi symbol of $c^2 - 4a$ over N, that is:

$$GT(a, c, N) = \left(\frac{c^2 - 4a}{N}\right)$$

Notice that when c is sampled from $S_a[N]$, the test $GT(a, c, N)$ will always return $+1$ given that $c^2 - 4a = (t - (a/t))^2$ is a square. However, if c is sampled from $S_b[N]$ the test is expected to return $+1$ with probability negligibly close to $1/2$ since, in this case, the distribution of the Jacobi symbol of the element $c^2 - 4a$ in $\mathbb{Z}_N^*$ follows the uniform distribution on $\{-1, +1\}$.

It is mentioned in [7] that since Cocks ciphertext is composed of several values sampled from either $S_a[N]$ (and $S_{-a}[N]$) or $S_b[N]$ (and $S_{-b}[N]$, respectively), then an adversary can repeatedly apply Galbraith's test to determine with overwhelming probability whether a given ciphertext is intended for a or b. However, one must first prove some meaningful results about the distribution of Jacobi symbols of elements of the form $c^2 - 4b$ in $\mathbb{Z}_N^*$, for *fixed* random elements $a, b \in \mathbb{Z}_N^*[+1]$ and for $c \in S_a[N]$. These results are reported in the next section.

2.3 Relevant Lemmata and Remarks

Damgård in [15] studied the distribution of Jacobi symbols of elements in $\mathbb{Z}_N^*$ in order to build pseudo-random number generators. In his paper, Damgård reports of a study performed in the 50s by Perron in which it is proven that for a prime p and for any a, the set $a + \mathbb{QR}(p)$ contains as many squares as non squares in $\mathbb{Z}_p^*$ when $p \equiv 1 \mod 4$, or the difference is just 1 when $p \equiv 3 \mod 4$. It might be possible to generalize Perron's result to study the set $a + \mathbb{QR}(N)$ in $\mathbb{Z}_N^*$ but we also point out that the security of Cocks IBE implicitly depends on the following Lemma:

Lemma 1. *The distribution* $\left\{\left(\frac{t^2+a}{N}\right) : N \xleftarrow{R} Gen(1^n),\ a \xleftarrow{R} \mathbb{Z}_N^*[+1],\ t \xleftarrow{R} \mathbb{Z}_N^*\right\}$ *is computationally indistinguishable from the uniform distribution on* $\{-1,+1\}$ *under the quadratic residuosity assumption.*

To prove the Lemma above it is enough to observe that if we compute the Jacobi symbol of a $c \in S_a[N]$ we obtain:

$$\left(\frac{c}{N}\right) = \left(\frac{(t^2+a)/t}{N}\right) = \left(\frac{t^2+a}{N}\right)\left(\frac{t}{N}\right)$$

However the Jacobi symbol of t over N is the plaintext in Cocks IBE and thus Lemma 1 must follow otherwise the CPA-security of Cocks IBE would not hold.

Remark. Let's pick c randomly in $\mathbb{Z}_N^*$. If $GT(a,c,N) = -1$, we can clearly conclude that $c \notin S_a[N]$. However, if $GT(a,c,N) = +1$, what is the probability that $c \in S_a[N]$? The answer is $1/2$ since a t can be found such that $c = t + a/t$ whenever $c^2 - 4a$ is a square (and this happens only half of the times). To summarize:

$$GT(a,c,N) = \begin{cases} +1 \Longrightarrow c \in S_a[N] \text{ with prob. } 1/2 \\ -1 \Longrightarrow c \notin S_a[N] \end{cases}$$

We argue next that there is no *better* test against anonymity over an encrypted bit. That is, we show that a test that returns $+1$ to imply that $a \in S_a[N]$ with probability $1/2 + \delta$ (for a non-negligible $\delta > 0$) cannot exist under the quadratic residuosity assumption. We first notice that $c \in S_a[N]$ if and only if $\Delta = c^2 - 4a$ is a square. Indeed, if $c = t + a/t$ then $\Delta = (t - a/t)^2$. If Δ is a square then the quadratic equation $t(t + a/t) = ct$ has solutions for t in $\mathbb{Z}_N^*$ with overwhelming probability. Intuitively, we can see Galbraith's test as an algorithm that checks whether the discriminant Δ has Jacobi symbol $+1$ or -1, and this is clearly *the best it can do* since the factors of the modulus N are unknown. (Remember that we do not consider cases where the Jacobi symbol is 0 since they occur with negligible probability.) We now prove it formally. We will denote the improved, or *ideal test*, with $IT(a,c,N)$ and define it such that it returns either $+1$ or -1. If it returns -1, it implies that $c \notin S_a[N]$. If it returns $+1$, it implies that $c \in S_a$ with probability $1/2 + \delta$, where $0 < \delta \leq 1/2$ (ideally $\delta = 1/2$). That is:

$$IT(a,c,N) = \begin{cases} +1 \Longrightarrow c \in S_a[N] \text{ with prob. } 1/2 + \delta \\ -1 \Longrightarrow c \notin S_a[N] \end{cases}$$

We make no assumptions on how $IT(\cdot)$ operates and we evaluate it via oracle access. Let x be a random element in $\mathbb{Z}_N^*[+1]$. The simulation proceeds as follows:

1. Find a random $r \in \mathbb{Z}_N^*$ such that $b = (r^2 - x)/4$ has Jacobi symbol $+1$ (see Lemma 1);
2. If $IT(b,r,N) = +1$ then output *"x is a square"* otherwise output *"x is not a square"*.

Notice that $r \in S_b[N]$ if and only if $r^2 - 4b$ is a square. But $r^2 - 4b = x$, so we proved that $IT(\cdot)$ cannot exist under the quadratic residuosity assumption.

We denote with $GT_a^N[+1]$ the set $\{x \in \mathbb{Z}_N^* \mid GT(a,x,N) = +1\}$. In other words, $GT_a^N[+1]$ is the set of all the elements in $\mathbb{Z}_N^*$ such that Galbraith's test for that element and public key a returns $+1$. Analogously, we can define $GT_a^N[-1]$ as the set $\{x \in \mathbb{Z}_N^* \mid GT(a,x,N) = -1\}$. The next Lemma follows directly from the discussion above:

Lemma 2. [VQR–Variable Quadratic Residuosity] *The distributions* $D_0(n) = \left\{(a,c,N) : N \xleftarrow{R} Gen(1^n),\ a \xleftarrow{R} \mathbb{Z}_N^*[+1],\ c \xleftarrow{R} S_a[N]\right\}$ *and* $D_1(n) = \left\{(a,c,N) : N \xleftarrow{R} Gen(1^n),\ a \xleftarrow{R} \mathbb{Z}_N^*[+1],\ c \xleftarrow{R} GT_a^N[+1] \setminus S_a[N]\right\}$ *are computationally indistinguishable under the quadratic residuosity assumption.*

The next Lemma easily follows from Lemma 1 since $c^2 - 4b$ can be written as $c^2 + h$ for a fixed $h \in \mathbb{Z}_N^*[+1]$.

Lemma 3. *The distribution* $\left\{GT(b,c,N) : N \xleftarrow{R} Gen(1^n),\ b \xleftarrow{R} \mathbb{Z}_N^*[+1],\ c \xleftarrow{R} \mathbb{Z}_N^*\right\}$ *is computationally indistinguishable from the uniform distribution on* $\{-1,+1\}$.

3 Our Basic Construction and Its Efficient Variants

We extend Cocks' scheme to support anonymity. Unlike previous proposals, our scheme UAnonIBE has efficiency, storage, and bandwidth requirements similar to those of the original scheme by Cocks (which is not anonymous). Our scheme is also the first universally anonymous IBE, according to the definition in [18].

3.1 The Basic Scheme

Let $H : \{0,1\}^* \rightarrow \mathbb{Z}_N^*[+1]$ a full-domain hash modeled as a random oracle. Let n and m be two security parameters. The algorithms which form UAnonIBE are defined as follows (all operations are performed modulo N):

Master Key: The public key of the trusted authority is the n-bit integer $N = pq$, where p and q are $n/2$-bit Blum-Williams primes (i.e., both congruent to 3 modulo 4).

Key Generation: Given the identity ID, the authority generates $a = H(ID)$ (thus the Jacobi symbol $\left(\frac{a}{N}\right)$ is $+1$). The secret key for the identity ID is a value r randomly chosen in $\mathbb{Z}_N^*$ such that $r^2 \equiv a \bmod N$ or $r^2 \equiv -a \bmod N$. This value r is stored and returned systematically.

Encryption: To encrypt a bit $b = \{-1, +1\}$ for identity ID, choose uniformly at random two values $t, v \in \mathbb{Z}_N^*$, such that $\left(\frac{t}{N}\right) = \left(\frac{v}{N}\right) = b$, and compute $(c, d) = \left(t + \frac{a}{t}\ , v - \frac{a}{v}\right)$.
Then, compute the *mask* to anonymize the ciphertext (c, d) as follows:

1. Pick two indices k_1 and k_2 independently from the geometric distribution D[1] with probability parameter $1/2$;
2. Select random T, V in $\mathbb{Z}_N^*$ and set $Z_1 = c + T$ and $Z_2 = d + V$;
3. For $1 \leq i < k_1$, select random values $T_i \in \mathbb{Z}_N^*$ s.t. $GT(a, Z_1 - T_i, N) = -1$;
4. For $1 \leq i < k_2$, select random values $V_i \in \mathbb{Z}_N^*$ s.t. $GT(-a, Z_2 - V_i, N) = -1$;
5. Set $T_{k_1} = T$ and $V_{k_2} = V$;
6. For $k_1 < i \leq m$, select random values $T_i \in \mathbb{Z}_N^*$;
7. For $k_2 < i \leq m$, select random values $V_i \in \mathbb{Z}_N^*$;

Finally, output $(Z_1, T_1, \ldots, T_m)$ and $(Z_2, V_1, \ldots, V_m)$.

Decryption: Given a ciphertext $(Z_1, T_1, \ldots, T_m)$ and $(Z_2, V_1, \ldots, V_m)$, first discard one of the two tuples based on whether a or $-a$ is a square. Let's assume we keep the tuple $(Z_1, T_1, \ldots, T_m)$ and we discard the other. In order to decrypt, find the smallest index $1 \leq i \leq m$ s.t. $GT(a, Z_1 - T_i, N) = +1$ and output:

$$\left(\frac{Z_1 - T_i + 2r}{N}\right) = b$$

We run the same procedure above if the second tuple is actually selected and the first is discarded. It is enough to replace a with $-a$, Z_1 with Z_2, and T_i with V_i.

3.2 Security Analysis

We need to show that our scheme, UAnonIBE, is ANON-IND-ID-CPA-secure [1, 9], that is, the ciphertext does not reveal any information about the plaintext and an adversary cannot determine the identity under which an encryption is computed, even thought the adversary selects the identities and the plaintext.

In [17], Halevi provides a sufficient condition for a CPA public-key encryption scheme to meet the notion of key-privacy, or anonymity, as defined by Bellare et al. in [4]. In [1], Abdalla et al. extend Halevi's condition

[1] The geometric distribution is a discrete memoryless random distribution for $k = 1, 2, 3, \ldots$ having probability function $P(k) = p(1-p)^{k-1}$ where $0 < p < 1$. Therefore, for $p = 1/2$ the probability that $k_1 = k$ is 2^{-k}. For more details see [27]

to identity-based encryption. In addition, their notion is defined within the random oracle model and Halevi's statistical requirement is weakened to a computational one. Informally, it was observed that if an IBE scheme is already IND-ID-CPA-secure then the oracle does not have to encrypt the message chosen by the adversary but can encrypt a random message of the same length. The game where the oracle replies with an encryption on a random message is called ANON-RE-CPA. In [1], it was shown that if a scheme is IND-ID-CPA-secure and ANON-RE-CPA-secure then it is also ANON-IND-ID-CPA-secure.

ANON-RE-CPA game. We briefly describe the security game introduced by Abdalla et al. in [1]. MPK represents the set of public parameters of the trusted authority. The adversary A has access to a random oracle H and to an oracle $KeyDer$ that given an identity ID returns the private key for ID according to the IBE scheme.

Experiment $\mathbf{Exp}^{\mathbf{anon-re-cpa-b}}_{\mathbf{IBE,A}}(n)$:
pick random oracle H
$(ID_0, ID_1, msg, state) \leftarrow A^{KeyDer(\cdot),H}(find, MPK)$
$W \xleftarrow{R} \{0,1\}^{|msg|}; C \leftarrow Enc^H(MPK, ID_b, W)$
$b' \leftarrow A^{KeyDer(\cdot),H}(guess, C, state)$
return b'

The adversary cannot request the private key for ID_0 or ID_1 and the message msg must be in the message space associated with the scheme. The ANON-RE-CPA-advantage of an adversary A in violating the anonymity of the scheme IBE is defined as

$$\mathbf{Adv}^{\mathbf{anon-re-cpa}}_{\mathbf{IBE,A}}(n) = P\left[\mathbf{Exp}^{\mathbf{anon-re-cpa-1}}_{\mathbf{IBE,A}}(n) = 1\right] - P\left[\mathbf{Exp}^{\mathbf{anon-re-cpa-0}}_{\mathbf{IBE,A}}(n) = 1\right]$$

A scheme is said to be ANON-RE-CPA-secure if the above advantage is negligible in n.

Theorem 1. UAnonIBE *is* ANON-IND-ID-CPA-*secure in the random oracle model under the quadratic residuosity assumption.*

In order to simplify the proof of theorem 1, we make and prove an important claim first. We will show that a ciphertext for a random $a \in \mathbb{Z}^*_N[+1]$ is indistinguishable from a sequence of random elements in $\mathbb{Z}^*_N$ to a PPT distinguisher $\mathcal{DS}$. In particular, let $\mathcal{O}\{S_a[N], GT^N_a[-1], GT^N_a[+1]\}$ be an oracle that returns UAnonIBE encryptions under public key a of random messages and let $\mathcal{O}^*$ an oracle that returns a $(m+1)$-tuple of elements picked uniformly at random from $\mathbb{Z}^*_N$. We prove the following:

Claim. The distinguisher $\mathcal{DS}$ has only negligible advantage in distinguishing the outputs of the oracles $\mathcal{O}\{S_a[N], GT^N_a[-1], GT^N_a[+1]\}$ and $\mathcal{O}^*$ under the quadratic residuosity assumption.

Proof. Let $(Z_1, T_1, \ldots, T_m)$ be the output of $\mathcal{O}\{S_a[N], GT^N_a[-1], GT^N_a[+1]\}$. In particular, a $c \in S_a[N]$ is randomly picked and Z_1 is set to $c + T_k$, where k is chosen according to the geometric distribution D defined in the UAnonIBE encryption algorithm. Let $(U_0, U_1, \ldots, U_m)$ be the output of $\mathcal{O}^*$.

First, notice that there exists a minimal index k' such that $GT(a, U_0 - U_{k'}, N) = +1$. Such an index exists with probability $1 - 2^{-m}$ because of Lemma 3. Second, it is easy to see that the distribution induced by the index k' is indistinguishable from the distribution D. This still follows from Lemma 3 since we know that the probability that $GT(a, U_0 - U_i, N) = +1$ is negligibly close to $1/2$, for $1 \leq i \leq m$. Hence, the probability that $k' = v$, for a positive integer $v \in \mathbb{N}$, is negligibly close to 2^{-v}. Thus, both indices k and k' determine the same distribution except with negligible probability (to account for the cases where Galbraith's test returns 0). Finally, because of Lemma 2, $\mathcal{DS}$ cannot determine whether $U_0 - U_{k'} \in GT^N_a[+1]$ is in $S_a[N]$ or not.

Remark. We point out an insightful analogy between the oracle $\mathcal{O}\{S_a[N], GT^N_a[-1], GT^N_a[+1]\}$ in the claim above and the oracle $\mathcal{O}\{\mathbb{QR}(N), \mathbb{Z}^*_N[-1], \mathbb{Z}^*_N[+1]\}$ which picks elements $c \in \mathbb{QR}(N)$ (rather than in $S_a[N]$) and sets $Z_1 = c + T_k$, where k is chosen according to D. Then, it generates elements $T_1, \ldots, T_m$ such that: (1) $Z_1 - T_i \in \mathbb{Z}^*_N[-1]$, for $1 \leq i < k$, (2) $Z_1 - T_k \in \mathbb{QR}(N)$, and (3) $Z_1 - T_i \in \mathbb{Z}^*_N$, for $k < i \leq m$. Evidently,

even the outputs of this oracle are indistinguishable from the outputs of $\mathcal{O}^*$ under the quadratic residuosity assumption.

Proof of Theorem 1. It must be clear that UAnonIBE is IND-ID-CPA-secure since Cocks IBE is IND-ID-CPA-secure in the random oracle model under the quadratic residuosity assumption and the mask is computed without knowing the plaintext or the secret key of the intended recipient. Thus, we only need to show that UAnonIBE is ANON-RE-CPA-secure. But, because of Claim 3.2, a PPT adversary A must have negligible advantage in determining whether the ciphertext C returned by $ENC^H(\cdot)$ is for ID_0 or ID_1 because C is (with overwhelming probability) a proper encryption for both ID_0 and ID_1 on two *random* bits. (It is equivalent to respond to A with two $(m+1)$-tuples of random elements in $\mathbb{Z}_N^*$.) □

3.3 A First Efficient Variant: Reducing Ciphertext Expansion

The obvious drawback of the basic scheme is its ciphertext expansion. Indeed, for each bit of the plaintext $2 \cdot (m+1)$ values in $\mathbb{Z}_N^*$ must be sent while in Cocks IBE each bit of the plaintext requires two values in $\mathbb{Z}_N^*$. Therefore, we need a total of $2 \cdot (m+1) \cdot n$ bits for a single bit in the plaintext, where n and m are the security parameters (e.g., $n = 1024$ and $m = 128$). However, this issue is easy to fix. Intuitively, since our scheme requires the random oracle model for its security, we could use another random oracle that expands a short seed into a value selected uniformly and independently in $\mathbb{Z}_N^*$.

Specifically, a function $G : \{0,1\}^* \to \mathbb{Z}_N^*$ is used, which we model as a random oracle, that maps a e-bit string α to a random value in $\mathbb{Z}_N^*$. The parameter e must be large enough, e.g., $e = 160$.

It is tempting to use the oracle G and a single short seed α plus a counter to generate all values $T_1, \ldots, T_m$ and $V_1, \ldots, V_m$. This first solution would provide minimal ciphertext expansion, since only the seed α must be sent, however it may turn out to be computationally expensive. To see this, consider for example that an α must be found such that $GT(a, Z_1 - T_i, N) = -1$ for $1 \leq i < k_1$. Now, if k_1 happens to be large, say $k_1 = 20$, then clearly finding a suitable α could be computationally intensive. Nevertheless, we prove that this scheme is secure as long as the basic UAnonIBE scheme is secure. More importantly, we emphasize that the proof of security of all other schemes proposed after this first one can easily be derived from the proof of the following theorem.

Theorem 2. *The first efficient variant of* UAnonIBE *is* ANON-IND-ID-CPA-*secure in the random oracle model under the quadratic residuosity assumption.*

Proof. We let the simulator S play the role of a man-in-the-middle attacker between two ANON-RE-CPA games: the first game is against the basic UAnonIBE and the second game is against an adversary A that has non-negligible advantage in breaking the first variant of UAnonIBE. We show that S can use A to win in the first ANON-RE-CPA game, thus violating the quadratic residuosity assumption. The simulation is straightforward: S forwards the H-queries and $KeyDer$-queries to the respective oracles. When A challenges for identities ID_0 and ID_1, S challenges on the same identities in the first ANON-RE-CPA game. Then S receives the ciphertext $(Z_1, T_1, \ldots, T_m)$, $(Z_2, V_1, \ldots, V_m)$. S sends to A, $(Z_1, \alpha), (Z_2, \beta)$ where α and β are chosen uniformly at random in $\{0,1\}^e$. At this point, the simulator responds to the G-queries as follows:

$$G(\alpha \,||\, i) = T_i \text{ and } G(\beta \,||\, i) = V_i \text{ , for } 0 < i \leq m,$$

and with random values in $\mathbb{Z}_N^*$ in any other cases. The adversary A eventually returns its guess which S uses in the first game in order to win with non-negligible advantage.

The obvious next-best solution is to use a single seed per value. Thus, rather than sending the ciphertext as per our basic scheme, that is $(Z_1, T_1, \ldots, T_m)$ and $(Z_2, V_1, \ldots, V_m)$, the following values could be sent:

$$(Z_1, \alpha_1, \ldots, \alpha_m) \text{ and } (Z_2, \beta_1, \ldots, \beta_m),$$

where each α_i, β_i are selected randomly in $\{0,1\}^e$ such that conditions 3. and 4. of the encryption algorithm of the basic scheme are satisfied. The recipient would then derive the intended ciphertext by computing

$T_i = G(\alpha_i)$ and $V_i = G(\beta_i)$, for $1 \leq i < m$. If we set e to be large enough, say $e = 160$, then clearly the security of this variant is equivalent to the one of the basic scheme in the random oracle model and a single bit of the plaintext would require $2 \cdot (m \cdot e + n)$ bits rather than $2 \cdot (m \cdot n + n)$, where $e < n$. Hence, for $n = 1024$, $m = 128$ and $e = 160$, we need to send $2 \cdot (160 \cdot 128 + 1024)$ bits while Cocks' scheme requires only $2 \cdot (1024)$ bits.

On a closer look however, it is easy to see that since G is a random oracle we just need to ensure that its inputs are repeated only with negligible probability. Let $X = x^{(1)}x^{(2)} \dots x^{(t)}$ be the plaintext of t bits. For each plaintext X, the sender selects a random message identifier $MID_X \in \{0,1\}^{e_1}$ which is sent along with the ciphertext. For bit $x^{(j)}$, the sender computes:

$$(Z_1^{(j)}, \alpha_1^{(j)}, \dots, \alpha_m^{(j)}) \text{ and } (Z_2^{(j)}, \beta_1^{(j)}, \dots, \beta_m^{(j)}),$$

where the coefficients $\alpha_i^{(j)}$, $\beta_i^{(j)}$ are selected randomly in $\{0,1\}^e$ such that conditions 3. and 4. of the encryption algorithm of the basic scheme are satisfied (thus notice that e can be small but still big enough to be able to find those values $T_i^{(j)}$ and $V_i^{(j)}$ that satisfy such conditions). The recipient will derive the intended ciphertext by computing:

$$T_i^{(j)} = G(MID_X \,||\, 0 \,||\, \alpha_i^{(j)} \,||\, i \,||\, j) \text{ or } V_i^{(j)} = G(MID_X \,||\, 1 \,||\, \beta_i^{(j)} \,||\, i \,||\, j),$$

where $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, t\}$. As an example, we can set $m = 128$, $e_1 = 160$, and $e = 8$. In this case the ciphertext expansion per single bit of the plaintext is only $2 \cdot (1024 + 1024)$ bits which is twice the amount required by Cocks IBE for $n = 1024$. (In addition an extra 160 bits are needed for MID_X but these bits are transmitted only once per message.)

3.4 A Second Efficient Variant: Trade-off Between Ciphertext Expansion and Performance

We propose a second variant of UAnonIBE which provides an optimal trade-off between efficiency and ciphertext expansion. We fix a new global parameter ℓ which is a small positive integer. Let $X = x^{(1)}x^{(2)} \dots x^{(t)}$ be the plaintext of t bits. For each plaintext X, the sender selects a random identifier $MID_X \in \{0,1\}^{e_1}$ which is sent along with the ciphertext. For bit $x^{(j)}$, the sender computes:

$$(Z_1^{(j)}, \alpha_1^{(j)}, \dots, \alpha_\ell^{(j)}) \text{ and } (Z_2^{(j)}, \beta_1^{(j)}, \dots, \beta_\ell^{(j)})$$

where α_i and β_i are both in $\{0,1\}^e$ when $i < \ell$, α_ℓ and β_ℓ are in $\{0,1\}^{e'}$, for some $e' > e$. The intended ciphertext is derived by the recipient by computing:

$$T_i^{(j)} = G(MID_X \,||\, 0 \,||\, \alpha_i^{(j)} \,||\, i \,||\, j) \text{ or } V_i^{(j)} = G(MID_X \,||\, 1 \,||\, \beta_i^{(j)} \,||\, i \,||\, j)$$

for $i < l$, and

$$T_i^{(j)} = G(MID_X \,||\, 0 \,||\, \alpha_\ell^{(j)} \,||\, i \,||\, j) \text{ or } V_i^{(j)} = G(MID_X \,||\, 1 \,||\, \beta_\ell^{(j)} \,||\, i \,||\, j)$$

for $i \geq \ell$. Note that in this variant both the sender and the receiver can generate an arbitrary number of $T_i^{(j)}$ and $V_i^{(j)}$ (i.e., there is no fixed global parameter m).

Given the distribution of k_1, k_2, for a large enough ℓ, we expect $k_1 \leq \ell$ or $k_2 \leq \ell$ with high probability. When $k_1 \leq \ell$ or $k_2 \leq \ell$ the scheme is as efficient as the first variant. When $k_1 > \ell$ (or $k_2 > \ell$) the computational cost of finding a value for $\alpha_\ell^{(j)}$ (or $\beta_\ell^{(j)}$) is exponential in $k_1 - \ell$ ($k_2 - \ell$, respectively).

As an example, we set the global parameter $\ell = 6$ and then $e_1 = 160$, $e = 8$, $e' = 80$, and $n = 1024$. The ciphertext expansion of this variant of UAnonIBE is $2 \cdot ((\ell - 1) \cdot e + e' + n)$, therefore, the ciphertext size for a single bit of the plaintext is now only $2 \cdot (120 + 1024)$ bits which is very close to the number of bits $(2 \cdot (1024))$ required by Cocks IBE (which is not anonymous). Note that for each message, the sender also transmits the random message identifier MID_X.

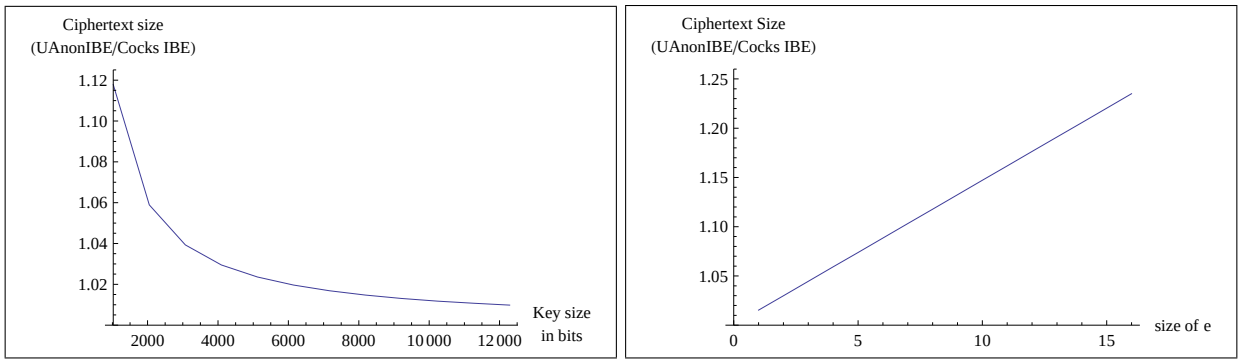

Fig. 1. The two graphs show UAnonIBE's ciphertext size relative to Cocks' scheme. In the first graph we show how the relative bandwidth overhead introduced by our solution decreases with the size of the master parameter. In the second graph we show how the size of the ciphertext increases, varying the size of e and fixing $e' = 10 \cdot e$, compared to Cocks' ciphertext.

4 Optimizations and Implementation

An important aspect that should be considered in order to implement UAnonIBE efficiently is the value of the parameters ℓ, e (which corresponds to the size of the values $\alpha_1^i, \ldots, \alpha_{\ell-1}^i$) and e' (the size of α_ℓ^i). These values affect both the ciphertext expansion and the encryption time significantly, therefore they must be selected carefully. Choosing e or e' to be too small can reduce the probability of encrypting to an unacceptable level. Choosing ℓ to be too small can make the encryption process very slow. If we set $e = 8$ and $e' = 80$, we can find a suitable value for each α_j^i, and therefore encrypt, with a probability of at least $1 - 2^{-80}$. We found that the value $\ell = 6$ is the best compromise between encryption time and ciphertext expansion. Setting $e = 8$, $e' = 80$ and $\ell = 6$, the ciphertext expansion for a 128-bit message is equal to 3840 bytes more than a Cocks encryption for both $+a$ and $-a$: for a 1024-bit modulus N the encrypted message size is about 36KB instead of 32KB with Cocks IBE.

Size of e	2	4	6	8	10	Cocks IBE
Ciphertext size (bytes)	33748	34708	35668	36628	37588	32768

We have implemented the second efficient variant of UAnonIBE and compared it with the original Cocks IBE [11] and the scheme proposed by Boneh and Franklin based on pairings [8]. We have two goals in mind. The first is to show that Cocks IBE and our schemes are practical when used as hybrid encryption algorithms (following the KEM-DEM paradigm, see Appendix B), even when compared with the Boneh-Franklin IBE, with the clear advantage compared to other IBE schemes of relying on a well-established assumption. Our second goal is to show that the efficiency of our scheme is comparable to that of the original scheme by Cocks.

For our performance analysis, we set the size of the values $\alpha_i^{(j)}$ and $\beta_i^{(j)}$ with $1 \leq i < \ell$ to 8 bits and the size of $\alpha_\ell^{(j)}$ and $\beta_\ell^{(j)}$ to 80 bits. However, our tests showed that the size of those parameters have no measurable impact on the performance of the scheme. In order to calculate the optimal value for ℓ, we measured the time required to anonymize a key of 128 bit (for a total of 256 encrypted bits, considering both cases $+a$ and $-a$). Figure 2 summarizes our results. The value $\ell = 6$ seems to be optimal, since further increasing ℓ does not noticeably affect the time required to anonymize a message or the decryption time.

Experimental Setup. We employed the MIRACL software package, developed by Shamus Software [24], to run our tests. MIRACL is a comprehensive library often used to implement cryptographic systems based on pairings. We used the optimized implementation of the Boneh-Franklin IBE provided by the library and

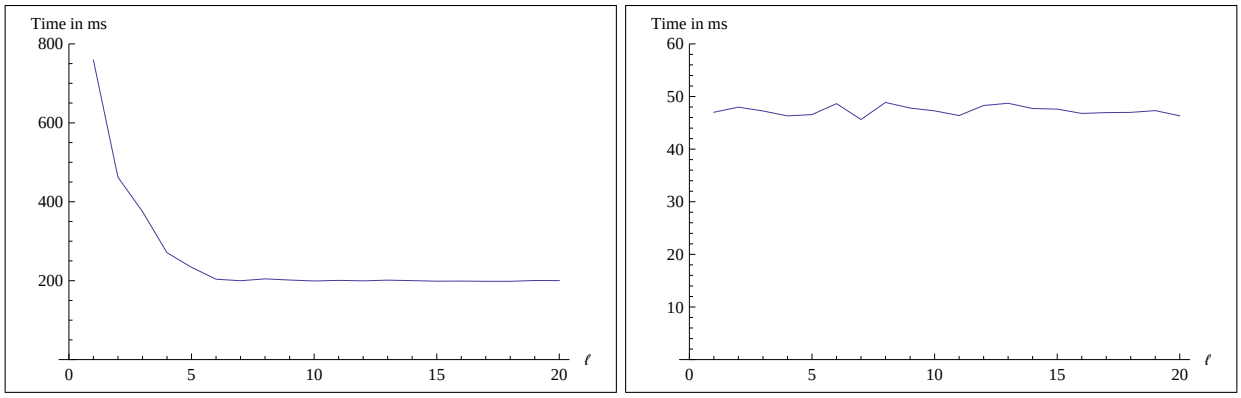

Fig. 2. The two graphs show the average time required respectively to anonymize and de-anonymize a 128-bit message encrypted with Cocks algorithm varying ℓ.

we implemented Cocks IBE and our scheme with an RSA modulus of 1024 bits. The implementation of the Boneh-Franklin IBE uses a 512-bit prime p, Tate pairing and a small 160-bit subgroup q. The curve used is $y^2 = x^3 + x$ instead of $y^2 = x^3 + 1$ because it allows for a faster implementation. Those two settings should provide the same level of security according to NIST [20]. The tests were run on a machine that consisted of an Intel Pentium 4 2.8GHz with 512MB RAM. The system was running Linux kernel 2.6.20 with the compiler GCC 4.1.2. We implemented the cryptographic primitives using version 5.2.3 of the MIRACL library. Every source file was compiled with optimization '-02', as suggested in the MIRACL documentation.

The table below shows average times over 1000 runs of the cryptographic operations specified in the first column. The symmetric key encrypted in our tests is of 128 bits.

	Extract	Encrypt	Decrypt	Anonymous	Universally Anonymous
Boneh-Franklin	9.1 ms	91.6 ms	85.4 ms	YES	NO
Cocks IBE	14.2 ms	115.3 ms	35.0 ms	NO	NO
Our Scheme	14.2 ms	319.4 ms	78.1 ms	YES	YES

In the table we also indicate whether a scheme is anonymous or not. Cocks IBE is not anonymous while Boneh-Franklin IBE is anonymous but not universally anonymous. One could try to turn Boneh-Franklin IBE into a universally anonymous scheme using for example the techniques in [18]. But, even assuming that this is possible, the new scheme would be different and more expensive than the original one and still depending on pairing-based assumptions.

5 Conclusions

We proposed UAnonIBE: the first IBE providing universal anonymity (thus key-privacy) and secure under the standard quadratic residuosity assumption. The efficiency and ciphertext expansion of our scheme are comparable to those of Cocks IBE. We showed that Cocks IBE and our anonymous variant are suitable in practice whenever hybrid encryption (KEM-DEM paradigm) is employed. We believe our schemes are valid alternatives to decidedly more expensive schemes introduced in [9] (which, in addition, are anonymous but not universally-anonymous). Incidentally, our results can be used to simplify existing PEKS constructions [14] and base their security on the standard quadratic residuosity assumption (which was left as an open problem in the area).

References

1. M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. MaloneLee, G. Neven, P. Paillier, and H. Shi. Searchable Encryption Revisited: Consistency Properties, Relation to Anonymous IBE, and Extensions. In *Advances in Cryptology, CRYPTO '05, volume 3621 of Lecture Notes in Computer Science*, pages 205–222. Springer-Verlag, 2005.
2. G. Ateniese, J. Camenisch, and B. de Medeiros. Untraceable RFID Tags via Insubvertible Encryption. In *CCS '05: Proceedings of the 12th ACM conference on Computer and communications security*, pages 92–101, New York, NY, USA, 2005. ACM.
3. L. Ballard, M. Green, B. de Medeiros, and F. Monrose. Correlation-Resistant Storage via KeywordSearchable Encryption. In *Cryptology ePrint Archive, Report 2005/417*, 2005. Available at `http://eprint.iacr.org/2005/417`.
4. M. Bellare, A. Boldyreva, A. Desai, and D. Pointcheval. Key-Privacy in Public-Key Encryption. In *Advances in Cryptology, ASIACRYPT '01, volume 2248 of Lecture Notes in Computer Science*, pages 566–582, London, UK, 2001. Springer-Verlag.
5. K. Bentahar, P. Farshim, J. Malone-Lee, and N. Smart. Generic Constructions of Identity-Based and Certificateless KEMs. In *Cryptology ePrint Archive, Report 2005/058*, 2005. Available at `http://eprint.iacr.org/2005/058`.
6. D. Boneh, X. Boyen, and H. Shacham. Short Group Signatures. In *Advances in Cryptology, CRYPTO '04, volume 3152 of Lecture Notes in Computer Science*, pages 41–55. Springer-Verlag, 2004.
7. D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano. Public Key Encryption with Keyword Search. In *Advances in Cryptology, EUROCRYPT '04, volume 3027 of Lecture Notes in Computer Science*, pages 506–522, Interlaken, Switzerland, 2004.
8. D. Boneh and M. Franklin. Identity-Based Encryption from the Weil Pairing. *SIAM Journal on Computing, volume 32-3*, pages 586–615, Philadelphia, PA, USA, 2003.
9. D. Boneh, C. Gentry, and M. Hamburg. Space-Efficient Identity Based Encryption Without Pairings. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 647–657, Washington, DC, USA, 2007. IEEE Computer Society.
10. X. Boyen and B. Waters. Anonymous Hierarchical Identity-Based Encryption (Without Random Oracles). In *Advances in Cryptology, CRYPTO '06, volume 4117 of Lecture Notes in Computer Science*, pages 290–307. Springer-Verlag, 2006.
11. C. Cocks. An Identity Based Encryption Scheme Based on Quadratic Residues. In *Proceedings of the 8th IMA International Conference on Cryptography and Coding*, pages 360–363, London, UK, 2001. Springer-Verlag.
12. R. Cramer and V. Shoup. A Practical Public Key Cryptosystem Provably Secure against Adaptive Chosen Ciphertext Attack. In *Advances in Cryptology, CRYPTO '98, volume 1462 of Lecture Notes in Computer Science*, pages 13–25. Springer-Verlag, 1998.
13. R. Cramer and V. Shoup. Design and Analysis of Practical Public-Key Encryption Schemes Secure against Adaptive Chosen Ciphertext Attack. *SIAM Journal on Computing, volume 33-1*, pages 167–226, Philadelphia, PA, USA, 2004.
14. G. Di Crescenzo and V. Saraswat. Public Key Encryption with Searchable Keywords Based on Jacobi Symbols. In *Progress in Cryptology, INDOCRYPT '07, volume 3797 of Lecture Notes in Computer Science*, pages 282–296, Chennai, India, December 9-13, 2007.
15. I. Damgård. On the Randomness of Legendre and Jacobi Sequences. In *Advances in Cryptology, CRYPTO '88, volume 403 of Lecture Notes in Computer Science*, pages 163–172, London, UK, 1990. Springer-Verlag.
16. E. Fujisaki and T. Okamoto. Secure Integration of Asymmetric and Symmetric Encryption Schemes. In *Advances in Cryptology, CRYPTO '99, volume 1666 of Lecture Notes in Computer Science*, pages 537–554, London, UK, 1999. Springer-Verlag.
17. S. Halevi. A Sufficient Condition for Key-Privacy. In *Cryptology ePrint Archive, Report 2005/05*, 2005. Available at `http://eprint.iacr.org/2005/005`.
18. R. Hayashi and K. Tanaka. Universally Anonymizable Public-Key Encryption. In *Advences in Cryptologu, ASIACRYPT '05, volume 3788 of Lecture Notes in Computer Science*, pages 293–312, London, UK, 2005.
19. K. Kurosawa and Y. Desmedt. A New Paradigm of Hybrid Encryption Scheme. In *Advances in Cryptology, CRYPTO '04, volume 3152 of Lecture Notes in Computer Science*, pages 426–442, London, UK, 2004.
20. NIST. The Case for Elliptic Curve Cryptography. Available at `http://www.nsa.gov/ia/industry/crypto_elliptic_curve.cfm`.
21. R. Sakai, K. Ohgishi, and M. Kasahara. Cryptosystems Based on Pairing. In *Symposium on Cryptography and Information Security (SCIS 2000), Okinawa, Japan*, 2000.

22. M. Scott. Authenticated ID-based Key Exchange and Remote Log-in With Insecure Token and PIN Number. In *Cryptology ePrint Archive, Report 2002/164*, 2002. Available at `http://eprint.iacr.org/2002/164`.
23. A. Shamir. Identity-Based Cryptosystems and Signature Schemes. In *Advances in Cryptology, CRYPTO '84, volume 196 of Lecture Notes in Computer Science*, pages 47–53, New York, NY, USA, 1985. Springer-Verlag.
24. Shamus Software. The MIRACL library. Available at `http://www.shamus.ie`.
25. V. Shoup. A Proposal for an ISO Standard for Public Key Encryption (Version 2.1), manuscript, December 20, 2001. Available at `http://www.shoup.net/papers/iso-2_1.pdf`.
26. V. Shoup. Using Hash Functions as a Hedge against Chosen Ciphertext Attack. In *Advances in Cryptology, EUROCRYPT '00, volume 1807 of Lecture Notes in Computer Science*, pages 275–288, 2000.
27. M. R. Spiegel. *Theory and Problems of Probability and Statistics*. McGraw-Hill, 1992.

A Comparison with Previous Works

Boneh et al. [9] have recently proposed a new IBE scheme, based on the quadratic residuosity assumption, that provides a very low ciphertext expansion: a t-bit message is encrypted into a $n + t + 1$-bit ciphertext, where n is the security parameter. Unfortunately, as they point out in their paper, the time efficiency of their scheme is far from ideal. The encryption time in all practical public-key and IBE systems, such as RSA, Cocks IBE [11] and Boneh and Franklin IBE [8], is cubic in the security parameter. The encryption time in the Boneh et al. scheme is quartic in the security parameter per message bit [9]. For each bit in the plaintext, the encryption algorithm must find the solution $(x, y) \in (\mathbb{Z}_N)^2$ to an equation of the form $Rx^2 + Sy^2 = 1$ and the solution $(\alpha, \beta) \in (\mathbb{Z}_N)^2$ to an equation $u\alpha^2 + S\beta^2 = 1$. For this reason, the encryption algorithm is particularly inefficient and may take several seconds to complete even on a fast machine.

Di Crescenzo et al. [14] provided the first Public-Key Encryption with Keyword Search (PEKS) scheme that does not use or depend on pairings. Their scheme is secure under a new assumption they introduced which they called the *Quadratic Indistinguishability* assumption[2]. (We remark that our Lemma 2 in Section 2.3 proves that their assumption is actually equivalent to the quadratic residuosity assumption.) The PEKS scheme in [14] is built out of an anonymous IBE scheme for keywords based on Cocks IBE. Although their scheme is suitable for PEKS, it is quite impractical when turned into an IBE scheme. In particular, for a t-bit plaintext, the ciphertext expansion in their scheme becomes $8t \cdot n$ (where n is the security parameter in Cocks IBE, e.g., $n = 1024$) and each recipient must store $4t$ secret keys. Indeed, if $W \in \{0,1\}^t$ is a keyword then the PEKS encryption algorithm outputs $4t$ elements in $\mathbb{Z}_N^*[+1]$, that is $a_1 = H(W \mid\mid 1)$, ... , $a_{4t} = H(W \mid\mid 4t)$, and releases the encrypted keyword as $(s_1, \ldots, s_{4t})$, where s_i is either in $S_{a_i}[N]$ or in $GT_a^N[-1]$. Whenever s_i is in $S_{a_i}[N]$, it will encrypt the bit "1". The decryption algorithm takes as input $4t$ secret keys and checks that all the $s_i \in S_{a_i}[N]$ encrypt the bit "1". When used as IBE for arbitrary messages, clearly any $s_i \in S_{a_i}[N]$ could also encrypt the bit "-1", but then the sender does not know whether a_i or $-a_i$ is a square and thus must encrypt for both (as in Cocks IBE), thus effectively doubling the amount of bandwidth required. We remark that thanks to our Lemma 3, a PEKS construction can be easily derived from the one in [14] that uses just one public key instead of $4t$.

Finally, we emphasize that our basic scheme and its efficient variants are all *universally anonymous*, and not just anonymous as other IBE schemes. Thus, our scheme can be seen as the first universally anonymous IBE (whether based on pairings or not) and the first efficient and anonymous IBE scheme based on the standard quadratic residuosity assumption.

B Hybrid Encryption and CCA-security

It is well-known that in order to encrypt long messages, asymmetric encryption can be used in combination with symmetric encryption for improved efficiency. This simple and well-known paradigm has been formalized

[2] Basically, their assumption states that it is hard to distinguish between the following two distributions (using our notation): $D_0 = \{(a, c, N) : a \stackrel{R}{\leftarrow} \mathbb{Z}_N^*[+1]; c \stackrel{R}{\leftarrow} \{x \mid x \in \mathbb{Z}_N^* \wedge GT(a, x) = -1\} \cup S_a[N]\}$ and $D_1 = \{(a, c, N) : a \stackrel{R}{\leftarrow} \mathbb{Z}_N^*[+1]; c \stackrel{R}{\leftarrow} \mathbb{Z}_N^*\}$ where N is the product of two Blum-Williams integers.

only recently by Cramer and Shoup [13, 12] and Shoup [25]. It is introduced as the KEM-DEM construction which consists of two parts: the key encapsulation mechanism (KEM), used to encrypt a symmetric key, and the data encapsulation mechanism (DEM) that is used to encrypt the plaintext via a symmetric cipher. The security of the two components, KEM and DEM, can be analyzed separately.

Cramer and Shoup [13] showed that a hybrid encryption scheme is CCA secure (i.e., secure against the adaptive chosen ciphertext attack) in the standard model if the KEM component is CCA secure and the DEM component is a CCA-secure one-time symmetric encryption. Later, Kurosawa and Desmedt [19] showed that the KEM component does not have to be CCA-secure as long as the CCA-secure one-time symmetric encryption satisfies an extra condition (which is satisfied by the DEM scheme proposed by Shoup [13, 25, 26]). The construction of CCA-secure one-time symmetric encryption is standard and it is usually accomplished by coupling a message authentication code (MAC) with a symmetric encryption. In particular, it can be shown that applying a one-time MAC on the output of a CPA-secure symmetric encryption results in a CCA-secure symmetric encryption scheme (see e.g. Cramer and Shoup [13]).

The focus of this paper is on variants of Cocks IBE which can be proven secure only in the random oracle model (ROM). Thus, it makes sense to consider KEM-DEM constructions that are CCA-secure in ROM. In this case, the most relevant work is the one from Fujisaki and Okamoto [16] that shows how to build CCA-secure hybrid encryption schemes and how to convert any CPA-secure asymmetric scheme into a CCA-secure one in ROM. Even more relevant is the work of Bentahar et al. [5] that formalizes the concept of id-based KEM-DEM and provides a generic transformation from any IBE scheme to CCA-secure ID-based KEM in ROM.

Given this state of affair, if we build a KEM based on a IBE scheme that requires the random oracle model for its security, say Cocks IBE, we can just focus on the CPA security of the encryption algorithm. This is justifiable because it is possible to show (see, e.g., Bentahar et al. [5]) that if a KEM returns $(\text{Encrypt}_{\mathsf{UAnonIBE}}(K), F(K))$, where $\text{Encrypt}_{\mathsf{UAnonIBE}}(K)$ is a one-way encryption for an identity and F is a hash function modeled as a random oracle, then the combination of this KEM with a CCA-secure DEM results in a CCA-secure hybrid encryption. (Note that one-way encryption is implied by CPA-security.) Since our scheme UAnonIBE and its efficient variants are CPA-secure in the random oracle model, the resulting hybrid encryption that follows from the paradigm above is a CCA-secure encryption in the random oracle model.