Fully Secure Accountable-Authority Identity-Based Encryption

Amit Sahai and Hakan Seyalioglu

The University of California - Los Angeles
sahai@cs.ucla.edu, hseyalioglu@ucla.edu

Abstract. The problem of trust is one of the biggest concerns in any identity-based infrastructure where the key-generation authority (called the PKG) must choose secret keys for participants and therefore be highly trusted by all parties. While some abilities of the PKG are intrinsic to this setting, reducing this trust as much as possible is beneficial to both user and authority as the less trust is placed in it, the less an honest authority can be accused of abusing that trust. Goyal (CRYPTO 2007) defined the notion of Accountable-Authority IBE in which a dishonest PKG who had leaked a user's private key could be proven guilty. Later, Goyal et al. (CCS 2008) asked whether it would be possible to implicate a PKG who produced an unauthorized decoder box, enabling decryption with a noticeable probability but which may not actually grant access to a well-formed key. Formally, would it be possible for a tracing algorithm to implicate a dishonest PKG given only black-box access to such a decoder? Goyal et al. could only provide such a scheme in the weaker setting of selective security, where an adversary must declare at the start of the game which identity it intends to target. In this work, we provide the first fully secure accountable-authority IBE scheme. We prove security from the standard DBDH assumption while losing none of the functionality or security of the original proposal.

Keywords: Identity-Based Encryption, Accountable Authority, Tracing.

1 Introduction

Since its introduction by Shamir [24] and first constructions by Boneh and Franklin [4] and Cocks [7], identity-based encryption (IBE) has been one of the most active areas of cryptographic research, with numerous applications to computer security and privacy (e.g. [3,2,9,25]). Many concepts of independent interest sprang from this topic such as Fuzzy IBE and Attribute Based Encryption [23,14,12,22] which allow encryption to groups of users whose credentials (a.k.a. attributes) satisfy a given access policy, and Hierarchical IBE [10,11,17] which allows key generation in a leveled fashion.

Despite its applications and practicality, the security of an IBE scheme relies heavily on trusting the key generation authority. Despite the tremendous practical security benefits of not having to manage a public-key infrastructure (PKI)

D. Catalano et al. (Eds.): PKC 2011, LNCS 6571, pp. 296–316, 2011.

based on individual public keys, trusting any single authority is very troubling. It is a natural and important question to ask how one can reduce the trust one must have in the private key generator. One proposed method to reduce this trust is to employ multiple PKGs such that no single entity can compromise security [2]; however, the question of widespread collusion naturally remained.

This left open the problem of reducing trust in a single central authority. Note that such a central authority can clearly decrypt all messages in the system. However, in many practical situations, users may not be worried that the authority would have the inclination or resources to specifically eavesdrop on their communications. On the other hand, users would want some assurance that the authority doesn't issue their secret keys to *other potentially malicious* users, who might have a lower profile and have less to lose if caught. This is precisely the problem proposed and addressed by the work of Goyal [13] by introducing the notion of Accountable-Authority IBE (A-IBE). If a user finds a key for his identity being disseminated without permission (e.g. the user finds that his secret key is being put up for auction on EBay), Goyal proposed a system where the user can prove that it would have been computationally impossible for the user to have created this compromised key. Since the only other entity with access to keys to the user's identity is the PKG, this allowed the user to successfully implicate a malicious PKG of malfeasance.

Goyal's first construction relied heavily on seeing the *actual* compromised key, and as-such this first construction is called only *white-box* secure. Such white-box security leaves a great deal to be desired, for instance, while able to still decrypt, the key may have been manipulated in some way. A natural extension is to ask whether it would be possible to implicate a PKG who has disseminated a decoder "box" which allows decryption of ciphertexts encrypted to some identity rather than simply leaking a key itself in the clear (e.g. selling such a box or a heuristically obfuscated piece of decryption software). Since the exact information used to create this decoder box may be hidden, this is called *black box* security and is the strongest current proposed notion of A-IBE. If we can trace a PKG who issues decoder boxes, it forces a user attempting to dishonestly acquire decryptions to continually interact with the PKG (who may, for example, be acting as a decryption oracle). Such a setting is much riskier for the dishonest PKG since it requires a much greater level of communication with the dishonest user.

Giving a black-box secure A-IBE scheme was specifically labeled the most important problem left open in Goyal's original work. The first work made some progress towards obtaining black-box security, which was later refined [15,19] (the latter of which proves security if a cheating authority is denied a decryption oracle); however, major shortcomings remained. The only known *black-box* secure constructions are in a very weak model of security called "selective security" where the adversary must designate the identity which it will create a decoder box for before even seeing the public parameters. While selective security was useful for exploring these concepts, it is a completely unreasonable requirement for security since real-world adversaries have the ability to pick targets adaptively.

In fact, selective security is especially troubling in the A-IBE setting since selective security is only a reasonable benchmark if there are only a few high value targets whom an adversary would be interested in attacking (making the prospect of picking a target adaptively during the attack less appealing). However, in a subscription service handled through decoder boxes (one of the most natural applications for a black-box A-IBE scheme), there would be many targets of equal value, making selective-security particularly troubling.

1.1 Our Contributions

In this work, we provide the first fully secure black-box accountable-authority IBE scheme. We are able to prove security based on the standard DBDH number-theoretic intractability assumption while losing none of the functionality or security of the original proposal of Goyal [13].

A limitation of previous work [15] which obtained selective black-box security was that it obtained functionality from established Attribute-Based Encryption schemes without modifying the scheme to better suit the A-IBE framework. A main contribution of this work is that we can obtain A-IBE by only moving "halfway" between basic IBE and Attribute-Based Encryption, in a setting where identities do have attributes, but these attributes are assigned randomly. We call such a scheme a *Dummy-IBE* scheme and use it to construct black-box A-IBE. We then give a construction of a *Dummy-IBE* scheme by combining mechanics of Waters' IBE scheme [25] and the Attribute-Based Encryption scheme due to Sahai and Waters [23]. We note that a fully secure ABE scheme would not directly suffice to make the previous construction work, in particular, ciphertext and key-sanity checks (to be defined later) would also be necessary. Indeed in contrast to recent findings in fully secure ABE [18,21], we are able to follow the framework of previous *selectively secure* ABE constructions [23] to achieve *full* black-box security, and as a result inherit these schemes' improved efficiency and analytic simplicity.

2 Preliminaries

We will often use the notation $[a, b]$ to denote all integers between a and b, inclusive and bold font (e.g. $\mathbf{x}$, $\mathbf{A}$) to denote vectors. A bracketed index (e.g. $\mathbf{x}[j]$, $\mathbf{A}[i]$) will denote the value at that index of the relevant vector. A negligible value is one that grows slower than any inverse polynomial of a certain parameter (usually the security parameter) and a probability is said to be overwhelming if it is within an additive negligible probability of 1.

2.1 Bilinear Maps

Let $\mathbb{G}, \mathbb{G}_T$ be two multiplicative groups of prime order p. Let g be a generator of $\mathbb{G}$ and $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$. We assume e satisfies bilinearity (for all $u, v \in \mathbb{G}_1$ and $a, b \in \mathbb{Z}_p$, we have $e(u^a, v^b) = e(u, v)^{ab}$) and non-degeneracy ($e(g, g) \neq 1$).

2.2 Security Assumption

Definition 1 (Decisional Bilinear Diffie-Hellman (DBDH) Assumption) *Suppose the values* $a, b, c, z \leftarrow \mathbb{Z}_p, \beta \leftarrow \{0,1\}$ *are chosen at random. The DBDH assumption states that no probabilistic polynomial-time algorithm can distinguish the tuple* $(g, A = g^a, B = g^b, C = g^c, e(g,g)^{abc})$ *from the tuple* $(g, A = g^a, B = g^b, C = g^c, e(g,g)^z)$ *with more than a negligible advantage.*

2.3 Fully-Simulatable Oblivious Transfer

Informally, a k-out-of-n oblivious transfer protocol [8] is a two-party protocol in which n strings are stored by one party and at the end, the other party has received exactly k of them with the requirements that the receiving party gains no more information than the k strings it received and the sending party gains no information about which k the receiving party acquired.

Formally, by a fully-simulatable OT protocol, we will refer to one that meets Canetti's requirements for universal composability [6]. If an adversary $\mathcal{A}'$ is interacting with a challenger and has non-negligible success probability ϵ in a game which includes some Fully-Simulatable OT protocol, there also exists an adversary $\mathcal{A}$ with probability of success within a negligible factor of ϵ which instead uses the ideal-world OT protocol where $\mathcal{A}$ directly queries the indices of the k values it will receive from the challenger. We will often transfer our adversaries into ones which engage in the ideal-world OT protocol. For examples of fully simulatable Oblivious Transfer using only the DDH and DBDH assumptions please refer to the following papers: [20,16,5].

2.4 Accountable Authority IBE (A-IBE)

We now recall what it means for an IBE scheme to be A-IBE secure. First, the scheme must function as a good IBE scheme. Traditionally, it is assumed that the Key Generation algorithm simply outputs a key, however, since the key received by each party should not be available to the central authority, we instead assume it is generated through an interactive protocol.

Definition 2 (Identity Based Encryption (IBE) Scheme). *An* Identity Based Encryption Scheme *consists of the following four probabilistic polynomial time algorithms:*

- **Setup**$(1^\lambda) \rightarrow$ PK, MK. The setup takes as input 1^λ with λ the security parameter and outputs public parameters PK and master secret key MK.
- **KeyGen**(PK, MK,ID)$(\rightarrow)$ $\mathrm{K_{ID}}$. The user key generation algorithm takes as inputs PK, MK and an identity ID and engages in an interactive protocol with the recipient. At the end, the recipient receives a key for ID, $\mathrm{K_{ID}}$. We use the notation $(\rightarrow)$ to highlight the fact that **KeyGen** may not know exactly which key the user receives (for example, if **KeyGen** is implemented using an oblivious transfer protocol).

- **Encrypt**(M, PK, ID)$\rightarrow$ C_{ID}. The encryption algorithm takes as input PK, ID and a message M and outputs a ciphertext C_{ID}.
- **Decrypt**(C_{ID}, K_{ID}, PK) $\rightarrow$ M. The decryption algorithm outputs the original message with overwhelming probability assuming it is given as input a key to the identity ID.

We will assume that ID is included in the plain as part of C_{ID} and K_{ID} during our analysis for notational simplicity.

Definition 3 (ϵ-Useful Decoder Box). *For non-negligible ϵ, a probabilistic polynomial time algorithm $\mathcal{D}_{ID}$ is an* **ϵ-useful decoder box** *for the identity ID if:*

$$\Pr[M \leftarrow \mathcal{M} : \mathcal{D}_{\mathrm{ID}}(\mathbf{Encrypt}(M, PK, ID) = M] \geq \epsilon$$

The main additional functionality required of an A-IBE scheme follows:

- **Trace**$^{\mathcal{D}_{\mathrm{ID}}}$(PK, ID, K_{ID}) outputs either User or PKG. Given oracle access to an ϵ-useful decoder box $\mathcal{D}_{\mathrm{ID}}$, PK, ID and a key K_{ID}, the tracing algorithm will decide if $\mathcal{D}_{\mathrm{ID}}$ was created by the trusted authority or by the user who supplied K_{ID}.

Intuitively the tracing algorithm allows a user who has recovered a decoder box for its identity to reveal its secret key to prove that it could not have generated this decoder box. This would be impossible if all keys could decrypt perfectly, which is why we allow decryption a small probability of failure. Additionally, it has proven very useful to include two more algorithms in most A-IBE schemes for use in tracing. A term is said to be valid if it is a possible output of the relevant generation algorithm (in other words, a key K_{ID} is valid if it is a possible output of the key generation algorithm on inputs PK, MK and a ciphertext is valid if it is the encryption of some M to the relevant identity):

- **KeySanity**(K_{ID},PK) tests if K_{ID} is a proper key for identity ID and outputs Valid or Not-Valid.
- **CiphertextSanity**(C_{ID},PK) tests if C_{ID} is a proper ciphertext for ID and outputs Valid or Not-Valid.

2.5 Security Requirements

An A-IBE scheme is called black-box secure if it satisfies the following requirements. First, it must satisfy IND-ID-CPA security. Second, if a decoder box $\mathcal{D}_{\mathrm{ID}}$ was created by the central authority, the tracing algorithm should implicate the PKG and when it was created by the colluding users, it should implicate the users. This security requirement is captured in the following games.

Definition 4 (IND-ID-CPA Security). *An A-IBE scheme is* IND-ID-CPA *secure if the advantage of any probabilistic polynomial time adversary $\mathcal{B}$ is negligible in the game below:*

1. **Setup.** The challenger runs the **Setup** algorithm and sends PK to $\mathcal{B}$.
2. **Phase 1.** $\mathcal{B}$ engages in a **KeyGen** protocol with the challenger with adaptively chosen ID_i and receives K_{ID_i}. This may be repeated multiple times for different identities.
3. **Challenge.** $\mathcal{B}$ submits two messages M_0, M_1 and an identity ID which it did not query during Phase 1. The challenger flips a coin b and encrypts M_b to the identity ID and sends the encryption C to $\mathcal{B}$.
4. **Phase 2.** Same as Phase 1 except $\mathcal{B}$ may not query a decryption key for ID.
5. **Guess.** The adversary outputs $b' \in \{0, 1\}$.

The advantage of $\mathcal{B}$ is defined as $\Pr[b' = b] - \frac{1}{2}$.

Definition 5 (Dishonest-User Security). *An A-IBE scheme is Dishonest-User secure if the advantage of any probabilistic polynomial time adversary $\mathcal{B}$ is negligible in the game below:*

1. **Setup.** The challenger runs the **Setup** algorithm and sends **PK** to $\mathcal{B}$.
2. **Key Generation.** The adversary adaptively queries keys for distinct ID_i, and receives K_{ID_i}. This may be repeated multiple times for different identities.
3. **Create Decoder Box.** The adversary outputs an ϵ-useful decoder box $\mathcal{D}_{\mathrm{ID}}$ and K'_{ID} for some identity ID queried during the Key-Generation phase.
4. **Tracing Failure.** Finally, success is defined as the event that: $\mathbf{Trace}^{\mathcal{D}_{\mathrm{ID}}}(\mathrm{PK}, \mathrm{ID}, K'_{\mathrm{ID}}) = \mathsf{PKG}$.

At this point it may seem artificial to have the requirement that ID was queried once in the Key Generation phase of the game, however since the scheme is required to be IND-ID-CPA secure, outputting a key for an identity which has not been queried would contradict the previous security requirement and therefore adding this additional requirement does not weaken security.

Definition 6 (Dishonest-PKG Security). *An A-IBE scheme is Dishonest-PKG secure if the advantage of any probabilistic polynomial time adversary $\mathcal{B}$ is negligible in the game below:*

1. **Setup.** The adversary $\mathcal{B}$ generates and passes public parameters PK and ID to the challenger. The challenger checks PK and ID are well-formed, aborts if not.
2. **Key Generation.** The challenger and $\mathcal{B}$ engage in the key generation protocol for an identity ID. If neither party aborts, the challenger receives K_{ID} as output.
3. **Decryption.** The adversary adaptively queries ciphertexts C_i to the challenger and the challenger replies with the decryption under K_{ID}. This may be repeated multiple times for different ciphertexts.
4. **Create Decoder Box.**. The adversary outputs an ϵ-useful decoder box $\mathcal{D}_{\mathrm{ID}}$.
5. **Tracing Failure.** Finally, success is defined as the event that $\mathbf{Trace}^{\mathcal{D}_{\mathrm{ID}}}(\mathrm{PK}, \mathrm{ID}, \mathrm{K}_{\mathrm{ID}}) = \mathsf{USER}$.

If our scheme has a **CiphertextSanity** check we can always assume decryption is preceded with verification that the ciphertext is well-formed and if it is not well-formed, the decryption process can just output $\perp$. Informally, this will allow us to argue that the PKG gains no information from its decryption queries since the decryption oracle will only decrypt well-formed ciphertexts, which the PKG could have decrypted without using the oracle (since the PKG can already generate a key for any identity).

3 Preliminary Reduction

The first key step in our proof is realizing that A-IBE can be built from an encryption scheme which falls somewhere between usual IBE and Attribute-Based Encryption. For this, we introduce the notion of *Dummy-IBE* in which every user is assigned a set of attributes which restricts the ciphertexts that can be decrypted but the user has no control over which attributes are assigned. We stress again that because of the importance of the sanity checks in our setting, a fully secure attribute based encryption scheme by itself would not suffice for our purposes.

3.1 Dummy Identity-Based Encryption

The intuition for *Dummy IBE* (*D-IBE*) comes from previous work [15] which achieves full *Dishonest-PKG* security (but not full *Dishonest-User* security). Keys for identities and ciphertexts will have k attributes and decryption will only succeed if the encryption was to the target identity and the attribute sets overlap in at least d indices.

Definition 7 (Dummy Identity-Based Encryption (D-IBE) Scheme). *A D-IBE Encryption scheme D with parameters $d \leq k \leq n \in \Theta(\lambda)$ consists of the following four poly-time algorithms:*

- **Setup**$(1^\lambda, d, k, n) \rightarrow$ PK, MK public and master keys.
- **KeyGen**(PK, MK) $\rightarrow K_{\mathrm{ID}}(S)$ The key generation algorithm selects $S \subset [1, n]$ a random subset of size k, generates[1] K^{α}_{ID} for all $\alpha \in [1, n]$ and outputs $K_{\mathrm{ID}}(S) = \{K^{\alpha}_{\mathrm{ID}} \mid \alpha \in S\}$.
- **Encrypt**(M, PK, ID, S)$\rightarrow C_{\mathrm{ID}}(S)$.
- **Decrypt**($\mathrm{C}_{\mathrm{ID}}(S)$,$\mathrm{K}_{\mathrm{ID}}(S')$,PK)$\rightarrow M$ where $C_{\mathrm{ID}}(S)$ is an encryption of M if $|S \cap S'| \geq d$.
- **KeySanity**$(K^{*}_{\mathrm{ID}}(S), PK)$ outputs Valid or Not-Valid depending on whether the key is a valid key for the implied identity and attribute set.

[1] The fact that the key generation algorithm is able to generate all key components will be important for our reduction. To formalize the above notion, we could have the key generation algorithm output all key components and only send the ones corresponding to members in S to the user but we present it as above for notational simplicity.

– **CiphertextSanity**$(C^*_{\mathrm{ID}}(S), PK)$ outputs either Valid or Not-Valid depending on whether the ciphertext is a valid encryption of some message for the implied identity and attribute set.

We will assume that correctly formatted keys and ciphertexts will include a description of the relevant identity and dummy-attribute set. In the above notation α are called the dummy attributes and S is called a dummy attribute set. Correctness of a Dummy-IBE scheme is as expected, decryption should succeed with overwhelming probability if the identities of the key and ciphertext match and the dummy attribute sets overlap in at least d indices.

Definition 8 (Dummy-IBE Security). *A Dummy-IBE scheme is said to be* D-IBE *secure if the advantage of any probabilistic polynomial time adversary $\mathcal{B}$ is negligible in the game below:*

1. **Setup.** The challenger runs the **Setup** algorithm and sends PK to $\mathcal{B}$.
2. **Queries.** $\mathcal{B}$ adaptively queries keys for distinct ID_i, the challenger returns $K_{\mathrm{ID}_i}(S_i)$. This may be repeated multiple times for different identities.
3. **Challenge.** The adversary specifies M_0, M_1 and an identity ID_j which has been queried during the **Queries** phase and a dummy attribute set S such $|S \cap S_j| < d$ where S_j is from the **Queries** phase. The challenger picks $b \in \{0, 1\}$ at random and sends $\mathcal{B}$, **Encrypt**(M_b, PK, ID_j, S).
4. **Guess.** The adversary outputs a guess b' for b.

3.2 D-IBE Implies Dishonest-User Security

We describe how to transform a Dummy-IBE scheme D into a *Dishonest-User* Secure A-IBE scheme A. The overall structure of our construction can be considered a generalization of the construction in [15] where the role of our Dummy-IBE scheme is instead replaced by a ABE scheme which is not able to satisfy all the required notions of security. The similarities in structure will allow us to reuse an information theoretic result for *Dishonest-PKG* security from previous work [15]. We will from now on assume that the message space is a group of size superpolynomial in the security parameter, as is the case for our construction (notice that making a ϵ-useful decoder box is trivial if the size is not superpolynomial).

Let $d \leq k \leq n \in \Theta(\lambda)$, we give concrete bounds k, d and n should satisfy at the end of this section.

– **Setup**$(1^\lambda) = D$.**Setup**$(1^\lambda, n, k, d)$
– **KeyGen**(MK, PK, ID) Using MK, PK, ID, generate $K = \{K^\alpha_{\mathrm{ID}} : \alpha \in [1, n]\}$ and randomly permute them (call this permutation σ) and engage in a non-adaptive k-out-of-n oblivious transfer protocol with the user querying the permuted set as the set of possible values to be transferred. After the OT protocol concludes send the user σ in the clear. The user now has access to $K_{\mathrm{ID}}(S)$ for some random $S \subset [1, n]$ and runs a **KeySanity** check on $K_{\mathrm{ID}}(S)$ and terminates the protocol if it fails.

- **Encryption**(M, PK, ID) Pick $S \subset [1, n]$ at random of size k, output D. **Encrypt**(M, PK, ID, S).
- **Decrypt**($C_{ID}(S)$, PK, K_{ID}(S')) If the ciphertext sanity check of $C_{\mathrm{ID}}(S)$ passes output:
  D.**Decrypt**($C_{ID}(S)$,$\mathrm{K}_{ID}(S')$,PK), else output $\perp$.

Notice that at the moment we only have a guarantee that the decryption algorithm will decrypt if $|S \cap S'| \geq d$. Therefore, pick k and d such that given some $S \subset [1, n], |S| = k$ (the decryption dummy-atributes) a random set $S' \subset [1, n]$ (the dummy-atributes used during encryption),

$$Pr[|S \cap S'| \geq d] > 1 - \mu(\lambda)$$

for some negligible function μ. If we pick $k = cn$, $d = (c^2 - \delta)n$ for some constant $c \in (0, 1)$ and $\delta \in (0, c^2)$ by Chernoff bounds, two sets of size k will intersect in at least d locations with overwhelming probability. From this point, assume k and d are initialized thusly and $c < 1/4$.

3.3 Tracing Algorithm

The overarching idea is that if a message has been encrypted to identity ID under dummy attribute set A, it should only be decryptable by someone holding a key $K_{ID}(A')$ where $|A \cap A'| \geq d$. Therefore, the tracing algorithm, which has oracle access to $\mathcal{D}_{\mathrm{ID}}$ will repeatedly query ciphertexts the user ID should not be able to decrypt given the attribute set that was assigned during the key generation phase. If the decoder box decrypts such a ciphertext, we will be able to conclude that the decoder box was not created by the user, proving malfeasance on the part of the PKG.

On input $K_{\mathrm{ID}}(S)$ with oracle access to $\mathcal{D}_{\mathrm{ID}}$, the tracing algorithm will run a **KeySanity** check to verify the validity of the input key. Then, it will repeat the following experiment $\nu(\epsilon) \in poly(\lambda)$ times (we will choose $\nu(\epsilon)$ after the analysis):

- Select a dummy attribute set S_i with $|S_i \cap S| < d$.
- Select a random message M and encrypt M using S_i as the dummy attributes.
- The decoder box outputs some $M' = \mathcal{D}_{\mathrm{ID}}(C_{\mathrm{ID}}(S_i))$.

If for any iteration $M' = M$, implicate the PKG otherwise, implicate the User.

Theorem 1. *If D is a secure Dummy-IBE scheme, the A-IBE construction A is Dishonest-User secure.*

Proof. Assume p.p.t. $\mathcal{A}'$ succeeds in the Dishonest-User game above with non-negligible probability δ. Since **KeyGen** is implemented with a fully-simulatable k-out-of-n oblivious transfer scheme, we can work in the OT-hybrid model by Canetti's [6] theorem on composability, which implies there exists p.p.t. $\mathcal{A}$ which also succeeds in the game with non-negligible probability such that $\mathcal{A}$ queries

indices during the OT protocol directly from the challenger and the challenger simply sends the values stored in these indices to $\mathcal{A}$. With this new $\mathcal{A}$ which simply requests indices for messages to receive from the challenger as in the OT-Hybrid model, the reduction is as follows:

- **Setup.** Send the public key of the D-IBE scheme to the adversary $\mathcal{A}$.
- **Key Generation.** On each of the adversary's queries for an identity to ID_j and indices (for the OT protocol) $I \subset [1, n]$, $|I| = k$, query the D-IBE scheme for a key on identity ID_j and receive $K_{\mathrm{ID}}(S_j) = \{K_{\mathrm{ID}}^{\alpha_i}\}_{i \in [1,k]}$. Now pick a random σ such that $\sigma(I) = S_j$. Send $\mathcal{A}$: $\{K_{\mathrm{ID}}^{\alpha}, \alpha \in S_j\}$ and σ.
- **Create Decoder Box.** The adversary outputs a decoder box $\mathcal{D}_{\mathrm{ID}_j}$ along with a key $K_{ID_j}(S)$ with an identity ID_j queried at some point in the key generation phase. Run a **KeySanity** check to make sure $K_{ID_j}(S)$ is a valid key. If $S \neq S_j$ run **New Attributes** phase, otherwise proceed to the **Tracing Failure** phase.
- **New Attributes.** Since $S \neq S_j$, pick a d element subset S^* of S which does not overlap with S_j in more than $d-1$ indices (recall both are of size k) and $k-d$ element subset $A \subset [1, n] \setminus (S \cup S_j)$ (this is why we assume $d \leq k < n/4$) and initiate the challenge query with $M_0, M_1 \leftarrow \mathcal{M}$ with identity ID_j and attribute set $S^* \cup A$. Since the challenger has a key which overlaps the challenge attribute set in more than $d-1$ indices with the same identity, it can trivially decrypt and output b. This is a valid challenge since $|(S^* \cup A) \cap S_j| < d$.
- **Tracing Failure.** Pick some $d \in [1, \nu(\epsilon)]$ at random. Then, run the tracing algorithm normally $d-1$ times. For the d^{th} iteration, choose M_0, M_1 at random and initiate the challenge phase of the D-IBE scheme by outputting both messages, the challenge identity ID and a random dummy attribute set S' such that $|S' \cap S| < d$. Receive C from the D-IBE scheme, either an encryption of M_0 or M_1 under identity ID under the attribute set A and compute $\mathcal{D}_{\mathrm{ID}}(C_{\mathrm{ID}}(S'))$. If $\mathcal{D}_{\mathrm{ID}}(C_{\mathrm{ID}}(S')) = M_0$ output 0, if $\mathcal{D}_{\mathrm{ID}}(C_{\mathrm{ID}}(S')) = M_1$ output 1, otherwise, guess randomly.

Now, notice that if the **New-Attributes** phase is initiated, the challenger decrypts and breaks the Dummy-IBE security trivially. Also, since the tracing algorithm only ever queries messages encrypted under dummy-attribute sets which do not overlap with the key dummy-atribute set in more than $d-1$ indices, if the tracing algorithm outputs PKG with non-negligible probability, for some $d \in [1, \nu(\epsilon)]$ it must succeed in correctly decrypting with non-negligible probability. Since the challenge ciphertext the challenger returns to it is from the exact distribution that it expects, with non-negligible probability (since $\nu(\epsilon)$ is polynomial in the security parameter) it will succeed in decrypting the challenge ciphertext, breaking semantic security. Since $\mathcal{D}_{\mathrm{ID}}$ is oblivious to the second message in the challenge phase, with probability within a negligible function of 1, whenever we do not guess randomly, we will be correct. Since we have at least a $1/\nu(\epsilon)$ probability of not guessing randomly if the decoder decrypts on a single query and we assume the decoder will decrypt some query with non-negligible probability, we succeed in breaking semantic security.

3.4 Running in Parallel for PKG Security

To achieve Dishonest-PKG security, we actually will have to modify our construction slightly by running it in parallel m times where $m \in \Theta(\lambda)$. In this setting, if we call the A-IBE candidate above A, we will call our new scheme $\mathbf{A}$, where $\mathbf{A}[1], \ldots, \mathbf{A}[\mathrm{m}]$ denote the m different instantiations of A. The public key and master secret key are simply the collection of public and master secret keys of each of the individual instantiations.

A private key is obtained for identity ID by running **KeyGen** in parallel m times and giving the output of each to the user. The user's dummy attribute set is now $\mathbf{S}$, with each $\mathbf{S}[i]$ a random k element subset of $[1, n]$. Encryption is handled by splitting M into m shares, $M = M_1 \oplus M_2 \oplus \ldots \oplus M_m$ where $M_1, \ldots M_{m-1}$ are chosen are random and each M_i is encrypted under $\mathbf{A}[i]$ with a dummy attribute set $\mathbf{S}'[i]$, with each $\mathbf{S}'[i]$ a k element subset of $[1, n]$. Now $C_{\mathrm{ID}}(\mathbf{S}')$ should be decrypted only by keys $K_{\mathrm{ID}}(\mathbf{S})$ where $|\mathbf{S}[i] \cap \mathbf{S}'[i]| \geq d$ for each $i \in [1, m]$. **KeySanity** and **CiphertextSanity** checks of $\mathbf{A}$ work by checking each of the individual components with the existing sanity checks for A. We now describe the tracing algorithm:

3.5 Tracing Algorithm for the Parallel Scheme

On input $K_{\mathrm{ID}}(\mathbf{S})$, the tracing algorithm will run a **KeySanity** check to verify the validity of the input key. Then, it will then repeat the following experiment $\nu \in poly_m(\lambda)$ times:

- Pick $j \in [1, m]$ at random,
- If $i \neq j$, choose $\mathbf{S}'[i]$ a k element subset of $[1, n]$,
- If $i = j$ include the additional restriction that $|\mathbf{S}'[i] \cap \mathbf{S}[i]| < d$,
- Select a random message M and encrypt M using $\mathbf{S}'$ as the dummy attributes,
- The decoder box outputs some $M' = \mathcal{D}_{\mathrm{ID}}(C_{\mathrm{ID}}(\mathbf{S}'))$.

If for any iteration $M' = M$ output PKG otherwise, output User.

Theorem 2. *If D is a secure Dummy-IBE scheme, $\mathbf{A}$ is Dishonest-User secure.*

Proof. The proof for this is identical to the proof for the proof of security for A, first pick $i \in [1, m]$ at random and for all $j \neq i$, the challenger will instantiate $\mathbf{A}[j]$ itself and use D, the Dummy-IBE scheme its trying to break to instantiate the j^{th} index. Since there are only polynomially many choices for j, with non-negligible probability, the adversary which outputs $K_{\mathrm{ID}}(\mathbf{S}')$ and $\mathcal{D}_{\mathrm{ID}}$ will either have $\mathbf{S}'[j]$ overlap the queried keys $\mathbf{S}[j]$ less than d locations or, the tracing algorithm will decrypt a message whose dummy attribute set $\mathbf{S}^*$ has $|\mathbf{S}^*[j] \cap \mathbf{S}[j]| < d$.

Theorem 3. *If D is a secure Dummy-IBE scheme, $\mathbf{A}$ is Dishonest-PKG secure.*

Proof. Since the setup of $\mathbf{A}$ is very closely related to the construction in Goyal et al.[15], (which is fully secure in the *Dishonest-PKG* game) we will be able

to reuse their combinatorial results for the security of $\mathbf{A}$. They also compose m sets of dummy-attributes in parallel and use fully-simulatable OT during the key generation phase in the same manner we do. In their result, by using the ciphertext and key sanity checks they are able to prove a purely information theoretic bound on the fraction of dummy-attribute sets that will implicate the PKG with access to a decoder box $\mathcal{D}_{\mathrm{ID}}$. They show that as long as **KeyGen** does not terminate with non-negligible probability, with overwhelming probability the PKG will not query a ciphertext the user is unable to decrypt. Since all ciphertexts which the user can decrypt decrypt to the same value by the ciphertext sanity check and the PKG can decrypt to this value by itself, this allows analysis in the Dishonest-PKG game where the PKG has no decryption capabilities.

Their result (Lemma 5 in the original paper)[2] gives us precisely the second information theoretic guarantee we need, namely that any tracing algorithm with an ϵ-useful decoder box will output PKG for all but a negligible fraction of possible dummy attributes. We refer the reader to the original paper for the proof of the lemma below. Assume m is super-logarithmic and n is linear in the security parameter λ, $k = c_1 n$, $d = c_2 n$ positive constants less than 1 such that $c_2 < c_1^2$ (which ensures decryption with overwhelming probability the Chernoff bound).

Lemma 1. *Let ϵ be non-negligible and $\mathcal{D}_{ID}$ an ϵ-useful decoder box. Let $\mathbf{S}$ be a dummy attribute set for the user and consider the following experiment:*

- *Select a dummy attribute set $\mathbf{S}'$ such that $|\mathbf{S}'[i] \cap \mathbf{S}[i]| < d$ for some $i \in [1, m]$*
- *Select a random message M and encrypt it using $\mathbf{S}'$ as the dummy attributes and outputs the ciphertext C*
- *The decoder box outputs $M' = \mathcal{D}(C)$*
- *Output* PKG *if $M' = M$*

  Then, for all but a negligible fraction of choices of $\mathbf{S}$, the above experiment has probability of outputting PKG *greater than $\epsilon/(24m)$.*

By the above lemma, if we consider ideal OT functionality and reduce to the case where the PKG has no decryption capabilities, the tracing algorithm will output PKG with overwhelming probability by only running in polynomial time in the *Dishonest-PKG* game as desired as long as $\mathbf{S}$ does not come from a exceptional set which forms a negligible fraction of all possible $\mathbf{S}$. Since full *Dishonest-PKG* security is obtained information theoretically for an identical abstract construction in [15] we refer to their paper for a rigorous treatment of the above argument.

3.6 A Modification for IND-ID-CPA Security

After we have a Dishonest-User and Dishonest-PKG secure A-IBE scheme, we may use the same method as the previous selective secure construction [15]. We will achieve IND-ID-CPA security by using the above A-IBE scheme and a secure

[2] Note that the proof of **Lemma 5** is not in the original, but instead in the full version of the cited paper.

IBE scheme together. For any M, we will secret share M into $M_1 \oplus M_2$ where M_1 was chosen at random and encrypt M_1 with the A-IBE scheme and M_2 with the IBE scheme. Since Waters' IBE scheme [25] achieves IND-ID-CPA security using the DBDH assumption, we may combine it with our Dummy-IBE scheme to achieve full A-IBE with no additional security assumptions.

4 A Dummy-IBE Scheme

The main technical contribution of this paper is the construction of a secure *Dummy-IBE* scheme, which as described, provides a fully secure *A-IBE* scheme. For our construction, we will reuse many of the methods introduced by Waters [25]. As such, we will use Waters' hash extensively. Given an identity $\mathrm{ID} \in \{0,1\}^n$ and $\mathbf{u}[i] \in \mathbb{G}$ for $i \in [0,n]$ where $\mathrm{ID}[j]$ is the j^{th} bit of ID:

$$H(\mathbf{u}, \mathrm{ID}) = \mathbf{u}[0] \prod_{j \in [1,n]} \mathbf{u}[j]^{\mathrm{ID}[j]}$$

We use $x \xleftarrow{\$} X$ to denote x being picked at random from a set X and $S \subset_\$ X$ a set S being picked at random as a subset of X. We describe the scheme below with $|p|, n, d, k \in \Theta(\lambda)$ and $d \leq k \leq n/4$. The terms $e, \mathbb{G}_1, \mathbb{G}_2, g, p$ are parameters of the common bilinear map.

Setup (1^λ)**.** Pick $a \xleftarrow{\$} \mathbb{Z}_p$, $g_1 = g^a$. $g_2 \xleftarrow{\$} \mathbb{G}$, $\mathbf{T}[i] \xleftarrow{\$} \mathbb{G}, \mathbf{u}[i] \xleftarrow{\$} \mathbb{G}$ for $i \in [0,n]$. The public key is $(g_1, g_2, \mathbf{u}, \mathbf{T})$. The master secret key is g_2^a.

KeyGen. (PK, ID) Pick $S \subset_\$ [1,n]$ of size k, $\forall\ i \in S$, $r_i \xleftarrow{\$} \mathbb{Z}_p$ and output[3]:

$$K_{\mathrm{ID}}(S) = \left(\mathrm{ID}, S, \left(g_2^a \left[H(\mathbf{u}, \mathrm{ID}) T_i\right]^{r_i}, g^{r_i}\right)_{i \in S}\right).$$

Encrypt. (M, ID, S) Pick $c \xleftarrow{\$} \mathbb{Z}_p$. Select a $d-1$-degree polynomial q in $\mathbb{Z}_p$ at random with $q(0) = c$,

$$C_{\mathrm{ID}}(S) = \left(\mathrm{ID}, S, Me(g_1,g_2)^c, \left(g^{q(i)}, \left[H(\mathbf{u}, \mathrm{ID}) T_i\right]^{q(i)}\right)_{i \in S}\right).$$

Where the entries in the third component are given in ascending order of $i \in [1,n]$ to avoid ambiguity.

Decrypt($C_{\mathbf{ID}}(S')$,$K_{ID}(S)$). Take $I \subset S' \cap S$ of size d. For all $i \in I$, compute:

$$\frac{e(\left[H(\mathbf{u}, \mathrm{ID}) T_i\right]^{q(i)}, g^{r_i})}{e(g_2^a \left[H(\mathbf{u}, \mathrm{ID}) T_i\right]^{r_i}, g^{q(i)})} = \frac{1}{e(g, g_2^a)^{q(i)}} = \frac{1}{e(g_1,g_2)^{q(i)}},$$

$$Me(g_1,g_2)^c \prod_{i \in I} \left(\frac{1}{e(g_1,g_2)^{q(i)}}\right)^{\Delta_{i,I}(0)} = M.$$

Where $\Delta_{i,I}(j)$ is the Lagrange coefficient ($\Delta_{i,I}(j) = \prod_{k \in I \setminus \{i\}} \frac{k-j}{i-j}$).

[3] Notice it is possible to generate K_{ID}^α for all $\alpha \in [1,n]$ as necessary for the previous reduction.

CiphertextSanity($C_{\mathrm{ID}}(S)$, PK). First check the input is formatted as a valid ciphertext ($|S| = k$):

$$\left(\mathrm{ID}, S, C, \left(C_i^{(1)}, C_i^{(2)}\right)_{i \in S}\right)$$

Assuming $H(\mathbf{u}, ID)T_i \neq 1 \in \mathbb{G}$, since $\mathbb{G}, \mathbb{G}_T$ are of prime order, write the last two components for a single $i \in S$ as $g^{w_i}, [H(\mathbf{u}, ID)T_i]^{y_i}$ for some constants $w_i, y_i \in \mathbb{Z}_p$. If $H(\mathbf{u}, ID)T_i = 1 \in \mathbb{G}$, unless the second component above is also 1 output Not-Valid. The ciphertext is well formed if for all components with second entry not 1 the $w_i = y_i$ and the w_i values fall on the same $d-1$ degree polynomial q (note that the value $q(0)$ is actually completely irrelevant to the ciphertext sanity check since as long as all the w_i are on the same $d-1$ degree polynomial, if $q(0) = c$, it will be a valid encryption of $Ce(g,g)^{-c}$). To check that $w_i = y_i$, notice:

$$e\left(H(\mathbf{u}, \mathrm{ID})T_j), C_i^{(1)}\right) \stackrel{?}{=} e\left(C_i^{(2)}, g\right) \Leftrightarrow$$

$$e\left((H(\mathbf{u}, \mathrm{ID})T_j)^{w_i}, g\right) \stackrel{?}{=} e\left(H(\mathbf{u}, \mathrm{ID})T_j, g^{y_i}\right) \Leftrightarrow$$

$$e\left(H(\mathbf{u}, \mathrm{ID})T_j), g\right)^{w_i} \stackrel{?}{=} e\left(H(\mathbf{u}, \mathrm{ID})T_j, g\right)^{y_i} \Leftrightarrow w_i \stackrel{?}{=} y_i.$$

The final implication follows from the fact that both numbers being paired are not 1 in their original group of prime order and so the target group element is not 1 with non-trivial pairing. Now, to check that all the w_i lie on the same $d-1$ degree polynomial, pick a subset $I \subset S$ with $|J| = d$. Then, interpolate $C_i^{(1)}$ to all the other $x \in S \setminus I$ values. In other words, for all $x \in S \setminus I$, check that:

$$\prod_{i \in I} {C_i^{(1)}}^{\Delta_{i,I}(x)} = \prod_{i \in I} g^{w_i \Delta_{i,I}(x)} \stackrel{?}{=} g^{w_x} = C_x^{(1)}.$$

It's now clear that this process will only accept if all w_i lie on some polynomial q of degree not exceeding $d-1$. This suffices to show that the interpolation to $q(0)$ is unique no matter which d points are picked, showing this is a valid ciphertext (can only be decrypted to one value).

KeySanity($K_{\mathrm{ID}}(S)$, PK). First check that the key is formatted correctly. If $H(\mathbf{u}, ID)T_i = 1 \in \mathbb{G}$, checking this component is valid is trivial (the user then should have access to the master secret key and can check this with one pairing from the PK). We can now write each key component as:

$$\left(H(\mathbf{u}, ID)T_i\right)^r g_2^a, g^{r'}.$$

This will now be a valid key component if $r = r'$ (for elements in $\mathbb{Z}_p$ we write equality as elements in this group). Check (recall $g_1 = g^a$):

$$e\left[\left(H(\mathbf{u}, ID)T_i\right)^r g_2^a, g\right] \stackrel{?}{=} e\left(H(\mathbf{u}, ID)T_i, g^{r'}\right) e(g_2, g_1) \Leftrightarrow$$

$$e\left[\left(H(\mathbf{u}, ID)T_i\right), g\right]^r \stackrel{?}{=} e\left[\left(H(\mathbf{u}, ID)T_i\right), g\right]^{r'}.$$

Which is equivalent with $r \stackrel{?}{=} r'$ since g, $H(\mathbf{u}, ID)T_i$ are not 1.

4.1 Proof of Security

We will now prove the *Dummy-IBE* security of the given construction. Let q be an upper bound on the number of queries an adversary $\mathcal{A}$ makes to succeed in the *Dummy-IBE* security game with non-negligible probability ϵ. We now describe how to use this simulator to succeed in the DBDH game with non-negligible probability. Below m will be a value polynomial in an upper bound of the number of queries the adversary makes that will be initialized later in the analysis. We now describe the simulator $\mathcal{S}$ that will be given as input (A, B, C, Z) where $(g, g_1, g_2, C) = (g, g^a, g^b, g^c)$ and $Z = e(g,g)^{abc}$ or $Z \xleftarrow{\$} \mathbb{G}_T$. Generate Φ a random k element subset of $[1, n]$, which will serve as the dummy attribute set of the challenge query and:

$$\mathbf{x}[i] \xleftarrow{\$} [0, m-1] \text{ for } i \in [1, n] \ , \ \mathbf{x}[0] \xleftarrow{\$} [-n(m-1), 0]$$

$$\mathbf{y}[i] \xleftarrow{\$} \mathbb{Z}_p \text{ for } i \in [0, m-1].$$

Define the following two functions:

$$F'(\mathbf{x}, \mathrm{ID}) = \mathbf{x}[0] + \sum_{i=1}^{n} \mathbf{x}[i]\mathrm{ID}[i]$$

$$G'(\mathbf{y}, I) = \mathbf{y}[0] + \sum_{i=1}^{n} \mathbf{y}[i]\mathrm{ID}[i].$$

Our simulator $\mathcal{S}$ will abort unless for exactly one key query $\mathbf{ID}[i]$, $F'(\mathbf{x}, \mathbf{ID}[i]) = 0$ and for all others $F'(\mathbf{x}, \mathbf{ID}[j]) \notin \{0, 1\}$. Additionally, $\mathcal{S}$ will abort and guess randomly unless $\mathbf{ID}[i]$ as specified above is not the challenge identity (recall that by the definition of *Dummy-IBE* security, we may assume that the challenge identity has been queried once during the key queries phase and that the challenge dummy-attribute set overlaps the dummy-attribute set the adversary received during the key generation phase in no more than $d-1$ indices). We now describe how $\mathcal{S}$ will generate the public key and answer key queries.

Simulated Public Key Generation. $\mathcal{S}$.**Setup**(1^λ) is defined as follows: For $i \in [1, n]$, $\mathbf{B}[i] \xleftarrow{\$} \mathbb{Z}_p$ and,

$$\mathbf{T}[j] = \begin{cases} g_2 g^{\mathbf{B}[j]}, \text{ if } j \in \Phi; \\ g^{\mathbf{B}[j]}, \quad \text{otherwise.} \end{cases}$$

Related to $\mathbf{T}$, define $G(\mathbf{y}, \mathrm{ID}, j) = G'(\mathbf{y}, \mathrm{ID}) + \mathbf{B}[j]$ and,

$$F(\mathbf{x}, \mathrm{ID}, \Phi, j) = \begin{cases} F'(\mathbf{x}, \mathrm{ID}) + 1, \text{ if } j \in \Phi; \\ F'(\mathbf{x}, \mathrm{ID}), \quad \text{otherwise.} \end{cases}$$

Output PK $= (g_1, g_2, \mathbf{u}, \mathbf{T})$. Notice if $\mathbf{u}[i] = g_2^{\mathbf{x}[i]} g^{\mathbf{y}[i]}$ then, $H(\mathbf{u}, \mathrm{ID})\mathbf{T}[i] = g_2^{F(\mathbf{x}, \mathrm{ID}, \Phi, i)} g^{G(\mathbf{y}, \mathrm{ID}, j)}$. Observe the output is completely independent of Φ and the

output distribution is precisely that of the true public key generation algorithm.

Simulated KeyGen Queries. If $F'(\mathbf{x}, \mathrm{ID}) \notin \{-1, 0\}$, $\mathcal{S}$.**KeyGen**(ID) is defined as follows: Pick $S \subset_{\$} [1, n]$ of size k. For $j \in S$: $r \xleftarrow{\$} \mathbb{Z}_p$ and let $w = F(\mathbf{x}, \mathrm{ID}, \Phi, j)^{-1}$ (which exists since $F'(\mathbf{x}, \mathrm{ID}) \neq 0$). Set:

$$(K_j^{(1)}, K_j^{(2)}) = (g_2^{F(\mathbf{x},\mathrm{ID},\Phi,j)\cdot r} g^{G(\mathbf{y},\mathrm{ID},\Phi)\cdot r} g_1^{-G(\mathbf{y},\mathrm{ID},\Phi)\cdot w}, g^r g_1^{-w})$$

For all $j \in S$. Output $K_{\mathrm{ID}}(S) = \left(\mathrm{ID}, S, (K_j^{(1)}, K_j^{(2)})_{j\in S}\right)$. Notice for $j \in S$:

$$\begin{aligned} g_2^{F(\mathbf{x},\mathrm{ID},\Phi,j)\cdot r} g^{G(\mathbf{y},\mathrm{ID},\Phi)\cdot r} g_1^{-G(\mathbf{y},\mathrm{ID},\Phi)\cdot w} &= g_2^a (g_2^{F(\mathbf{x},\mathrm{ID},\Phi,j)} g^{G(\mathbf{y},\mathrm{ID},\Phi)})^{-a\cdot w} (g_2^{F(\mathbf{x},\mathrm{ID},\Phi,j)} g^{G(\mathbf{y},\mathrm{ID},\Phi)})^r \\ &= g_2^a (g_2^{F(\mathbf{x},\mathrm{ID},\Phi,j)} g^{G(\mathbf{y},\mathrm{ID},\Phi)})^{r-a\cdot w}. \end{aligned}$$

And similarly $g^r g_1^{-w} = g^{r-a\cdot w}$. Therefore, this is a valid key component with randomness $r - a \cdot w$, which is also distributed uniformly over $\mathbb{Z}_p$. So, every key component is generated from the correct distribution if $F(\mathbf{x}, \mathrm{ID}) \notin \{-1, 0\}$. This is where the main divergence with Waters' proof occurs, we must be able to handle one query to the identity that will later be challenged on. For this, we notice that if $j \in \Phi$, and $F'(\mathbf{x}, \mathrm{ID}) = 0$ then, $F(\mathbf{x}, \mathrm{ID}, \Phi, j) \neq 0$ and we can repeat the above process by taking $S = \Phi$.

If $F'(\mathbf{x}, \mathrm{ID}) = 0$, $\mathcal{S}$.**KeyGen**(ID) is the same as the above except instead of $S \subset_{\$} [1, n]$, take $S = \Phi$.

Recall that $\mathcal{S}$ aborts if there is ever more than one query with $F'(\mathbf{x}, \mathrm{ID}) = 0$ and that the output of the simulated public key is independent of Φ. Therefore, if Φ is only used once during the key queries phase, as the dummy-attribute set of a single query, the view of the adversary interacting with $\mathcal{S}$ is still identical to the view of the adversary interacting with the true *Dummy-IBE* scheme.

Simulated Challenge Ciphertext. Assume $F'(\mathbf{x}, \mathrm{ID}) = 0$ and $|S \cap \Phi| < d$ then,

$\mathcal{S}$.**Challenge**$(M_0, M_1, \mathrm{ID}, S)$ is defined as follows: Pick $\gamma \xleftarrow{\$} \{0, 1\}$ and let $C_1 = Z \cdot M_\gamma$ and choose $S \supset K \supset S \cap \Phi$, with $|K| = d - 1$ and:

For $i \in K$:

$$r_i \xleftarrow{\$} \mathbb{Z}_p \ , \ C_i^{(1)} = g^{r_i} \ , \ C_i^{(2)} = \left(g_2^{F(\mathbf{x},\mathrm{ID},\Phi,i)} g^{G(\mathbf{y},\mathrm{ID},i)}\right)^{r_i}.$$

For $i \notin K$:

$$C_i^{(1)} = (g^c)^{\Delta_{0,K}(i)} \prod_{j\in K} (g^{r_j})^{\Delta_{j,K}(i)} \, , C_i^{(2)} = \left(C_i^{(1)}\right)^{G(\mathbf{y},\mathrm{ID},i)}.$$

Respond to the query with:

$$C_{\mathrm{ID}}(S) = \left(\mathrm{ID}, S, C_1, \left(C_i^{(1)}, C_i^{(2)}\right)_{i\in S}\right).$$

Remark (1). Notice the encryption is valid with the implied polynomial $q(x)$ in the exponent being defined by $q(i) = r_i \xleftarrow{\$} \mathbb{Z}_p$ for all $i \in K$ and $q(0) = c$ where $C = g^c$. Since choosing a $d-1$ degree polynomial at random is equivalent to specifying its value randomly at d points, the encryption is valid and uniform (since r_i are generated at random during the simulated challenge and the view of the adversary has been independent of c which was chosen at random in the DBDH initialization). This shows that the $C_i^{(1)}$ components are correctly generated from the stated distribution for $i \in K$.

Remark (2). Additionally, since $F'(\mathbf{x}, \mathrm{ID}) = 0$, we have that for all $i \notin K (\Rightarrow i \notin \Phi$ if $i \in S)$, $F(\mathbf{x}, \mathrm{ID}, \Phi, i) = 0$ and therefore $H(\mathbf{u}, \mathrm{ID})T_i = g^{G(\mathbf{y}, \mathrm{ID}, i)}$ which shows the second component is also generated correctly with implied polynomial q. Therefore, as long as $F'(\mathbf{x}, ID) = 0$ the $C_i^{(b)}$ values are drawn from the correct distribution.

Therefore, if Z is $e(g_1, g_2)^c$ we have the above is a uniform encryption of M_γ. Otherwise, Z is a random element of $\mathbb{G}_T$ and the ciphertext gives no information on the choice of γ.

To remind the reader of the relevant security game, the adversary $\mathcal{A}$ will make q distinct key queries $\mathbf{ID}[i]$ to which the simulator will respond with keys $K_{\mathbf{ID}[i]}(\mathbf{S}[i])$ where $\mathbf{S}[i]$ is a k element subset of $[1, n]$ consisting of the dummy attributes of the key. Finally, the adversary will make a challenge query with $M_0, M_1, S^*, \mathbf{ID}[j]$ where $\mathbf{ID}[j]$ with $j \in [1, q]$ was queried before and $|S^* \cap \mathbf{S}[j]| < d$. Our simulator will not abort if and only if:

1. $F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0, 1\}$ for all $i \in [1, q] \setminus \{j\}$ for some $j \in [1, q]$,
2. $F'(\mathbf{x}, \mathbf{ID}[j]) = 0$,
3. The challenge identity is $\mathbf{ID}[j]$.

This is very similar to the requirement in Waters' IBE, where it is required that $F'(\mathbf{x}, \mathbf{ID}[i]) \neq 0$ for all key queries and $F'(\mathbf{x}, \mathbf{ID}[j]) = 0$ for the challenge identity $\mathbf{ID}[j]$. Here, we require that the challenge identity be queried once before, with the small caveat that for most key queries we are aborting on two values instead of one, but the impact of this on analysis is minimal. Notice that if our algorithm does not abort, it is drawing from the same distribution as the real *Dummy-IBE* scheme since while Φ is used as the dummy attribute set for a single query, the view of the adversary otherwise is independent of Φ (the requirement that $|\Phi \cap S| < d$ is automatically fulfilled since $\mathcal{A}$ was given Φ as the dummy attribute set of $\mathbf{ID}[j]$ and therefore, by the rules of the *Dummy-IBE* security game, the dummy attribute set of the challenge must overlap the dummy attribute set received during the key generation phase in less than d indices). It only remains to bound below the probability of aborting naturally and introduce an artifical abort step to make sure the success probability of the simulator is not correlated with the probability of aborting. We defer bounding the abord probability to **Appendix A**. If for any sequence of queries, our simulation has a non-negligible probability of perfectly simulating the behavior of *Dummy-IBE* security game in the case where the fourth member of the DBDH tuple is $Z = e(g, g)^{abc}$ and

in the other case, all information on the challenge message is lost, we can use the advantage of a distinguishing adversary to discern which of the above two possibilities the fourth member of the DBDH tuple is, since its advantage can only be maintained (information theoretically) in the case where $Z = e(g, g)^{abc}$.

5 Conclusion

We have demonstrated the first provably secure black-box accountable authority IBE scheme, answering the primary question posed in multiple previous works [13,15] using the standard DBDH assumption. To achieve this goal, we introduced the notion of *Dummy-IBE* encryption, a hybrid between usual IBE and *Attribute-Based Encryption* where the exact attributes given to an identity are not important but which should exist for some added functionality (in this case tracing), which may be of independent interest.

References

1. Bellare, M., Ristenpart, T.: Simulation without the Artificial Abort: Simplified Proof and Improved Concrete Security for Waters' IBE Scheme. In: Joux, A. (ed.) EUROCRYPT 2009. LNCS, vol. 5479, pp. 407–424. Springer, Heidelberg (2009)
2. Boneh, D., Franklin, M.: Identity-based encryption from the Weil pairing. SIAM Journal on Computing 32(3), 586–615 (2003)
3. Boneh, D., Boyen, X.: Secure Identity Based Encryption Without Random Oracles. In: Franklin, M.K. (ed.) CRYPTO 2004. LNCS, vol. 3152, pp. 443–459. Springer, Heidelberg (2004)
4. Boneh, D., Franklin, M.K.: Identity-Based Encryption from the Weil Pairing. In: Kilian, J. (ed.) CRYPTO 2001. LNCS, vol. 2139, pp. 213–229. Springer, Heidelberg (2001)
5. Camenisch, J., Neven, G., Shelat, A.: Simulatable Adaptive Oblivious Transfer. In: Naor, M. (ed.) EUROCRYPT 2007. LNCS, vol. 4515, pp. 573–590. Springer, Heidelberg (2007)
6. Canetti, R.: Security and composition of multiparty cryptographic protocols. Journal of Cryptology 13(1), 143–202 (2000)
7. Cocks, C.: An Identity Based Encryption Scheme Based on Quadratic Residues. In: Honary, B. (ed.) Cryptography and Coding 2001. LNCS, vol. 2260, pp. 360–363. Springer, Heidelberg (2001)
8. Even, S., Goldreich, O., Lempel, A.: A randomized protocol for signing contracts. Communications of the ACM 28(6), 647 (1985)
9. Gentry, C.: Practical identity-based encryption without random oracles. In: Vaudenay, S. (ed.) EUROCRYPT 2006. LNCS, vol. 4004, pp. 445–464. Springer, Heidelberg (2006)
10. Gentry, C., Halevi, S.: Hierarchical Identity Based Encryption with Polynomially Many Levels. In: Reingold, O. (ed.) TCC 2009. LNCS, vol. 5444, pp. 437–456. Springer, Heidelberg (2009)
11. Gentry, C., Silverberg, A.: Hierarchical ID-Based Cryptography. In: Zheng, Y. (ed.) ASIACRYPT 2002. LNCS, vol. 2501, pp. 548–566. Springer, Heidelberg (2002)

12. Goyal, V., Pandey, O., Sahai, A., Waters, B.: Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM Conference on Computer and Communications Security, p. 98. ACM, New York (2006)
13. Goyal, V.: Reducing Trust in the PKG in Identity Based Cryptosystems. In: Menezes, A. (ed.) CRYPTO 2007. LNCS, vol. 4622, pp. 430–447. Springer, Heidelberg (2007)
14. Goyal, V., Jain, A., Pandey, O., Sahai, A.: Bounded Ciphertext Policy Attribute Based Encryption. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part II. LNCS, vol. 5126, pp. 579–591. Springer, Heidelberg (2008)
15. Goyal, V., Lu, S., Sahai, A., Waters, B.: Black-box accountable authority identity-based encryption. In: Ning, P., Syverson, P.F., Jha, S. (eds.) ACM Conference on Computer and Communications Security, pp. 427–436. ACM, New York (2008)
16. Green, M., Hohenberger, S.: Blind identity-based encryption and simulatable oblivious transfer. In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 265–282. Springer, Heidelberg (2007)
17. Horwitz, J., Lynn, B.: Toward Hierarchical Identity-Based Encryption. In: Knudsen, L.R. (ed.) EUROCRYPT 2002. LNCS, vol. 2332, pp. 466–481. Springer, Heidelberg (2002)
18. Lewko, A., Okamoto, T., Sahai, A., Takashima, K., Waters, B.: Fully Secure Functional Encryption: Attribute-Based Encryption and (Hierarchical) Inner Product Encryption. In: Gilbert, H. (ed.) EUROCRYPT 2010. LNCS, vol. 6110, pp. 62–91. Springer, Heidelberg (2010)
19. Libert, B., Vergnaud, D.: Towards Black-Box Accountable Authority IBE with Short Ciphertexts and Private Keys. In: Jarecki, S., Tsudik, G. (eds.) PKC 2009. LNCS, vol. 5443, pp. 235–255. Springer, Heidelberg (2009)
20. Lindell, A.Y.: Efficient Fully-Simulatable Oblivious Transfer. In: Malkin, T. (ed.) CT-RSA 2008. LNCS, vol. 4964, pp. 52–70. Springer, Heidelberg (2008)
21. Okamoto, T., Takashima, K.: Fully Secure Functional Encryption with General Relations from the Decisional Linear Assumption. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 191–208. Springer, Heidelberg (2010)
22. Ostrovsky, R., Sahai, A., Waters, B.: Attribute-based encryption with non-monotonic access structures. In: Proceedings of the 14th ACM Conference on Computer and Communications Security, p. 203. ACM, New York (2007)
23. Sahai, A., Waters, B.: Fuzzy identity-based encryption. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 457–473. Springer, Heidelberg (2005)
24. Shamir, A.: Identity-based cryptosystems and signature schemes. In: Blakely, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 47–53. Springer, Heidelberg (1985)
25. Waters, B.: Efficient identity-based encryption without random oracles. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 114–127. Springer, Heidelberg (2005)

A Bounding the Abort Probability

Artificial Abort. Additionally to the probability of aborting naturally, we must make sure that if the simulator $\mathcal{S}$'s probability of success is not correlated with its probability of aborting. Since this method is standard by now, we refer the reader to Waters' original work [25] for more in depth analysis of the necessity

of this step. Our application differs little from the analysis in [25] except for a factor of 2 due to the fact that we have two restricted values but include the analysis for completeness[4].

As in Waters' scheme we can now define a second simulator which first generates the secret key g_2^a before initializing $\mathbf{x}$ and $\mathbf{y}$. With the master secret key, it is now able to perfectly respond to all key generation queries. Now, this second simulator will abort and guess randomly unless $\mathbf{ID}[j]$, the challenge identity (which has been queried once before) satisfies $F'(\mathbf{x}, \mathbf{ID}[j]) = 0$ and for all other identities keys are queried for $\mathbf{ID}[i] \neq \mathbf{ID}[j]$, $F'(\mathbf{x}, \mathbf{ID}[j]) \notin \{0, 1\}$.

As long as we can show the second simulator has a non-negligible probability of not aborting for any set of queries, by using artificial aborts, we can get a near-uniform, non-negligible probability to abort and guess randomly for any set of identity queries. Then, our first simulator works as a distinguisher in the DBDH game since if for a given set of queries our simulator has not aborted, the adversary will have a non-negligible advantage of guessing the correct b if $Z = e(g, g)^{abc}$ and no advantage in guessing b if Z is random since b is information theoretically masked in the ciphertext.

Analysis of Abort Probability: We now give a lower bound on the probability for a set of identities $\{\mathbf{ID}[i]\}_{i \in [1,q]}$ with $\mathbf{ID}[j]$, $j \in [1, q]$ the challenge identity that the second simulator does not abort.

$$\begin{aligned}
\Pr[\overline{\texttt{abort}}] &= \Pr[\bigwedge_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0,1\}] \ \wedge \ F'(\mathbf{x}, \mathbf{ID}[j]) = 0 \\
&= (1 - \Pr[\bigvee_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\}]) \times \Pr[F'(\mathbf{x}, \mathbf{ID}[j]) = 0 | \bigwedge_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0,1\}] \\
&\geq (1 - \sum_{i \in [1,q] \setminus \{j\}} \Pr[F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\}]) \times \Pr[F'(\mathbf{x}, \mathbf{ID}[j]) = 0 | \bigwedge_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0,1\}] \\
&\geq (1 - \frac{2q}{(n+1)m}) \Pr[F'(\mathbf{x}, \mathbf{ID}[j]) = 0 | \bigwedge_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0,1\}]
\end{aligned}$$

As in lines (1e) through (1i) in Waters' derivation [25] we can simplify the probability on the right to:

$$= \frac{\Pr[F'(\mathbf{x}, \mathbf{ID}[j]) = 0]}{\Pr[\bigwedge_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \notin \{0,1\}]} \times \left(1 - \Pr[\bigvee_{i \in [1,q] \setminus \{j\}} F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\} | F'(\mathbf{x}, \mathbf{ID}[j]) = 0]\right)$$

Which can be bounded below by:

$$\frac{1}{(n+1)m}(1 - \sum_{i \in [1,q] \setminus \{j\}} \Pr[F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\} | F'(\mathbf{x}, \mathbf{ID}[j]) = 0]).$$

[4] Note that it is possible to prove our main result in a fashion similar to Bellare and Ristenpart's recent simplification [1] of Waters' proof using game playing techniques but we include this proof due to its transparency. A version of this paper which includes a game playing proof of security is available from the authors.

Notice we can replace $F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\}$ with $F'(\mathbf{x}, \mathbf{ID}[i]) \in \{0,1\} \bmod m$ to get a lower bound. Now, for each $\mathbf{ID}[i]$ with $i \neq j$, $F'(\mathbf{x}, \mathbf{ID}[i])$ modulo m is independent from $F'(\mathbf{x}, \mathbf{ID}[j])$ by the properties of the Waters hash, and therefore, using this substitution the above is bounded below by:

$$\frac{1}{(n+1)m}\left(1 - \frac{2q}{m}\right).$$

Substituting back to the original inequality yields:

$$\left(1 - \frac{2q}{(n+1)m}\right)\frac{1}{(n+1)m}\left(1 - \frac{2q}{m}\right) \geq \frac{1}{(n+1)m}\left(1 - \frac{4q}{m}\right).$$

Which can be taken to be non-negligible with $m = 8q$. This shows for any sequence of queries, the probability that the simulator does not abort can be bounded below by a non-negligible amount. As discussed before, using artificial aborts, this suffices to prove security.