An extended abstract of this paper appears in Kaoru Kurosawa (Ed.): Advances in Cryptology - ASIACRYPT 2007, volume 4833 of Lecture Notes in Computer Science, pages 265–282, Springer-Verlag, 2007. This is the full version.

Blind Identity-Based Encryption and Simulatable Oblivious Transfer

Matthew Green Susan Hohenberger

Information Security Institute
The Johns Hopkins University
3400 N. Charles St.
Baltimore, MD 21218
{mgreen,susan}@cs.jhu.edu

Abstract

In an identity-based encryption (IBE) scheme, there is a *key extraction* protocol where a user submits an identity string to a master authority who then returns the corresponding secret key for that identity. In this work, we describe how this protocol can be performed efficiently and in a *blind* fashion for several known IBE schemes; that is, a user can obtain a secret key for an identity without the master authority learning anything about this identity.

We formalize this notion as *blind IBE* and discuss its many practical applications. In particular, we build upon the recent work of Camenisch, Neven, and shelat [CNS07] to construct oblivious transfer (OT) schemes which achieve full simulatability for both sender and receiver. OT constructions with comparable efficiency prior to Camenisch *et al.* were proven secure in the weaker half-simulation model. Our OT schemes are constructed from the blind IBE schemes we propose, which require only static complexity assumptions (*e.g.*, DBDH) whereas prior comparable schemes require dynamic assumptions (*e.g.*, q-PDDH).

1 Introduction

In an oblivious transfer (OT_k^N) protocol, introduced by Rabin [Rab81] and generalized by Even, Goldreich and Lempel [EGL82] and Brassard, Crépeau and Robert [BCR86], a Sender with messages $M_1, \ldots, M_N$ and a Receiver with indices $\sigma_1, \ldots, \sigma_k \in [1, N]$ interact in such a way that at the end the Receiver obtains $M_{\sigma_1}, \ldots, M_{\sigma_k}$ without learning anything about the other messages and the Sender does not learn anything about $\sigma_1, \ldots, \sigma_k$. Naor and Pinkas were the first to consider an *adaptive* setting, $\mathsf{OT}_{k\times 1}^N$, where the sender may obtain $M_{\sigma_{i-1}}$ before deciding on σ_i [NP99b]. Oblivious transfer is a useful, interesting primitive in its own right, but it has even greater significance as OT_1^4 is a key building block for secure multi-party computation [Yao86, GMW87, Kil88]. Realizing efficient protocols under modest complexity assumptions is therefore an important goal.

The definition of security for oblivious transfer has been evolving. Informally, security is defined with respect to an ideal-world experiment in which the Sender and Receiver exchange messages via

This work was supported in part by the NSF under grant CT-0716142.

a trusted party. An OT protocol is secure if, for every real-world cheating Sender (resp., Receiver) we can describe an ideal-world counterpart who gains as much information from the ideal-world interaction as from the real protocol. Bellare and Micali [BM89] presented the first practical OT_1^2 protocol to satisfy this intuition in the honest-but-curious model. This was followed by practical OT protocols due to Naor and Pinkas [NP99a, NP99b, NP01] in the "half-simulation" model where the simulation-based model (described above) is used only to show Sender security and Receiver security is defined by a simpler game-based definition. Almost all efficient OT protocols are proven secure with respect to the half-simulation model, *e.g.,* [NP99b, NP99a, NP01, DHRS04, OK04, Kal05, CT05]. Unfortunately, Naor and Pinkas demonstrated that this model permits *selective-failure* attacks, in which a malicious Sender can induce transfer failures that are dependent on the message that the Receiver requests [NP99b].

Recently, Camenisch, Neven, and shelat [CNS07] proposed practical $\mathsf{OT}_{k\times 1}^N$ protocols that are secure in the "full-simulation" model, where the security of both the Sender and Receiver are simulation-based. These simulatable OT protocols are particularly nice because they can be used to construct other cryptographic protocols in a simulatable fashion. More specifically, Camenisch *et al.* [CNS07] provide two distinct results. First, they show how to efficiently construct $\mathsf{OT}_{k\times 1}^N$ generically from any unique blind signature scheme in the random oracle model. The two known efficient unique blind signature schemes due to Chaum [Cha82] and Boldyreva [Bol03] both require *interactive* complexity assumptions: one-more-inversion RSA and chosen-target CDH, respectively. (Interestingly, when instantiated with Chaum signatures, this construction coincides with a prior one of Ogata and Kurosawa [OK04] that was analyzed in the half-simulation model.) Second, they provide a clever $\mathsf{OT}_{k\times 1}^N$ construction in the standard model based on dynamic complexity assumptions, namely the q-Power Decisional Diffie-Hellman (*i.e.*, in a bilinear setting $e : \mathbb{G}\times\mathbb{G} \to \mathbb{G}_T$, given $(g, g^x, g^{x^2}, \ldots, g^{x^q}, H)$ where $g \leftarrow \mathbb{G}$ and $H \leftarrow \mathbb{G}_T$, distinguish $(H^x, H^{x^2}, \ldots, H^{x^q})$ from random values) and q-Strong Diffie-Hellman (q-SDH) assumptions. (Unfortunately, Cheon showed that q-SDH requires larger than commonly used security parameters [Che06]). These dynamic (including interactive) assumptions seem significantly stronger than those, such as DDH and quadratic residuosity, used to construct efficient OT schemes in the half-simulation model. Thus, a well-motivated problem is to find efficient, fully-simulatable and *adaptive* OT schemes under weaker complexity assumptions.

Our Contributions. In this work, we provide, to our knowledge, the first efficient and fully-simulatable OT_k^N and $\mathsf{OT}_{k\times 1}^N$ schemes secure under *static* complexity assumptions (*e.g.*, DBDH, where given (g, g^a, g^b, g^c), it is hard to distinguish $e(g, g)^{abc}$ from random). We summarize our results as follows.

First, we introduce a building block, which is of independent interest. In identity-based encryption (IBE) [Sha84], there is an *extraction* protocol where a user submits an identity string to a master authority who then returns the corresponding decryption key for that identity. We formalize the notion of *blindly* executing this protocol, in a strong sense; where the authority does not learn the identity nor can she cause failures dependent on the identity, and the user learns nothing beyond the normal extraction protocol. This concept has similarities to recent work by Goyal [Goy07], in which a user wishes to hide certain characteristics of an extracted IBE key from the authority. In §3.1, we describe efficient *blind extraction* protocols satisfying this definition for the IBE schemes due to Boneh and Boyen [BB04] and Waters [Wat05] (using a generalization proposed independently by Naccache [Nac05] and Chatterjee and Sarkar [CS05]). The latter protocol is similar to

a blind signature scheme proposed by Okamoto [Oka06]. We call IBE schemes supporting efficient blind extraction protocols: *blind IBE*, for short.

Second, we present an efficient and fully-simulatable OT_k^N protocol constructed from any of the proposed blind IBE schemes (without requiring additional assumptions), and thus our constructions are secure under only DBDH. Intuitively, consider the following OT_k^N construction. The Sender runs the IBE setup algorithm and sends the corresponding public parameters to the Receiver. Next, for $i = 1$ to N, the Sender encrypts M_i under identity "i" and sends this ciphertext to the Receiver. To obtain k messages, the Receiver blindly extracts k decryption keys for identities of his choice and uses these keys to decrypt and recover the corresponding messages. While this simple protocol does not appear to be simulatable, we are able to appropriately modify it. (Indeed, one must also be cautious of possibly malformed ciphertexts, as we discuss later.) Our constructions from blind IBE are inspired by the Camenisch *et al.* [CNS07] generic construction from unique blind signatures. Indeed, recall that the secret keys sk_{id} of any fully-secure IBE can be viewed as signatures by the authority on the message id [BF01]. Camenisch *et al.* [CNS07] require *unique* blind signatures, whereas we do not; however, where they require unforgeability, we require that our "blind key extraction" protocol does not jeopardize the semantic security of the IBE.

Third, we present an efficient and fully-simulatable $\mathsf{OT}_{k\times 1}^N$ protocol constructed from our proposed blind IBE schemes in the random oracle model. We discuss how to remove these oracles at an additional cost. This improves on the complexity assumptions required by the comparable random-oracle scheme in Camenisch *et al.* [CNS07], although we leave the same improvement for their adaptive construction without random oracles as an open problem. Finally, in §5, we discuss the independent usefulness of blind IBE to other applications, such as blind signatures, anonymous email, and encrypted keyword search.

Subsequent Work. The original publication of this work left open some problems which have since been addressed. We now briefly summarize these developments. First, Lindell [Lin08] proposed several efficient and simulatable OT_k^N protocols secure under weaker assumptions than those used in this work, *e.g.,* DDH and Quadratic Residuosity. While these protocols have efficiency comparable to (or better than) the non-adaptive OT_k^N protocols in this work, they do not address the *adaptive* case ($\mathsf{OT}_{k\times 1}^N$). Thus it remains open to devise efficient and simulatable $\mathsf{OT}_{k\times 1}^N$ under weaker assumptions than those used by this work and [CNS07].

Other works have proposed efficient OT protocols secure under *concurrent* execution. Peikert, Vaikuntanathan and Waters [PVW07] recently proposed an elegant framework for constructing (non-adaptive) OT_k^N using "messy keys", and showed how to realize these in the Universal Composability (UC) model of Canetti [Can01] under DDH, Quadratic Residuosity, or lattice assumptions. Independently, Green and Hohenberger [GH08] proposed a UC-secure adaptive $\mathsf{OT}_{k\times 1}^N$ using assumptions in bilinear groups. While each of these works represents a theoretical advance, we note that the adaptive protocols of this work and [CNS07] are still the most efficient choice for applications where concurrent execution is not required.

2 Technical Preliminaries

Let $\mathsf{BMsetup}$ be an algorithm that, on input the security parameter 1^κ, outputs the parameters for a bilinear mapping as $\gamma = (q, g, \mathbb{G}, \mathbb{G}_T, e)$, where g generates $\mathbb{G}$, both $\mathbb{G}$ and $\mathbb{G}_T$ have prime order q, and $e : \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T$. In our schemes, we will require that the correctness of these parameters be

publicly verifiable (Chen *et al.* [CCS07] describe efficient techniques for verifying these parameters in a typical instantiation). We will refer to the following complexity assumption made in these groups.

Decisional Bilinear Diffie-Hellman (DBDH) [BF01]: Let $\mathsf{BMsetup}(1^\kappa) \rightarrow (q, g, \mathbb{G}, \mathbb{G}_T, e)$. For all p.p.t. adversaries Adv, the following probability is strictly less than $1/2 + 1/\mathrm{poly}(\kappa)$: $\Pr[a, b, c, d \leftarrow \mathbb{Z}_q;\ x_0 \leftarrow e(g,g)^{abc};\ x_1 \leftarrow e(g,g)^d;\ z \leftarrow \{0,1\};\ z' \leftarrow \mathsf{Adv}(g, g^a, g^b, g^c, x_z) : z = z']$.

Known Discrete-Logarithm-Based, Zero-Knowledge Proofs. We use known techniques for proving statements about discrete logarithms, such as (1) proof of knowledge of a discrete logarithm modulo a prime [Sch91], (2) proof that a committed value lies in a given integer interval [CFT98, CM99, Bou00], and also (3) proof of the disjunction or conjunction of any two of the previous [CDS94]. These protocols are secure under the discrete logarithm assumption, although some implementations of (2) require the Strong RSA assumption.

When referring to the proofs above, we will use the notation of Camenisch and Stadler [CS97]. For instance, $PoK\{(x, r) : y = g^x h^r \wedge (1 \leq x \leq n)\}$ denotes a zero-knowledge proof of knowledge of integers x and r such that $y = g^x h^r$ holds and $1 \leq x \leq n$. All values not in enclosed in ()'s are assumed to be known to the verifier. We can apply the Fiat-Shamir heuristic [FS86] to make such proofs non-interactive in the random oracle model.

Commitments. Let $(\mathsf{CSetup}, \mathsf{Commit}, \mathsf{Decommit})$ be a commitment scheme where CSetup generates public parameters ρ; on input a message M, $\mathsf{Commit}(\rho, M)$ outputs a pair $(\mathcal{C}, \mathcal{D})$; and $\mathsf{Decommit}(\rho, M, \mathcal{C}, \mathcal{D})$ outputs 1 if $\mathcal{D}$ decommits $\mathcal{C}$ to M, or 0 otherwise. Our subsequent constructions require an efficient protocol for proving knowledge of a decommitment $\mathcal{D}$ with respect to $(\rho, M, \mathcal{C})$. We recommend using the Pedersen commitment scheme [Ped92] based on the discrete logarithm assumption, in which the public parameters are a group of prime order q, and random generators $(g_0, \ldots, g_m)$. In order to commit to the values $(v_1, \ldots, v_m) \in \mathbb{Z}_q^m$, pick a random $r \in \mathbb{Z}_q$ and set $\mathcal{C} = g_0^r \prod_{i=1}^m g_i^{v_i}$ and $\mathcal{D} = r$. Schnorr's technique [Sch91] is used to efficiently prove knowledge of the value $\mathcal{D} = r$.

3 Blind Identity-Based Encryption

An identity-based encryption (IBE) scheme supports two types of players: a single master authority and multiple users; together with the algorithms Setup, $\mathsf{Encrypt}$, $\mathsf{Decrypt}$ and the protocol $\mathsf{Extract}$. Let us provide some input/output specification for these protocols with intuition for what they do.

Notation: Let $\mathcal{I}$ be the identity space and $\mathcal{M}$ be the message space. We write $P(\mathcal{A}(a), \mathcal{B}(b)) \rightarrow (c, d)$ to indicate that protocol P is between parties $\mathcal{A}$ and $\mathcal{B}$, where a is $\mathcal{A}$'s input, c is $\mathcal{A}$'s output, b is $\mathcal{B}$'s input and d is $\mathcal{B}$'s output.

- In the $\mathsf{Setup}(1^\kappa, c(\kappa))$ algorithm, on input a security parameter 1^κ and a description of an the identity space $|\mathcal{I}| \leq 2^{c(\kappa)}$ where $c(\cdot)$ is a computable, polynomially-bounded function, the master authority $\mathcal{P}$ outputs master parameters *params* and a master secret key *msk*.
- In the $\mathsf{Extract}(\mathcal{P}(params, msk), \mathcal{U}(params, id)) \rightarrow (id, sk_{id})$ protocol, an honest user $\mathcal{U}$ with identity $id \in \mathcal{I}$ obtains the corresponding secret key sk_{id} from the master authority $\mathcal{P}$ or

outputs an error message. The master authority's output is **the identity** id or an error message.[1] (Note that $\mathcal{P}$ is permitted to abort the protocol selectively based on id.)
- In the $\mathsf{Encrypt}(params, id, m)$ algorithm, on input identity $id \in \mathcal{I}$ and message $m \in \mathcal{M}$, any party can output ciphertext C.
- In the $\mathsf{Decrypt}(params, id, sk_{id}, C)$ algorithm, on input a ciphertext C, the user with sk_{id} outputs a message $m \in \mathcal{M}$ or the distinguished symbol ϕ.

Throughout the remainder of the text we will assume that *params* defines $\mathcal{I}$ and $\mathcal{M}$.

Definition 3.1 (Selective-Identity Secure IBE (IND-sID-CPA) [CHK04]) Let κ be a security parameter, $c(\cdot)$ be a polynomially-bounded function, $|\mathcal{I}| \leq 2^{c(\kappa)}$ and $\mathcal{M}$ be the message space. An IBE is IND-sID-CPA-secure if every p.p.t. adversary $\mathcal{A}$ has an advantage negligible in κ for the following game: (1) $\mathcal{A}$ outputs a target identity $id^* \in \mathcal{I}$. (2) Run $\mathsf{Setup}(1^\kappa, c(\kappa))$ to obtain $(params, msk)$, and give *params* to $\mathcal{A}$. (3) $\mathcal{A}$ may run the Extract protocol with an oracle $O_{params,msk}(\cdot)$ polynomially many times, where on any input $id \neq id^*$ in $\mathcal{I}$, the oracle returns sk_{id}, and on any other input, the oracle returns an error message. (4) $\mathcal{A}$ outputs two messages $m_0, m_1 \in \mathcal{M}$ where $|m_0| = |m_1|$. Select a random bit b and give $\mathcal{A}$ the challenge ciphertext $c^* \leftarrow \mathsf{Encrypt}(params, id^*, m_b)$. (5) $\mathcal{A}$ may continue to query oracle $O_{params,msk}(\cdot)$ under the same conditions as before. (6) $\mathcal{A}$ outputs $b' \in \{0, 1\}$. We define $\mathcal{A}$'s advantage in the above game as $|\Pr[b' = b] - 1/2|$.

On stronger notions of ciphertext security for IBE. A stronger notion of ciphertext security for IBE schemes is adaptive-identity security (IND-ID-CPA) [BF01], which strengthens the IND-sID-CPA definition by allowing $\mathcal{A}$ to select the target identity id^* at the start of step (4) in the above game. In §3.1, we show blind IBE schemes satisfying both IND-sID-CPA and IND-ID-CPA security. Fortunately, our oblivious transfer applications in §4 require only IND-sID-CPA-security (because the "identities" will be fixed integers from 1 to poly(κ)), some additional applications in §5 require the stronger IND-ID-CPA-security.

Blind IBE. So far, we have only described traditional IBE schemes. A *blind IBE* scheme consists of the same players, together with the same algorithms Setup, Encrypt, Decrypt and yet we replace the protocol Extract with a new protocol BlindExtract which differs only in the authority's output:

- In the $\mathsf{BlindExtract}(\mathcal{P}(params, msk), \mathcal{U}(params, id)) \rightarrow (\text{nothing}, sk_{id})$ protocol, an honest user $\mathcal{U}$ with identity $id \in \mathcal{I}$ obtains the corresponding secret key sk_{id} from the master authority $\mathcal{P}$ or outputs an error message. The master authority's output is **nothing** or an error message.

We now define security for blind IBE, which informally is any IND-sID-CPA-secure IBE scheme with a BlindExtract protocol that satisfies two properties:

1. **Leak-free Extract:** a potentially malicious user cannot learn anything by executing the BlindExtract protocol with an honest authority which she could not have learned by executing the Extract protocol with an honest authority; moreover, as in Extract, the user must know the identity for which she is extracting a key.

[1] The standard definition of IBE [BF01] specifies an extraction *algorithm*. Note however that given such an algorithm, one can define a simple Extract protocol as: (1) $\mathcal{U}$ transmits id, (2) if $id \in \mathcal{I}$, $\mathcal{P}$ runs the extraction algorithm on $(params, msk, id)$ to obtain sk_{id} and returns this value (or an error), (3) user checks the validity of sk_{id} by encrypting a polynomially-bounded number of random messages and verifying their correct decryption.

2. Selective-failure Blindness: a potentially malicious authority cannot learn anything about the user's choice of identity during the BlindExtract protocol; moreover, the authority cannot cause the BlindExtract protocol to fail in a manner dependent on the user's choice.

Of course, a protocol realizing the functionality BlindExtract (in a fashion that satisfies the properties above) is a special case of secure two-party computation [Yao86, GMW87, Kil88]. However, using generic tools may be inefficient, so as in the case of blind signature protocols, we seek to optimize this specific computation. Indeed, recall that sk_{id} in an adaptive-identity secure IBE can be viewed as a signature by the authority on message id (see §5). Thus, our BlindExtract protocol (for an adaptive-identity secure IBE) *is* a blind signature scheme, but the converse implication is not necessarily true. Our leak-free extraction property is much stronger than the common *one-more unforgeability* requirement of blind signatures. Moreover, we will not require adaptive-identity security for the IBE in our OT applications. Let us now formally state these properties.

Definition 3.2 (Leak-Free Extract) A protocol $\mathsf{BlindExtract} = (\mathcal{P},\mathcal{U})$ associated with an IBE scheme $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ is *leak free* if for all efficient adversaries $\mathcal{A}$, there exists an efficient simulator $\mathcal{S}$ such that for every value κ and polynomial $c(\cdot)$, no efficient distinguisher D can distinguish the output of Game Real from Game Ideal with non-negligible advantage:

Game Real: Run $(params, msk) \leftarrow \mathsf{Setup}(1^\kappa, c(\kappa))$ and give $params$ to D. As many times as D wants, $\mathcal{A}$ chooses an identity id and atomically executes the BlindExtract protocol with $\mathcal{P}$: $\mathsf{BlindExtract}(\mathcal{P}(params, msk), \mathcal{A}(params, id))$. $\mathcal{A}$'s output (which is the output of the game) includes the list of identities and extracted keys.

Game Ideal: Run $(params, msk) \leftarrow \mathsf{Setup}(1^\kappa, c(\kappa))$. As many times as D wants, $\mathcal{S}$ chooses an identity id and queries a trusted party to obtain the output of $\mathsf{Extract}(params, msk, id)$, if $id \in \mathcal{I}$ and $\perp$ otherwise. $\mathcal{S}$'s output (which is the output of the game) includes the list of identities and extracted keys.

In the games above, BlindExtract and Extract are treated as atomic operations. Hence D and $\mathcal{A}$ (or $\mathcal{S}$) may communicate at any time except during the execution of those protocols. Additionally, while we do not explicitly specify that auxiliary information is given to the parties, this information must be provided in order to achieve the sequential composition property required by our OT protocols in §4.

This definition implies that the identity id (for the key being extracted) is *extractable* from the BlindExtract protocol— with all but negligible probability— since for every adversary there exists a $\mathcal{S}$ that must be able to interact with $\mathcal{A}$ to learn which identities to submit to the trusted party. We will make use of this observation later. Another nice property of this definition is that any key extraction protocol with leak-freeness (regardless of whether blindness holds or not) composes into the existing security definitions for IBE. (This would not necessarily be true of a blind signature protocol for the same type of signatures.) We state this formally below.

Lemma 3.3 *If* $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ *is an* IND-sID-CPA*-secure (resp.,* IND-ID-CPA*) IBE scheme and* BlindExtract *associated with* Π *is leak-free, then* $\Pi' = (\mathsf{Setup}, \mathsf{BlindExtract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ *is an* IND-sID-CPA*-secure (resp.,* IND-ID-CPA*) IBE scheme.*

Next, we define the second property of *blindness*. We use a strong notion of blindness called *selective-failure blindness* proposed recently by Camenisch et al. [CNS07], ensuring that even a malicious authority is unable to induce BlindExtract protocol failures that are dependent on the identity being extracted.

Definition 3.4 (Selective-Failure Blindness (SFB) [CNS07]) *A protocol $P(\mathcal{A}(\cdot), \mathcal{U}(\cdot,\cdot))$ is said to be selective-failure blind if every p.p.t. adversary $\mathcal{A}$ has a negligible advantage in the following game: First, $\mathcal{A}$ outputs params and a pair of identities $id_0, id_1 \in \mathcal{I}$. A random $b \in \{0,1\}$ is chosen. $\mathcal{A}$ is given black-box access to two oracles $\mathcal{U}(params, id_b)$ and $\mathcal{U}(params, id_{b-1})$. The $\mathcal{U}$ algorithms produce local output sk_b and sk_{b-1} respectively. If $sk_b \neq \perp$ and $sk_{b-1} \neq \perp$ then $\mathcal{A}$ receives (sk_0, sk_1). If $sk_b = \perp$ and $sk_{b-1} \neq \perp$ then $\mathcal{A}$ receives $(\perp, \varepsilon)$. If $sk_b \neq \perp$ and $sk_{b-1} = \perp$ then $\mathcal{A}$ receives $(\varepsilon, \perp)$. If $sk_b = \perp$ and $sk_{b-1} = \perp$ then $\mathcal{A}$ receives $(\perp, \perp)$. Finally, $\mathcal{A}$ outputs its guess b'. We define $\mathcal{A}$'s advantage in the above game as $|\Pr[b' = b] - 1/2|$.*

We thus arrive at the following definition.

Definition 3.5 (Secure Blind IBE) *A blind IBE Π = (Setup, BlindExtract, Encrypt, Decrypt) is called* IND-sID-CPA-secure *(resp.* IND-ID-CPA*) if and only if: (1) Π is* IND-sID-CPA*-secure (resp.* IND-ID-CPA*), and (2)* BlindExtract *is leak free and selective-failure blind.*

3.1 IBE Schemes with Efficient BlindExtract Protocols

In this section, we describe efficient BlindExtract protocols for: (1) the IND-sID-CPA-secure IBE due to Boneh and Boyen [BB04] and (2) the IND-ID-CPA-secure IBE proposed independently by Naccache [Nac05] and Chatterjee-Sarkar [CS05] which is a generalized version of Waters IBE [Wat05]. Note that in §3.3 we will be adding some additional features to these IBE schemes; these will help us to construct oblivious transfer protocols in §4. Since all of these schemes share a similar structure, we'll begin by describing their common elements.

Setup$(1^\kappa, c(k))$**:** Let $\gamma = (q, g, \mathbb{G}, \mathbb{G}_T, e)$ be the output of BMsetup(1^κ). Choose random elements $h, g_2 \in \mathbb{G}$ and a random value $\alpha \in \mathbb{Z}_q$. Set $g_1 = g^\alpha$. Finally, select a function $F : \mathcal{I} \to \mathbb{G}$ that maps identities to group elements. (The descriptions of F and $\mathcal{I}$ will be defined specific to the schemes below.) Output $params = (\gamma, g, g_1, g_2, h, F)$ and $msk = g_2^\alpha$.

Extract**:** Identity secret keys are of the form: $sk_{id} = (d_0, d_1) = (g_2^\alpha \cdot F(id)^r, g^r)$, where $r \in \mathbb{Z}_q$ is randomly chosen by the master authority. Note that the correctness of these keys can be publicly verified using a test described below.

Encrypt$(params, id, M)$**:** Given an identity $id \in \mathcal{I}$, and a message $M \in \mathbb{G}_T$, select a random $s \in \mathbb{Z}_q$ and output the ciphertext $C = (e(g_1, g_2)^s \cdot M, g^s, F(id)^s)$.

Decrypt$(params, id, sk_{id}, c_{id})$**:** On input a decryption key $sk_{id} = (d_0, d_1) \in \mathbb{G}^2$ and a ciphertext $C = (X, Y, Z) \in \mathbb{G}_T \times \mathbb{G}^2$, output $M = X \cdot e(Z, d_1)/e(Y, d_0)$.

Next, we'll describe the precise format of the secret keys sk_{id} and corresponding BlindExtract protocols for particular IBEs.

$\mathcal{P}(params, msk)$	$\mathcal{U}(params, id)$
	1. Choose $y \xleftarrow{\$} \mathbb{Z}_q$.
	2. Compute $h' \leftarrow g^y g_1^{id}$ and send h' to $\mathcal{P}$.
	3. Execute $PoK\{(y, id) : h' = g^y g_1^{id}\}$.
4. If the proof fails to verify, abort.	
5. Choose $r \xleftarrow{\$} \mathbb{Z}_q$.	
6. Compute $d_0' \leftarrow g_2^\alpha \cdot (h'h)^r$.	
7. Compute $d_1' \leftarrow g^r$.	
8. Send (d_0', d_1') to $\mathcal{U}$.	
	9. Check that $e(g_1, g_2) \cdot e(d_1', h'h) = e(d_0', g)$.
	10. If the check passes, choose $z \xleftarrow{\$} \mathbb{Z}_q$; otherwise, output $\perp$ and abort.
	11. Compute $d_0 \leftarrow (d_0'/(d_1')^y) \cdot F(id)^z$ and $d_1 \leftarrow d_1' \cdot g^z$.
	12. Output $sk_{id} = (d_0, d_1)$.

Figure 1: A BlindExtract protocol for the Boneh-Boyen IBE.

3.1.1 A BlindExtract Protocol for an IND-sID-CPA-Secure IBE

In the Boneh-Boyen IBE [BB04], $\mathcal{I} \subseteq \mathbb{Z}_q$ and the function $F : \mathcal{I} \to \mathbb{G}$ is defined as $F(id) = h \cdot g_1^{id}$. A secret key for identity id, where $r \in \mathbb{Z}_q$ is random, is:

$$sk_{id} = (d_0, d_1) = (g_2^\alpha \cdot F(id)^r, g^r) = (g_2^\alpha \cdot (h \cdot g_1^{id})^r, g^r).$$

The protocol BlindExtract$(\mathcal{P}(params, msk), \mathcal{U}(params, id))$ is described in Figure 1. Recall that $\mathcal{U}$ wants to obtain sk_{id} without revealing id, and $\mathcal{P}$ wants to reveal no more than sk_{id}. Let Π_1 be the blind IBE that combines algorithms Setup, Encrypt, Decrypt with the protocol BlindExtract in Figure 1.

Theorem 3.6 *Under the DBDH assumption, blind IBE Π_1 is secure (according to Definition 3.5); i.e.,* BlindExtract *is both leak-free and selective-failure blind.*

See Appendix A for proof of Theorem 3.6.

3.1.2 A BlindExtract Protocol for an IND-ID-CPA-Secure IBE

In the generalized version of Waters IBE [Wat05], proposed independently by Naccache [Nac05] and Chatterjee and Sarkar [CS05], the identity space $\mathcal{I}$ is the set of bit strings of length N, where N is polynomial in κ, represented by n blocks of ℓ bits each. The function $F : \{0,1\}^N \to \mathbb{G}$ is defined as $F(id) = h \cdot \prod_{j=1}^n u_j^{a_j}$, where each $u_j \in \mathbb{G}$ is randomly selected by the master authority and each a_j is an ℓ-bit segment of id. Naccache discusses practical IBE deployment with $N = 160$ and $\ell = 32$ [Nac05]. A secret key for identity id, where $r \in \mathbb{Z}_q$ is random, is:

$$sk_{id} = (d_0, d_1) = (g_2^\alpha \cdot F(id)^r, g^r) = (g_2^\alpha \cdot (h \cdot \prod_{j=1}^n u_j^{a_j})^r, g^r).$$

The protocol $\mathsf{BlindExtract}(\mathcal{P}(params, msk), \mathcal{U}(params, id))$ is described in Figure 1, with the following alterations. Parse the identity as $id = (a_1, \ldots, a_n)$, where each a_i is ℓ bits. In line 2, compute h' as $g^y \cdot \prod_{j=1}^{n} u_j^{a_j}$. In line 3, execute the proof $PoK\{(y, a_1, \ldots, a_n) : h' = g^y \cdot \prod_{j=1}^{n} u_j^{a_j} \wedge 0 \leq a_i < 2^\ell$, for $i = 1$ to $n\}$. The range part of this proof (*e.g.*, $0 \leq a_i < 2^\ell$) can be performed exactly or, by shortening each a_i by a few bits, can be done at almost no additional cost [CFT98, CM99, Bou00]. Follow the rest of the protocol as is. Let Π_2 be the blind IBE that combines Setup, $\mathsf{Encrypt}$, $\mathsf{Decrypt}$ with the $\mathsf{BlindExtract}$ protocol described above.

Theorem 3.7 *Under the DBDH assumption, blind IBE* Π_2 *is secure (according to Definition 3.5); i.e.,* $\mathsf{BlindExtract}$ *is both leak-free and selective-failure blind.*

See Appendix A for proof of Theorem 3.7.

3.2 On Other IBEs and HIBEs

Let us briefly summarize what we know about efficient $\mathsf{BlindExtract}$ protocols for other IBE schemes and hierarchical IBE (HIBE) schemes. First, random oracle based IBEs [BF01, Coc01] appear to be less suited to developing efficient $\mathsf{BlindExtract}$ protocols than their standard model successors. This is in part due to the fact that the identity string is hashed into an element in $\mathbb{G}$ in these schemes, instead of represented as an integer exponent, which makes our proof of knowledge techniques unwieldy. We were not able to find $\mathsf{BlindExtract}$ protocols for the Boneh and Franklin [BF01], Cocks [Coc01], or the recent Boneh-Gentry-Hamburg [BGH07] IBEs with running time better than $O(|\mathcal{I}|)$, where $\mathcal{I}$ is the identity space. Additionally, we did not consider the efficient IBE of Gentry [Gen06], as our focus was on schemes with *static* complexity assumptions.

We additionally considered hierarchical IBE schemes, such as those due to Boneh and Boyen [BB04], Waters [Wat05] and Chatterjee and Sarkar [CS06]. For all of these HIBEs, the number of elements comprising an identity secret key grow with the depth of the hierarchy, but each piece is similar in format to the original keys and our same techniques would apply.

3.3 Additional Properties for a Blind IBE

In §4, we use blind IBE as a tool for constructing oblivious transfer protocols. We can use either of the efficient blind IBEs Π_1 and Π_2 defined above together with the following observations about efficient protocols relating to them.

Efficient PoK for master secret. Our OT constructions in §4 require an efficient zero-knowledge proof of knowledge protocol for the statement $PoK\{(msk) : (params, msk) \in \mathsf{Setup}(1^\kappa, c(\kappa))\}$. If we were not concerned about efficiency, we could accomplish this proof using general techniques [Yao86, GMW87, Kil88]. However, for the parameters used in Π_1, Π_2, this proof can be conducted efficiently in a number of ways; one technique is to redefine $msk = \alpha \in \mathbb{Z}_q$ and prove the equivalent statement $PoK\{(\alpha) : g_1 = g^\alpha\}$ using a standard Schnorr proof [Sch91].

Committing IBE. To construct our OT protocols, we will require that our blind IBE schemes be *committing*. This property is related to committing encryption [CFGN96], but deals with the fact that IBE decryption keys may be extracted from malicious parties. Intuitively, we want to ensure that a given ciphertext C_{id} always decrypts to a single plaintext, even when the decryption keys are extracted from a malicious PKG.

Somewhat more formally, we require that an adversary playing the role of the PKG is unable to generate an identity/ciphertext pair (id, C) and— by conducting the extraction protocol with an honest party— any two keys sk_{id}, sk'_{id} such that $\mathsf{Decrypt}(params, id, sk_{id}, C) \neq \mathsf{Decrypt}(params, id, sk'_{id}, C)$. We observe that this property holds trivially for any IBE scheme where identity keys are "unique" (there is at most one decryption key per identity). However, in the schemes Π_1 and Π_2, there are many valid decryption keys for a given identity. This may lead to a condition where some incorrectly-formed ciphertexts will decrypt to different values depending on which secret key is used.

Fortunately, this issue can be addressed via a publicly-computable ciphertext correctness check, which we denote by $\mathsf{IsValid}(params, id, C)$. In the case of blind IBE schemes Π_1 and Π_2, we implement this check by first verifying the group parameters γ are valid (see [CCS07]), then verifying that for any *params* and $C = (X, Y, Z)$, all the values are in the correct groups and the following relation holds:

$$e(Y, F(id)) = e(Z, g)$$

The *correctness* property for the IsValid algorithm is that it outputs 1 for all honestly-generated parameters and ciphertexts. From the description of Π_1 and Π_2, it is easy to see that IsValid is correct. The algorithm's behavior in the case of maliciously-generated input is implicitly contained within the following definition:

Definition 3.8 (Committing IBE) *An IBE scheme (resp., blind IBE)* Π *is* committing *if and only if: (1) it is* IND-sID-CPA*-secure (resp., secure in the sense of definition 3.5) and (2) every p.p.t. adversary* $\mathcal{A}$ *has an advantage negligible in* κ *for the following game: First,* $\mathcal{A}$ *outputs* $params, id \in \mathcal{I}$ *and a ciphertext* C*. If* $\mathsf{IsValid}(params, id, C) \neq 1$ *then abort. Otherwise, the challenger, on input* $(params, id)$*, runs the* Extract *(resp.,* BlindExtract*) protocol with* $\mathcal{A}$ *twice to obtain purported keys* sk_{id}, sk'_{id}*.* $\mathcal{A}$*'s advantage is defined as:*

$$\left|\Pr\left[\mathsf{Decrypt}(params, id, sk_{id}, C) \neq \mathsf{Decrypt}(params, id, sk'_{id}, C)\right]\right|$$

Theorem 3.9 *Combined with the* IsValid *algorithm defined above, both* Π_1 *and* Π_2*, are committing blind IBE schemes (in the sense of definition 3.8).*

We present a proof of theorem 3.9 in Appendix D.

4 Simulatable Oblivious Transfer

We now turn our attention to constructing efficient and fully-simulatable oblivious transfer protocols. We'll use any of the efficient blind IBEs presented in the previous section as a building block. In particular, we focus on building (non-adaptive) OT_k^N and (adaptive) $\mathsf{OT}_{k\times 1}^N$ protocols, in which a Sender and Receiver transfer up to k messages out of an N-message set. In the non-adaptive model [BCR86, NP99a], the Receiver requests all k messages simultaneously. In the adaptive model [NP99b], the Receiver may request the messages one at a time, using the result of previous transfers to inform successive requests. Intuitively, the Receiver should learn only the messages it requests (and nothing about the remaining messages), while the Sender should gain no information about *which* messages the Receiver selected.

Full-simulation vs. half-simulation security. Security for oblivious transfer is defined using the real-world/ideal-world paradigm. In the real world, a Sender and Receiver interact directly according to the protocol, while in the ideal world, the parties interact via a trusted party. Informally, a protocol is secure if, for every real-world cheating Sender (resp., Receiver) we can describe an ideal-world counterpart who gains as much information from the ideal-world interaction as from the real protocol. Much of the oblivious transfer literature uses the simulation-based definition only to show *Sender* security, choosing to define Receiver security by a simpler game-based definition. Naor and Pinkas demonstrated that this weaker "half-simulation" approach permits *selective-failure* attacks, in which a malicious Sender induces transfer failures that are dependent on the message that the Receiver requests [NP99b]. Recently, Camenisch *et al.* [CNS07] proposed several practical $\mathsf{OT}^N_{k\times 1}$ protocols that are secure under a "full-simulation" definition, using adaptive (*e.g.*, q-PDDH) or interactive (*e.g.*, one-more-inversion RSA) assumptions. We now enhance their results by demonstrating efficient full-simulation OT^N_k and $\mathsf{OT}^N_{k\times 1}$ protocols secure under static complexity assumptions (*e.g.*, DBDH).

4.1 Definitions of OT^N_k, $\mathsf{OT}^N_{k\times 1}$

For consistency with earlier work, we quote the definitions of Camenisch *et al.* [CNS07] with the following modification: while that work focuses solely on *adaptive* OT, our definitions also consider the non-adaptive version of the primitive.

Definition 4.1 (k-out-of-N Oblivious Transfer (OT^N_k, $\mathsf{OT}^N_{k\times 1}$)) An oblivious transfer scheme is a tuple of algorithms $(\mathsf{S_I}, \mathsf{R_I}, \mathsf{S_T}, \mathsf{R_T})$. During the initialization phase, the Sender and the Receiver run an interactive protocol, where the Sender runs $\mathsf{S_I}(M_1, \ldots, M_N)$ to obtain state value S_0, and the Receiver runs $\mathsf{R_I}()$ to obtain state value R_0. Next, during the transfer phase, the Sender and Receiver interactively execute $\mathsf{S_T}, \mathsf{R_T}$, respectively, k times as described below.

Adaptive OT. In the adaptive $\mathsf{OT}^N_{k\times 1}$ case, for $1 \le i \le k$, the i^{th} transfer proceeds as follows: the Sender runs $\mathsf{S_T}(S_{i-1})$ to obtain state value S_i, and the Receiver runs $\mathsf{R_T}(R_{i-1}, \sigma_i)$ where $1 \le \sigma_i \le N$ is the index of the message to be received. The receiver obtains state information R_i and the message M'_{σ_i} or $\perp$ indicating failure.

Non-adaptive OT. In the non-adaptive OT^N_k case the parties execute the protocol as in the previous case; however, for each round $i < k$ the algorithm $\mathsf{R_T}(R_{i-1}, \sigma_i)$ **does not** output a message. At the end of the the k^{th} transfer $\mathsf{R_T}(R_{k-1}, \sigma_k)$ outputs the full collection $(M'_{\sigma_1}, \ldots, M'_{\sigma_k})$ where for $j = 1, \ldots, N$ each M'_{σ_j} is a valid message or the symbol $\perp$ indicating protocol failure. (In a non-adaptive scheme, the k transfers do not necessarily require a corresponding number of communication rounds.)

We now address the security definition for Oblivious Transfer.

Definition 4.2 (Full Simulation Security.) Full-simulation security for OT^N_k, $\mathsf{OT}^N_{k\times 1}$ is defined according to the following experiments. Note that, as in [CNS07] we do not explicitly specify auxiliary input to the parties, but note that this information can be provided in order to achieve sequential composition.

Real experiment. In experiment $\mathbf{Real}_{\hat{\mathsf{S}},\hat{\mathsf{R}}}(N, k, M_1, \ldots, M_N, \Sigma)$ the possibly cheating sender $\hat{\mathsf{S}}$ is given messages $(M_1, \ldots, M_N)$ as input and interacts with possibly cheating receiver $\hat{\mathsf{R}}(\Sigma)$, where

Σ is a selection algorithm that on input the full collection of messages thus far received, outputs the index σ_i of the next message to be queried. At the beginning of the experiment, both $\hat{\mathsf{S}}$ and $\hat{\mathsf{R}}$ output initial states (S_0, R_0). In the adaptive case, for $1 \leq i \leq k$ the sender computes $S_i \leftarrow \hat{\mathsf{S}}(S_{i-1})$, and the receiver computes $(R_i, M'_i) \leftarrow \hat{\mathsf{R}}(R_{i-1})$, where M'_i may or may not be equal to M_i. In the non-adaptive case, the Receiver obtains no messages until the k^{th} round, and therefore the selection strategy Σ must be non-adaptive. At the end of the k^{th} transfer the output of the experiment is (S_k, R_k).

We will define the *honest* Sender algorithm S as one that runs $\mathsf{S}_{\mathsf{I}}(M_1, \ldots, M_N)$ in the first phase, during each transfer runs $\mathsf{S}_{\mathsf{T}}()$ and outputs $S_k = \varepsilon$ as its final output. The honest Receiver R runs R_{I} in the first phase, and $\mathsf{R}_{\mathsf{T}}(R_{i-1}, \sigma_i)$ at the i^{th} transfer, and outputs $R_k = (M'_{\sigma_1}, \ldots, M'_{\sigma_k})$ as its final output.

Ideal experiment. In experiment $\mathbf{Ideal}_{\hat{\mathsf{S}}', \hat{\mathsf{R}}'}(N, k, M_1, \ldots, M_N, \Sigma)$ the possibly cheating sender algorithm $\hat{\mathsf{S}}'$ generates messages $(M_1^*, \ldots, M_N^*)$ and transmits them to a trusted party T. In the i^{th} round $\hat{\mathsf{S}}'$ sends a bit b_i to T; the possibly cheating receiver $\hat{\mathsf{R}}'(\Sigma)$ transmits σ_i^* to T. In the adaptive case, if $b_i = 1$ and $\sigma_i^* \in \{1, \ldots, N\}$ then T hands $M_{\sigma_i^*}^*$ to $\hat{\mathsf{R}}'$. If $b_i = 0$ then T hands $\perp$ to $\hat{\mathsf{R}}'$. Note that in the non-adaptive case, T caches its responses to $\hat{\mathsf{R}}'$ and delivers the full collection at the conclusion of the k^{th} round. After the k^{th} transfer the output of the experiment is (S_k, R_k).

Let $\ell(\cdot)$ be a polynomially-bounded function. We now define Sender and Receiver security in terms of the experiments above.

Sender Security. An $\mathsf{OT}_{k\times 1}^N$ provides Sender security if for every real-world p.p.t. receiver $\hat{\mathsf{R}}$ there exists a p.p.t. ideal-world receiver $\hat{\mathsf{R}}'$ such that $\forall N = \ell(\kappa)$, $k \in [1, N]$, $(M_1, \ldots, M_N)$, Σ, and every p.p.t. distinguisher:

$$\mathbf{Real}_{\mathsf{S}, \hat{\mathsf{R}}}(N, k, M_1, \ldots, M_N, \Sigma) \overset{c}{\approx} \mathbf{Ideal}_{\mathsf{S}', \hat{\mathsf{R}}'}(N, k, M_1, \ldots, M_N, \Sigma)$$

Receiver Security. $\mathsf{OT}_{k\times 1}^N$ provides Receiver security if for every real-world p.p.t. sender $\hat{\mathsf{S}}$ there exists a p.p.t. ideal-world sender $\hat{\mathsf{S}}'$ such that $\forall N = \ell(\kappa)$, $k \in [1, N]$, $(M_1, \ldots, M_N)$, Σ, and every p.p.t. distinguisher:

$$\mathbf{Real}_{\hat{\mathsf{S}}, \mathsf{R}}(N, k, M_1, \ldots, M_N, \Sigma) \overset{c}{\approx} \mathbf{Ideal}_{\hat{\mathsf{S}}', \mathsf{R}'}(N, k, M_1, \ldots, M_N, \Sigma)$$

4.2 Constructions

Next, we present several generic fully-simulatable oblivious transfer schemes built from any secure, committing blind IBE scheme.We approach this task in two phases: first, we propose a *non-adaptive* OT_k^N protocol and prove its security in the standard model. We next propose an adaptive $\mathsf{OT}_{k\times 1}^N$ secure the random oracle model. Finally, we sketch a path to removing the random oracles, by adapting techniques proposed in [CNS07].

4.2.1 Non-adaptive OT_k^N without Random Oracles

Given a committing blind IBE scheme Π, it is tempting to consider the following "intuitive" protocol: First, the Sender runs the IBE Setup algorithm and sends *params* to the Receiver. Next, for $i = 1, \ldots, N$ the Sender transmits an encryption of message M_i under identity "i". To obtain

k messages, the Receiver extracts decryption keys for identities $(\sigma_1, \ldots, \sigma_k)$ via k distinct executions of $\mathsf{BlindExtract}$, and uses these keys to decrypt the corresponding ciphertexts. If Π is a blind IBE secure in the sense of definition 3.5, then a cheating Receiver gains no information about the messages corresponding to secret keys he did not extract. Similarly, with additional precautions, a cheating Sender does not learn the identities extracted. However, it seems difficult to show this protocol is fully-simulatable, because the ideal Sender would have to form the N ciphertexts *before* learning the messages that k of them must decrypt to!

Fortunately, we are able to convert this simple idea into the fully-simulatable OT_k^N protocol shown in Figure 2. We require only the following modifications: first, we have the Sender prove knowledge of the value *msk* using appropriate zero-knowledge techniques.[2] Then, rather than transmitting the ciphertext vector during the first phase of the protocol, the Sender transmits only a *commitment* to a collision-resistant hash of the ciphertext vector, and sends the actual ciphertexts at the end of the k^{th} round together with a proof that she can open the commitment to the hash of the ciphertexts. (She does *not* open the commitment; she only proves that she knows how to do so.)

Theorem 4.3 (Full-simulation Security of the OT_k^N Scheme) *If* Π *is a* committing *blind IBE scheme secure in the sense of definition 3.8, and* $(\mathsf{CSetup}, \mathsf{Commit}, \mathsf{Decommit})$ *is a secure commitment scheme, then the* OT_k^N *protocol of figure 2 is sender-secure and receiver-secure in the full-simulation model. Furthermore, when* $\Pi \in (\Pi_1, \Pi_2)$ *the resulting* OT *scheme is secure under DBDH.*

We present a full proof of Theorem 4.3 in Appendix B.

4.2.2 Adaptive $\mathsf{OT}_{k\times 1}^N$ in the Random Oracle Model

While our first protocol is efficient and full-simulation secure, it permits only *non-adaptive* queries. For many practical applications (*e.g.*, oblivious retrieval from a large database), we desire a protocol that supports an adaptive query pattern. We approach this goal by first proposing an efficient $\mathsf{OT}_{k\times 1}^N$ protocol secure in the random oracle model. The protocol, which we present in Figure 3, requires an IBE scheme with a super-polynomial message space (as in the constructions of §3.1), and has approximately the same efficiency as the construction with random oracles of Camenisch *et al.* [CNS07]. However, their construction requires unique blind signatures and the two known options due to Chaum [Cha82] and Boldyreva [Bol03] both require interactive complexity assumptions. By using the blind IBE schemes in §3.1, our protocols can be based on the DBDH assumption.

Theorem 4.4 (Full-simulation Security of the $\mathsf{OT}_{k\times 1}^N$ Scheme) *If* Π *is a committing blind IBE scheme secure in the sense of 3.8, and* $H(\cdot)$ *is modeled as a random oracle, then the* $\mathsf{OT}_{k\times 1}^N$ *protocol of figure 3 is sender-secure and receiver-secure in the full-simulation model. Furthermore, when* $\Pi \in (\Pi_1, \Pi_2)$*, the* OT *protocol is secure under the DBDH assumption.*

We sketch a proof of theorem 4.4 in Appendix C. A nice feature of this proof is our ability to use the random oracle $H(\cdot)$ in place of the extractor for $\mathsf{BlindExtract}$.

[2] In §3.3, we describe how to conduct these proofs efficiently for the practical blind IBE constructions we consider.

$\underline{\mathsf{S}_\mathsf{I}(M_1,\ldots,M_N),\mathsf{S}_\mathsf{T}()}$ $\underline{\mathsf{R}_\mathsf{I}(),\mathsf{R}_\mathsf{T}(\sigma_1,\ldots,\sigma_k)}$

Sender and Receiver agree on parameters for a commitment scheme and a collision-resistant hash function H.[a]

1. Generate $(params, msk) \leftarrow \mathsf{Setup}(1^\kappa, c(\kappa))$.
2. For $j = 1,\ldots,N$, set $C_j \leftarrow \mathsf{Encrypt}(params, j, M_j)$.
3. Compute $(\mathcal{C},\mathcal{D}) \leftarrow \mathsf{Commit}\,(H(C_1,\ldots,C_N))$.
4. Send $(params,\mathcal{C})$ to Receiver.
5. Conduct $PoK\{(msk) : (params, msk) \in \mathsf{Setup}(1^\kappa, c(\kappa))\}$.

6. If the proof does not verify, abort.

For $i = 1, ..., k$, run $\mathsf{BlindExtract}$ on identity σ_i for Receiver to obtain sk_{σ_i}.

Following the k^{th} extraction:

1. Send the ciphertexts $(C_1,\ldots,C_N)$ to the Receiver.
2. Conduct $PoK\{(\mathcal{D}) : \mathsf{Decommit}\,(H(C_1,\ldots,C_N),\mathcal{C},\mathcal{D}) = 1\}$.

3. If the proof does not verify, or for any i $\mathsf{IsValid}(params, i, C_i) \neq 1$, abort and output $\perp$.
4. For $i = 1$ to k: If $\mathsf{BlindExtract}$ on σ_i failed, set $M'_{\sigma_i} \leftarrow \perp$; else, set M'_{σ_i} to the value $\mathsf{Decrypt}(params, \sigma_i, sk_{\sigma_i}, C_{\sigma_i})$.

Output $(msk,\mathcal{D})$ Output $(params,\mathcal{C},C_1,\ldots,C_N,M'_{\sigma_1},\ldots,M'_{\sigma_k})$.

[a] In the case of Pedersen's commitment scheme, the parameters may be generated by the Receiver. H may also be selected by the Receiver.

Figure 2: OT^N_k from any of the committing blind IBEs in §3, with input messages $M_1,\ldots,M_N \in \mathcal{M}$. We present the $\mathsf{S}_\mathsf{I},\mathsf{R}_\mathsf{I},\mathsf{S}_\mathsf{T},\mathsf{R}_\mathsf{T}$ algorithms in a single protocol flow.

4.2.3 Adaptive $\mathsf{OT}^N_{k\times 1}$ without Random Oracles

The random-oracle $\mathsf{OT}^N_{k\times 1}$ presented above is reasonably efficient both in terms of communication cost and round-efficiency. Ideally, we would like to construct a protocol of comparable efficiency in the standard model. We could construct an $\mathsf{OT}^N_{k\times 1}$ protocol by compiling k instances of the non-adaptive OT^N_k from §4.2.1. Each protocol round would consist of a 1-out-of-N instance of the protocol, with new IBE parameters and new a vector of ciphertexts $(C_1,\ldots,C_N)$. To ensure that each round is consistent with the previous rounds, the Sender would need to prove that the underlying plaintexts remain the same from round to round. This can be achieved using standard proof techniques, but is impractical for large values of k or N.

Alternatively, we could combine our scheme with the standard model $\mathsf{OT}^N_{k\times 1}$ of Camenisch *et al.* [CNS07]. Their efficient $\mathsf{OT}^N_{k\times 1}$, for example, incurs only a constant cost per transfer phase. However, the protocol relies on the dynamic q-Strong DH and q-Power Decisional DH assumptions,

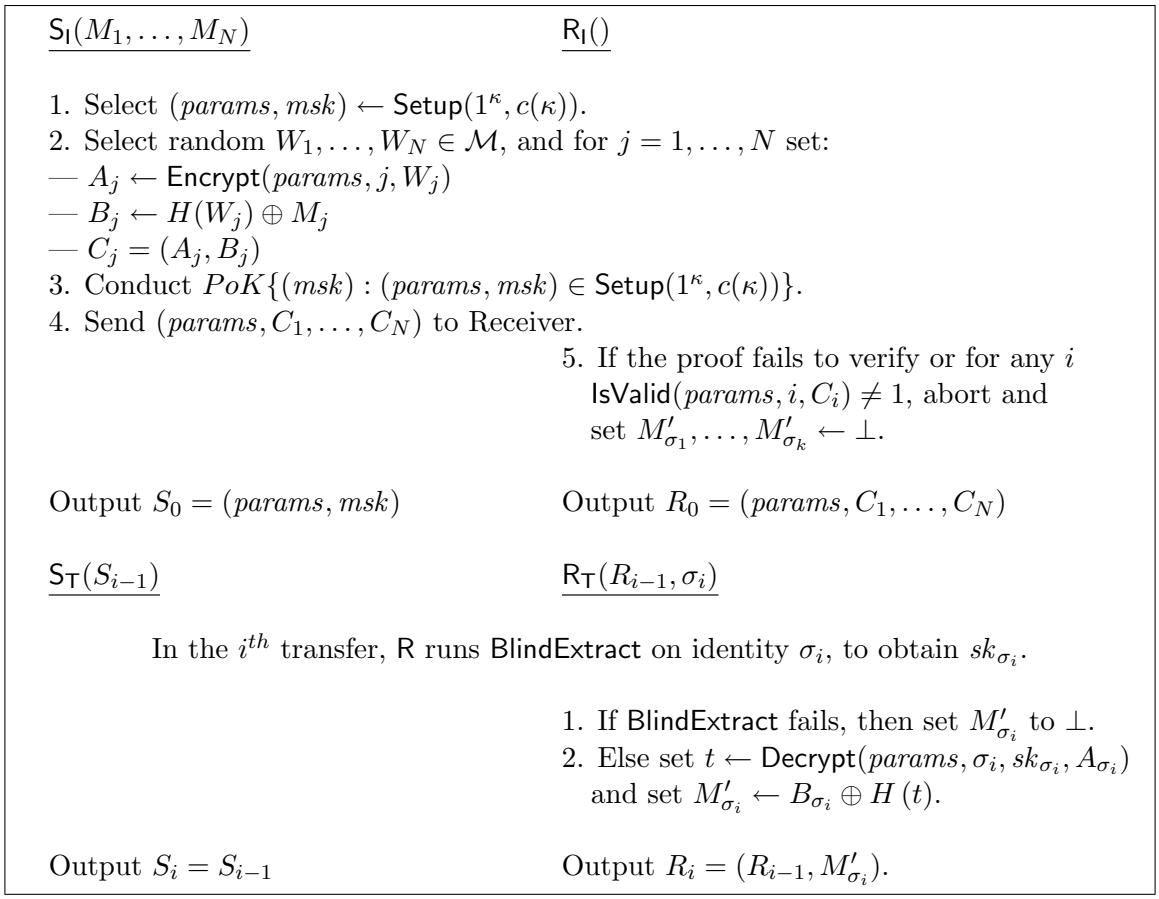

Figure 3: Adaptive $\mathsf{OT}^N_{k\times 1}$ from any of the committing blind IBEs in §3, with $M_1,\ldots,M_N \in \{0,1\}^n$. Let hash $H : \mathcal{M} \to \{0,1\}^n$ be modeled as a random oracle.

where large values of q require larger than normal security parameters [Che06]. Fortunately, one might be able to keep q small (on the order of k rather than N) by combining the Camenisch *et al.* scheme with ours as follows: in their initialization, the Sender releases N values corresponding to the messages that require $q = N$. Instead, we could use a blind IBE scheme to encrypt these N values during initialization, and then during the adaptive transfer phase, a Receiver could request the decryption key of his choice along with the information required in the Camenisch *et al.* scheme. Thus, reducing the values available to an adversary to $q = k$.

5 Other Applications of Blind IBE

Privacy-preserving delegated keyword search. Several works use IBE as a building-block for *public-key searchable encryption* [BCOP04, WBDS04]. These schemes permit a keyholder to delegate search capability to other parties. For example, Waters *et al.* [WBDS04] describe a searchable encrypted audit log in which a third party auditor is granted the ability to independently search the encrypted log for specific keywords. To enable this function, a central authority generates "trapdoors" for the keywords that the auditor wishes to search on. In this scenario, the trapdoor

generation authority necessarily learns each of the search terms. This may be problematic in circumstances where the pattern of trapdoor requests reveals sensitive information (*e.g.*, the name of a user under suspicion). By using blind and partially-blind IBE, we permit the authority to generate trapdoors, yet learn no information (or only partial information) about the search terms.[3]

Blind and partially-blind signature schemes. Moni Naor observed that each adaptive-identity secure IBE implies an existentially unforgeable signature scheme [BF01]. By the same token, an adaptive-identity secure blind IBE scheme implies an unforgeable, selective-failure blind signature scheme. This result applies to the adaptive-identity secure Π_2 protocol of §3.1.2, and to the selective-identity secure protocol Π_1 when that scheme is instantiated with appropriately-sized parameters and a hash function (see §7 of [BB04]). The efficient BlindExtract protocol for the adaptive-identity secure Π_2 scheme can also be used to construct a *partially-blind* signature, by allowing the signer (the master authority) to supply a portion of the input string. Partially-blind signatures have many applications, such as document timestamping and electronic cash [MS98].

Temporary anonymous identities. In a typical IBE, the master authority can link users to identities. For some applications, users may wish to remain anonymous or pseudonymous. By employing (partially-)blind IBE, an authority can grant temporary credentials without linking identities to users or even learning which identities are in use.

Acknowledgments

The authors are grateful to abhi shelat for helpful discussions, and to an anonymous ASIACRYPT 2007 reviewer for pointing out an issue regarding malformed ciphertexts.

References

[BB04] Dan Boneh and Xavier Boyen. Efficient selective-ID secure Identity-Based Encryption without random oracles. In *EUROCRYPT '04*, volume 3027 of LNCS, pages 223–238, 2004.

[BCOP04] Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky, and Giuseppe Persiano. Public key encryption with keyword search. In *EUROCRYPT '04*, volume 3027 of LNCS, pages 506–522, 2004.

[BCR86] Gilles Brassard, Claude Crépeau, and Jean-Marc Robert. All-or-nothing disclosure of secrets. In *CRYPTO '86*, volume 263 of LNCS, pages 234–238, 1986.

[BF01] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the Weil Pairing. In *CRYPTO '01*, volume 2139 of LNCS, pages 213–229, 2001.

[BGH07] Dan Boneh, Craig Gentry, and Michael Hamburg. Space-efficient identity based encryption without pairings, 2007. Available at `http://crypto.stanford.edu/~dabo/pubs.html`.

[BM89] Mihir Bellare and Silvio Micali. Non-interactive oblivious transfer and applications. In *CRYPTO '89*, volume 435 of LNCS, pages 547–557, 1989.

[Bol03] Alexandra Boldyreva. Threshold, Multisignature and Blind Signature Schemes Based on the Gap-Diffie-Hellman-Group Signature Scheme. In *PKC '03*, volume 2139 of LNCS, pages 31–46, 2003.

[Bou00] Fabrice Boudot. Efficient proofs that a committed number lies in an interval. In *EUROCRYPT '00*, volume 1807 of LNCS, pages 431–444, 2000.

[BW06] Xavier Boyen and Brent Waters. Anonymous hierarchical identity-based encryption (without random oracles). In *CRYPTO '06*, volume 4117 of LNCS, pages 290–307, 2006.

[3] Boneh *et al.* [BCOP04] note that keyword search schemes can be constructed from any *key anonymous* IBE scheme. While the schemes of §3 are not key anonymous, Boyen and Waters remark that key anonymity in similar schemes might be acheived by implementing them in *asymmetric* bilinear groups [BW06].

[Can01] Ran Canetti. Universally Composable Security: A new paradigm for cryptographic protocols. In *FOCS '01*, page 136. IEEE Computer Society, 2001. `http://eprint.iacr.org/2000/067`.

[CCS07] L. Chen, Z. Cheng, and Nigel Smart. Identity-based key agreement protocols from pairings. *International Journal of Information Security*, 6:213–241, August 2007.

[CDS94] Ronald Cramer, Ivan Damgård, and Berry Schoenmakers. Proofs of partial knowledge and simplified design of witness hiding protocols. In *CRYPTO '94*, volume 839 of LNCS, pages 174–187, 1994.

[CFGN96] Ran Canetti, Uri Feige, Oded Goldreich, and Moni Naor. Adaptively secure multi-party computation. In *Proc. of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing*, pages 639–648, 1996.

[CFT98] Agnes Chan, Yair Frankel, and Yiannis Tsiounis. Easy come – easy go divisible cash. In *EUROCRYPT '98*, volume 1403 of LNCS, pages 561–575, 1998.

[Cha82] David Chaum. Blind signatures for untraceable payments. In *CRYPTO '82*, pages 199–203. Plenum Press, 1982.

[Che06] Jung Hee Cheon. Security analysis of the strong diffie-hellman problem. In *EUROCRYPT '06*, volume 4004 of LNCS, pages 1–11, 2006.

[CHK04] Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from Identity Based Encryption. In *EUROCRYPT '04*, volume 3027 of LNCS, pages 207–222, 2004.

[CM99] Jan Camenisch and Markus Michels. Proving in zero-knowledge that a number n is the product of two safe primes. In *EUROCRYPT '99*, volume 1592 of LNCS, pages 107–122, 1999.

[CNS07] Jan Camenisch, Gregory Neven, and Abhi Shelat. Simulatable adaptive oblivious transfer. In *EUROCRYPT '07*, volume 4515 of LNCS, pages 573–590, 2007.

[Coc01] Clifford Cocks. An identity based encryption scheme based on Quadratic Residues. In *Cryptography and Coding, IMA International Conference*, volume 2260 of LNCS, pages 360–363, 2001.

[CS97] Jan Camenisch and M. Stadler. Efficient group signature schemes for large groups. In *CRYPTO '97*, volume 1296 of LNCS, pages 410–424, 1997.

[CS05] Sanjit Chatterjee and Palash Sarkar. Trading time for space: Towards an efficient IBE scheme with short(er) public parameters in the standard model. In *ICISC 2005*, volume 3935 of LNCS, pages 424–440, 2005.

[CS06] Sanjit Chatterjee and Palash Sarkar. HIBE with Short Public Parameters without Random Oracle. In *ASIACRYPT '06*, volume 4284 of LNCS, pages 145–160, 2006.

[CT05] Cheng-Kang Chu and Wen-Guey Tzeng. Efficient k-out-of-n oblivious transfer schemes with adaptive and non-adaptive queries. In *PKC '05*, volume 3386 of LNCS, pages 172–183. Springer, 2005.

[DHRS04] Yan Zong Ding, Danny Harnik, Alon Rosen, and Ronen Shaltiel. Constant-round oblivious transfer in the bounded storage model. In *TCC '04*, volume 2951 of LNCS, pages 446–472, 2004.

[EGL82] Shimon Even, Oded Goldreich, and Abraham Lempel. A randomized protocol for signing contracts. In *CRYPTO '82*, pages 205–210, 1982.

[FS86] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In *CRYPTO '86*, volume 263 of LNCS, pages 186–194, 1986.

[Gen06] Craig Gentry. Practical identity-based encryption without random oracles. In *EUROCRYPT '06*, volume 4004 of LNCS, pages 445–464, 2006.

[GH08] Matthew Green and Susan Hohenberger. Universally composable adaptive oblivious transfer. Cryptology ePrint Archive, Report 2008/163, 2008. `http://eprint.iacr.org/`.

[GMW87] Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game or a completeness theorem for protocols with honest majority. In *STOC '87*, pages 218–229, 1987.

[Goy07] Vipul Goyal. Reducing trust in the PKG in identity based cryptosystems. In *CRYPTO '07*, volume 4622 of LNCS, pages 430–447, 2007.

[Kal05] Yael Tauman Kalai. Smooth projective hashing and two-message oblivious transfer. In *EUROCRYPT '05*, volume 3494 of LNCS, pages 78–95, 2005.

[Kil88] Joe Kilian. Founding cryptography on oblivious transfer. In *STOC '88*, pages 20–31, 1988.

[Lin08] Andrew Lindell. Efficient fully-simulatable oblivious transfer. In *CT-RSA '08*, volume 4964 of LNCS, pages 52–70. Springer, 2008.

[MS98] Shingo Miyazaki and Kouichi Sakurai. A more efficient untraceable e-cash system with partially blind signatures based on the discrete logarithm problem. In *Financial Cryptography '98*, volume 1465 of LNCS, pages 296–308, 1998.

[Nac05] David Naccache. Secure and *practical* identity-based encryption. Cryptology ePrint Archive, Report 2005/369, 2005. `http://eprint.iacr.org/`.

[NP99a] Moni Naor and Benny Pinkas. Oblivious transfer and polynomial evaluation. In *STOC '99*, pages 245—254, 1999.

[NP99b] Moni Naor and Benny Pinkas. Oblivious transfer with adaptive queries. In *CRYPTO '99*, volume 1666 of LNCS, pages 573–590, 1999.

[NP01] Moni Naor and Benny Pinkas. Efficient oblivious transfer protocols. In *SODA '01*, pages 448–457, 2001.

[OK04] Wakaha Ogata and Kaoru Kurosawa. Oblivious keyword search. *Special issue on coding and cryptography Special issue on coding and cryptography Journal of Complexity*, 20(2-3):356–371, 2004.

[Oka06] Tatsuaki Okamoto. Efficient blind and partially blind signatures without random oracles. In *Theory of Cryptography (TCC)*, volume 3876 of LNCS, pages 80–99, 2006.

[Ped92] Torben Pryds Pedersen. Non-interactive and information-theoretic secure verifiable secret sharing. In *CRYPTO '92*, volume 576 of LNCS, pages 129–140, 1992.

[PVW07] Chris Peikert, Vinod Vaikuntanathan, and Brent Waters. A framework for efficient and composable oblivious transfer. Cryptology ePrint Archive, Report 2007/348, 2007. `http://eprint.iacr.org/2007/348.pdf`.

[Rab81] Michael Rabin. How to exchange secrets by oblivious transfer. Technical Report TR-81, Aiken Computation Laboratory, Harvard University, 1981.

[Sch91] Claus-Peter Schnorr. Efficient signature generation for smart cards. *Journal of Cryptology*, 4(3):239–252, 1991.

[Sha84] Adi Shamir. Identity-based cryptosystems and signature schemes. In *CRYPTO '84*, volume 196 of LNCS, pages 47–53, 1984.

[Wat05] Brent Waters. Efficient Identity-Based Encryption without random oracles. In *EUROCRYPT '05*, volume 3494 of LNCS, pages 114–127, 2005.

[WBDS04] Brent R. Waters, Dirk Balfanz, Glenn Durfee, and D. K. Smetters. Building an encrypted and searchable audit log. In *NDSS'04*, 2004.

[Yao86] Andrew Yao. How to generate and exchange secrets. In *FOCS '86*, pages 162–167, 1986.

A Security Proofs for Blind IBE Schemes

A.1 Proof of Theorem 3.6

Proof sketch. We begin by observing that the $\mathsf{Setup}, \mathsf{Encrypt}, \mathsf{Decrypt}$ algorithms of Π_1 are identical to the original Boneh-Boyen (H)IBE [BB04] instantiated with only one level. Thus, when Π_1 is considered with the key extraction algorithm of [BB04], it is $\mathsf{IND\text{-}sID\text{-}CPA}$-secure by the original proof of security. To prove the remaining properties, we must show that the $\mathsf{BlindExtract}$ protocol in Figure 1 is both leak free and selective-failure blind. We begin with leak freeness, which requires the existence of an efficient simulator $\mathcal{S}$ such that no efficient distinguisher D can distinguish Game Real (where $\mathcal{A}$ is interacting with an honest $\mathcal{P}$ running the $\mathsf{BlindExtract}$ protocol) from Game Ideal (where the ideal adversary $\mathcal{S}$ is given access to a trusted party executing the ideal $\mathsf{Extract}$ protocol).

We describe the ideal adversary $\mathcal{S}$ as follows:

1. On input *params* from the trusted party, $\mathcal{S}$ hands *params* to a copy of $\mathcal{A}$ it runs internally.

2. Each time $\mathcal{A}$ engages $\mathcal{S}$ in a $\mathsf{BlindExtract}$ protocol, $\mathcal{S}$ behaves as follows. In the first message of the protocol, $\mathcal{A}$ must send to $\mathcal{S}$ a value h' and prove knowledge of values (y, id) such that $h' = g^y \cdot g_1^{id}$. If the proof fails to verify, $\mathcal{S}$ aborts. Since this proof of knowledge is implemented using the *extractable* techniques mentioned in §2, $\mathcal{S}$ can efficiently extract the values (y, id).
3. Next, $\mathcal{S}$ submits id to the trusted party, who returns the valid secret key for this identity $sk_{id} = (d_0, d_1) = (g_2^{\alpha} \cdot F(id)^r, g^r)$ for some random $r \in \mathbb{Z}_q$.
4. Finally, $\mathcal{S}$ computes the pair $(d_0', d_1') = (d_0 \cdot d_1^y, d_1)$ and returns these values to $\mathcal{A}$.

Observe that the responses of $\mathcal{S}$ are always correctly formed (as $\mathcal{A}$ can verify) and drawn from the same distribution as those of $\mathcal{P}$. Thus, Game Real and Game Ideal are indistinguishable to both $\mathcal{A}$ and D. We also note (as above) that the identity id being requested by $\mathcal{A}$ is efficiently *extractable* (by an extractor with special rewind capabilities not available to $\mathcal{P}$).

Next, we turn out attention to selective-failure blindness for protocol $\mathsf{BlindExtract} = (\mathcal{P}, \mathcal{U})$. Here $\mathcal{A}$ outputs *params* and two identities $id_0, id_1 \in \mathcal{I}$. Then a random bit b is chosen. Next, $\mathcal{A}$ is given black-box access to two oracles $\mathcal{U}(params, id_b)$ and $\mathcal{U}(params, id_{b-1})$. The $\mathcal{U}$ algorithms conduct the $\mathsf{BlindExtract}$ protocol (with $\mathcal{A}$ playing the role of $\mathcal{P}$), and produce local output sk_b and sk_{b-1} respectively. If $sk_b \neq \perp$ and $sk_{b-1} \neq \perp$ then $\mathcal{A}$ receives (sk_0, sk_1). If $sk_b = \perp$ and $sk_{b-1} \neq \perp$ then $\mathcal{A}$ receives $(\perp, \varepsilon)$. If $sk_b \neq \perp$ and $sk_{b-1} = \perp$ then $\mathcal{A}$ receives $(\varepsilon, \perp)$. If $sk_b = \perp$ and $sk_{b-1} = \perp$ then $\mathcal{A}$ receives $(\perp, \perp)$. $\mathcal{A}$ then tries to predict b, which we want to argue he cannot do with non-negligible advantage over guessing.

First, we observe that in this protocol, $\mathcal{U}$ speaks first and sends to $\mathcal{A}$ a value h' uniformly distributed in $\mathbb{G}$ and then performs a zero-knowledge proof of knowledge $PoK\{(y, id) : h' = g^y \cdot g_1^{id}\}$. Suppose that $\mathcal{A}$ runs one or both of his oracles up to this point. Now, it is $\mathcal{A}$'s turn to speak, and at this point, his views so far are computationally indistinguishable. Let's assume that $\mathcal{A}$ must now return two values $(d_0', d_1') \in \mathbb{G}^2$ to the first oracle. Suppose $\mathcal{A}$ chooses this pair using any strategy he wishes. At the point $\mathcal{A}$ fixes on two values, he is able to *predict* the output sk_i of these oracles $\mathcal{U}(params, id_b)$ with non-negligible advantage as follows:

1. $\mathcal{A}$ checks if $e(g_1, g_2) \cdot e(d_1', h' \cdot h) = e(d_0', g)$ holds. If the test fails, record $sk_0 \leftarrow \perp$.
2. Next, $\mathcal{A}$ choses any two values $(d_0', d_1') \in \mathbb{G}^2$ for the second oracle, performs the same check and, in the event of failure, records $sk_1 \leftarrow \perp$.
3. If either test failed, then: if $sk_0 = \perp$ and $sk_1 \neq \perp$, output $(\perp, \varepsilon)$. If $sk_0 \neq \perp$ and $sk_1 = \perp$, output $(\varepsilon, \perp)$. If both tests failed, output $(\perp, \perp)$.
4. If both test succeeded, then: $\mathcal{A}$ initiates $\mathsf{BlindExtract}$ with itself on (id_0, id_1) (playing the roles of $\mathcal{U}$ and $\mathcal{P}$). If either protocol run fails, abort.[4] Otherwise output the returned keys (sk_0, sk_1).

This prediction is correct, because $\mathcal{A}$ is performing the same check as the honest $\mathcal{U}$, and when both tests succeed, outputting a pair of valid secret keys obtained via $\mathsf{BlindExtract}(params, id)$, as does $\mathcal{U}$. But at a higher-level, note that if $\mathcal{A}$ is able to predict the final output of its oracles accurately, then $\mathcal{A}$'s advantage in distinguishing $\mathcal{U}(params, id_b)$ and $\mathcal{U}(params, id_{b-1})$ is the same without this final output. Thus, all of $\mathcal{A}$'s advantage must come from distinguishing the earlier messages of the oracles. Since these oracles only send one uniformly random value $h' \in \mathbb{G}$ and then perform a zero-knowledge proof of knowledge about the representation of h' with respect to public

[4]Note that $\mathcal{A}$ only reaches this step if $\mathcal{U}$'s two previous executions of the protocol have succeeded. If that event occurs with non-negligible probability, then $\mathcal{A}$ successfully obtains (sk_0, sk_1) with non-negligible probability.

values, we know from the security of the underlying proof that $\mathcal{A}$ cannot distinguish between them with non-negligible probability.

□

A.2 Proof of Theorem 3.7

Proof sketch. This proof follows the outline of the proof of Theorem 3.6 almost identically. Again we observe that IND-ID-CPA security can be shown via the original proof by Naccache [Nac05]. To satisfy leak freeness, the simulator $\mathcal{S}$ operates exactly as before: starting up an internal copy of $\mathcal{A}$ in step (1), extracting the values (y, id) from $\mathcal{A}$ in step (2), querying the trusted party for $sk_{id} = (d_0, d_1) \leftarrow \mathsf{Extract}(msk, id)$ in step (3), and returning the pair $(d'_0, d'_1) = (d_0 \cdot d_1^y, d_1)$ to $\mathcal{A}$ in step (4). Although the internal structure of the secret keys in the Naccache-Waters IBE differ from those of the Boneh-Boyen IBE, the key observation here is that $\mathcal{S}$ doesn't need to know anything about this structure to compute the correct response in step (4).

To satisfy selective-failure blindness, we first observe that the prediction of $\mathcal{U}$'s final output is done exactly as before. Thus, $\mathcal{A}$ must be able to distinguish the oracles after seeing only a value h' again uniformly distributed in $\mathbb{G}$ and a zero-knowledge proof of knowledge about the representation of h' with respect to public values. We conclude that this advantage must be negligible. □

B Proofs of Security for the OT_k^N Protocol

We now prove Theorem 4.3. The proof is divided into two parts, one to show that the OT scheme meets the *sender security* property, and a second to show *receiver security.*

Proof of Sender Security (Theorem 4.3). For any real-world cheating receiver $\hat{\mathsf{R}}$ we can construct an ideal-world receiver $\hat{\mathsf{R}}'$ such that the "real" and "ideal" experiments are computationally indistinguishable. More formally, define some set of negligible functions where $\nu_n(\cdot)$ indicates the n^{th} function. Then $\forall$ p.p.t. D:

$$\mathbf{Pr}\left[D(\mathbf{Real}_{\mathsf{S},\hat{\mathsf{R}}}(N,k,M_1,\ldots,M_N,\Sigma)) = 1 \right] - \mathbf{Pr}\left[D(\mathbf{Ideal}_{\mathsf{S}',\hat{\mathsf{R}}'}(N,k,M_1,\ldots,M_N,\Sigma)) = 1 \right] \leq \nu_1(\kappa)$$

To describe the construction of $\hat{\mathsf{R}}'$ we will begin with the real-world experiment, and modify elements via a series of games until we arrive at the ideal-world experiment. For notational convenience, let $\mathbf{Adv}\,[\,\mathbf{Game\ i}\,]$ be D's advantage in distinguishing the output of **Game i** from the **Real** distribution.

> **Game 0**. In this game the honest real-world sender $\mathsf{S}(M_1,\ldots,M_N)$ interacts with the real-world cheating receiver $\hat{\mathsf{R}}$. Clearly $\mathbf{Adv}\,[\,\mathbf{Game\ 0}\,] = 0$.
>
> **Game 1**. In this game, we employ the knowledge extractor for BlindExtract to extract from $\hat{\mathsf{R}}$ each of the identities $(\sigma_1,\ldots,\sigma_k)$ from the k sequential executions of the BlindExtract protocol.[5] If the knowledge extractor fails for any execution, set $\hat{\mathsf{R}}'$'s output to $\perp$. Let $\mathbf{Pr}\,[\,error\,]$ be the probability that the knowledge extractor fails during any given execution,

[5]Note that the leak-freeness definition implies that for every adversary adversary $\mathcal{A}$, there exists a simulator that queries the trusted party and produces indistinguishable output (including the extracted identities). We can use this simulator as a black box to construct our extractor, which must fail with at most negligible probability.

then $\mathbf{Adv}[\mathbf{Game\ 1}] - \mathbf{Adv}[\mathbf{Game\ 0}] \leq (k \cdot \mathbf{Pr}[error])$. Since Π is *leak-free*, it must hold that $k \cdot \mathbf{Pr}[error] \leq \nu_2(\kappa)$, and thus, $\mathbf{Adv}[\mathbf{Game\ 1}] \leq \nu_2(\kappa)$.

Game 2. In this game, the commitment $\mathcal{C}$ is replaced with a commitment to a random value, and the final proof-of-knowledge for decommitment is replaced with a simulated proof. The difference between this game and **Game 2** is equal to D's advantage in correctly distinguishing $\mathcal{C}$ from a valid commitment on $H(C_1, \ldots, C_N)$, and the proof simulation from a valid proof. We define this probability as $\mathbf{Adv}[dec]$, and note that $\mathbf{Adv}[dec] \leq \nu_3(\kappa)$ for a secure commitment scheme and zero-knowledge proof.

Game 3. In the final game, we alter the ciphertext vector $(C_1, \ldots, C_N)$ to produce a new vector $(C'_1, \ldots, C'_N)$ as follows: for $j = 1, \ldots, N$ if $j \notin (\sigma_1, \ldots, \sigma_k)$, set $C'_j \leftarrow \mathsf{Encrypt}(params, j, M' \xleftarrow{\$} \mathcal{M})$, and otherwise set $C'_j \leftarrow C_j$. By Lemma B.1 below, the security properties of Π imply that $\mathbf{Adv}[\mathbf{Game\ 3}] - \mathbf{Adv}[\mathbf{Game\ 2}] \leq \nu_4(\kappa)$.

Summing the differences between the above games, it is clear that $\mathbf{Adv}[\mathbf{Game\ 3}]$ is negligible, and therefore no p.p.t. algorithm can distinguish the distribution of **Game 3** from **Game 0**. The ideal-world receiver $\hat{\mathsf{R}}'$ is an algorithm that runs $\hat{\mathsf{R}}$, and (1) issues it a random commitment, (2) extracts the values $(\sigma_1, \ldots, \sigma_k)$ from $\hat{\mathsf{R}}$'s executions of the $\mathsf{BlindExtract}$ protocol, (3) transmits these values to the trusted party T to receive $(M_{\sigma_1}, \ldots, M_{\sigma_k})$. Next, (4) for $i = 1, \ldots, k$, $\hat{\mathsf{R}}'$ sets $C'_{\sigma_i} = \mathsf{Encrypt}(params, \sigma_i, M_{\sigma_i})$ and each of the remaining ciphertexts to encryptions of a random message, and (5) sends $(C'_1, \ldots, C'_N)$ to $\hat{\mathsf{R}}$ along with a simulated proof of knowledge of the opening of the commitment.

Lemma B.1 (Indistinguishability of Ciphertexts) $\mathbf{Adv}[\mathbf{Game\ 3}] - \mathbf{Adv}[\mathbf{Game\ 2}] \leq \nu_4(\kappa)$ if Π is a blind IBE scheme secure in the sense of definition 3.5 (or definition 3.8).

Proof sketch. We show, via a series of hybrids, that no p.p.t. D distinguishes **Game 2** from **Game 3** except with negligible probability, as long as (1) the PoK of *msk* is zero-knowledge, and (2) Π is both leak-free and $\mathsf{IND\text{-}sID\text{-}CPA}$-secure.

Zero-Knowledge and Leak-freeness. Consider a pair of hybrid games. Hybrid 0 is identical to **Game 2**, except that S *simulates* the PoK of *msk*. Clearly the zero-knowledge property of Π ensures that this hybrid is indistinguishable from **Game 2**. Hybrid 1 extends the previous hybrid as follows: S does not run Setup, but is instead given *params* and an oracle $O_{params,msk}(\cdot)$ and answers $\hat{\mathsf{R}}$'s $\mathsf{BlindExtract}$ queries by running $\mathsf{Extract}$ with $O_{params,msk}$ and simulating the response. This hybrid is clearly indistinguishable from the previous by the leak-freeness property of Π.

$\mathsf{IND\text{-}sID\text{-}CPA}$ **security.** Now assume by contradiction that some D distinguishes hybrid 1 from **Game 3**. If this is the case, then we show how to construct an adversary $\mathcal{A}$ that wins the $\mathsf{IND\text{-}sID\text{-}CPA}$ game against Π with non-negligible advantage. This proof is just a standard hybrid argument, but we provide it for completeness. Beginning with hybrid 1 from above, we describe an additional $(N - k)$ hybrids, where the final hybrid is **Game 3**. Each hybrid $j \in [2, (N - k)]$ is identical to hybrid $(j - 1)$ except that the distribution of the ciphertext vector is different at position ℓ (C_ℓ is replaced with the encryption of a random message). If D distinguishes the first and last hybrids with non-negligible probability, then clearly there must exist a D' that distinguishes some pair of consecutive hybrids $(j, j - 1)$ with non-negligible probability.

Consider these two hybrids, and let ℓ be the position at which the ciphertext vectors differ. The IND-sID-CPA adversary $\mathcal{A}$ outputs $id^* = \ell$ and receives *params*. It then runs D' (which controls $\hat{\mathsf{R}}$) and conducts the initial stage of the OT protocol as in hybrid 1. Select $M^* \xleftarrow{\$} \mathcal{M}$ and output (M_ℓ, M^*) to obtain the challenge ciphertext C^*. Construct a ciphertext vector $\vec{C}$ with the correct distribution for hybrid $(j-1)$ (by encrypting either a real message or a random message at each position as appropriate)— however, at the ℓ^{th} position, set $C_\ell \leftarrow C^*$. Send $\vec{C}$ to $\hat{\mathsf{R}}$ and complete the protocol. Let b' be D''s output. Output b'.

Note that when C^* encrypts M_ℓ, D's view is that of hybrid $j-1$, and when C^* encrypts M^*, D's view is that of hybrid j. Thus, if D outputs 1 with probability α in the first, case, and probability β in the second, then $\mathcal{A}$ guesses correctly with probability $\frac{|\beta-\alpha|}{2}$. Since $|\beta-\alpha|$ is non-negligible in κ, then $\mathcal{A}$ wins the IND-sID-CPA game with non-negligible advantage.

□

□

Proof of Receiver Security (Theorem 4.3). For any real-world cheating sender $\hat{\mathsf{S}}$ we can construct an ideal-world sender $\hat{\mathsf{S}}'$ such that no p.p.t. algorithm D can distinguish the distributions $\mathbf{Real}_{\hat{\mathsf{S}},\mathsf{R}}(N, k, M_1, \ldots, M_N, \Sigma)$ and $\mathbf{Ideal}_{\hat{\mathsf{S}}',\mathsf{R}'}(N, k, M_1, \ldots, M_N, \Sigma)$. We arrive at the ideal-world sender via a series of games. Again let $\mathbf{Adv}\,[\,\mathbf{Game\ i}\,]$ be D's advantage in distinguishing the output of **Game i** from the **Real** distribution.

> **Game 0**. In this game the honest real-world receiver R interacts with the real-world cheating sender $\hat{\mathsf{S}}$. Clearly $\mathbf{Adv}\,[\,\mathbf{Game\ 0}\,] = 0$.
>
> **Game 1**. In this game, use the knowledge extractor for $PoK\{msk : (params, msk) \in \mathsf{Setup}(1^\kappa, c(\kappa))\}$ to extract *msk*. If the extractor fails or outputs an invalid *msk*, set R's output to $\perp$. Since this extractor fails with probability negligible in κ, then $\mathbf{Adv}\,[\,\mathbf{Game\ 1}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 0}\,] \leq \nu_1(\kappa)$.
>
> **Game 2**. In this game, replace the k executions of BlindExtract with executions on random identities $(\sigma'_1, \ldots, \sigma'_k)$. If the i^{th} execution fails, record $b_i \leftarrow 0$, otherwise set $b_i \leftarrow 1$. By Lemma B.2, $\mathbf{Adv}\,[\,\mathbf{Game\ 2}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 1}\,] \leq (k \cdot \nu_2(\kappa))$ if Π is selective-failure blind.
>
> **Game 3**. Verify that for all $j \in (\sigma'_1, \ldots, \sigma'_k)$, the condition $\mathsf{Decrypt}(sk_{\sigma_j}, C_{\sigma_j}) = \mathsf{Decrypt}(\mathsf{Extract}(msk, \sigma_j), C_{\sigma_j})$ holds. If this does not hold, then set R's output to $\perp$. By Lemma B.3, $\mathbf{Adv}\,[\,\mathbf{Game\ 2}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 1}\,] \leq \nu_3(\kappa)$ if Π is a *committing* blind IBE.

Summing the differences beteen the above games, it is clear that $\mathbf{Adv}\,[\,\mathbf{Game\ 3}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 0}\,]$ is negligible, and therefore no p.p.t. algorithm can distinguish the distribution of **Game 3** from **Game 0**. The ideal-world sender $\hat{\mathsf{S}}'$ is an algorithm that performs all of the changes between the games above, and on learning $(M_1, \ldots, M_N, b_1, \ldots, b_k)$ transmits these values to the trusted party T.

Lemma B.2 (Blindness of Extractions) $\mathbf{Adv}\,[\,\mathbf{Game\ 2}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 1}\,] \leq (k \cdot \nu_2(\kappa))$ if Π is selective-failure blind in the sense of definition 3.5.

Proof sketch. By contradiction, let D be a p.p.t. distinguisher that controls $\hat{\mathsf{S}}$ and distinguishes the distributions of **Game 2** and **Game 1** with advantage $> k \cdot \nu_2(\kappa)$. This implies that D

can distinguish two experiments that differ only in the distribution of extracted identities. We conduct our proof using a standard hybrid argument: beginning with **Game 1** define a series of k intermediate hybrids during each of which a single execution of $\mathsf{BlindExtract}$ is altered from using a "real" identity σ_j to some random $\sigma'_j \xleftarrow{\$} [1, N]$. The last hybrid is equivalent to **Game 2**. If D successfully distinguishes the first and last hybrids, then $\exists D', j$ such that D' distinguishes hybrid $(j-1)$ from hybrid j with maximal probability $> \nu_2(\kappa)$. We use D' to construct an adversary $\mathcal{A}$ with non-negligible advantage in winning the selective-failure blindness game against Π.

$\mathcal{A}$ runs D' and conducts the protocol with $\hat{\mathsf{S}}$ as in **Game 1** up to the point where R initiates the $\mathsf{BlindExtract}$ protocol. At all but the ℓ^{th} execution of $\mathsf{BlindExtract}$, $\mathcal{A}$ selects the appropriate identity distribution (σ_k or σ'_k) for hybrid $(j-1)$. At the ℓ^{th} execution, $\mathcal{A}$ selects $\sigma'_\ell \xleftarrow{\$} [1, N]$ and outputs $(params, \sigma_\ell, \sigma'_\ell)$ as the first move of the selective-failure blindness game. Now $\mathcal{A}$ forwards the messages from the first oracle, $\mathcal{U}_b$ directly to $\hat{\mathsf{S}}$, returning $\hat{\mathsf{S}}$'s responses until the $\mathsf{BlindExtract}$ protocol run is complete. When D' ultimately outputs a bit b', $\mathcal{A}$ outputs b' as its guess.

Note that when $b = 0$, the ℓ^{th} extraction is conducted on σ_ℓ, and thus the game has the correct distribution for hybrid $(j-1)$. When $b = 1$, the extraction is conducted on random σ'_ℓ and thus the game has the correct distribution for hybrid j. If D' outputs 1 with probability α when presented with hybrid $(j-1)$ and probability β when presented with hybrid j, then $\mathcal{A}$ guesses correctly and wins the selective-failure blindness game with probability $\frac{|\beta-\alpha|}{2}$. If we assume that $|\beta - \alpha| > \nu_2(\kappa)$ then $\mathcal{A}$ wins with non-negligible advantage. Since contradicts our assumption about Π, then D' succeeds with probability $\leq \nu_2(\kappa)$ and thus D succeeds with probability $\leq k \cdot \nu_2(\kappa)$.

We conclude our sketch by observing that $\hat{\mathsf{S}}$ commits to the ciphertext vector in the first stage of the protocol. Assuming that $H(\cdot)$ is collision resistant, and the commitment scheme is binding, then $\hat{\mathsf{S}}$'s choice of $(C_1, \ldots, C_N)$ is independent of all subsequent actions including executions of $\mathsf{BlindExtract}$. □

Lemma B.3 (Committing IBE) $\mathbf{Adv}\,[\,\mathbf{Game\ 3}\,] - \mathbf{Adv}\,[\,\mathbf{Game\ 2}\,] \leq \nu_3(\kappa)$ if Π is committing in the sense of definition 3.8.

Proof sketch. Let D be a p.p.t. distinguisher that distinguishes the distributions of **Game 3** and **Game 2** with non-negligible advantage. This implies that for some j it is the case that with non-negligible probability $\hat{\mathsf{S}}$ (in cooperation with D) outputs at least one ciphertext C_{σ_j} such that $\mathsf{Decrypt}(sk_{\sigma_j}, C_{\sigma_j}) \neq \mathsf{Decrypt}(\mathsf{Extract}(msk, \sigma_j), C_{\sigma_j})$, while simultaneously the statement $\mathsf{IsValid}(params, \sigma_j, C_{\sigma_j}) = 1$ (since this condition is ensured by the protocol). Thus, by definition the algorithm $\hat{\mathsf{S}}$ must succeed in the game of definition 3.8 with non-negligible probability. Since Π is a committing IBE scheme, then we can bound D's advantage as $\leq \nu_3(\kappa)$. □

□

C Proof of Security for the $\mathsf{OT}^N_{k\times 1}$ Protocol

Below, we sketch a proof of Theorem 4.4 in the random oracle model.

Proof sketch. **Sender Security.** For any real-world cheating receiver $\hat{\mathsf{R}}$ we can construct an ideal-world receiver $\hat{\mathsf{R}}'$ such that no p.p.t. algorithm D can distinguish the distributions $\mathbf{Real}_{\mathsf{S},\hat{\mathsf{R}}}(N, k, M_1, \ldots, M_N, \Sigma)$ and $\mathbf{Ideal}_{\mathsf{S}',\hat{\mathsf{R}}'}(N, k, M_1, \ldots, M_N, \Sigma)$. $\hat{\mathsf{R}}'$ interacts with $\hat{\mathsf{R}}$ and the

trusted party as follows. $\hat{\mathsf{R}}'$ first runs $\mathsf{Setup}(1^\kappa, c(\kappa))$ to generate the scheme parameters, proves knowledge of msk, and sends $(C_1, \ldots, C_N)$ formed by setting $(B_1, \ldots, B_N)$ to be random bitstrings and computing $(A_1, \ldots, A_N)$ as usual. $\hat{\mathsf{R}}'$ now simulates the random oracle $H : \mathcal{M} \to \{0,1\}^{|M_1|}$, observing $\hat{\mathsf{R}}$'s queries. Whenever $\hat{\mathsf{R}}$ calls $H(\cdot)$ on a value W_{σ_i} (for some $i \in [1, N]$), $\hat{\mathsf{R}}'$ queries the trusted party to obtain M_{σ_i}. If the trusted party outputs $\perp$, then $\hat{\mathsf{R}}'$ causes the $\mathsf{BlindExtract}$ protocol to fail. Otherwise, $\hat{\mathsf{R}}'$ now programs the random oracle so that $H(W_{\sigma_i}) = B_{\sigma_i} \oplus M_{\sigma_i}$. If a p.p.t. D can distinguish the real and ideal-world distributions then it must be the case that either (a) D breaks the $\mathsf{IND\text{-}sID\text{-}CPA}$ or Leak-Free security of the IBE scheme Π, or (b) the proof-of-knowledge on msk is not zero knowledge.

Receiver Security. Our proof of receiver security is almost identical to that of the non-adaptive OT protocol (Appendix B). For any real-world cheating sender $\hat{\mathsf{S}}$ we can construct an ideal-world sender $\hat{\mathsf{S}}'$ such that no p.p.t. D can distinguish the distributions $\mathbf{Real}_{\hat{\mathsf{S}},\mathsf{R}}(N, k, M_1, \ldots, M_N, \Sigma)$ and $\mathbf{Ideal}_{\hat{\mathsf{S}}',\mathsf{R}'}(N, k, M_1, \ldots, M_N, \Sigma)$. $\hat{\mathsf{S}}'$ interacts with $\hat{\mathsf{S}}$ and the trusted party as follows. When $\hat{\mathsf{S}}$ proves knowledge of the value msk, use the appropriate knowledge extractor to obtain msk. Use msk to decrypt the ciphertext vector $(C_1, \ldots, C_N)$ as per the protocol, and transmit the resulting messages $(M_1, \ldots, M_N)$ to the trusted party T. At the i^{th} protocol round, run $\mathsf{BlindExtract}$ on a random identity σ_i'. If $\mathsf{BlindExtract}$ fails, send $b_i = 0$ to T, otherwise send $b_i = 1$. Based on the selective-failure blindness property of the IBE scheme Π, any failures in the $\mathsf{BlindExtract}$ protocol are independent of the values $(\sigma_1, \ldots, \sigma_k)$ actually extracted by an ideal-world honest receiver. If a p.p.t. D can distinguish the real and ideal-world distributions then it must be the case that either (a) $\hat{\mathsf{R}}$ breaks the selective-failure blindness property of Π, (b) Π is not committing, or (c) the extractor for msk failed.

□

D Committing IBE

In this section we sketch a proof of theorem 3.9, which states that both Π_1 and Π_2 are *committing* blind IBE schemes in the sense of definition 3.8. Note that we have not defined a (non-blind) $\mathsf{Extract}$ protocol for Π_1 and Π_2; thus, for the purposes of this proof, we will assume that all key extraction is performed via the $\mathsf{BlindExtract}$ protocol. (An $\mathsf{Extract}$ protocol could use the extraction algorithms defined by Boneh-Boyen [BB04] and Naccache [Nac05], but would include an additional check conducted by the user, to verify the correctness of the obtained secret key.) We now proceed with the sketch.

Proof sketch. Recall that for $params = (\gamma, g, g_1, g_2, h, F)$, $msk = g_2^\alpha$, message $M \in \mathbb{G}_T$, identity $id \in \mathcal{I}$ and random elements $s, r \in \mathbb{Z}_q$, well-formed ciphertexts and keys have the following structure:

$$C_{id} = (X, Y, Z) = (e(g_1, g_2)^s \cdot M, g^s, F(id)^s) \tag{1}$$

$$sk_{id} = (d_0, d_1) = (g_2^\alpha \cdot F(id)^r, g^r) \tag{2}$$

In both Π_1 and Π_2, the $\mathsf{BlindExtract}$ protocol includes a correctness check on the returned secret key. When the group parameters γ are valid, this check ensures that the user's output will either be $\perp$, or a key of the form shown in equation 2 above.[6] Similarly, the $\mathsf{Decrypt}$ algorithm includes a

[6] For some known $y \in \mathbb{Z}_q$ selected by the user, this test can be written as the comparison $e(g_1, g_2) \cdot e(d_1', F(id)g^y) = e(d_0 g^{yr}, g)$.

validity check to ensure that (a) the group parameters γ are correct (this check may be probabilistic, but is inaccurate with at most negligible probability), and (b) ciphertexts are of the form shown in equation 1. A failure in the BlindExtract check causes that protocol to output $\perp$, and a failed ciphertext check will cause Decrypt to output ϕ regardless of which secret key is used.

Now consider a malicious master authority $\mathcal{A}$ with non-negligible advantage in the game of 3.8. For $\mathcal{A}$ to succeed, it must hold that neither execution of BlindExtract with $\mathcal{A}$ outputs $\perp$, and $\mathsf{Decrypt}(params, id, sk_{id}, C) \neq \mathsf{Decrypt}(params, id, sk'_{id}, C)$. This implies that the (possibly probabilistic) group parameter check was conducted twice on γ, and succeeded at least once (else both calls to Decrypt would output ϕ). We denote by β the probability that $\mathcal{A}$ succeeds when the parameters γ are *not* valid.

In the event that the group parameters are valid and $\mathcal{A}$ succeeds, then by the ciphertext/key validity checks in BlindExtract and Decrypt, it must be the case that C, sk_{id}, sk_{id} all have the correct form for (respectively) some values $s, r_1, r_2 \in \mathbb{Z}_q$ and yet $\mathsf{Decrypt}(params, id, sk_{id}, C) \neq \mathsf{Decrypt}(params, id, sk'_{id}, C)$. Yet, by examining the math of the decryption algorithm we see that this cannot be the case. The following equation *must* hold for every tuple $s, r_1, r_2 \in \mathbb{Z}_q$:

$$\mathsf{Decrypt}(params, id, sk_{id}, C) = \mathsf{Decrypt}(params, id, sk'_{id}, C)$$
$$X \cdot \frac{e(F(id)^s, g^{r_1})}{e(g^s, g_2^\alpha \cdot F(id)^{r_1})} = X \cdot \frac{e(F(id)^s, g^{r_2})}{e(g^s, g_2^\alpha \cdot F(id)^{r_2})} = \frac{X}{e(g^s, g_2^\alpha)}$$

$\mathcal{A}$'s advantage in the game as therefore bounded by β, the probability that at least one execution of the group parameter check incorrectly accepts γ as valid. Since the definition of the group parameter check ensures that β is negligible in κ, we conclude our proof. □

On other committing blind IBE schemes. We conclude with a general observation: that any "unique" (or "functionally-unique") secure blind IBE is implicitly committing. Borrowing from the language of signatures, we define a *unique* IBE as having one valid identity secret key for each identity in the system, while a *functionally-unique* IBE may possess multiple keys per identity, but it is hard for even the master authority to compute more than one. Since the schemes presented herein are not unique, we might simplify our constructions by looking for such schemes.