Definitional Issues in Functional Encryption

Adam O'Neill*

Abstract

We formalize the emergent notion of "functional encryption," as well as introduce various security notions for it, and study relations among the latter. In particular, we show that indistinguishability and semantic security based notions of security are *inequivalent* for functional encryption in general. This is alarming given the large body of work employing (special cases of) the former. We go on to show, however, that an equivalence does hold between indistinguishability and some form of semantic security for what we call *multi-preimage sampleable* schemes. Our interpretation is that for multi-preimage sampleable schemes an indistinguishability based notion is probably fine in practice. We show that common functionalities considered in the literature satisfy this requirement.

1 Introduction

Functional encryption. In recent years, a notion of "functional encryption" (FE) has emerged as a new paradigm for public-key encryption, wherein a receiver, given a ciphertext, is able to learn certain functions of the underlying message based on its secret keys (not necessarily the decryption). Special cases of FE include (anonymous) identity-based encryption [BF03, ABC$^+$08], public-key encryption with keyword search [BCOP04, ABC$^+$08], attribute-based encryption [SW05, GPSW06, BSW07, OT10], and predicate encryption [BW07, KSW08, LOS$^+$10, OT10]. However, a general study of FE and its security seems not to have appeared. Here we initiate one, and in doing so we uncover some interesting definitional issues that have important implications for work in this area.

Syntax and security notions. First we give a syntactical definition of FE, which extends that for predicate encryption introduced by Boneh and Waters [BW07]. We then formulate an "indistinguishability based" notion of privacy (IND), which again extends the security notion for predicate encryption introduced in [BW07]. Informally, the IND notion asks that it be hard for an adversary to distinguish between the encryptions of any two messages that agree on all the functions corresponding to the secret keys it requested. We go on to introduce a more complicated but more natural "semantic security based" (SS) notion of privacy in the spirit of the classical notion for public-key encryption [GM84], to capture the intuition that anything the adversary can compute from a ciphertext it could as well compute from the evaluations of the functions corresponding to the secret keys it requested on underlying message. We note that a novel feature of our definitions, which turns out to be important when considering relations among them, is that they distinguish between *adaptive* and *non-adaptive* access to the secret-key derivation oracle to

*University of Texas at Austin. Work done in part while the author was a Ph.D. student at Georgia Institute of Technology.

aid in the adversary's task (roughly, this distinction is analogous to that between access to the decryption oracle in adaptive versus non-adaptive chosen-ciphertext attack in the classical sense).

Relations among security notions. In the classical setting of public-key encryption, semantic security and indistinguishability based formulations of security are well-known to be equivalent [MRS88]. We ask whether the same is true for FE. Surprisingly, we show that IND under adaptive access to the secret-key derivation oracle does not imply SS even under *non-adaptive* such access. To see why, consider a functional encryption scheme for a single function f. But suppose there is another function g that has the same "equality pattern" as f on the message space (i.e., two messages have the same f-value just when they have the same g-value). Furthermore, suppose $g(m)$ is hard to compute given $f(m)$. Now, if the functional encryption scheme is such that the secret keys created by the scheme, which are supposed to allow computing f, also allow computing g, the scheme is certainly not semantically secure. However, an IND adversary is "bound" to choosing messages that agree on f, hence also on g, and so cannot use computing g to its advantage. Our counter-example formalizes this intuition. Another shortcoming of the IND notion we observe is that it is in essentially vacuous[1] for some functions, such as a collision-resistant hash function. Then, it is hard for the adversary to find two messages that agree on the function.

Achievability. Finally, we ask the question of whether the SS notion for FE is *achievable*. In particular, we note that achieving SS under adaptive access to the key derivation oracle seems difficult. In the proof, the simulator seemingly must choose a "dummy" ciphertext on which to run the adversary *before* knowing what values the challenge message should have when evaluated under the functions for which the adversary will later request secret keys. Intuitively, this means the number of possible keys for a given function should at least be as large as the number of possible outputs of the latter. This situation is reminiscent of that for (non-interactive) non-committing encryption, for which impossiblity results are known [Nie02].

Fortunately, we also bring some good news. In the case of non-adaptive SS, we identify a key property of functional encryption schemes that we call *multi-preimage sampleability*. Intuitively, this means that the functions of the messages an adversary is allowed to compute does not "narrow down" the message space too severely; given the function values of some underlying message it should always be possible to find two different messages consistent with them. We show that for multi-preimage sampleable FE schemes, IND is *equivalent* to SS (both under non-adaptive access to the key-derivation oracle). The reason we believe this is important is that non-adaptive SS suffices to rule out the "pathological" examples of schemes we gave that meet IND but not SS.[2] Thus, our interpretation is that for multi-preimage sampleable schemes, IND (under adaptive access to the key-derivation oracle) is probably fine in practice. We conclude by showing that some common function classes considered in the literature, including the powerful inner-product predicates realized in [KSW08, LOS+10, OT10], are multi-preimage sampleable.

Concurrent and independent work. Independently of our work, Boneh et al. [BSW10] also undertook a general study of FE. In particular they gave a syntactical definition as well as indistinguishability and semantic security based formulations of privacy (their formulation of the latter

[1] At least, it is vacuous with respect to attacks that require the adversary to query its key derivation oracle; i.e., attacks where the adversary actually uses the secret keys. A functional encryption scheme may of course already not be semantically secure in the classical sense.

[2] On the other hand, it is possible to extend them to even more extreme examples that violate SS only under adaptive access to the key-derivation oracle, but these start to really stretch plausibility. In any case, for SS non-adaptivity may be what we are stuck with.

differs somewhat from ours; see the discussion in Section 2). Under their formulations, they give a (again, somewhat different[3]) example showing that IND does not imply SS. They further showed that SS under adaptive access to the key derivation oracle (although they do not distinguish between adaptive versus non-adaptive here) is not achievable at all without (programmable) random oracles (following [Nie02]), but is achievable in the random oracle model. We feel this further highlights the importance of our results on the (standard model) achievability of non-adaptive SS.

2 Functional Encryption and its Security

We define the syntax of functional encryption and various security notions for it.

2.1 Syntax

A *functional encryption scheme* for the class of PT functions (aka. functionality) $\mathcal{F}$ on message-space Σ is a tuple of algorithms $\mathcal{FE} = (\mathsf{Setup}, \mathsf{KDer}, \mathsf{Enc}, \mathsf{Eval})$ such that:

- Setup on input 1^k outputs a *master public key* pk and *master secret key* sk.
- KDer on input the master secret key sk and a (description of a) function $f \in \mathcal{F}$ outputs an *evaluation token* (aka. secret key) sk_f for f.
- Enc on input a public key pk and a message (aka. attribute) $m \in \Sigma$ outputs a ciphertext c.
- Eval on input an evaluation token sk_f and a ciphertext c outputs a string y or $\perp$.

For correctness we require that for all $k \in \mathbb{N}$, all $f \in \mathcal{F}$, and all $m \in \Sigma$,

$$\mathsf{Eval}(sk_f, \mathsf{Enc}(pk, m)) = f(m)$$

with probability 1 over $(pk, sk) \xleftarrow{\$} \mathsf{Setup}(1^k)$ and $sk_f \xleftarrow{\$} \mathsf{KDer}(sk, f)$.

Note that this notion is in particular a generalization of (anonymous) identity-based encryption [BF03, ABC+08] (IBE), public-key encryption with keyword search [BCOP04, ABC+08], attribute-based encryption [SW05, GPSW06, BSW07, OT10], and predicate encryption [BW07, KSW08, LOS+10, OT10].[4] For example, in the case of identity-based encryption, the "message" would consist of the identity concatenated with the actual payload, and the secret key would be associated with the function $f_{ID}(ID' \| x) = x$ if $ID = ID'$ and $\perp$ otherwise.

2.2 Security Definitions

We present various formulations of privacy for functional encryption. Broadly, the definitions are either *indistinguishability based* or *semantic-security based*. In each case we also define a *token non-adaptive* (TNA) variant, where the adversary gets access to a token derivation oracle only before it sees the challenge ciphertext.

[3] In fact, their counter-example does not seem correct, since their scheme is not even semantically secure and therefore can be broken under the IND definition by an adversary that does not request any secret keys (and hence is unrestricted in its challenge messages). Even disregarding this, their counter-example would show that an FE scheme may meet the IND notion but not SS because it is hard for an adversary to even find two messages that agree on the functionality, whereas ours shows a separation even for schemes where the adversary can find such messages.

[4] To capture *non-anonymous* IBE or attribute-based encryption, we would need to enhance our definition to output a special part of the message in the clear, which we omit for simplicity.

INDISTINGUISHABILITY BASED PRIVACY. The indistinguishability-based formulation follows [BW07] and tries to capture the intuition that the adversary is unable to distinguish between the encryptions of two different messages that it cannot trivially distinguish using its tokens. Let $\mathcal{FE} = (\mathsf{Setup}, \mathsf{KDer}, \mathsf{Enc}, \mathsf{Eval})$ be a functional encryption scheme for the class of functions $\mathcal{F}$ over message-space Σ and let $A = (A_1, A_2)$ be an adversary. For mode $\in \{\text{full}, \text{tna}\}$ [5] and $k \in \mathbb{N}$ we associate to $\mathcal{FE}$ and A the experiments

$$\begin{array}{l}\textbf{Experiment } \mathbf{Exp}^{\text{ind-mode}}_{\mathcal{FE},A}(k)\text{:}\\ \quad b \xleftarrow{\$} \{0,1\}\\ \quad (pk, sk) \xleftarrow{\$} \mathsf{Setup}(1^k)\\ \quad (m_0, m_1, st) \xleftarrow{\$} A_1^{\mathsf{KDer}(sk,\cdot)}(pk)\\ \quad c \xleftarrow{\$} \mathsf{Enc}(pk, m_b)\\ \quad b' \xleftarrow{\$} A_2^{\mathcal{O}(sk,\cdot)}(pk, c, st)\\ \quad \text{If } b = b' \text{ return 1 else return 0}\end{array}$$

where if mode = full then $\mathcal{O}(sk, \cdot) = \mathsf{KDer}(sk, \cdot)$ and if mode = tna then $\mathcal{O}(sk, \cdot) = \varepsilon$ (the empty oracle). We require that every query f that A_1 or A_2 makes to its oracle satisfies $f(m_0) = f(m_1)$. Denote by $\Pr\left[\, \mathbf{Exp}^{\text{ind-mode}}_{\mathcal{FE},A}(k) = 1 \,\right]$ the probability that the corresponding IND-MODE experiment outputs 1, and define

$$\mathbf{Adv}^{\text{ind-mode}}_{\mathcal{FE},A}(k) = 2 \cdot \Pr\left[\, \mathbf{Exp}^{\text{ind-mode}}_{\mathcal{FE},A}(k) = 1 \,\right] - 1\ .$$

We say that $\mathcal{FE}$ is *IND-MODE secure* if $\mathbf{Adv}^{\text{ind-mode}}_{\mathcal{FE},A}(\cdot)$ is negligible for all PPT adversaries A.

SEMANTIC-SECURITY BASED PRIVACY. The semantic-security formulation is new and tries to capture the intuition that anything the adversary can compute from a ciphertext and the tokens it can compute from the tokens and the values of the corresponding functions on the underlying message. Let $\mathcal{FE} = (\mathsf{Setup}, \mathsf{KDer}, \mathsf{Enc}, \mathsf{Eval})$ be a functional encryption scheme for the class of functions $\mathcal{F}$ over message-space Σ, let $A = (A_1, A_2, A_3)$ be an adversary, let S be a simulator. For mode $\in \{\text{full}, \text{tna}\}$ and $k \in \mathbb{N}$ we associate to $\mathcal{FE}$, A, and S the experiments

$$\begin{array}{l|l}\textbf{Experiment } \mathbf{Exp}^{\text{ss-real-mode}}_{\mathcal{FE},A}(k)\text{:} & \textbf{Experiment } \mathbf{Exp}^{\text{ss-ideal-mode}}_{\mathcal{FE},A,S}(k)\text{:}\\ \quad (pk, sk) \xleftarrow{\$} \mathsf{Setup}(1^k) & \quad (pk, sk) \xleftarrow{\$} \mathsf{Setup}(1^k)\\ \quad st \xleftarrow{\$} A_1^{\mathsf{KDer}(sk,\cdot)}(pk) & \quad st \xleftarrow{\$} A_1^{\mathsf{KDer}(sk,\cdot)}(pk)\\ \quad (m, t) \xleftarrow{\$} A_2(pk, st) & \quad (m, t) \xleftarrow{\$} A_2(pk, st)\\ \quad c \xleftarrow{\$} \mathsf{Enc}(pk, m) & \quad \text{Let } f_1, \ldots, f_q \text{ be the queries made by } A_1\\ \quad t' \xleftarrow{\$} A_3^{\mathcal{O}(sk,\cdot)}(pk, c, st) & \quad t' \xleftarrow{\$} S^{\mathcal{O}'(sk,\cdot)}(pk, f_1(m), \ldots, f_q(m), st)\\ \quad \text{If } t = t' \text{ return 1 else return 0} & \quad \text{If } t = t' \text{ return 1 else return 0}\end{array}$$

where if mode = full then $\mathcal{O}(sk, \cdot) = \mathsf{KDer}(sk, \cdot)$ and for any f oracle $\mathcal{O}'(sk, f)$ returns $(sk_f, f(m))$ where $sk_f \xleftarrow{\$} \mathsf{KDer}(sk, f)$, and if mode = tna then $\mathcal{O}(sk, \cdot) = \mathcal{O}'(sk, \cdot) = \varepsilon$ (the empty oracle). We assume for simplicity that A_1's output (the state st) includes its oracle queries and the responses. Think of the string $t \in \{0,1\}^*$ in the output of A_2 as partial information on m. Note that in

[5] We stress that our use of the terminology "full" security differs from the literature in that it refers to adaptive access to the key derivation oracle rather than adaptive choice of the challenge messages.

the above formalization of semantic security, even in the ideal experiment we run A_1 and A_2. A more standard formalization would have the simulator also run at these stages. However, we want to "bind" the simulator to making the same key derivation queries as the adversary.[6] Denote by $\Pr\left[\,\mathbf{Exp}^{\text{ss-real-mode}}_{\mathcal{FE},A}(k)=1\,\right]$ the probability that the SS-REAL-MODE experiment outputs 1 and by $\Pr\left[\,\mathbf{Exp}^{\text{ss-ideal-mode}}_{\mathcal{FE},A,S}(k)=1\,\right]$ the probability that the SS-IDEAL-MODE experiment outputs 1. Define

$$\mathbf{Adv}^{\text{ss-mode}}_{\mathcal{FE},A,S}(k)=\Pr\left[\,\mathbf{Exp}^{\text{ss-real-mode}}_{\mathcal{FE},A}(k)=1\,\right]-\Pr\left[\,\mathbf{Exp}^{\text{ss-ideal-mode}}_{\mathcal{FE},A,S}(k)=1\,\right]\ .$$

We say that $\mathcal{FE}$ is *SS-MODE secure* if for every PPT adversary A there exists a PPT simulator S such that $\mathbf{Adv}^{\text{ss-mode}}_{\mathcal{FE},A,S}(\cdot)$ is negligible.

3 Inequivalence of the Definitions in General

We investigate relations among the notions of security we introduced for FE. First, we note that when giving the adversary adaptive access to the token derivation oracle (i.e., what we call FULL security), one reason semantic security seems stronger than indistinguishability is that the simulator apparently needs to commit to a "dummy" ciphertext on which to run the adversary *before* knowing what values the challenge message should have when evaluated under the functions for which the adversary will later request tokens.

But we show that there is actually a more subtle reason for inequivalence of the definitions. In fact, we show that in general IND-FULL security does not even imply SS-*TNA* security. To show the separation we start with a IND-FULL secure functional encryption scheme for any class of functions $\mathcal{F}$ of a certain form. We then modify it to construct a new scheme that is still IND-FULL secure for $\mathcal{F}$ but *not* SS-TNA secure. We show the latter by presenting a concrete attack. We first describe a concept our counter-example scheme employs.

HIDDEN FUNCTIONS. Let $\mathcal{G}=\{g_k\}_{k\in\mathbb{N}}$ and $\mathcal{F}=\{f_k\}_{k\in\mathbb{N}}$ be families of functions on a common domain $D=D(k)$. We say $\mathcal{G}$ is *hidden* by $\mathcal{F}$ if any PPT adversary A on inputs $f_k, f_k(x)$ where $x\xleftarrow{\$}D$ outputs $g_k(x)$ with only negligible probability in k. Note that such functions can be constructed under standard assumptions; for example, let f_k be a one-way function applied to the first half of the bits of the input and let g_k just output these bits (that is, the first half of the bits of the input).[7] We say $\mathcal{F}$ and $\mathcal{G}$ are *isomorphic* if f_k and g_k are isomorphic for every k, meaning

$$f_k(d_1)=f_k(d_2)\Leftrightarrow g_k(d_1)=g_k(d_2)\ .$$

for all $d_1,d_2\in D$. In other words, f_k and g_k have the same equality pattern across the domain. This is the case, for example, if f_k in the example above is an *injective* one-way function on the first

[6] On the other hand, in the case of FULL security this is not enforced by the definition. This was an oversight on our part that we leave in here, since we only noticed it after [BSW10] appeared; following [BSW10] one might ask that A_3 and S have the same query distributions in thie case. Other differences between their SS definition and ours include: theirs considers only adaptive access to the key derivation oracle whereas ours distinguishes between adaptive and non-adaptive, theirs allows the challenge message to depend only on the security parameter whereas ours also allows it to depend on public parameters and key derivation queries, and theirs considers the encryption of multiple messages whereas ours considers only a single message (in particular, the former is important for the proof of impossibility they give for meeting their notion).

[7] Indeed, a simpler example is to take f_k to be a one-way function and g_k to be the identity. However, in our counter-example this will prevent the adversary from even being able to *find* two messages that agree on f_k. We believe this points to a separate shortcoming of the IND definition.

half of the input bits. We supress dependence on k below for convenience, just talking of functions rather than function families.

THE COUNTER-EXAMPLE SCHEME. Let $\mathcal{AE}^* = (\mathsf{KDer}^*, \mathsf{Enc}^*, \mathsf{Dec}^*)$ be a (standard) public-key encryption scheme, and let $\mathcal{FE}' = (\mathsf{Setup}', \mathsf{KDer}', \mathsf{Enc}', \mathsf{Eval}')$ be a functional encryption scheme over message-space Σ for a class of functions $\mathcal{F} = \{f_1, \ldots, f_n\}$ satisfying the following: there is a function g on Σ such that the pointwise concatenated function[8] $f = f_1 \| \ldots \| f_n$ is isomorphic to g and moreover g is hidden by f. (For simplicity, we assume here that n is polynomial in k. The counter-example can easily be extended to larger function sets by instead requiring the forgoing condition on some fixed *subset* of the f_i's.) Then we define a new functional encryption scheme $\mathcal{FE} = (\mathsf{Setup}, \mathsf{KDer}, \mathsf{Enc}, \mathsf{Eval})$ over Σ for $\mathcal{F}$ as follows.

- Setup on input 1^k first runs $(pk', sk') \xleftarrow{\$} \mathsf{Setup}'(1^k)$, and $(pk^*, sk^*) \xleftarrow{\$} \mathsf{KDer}^*(1^k)$. It then selects $w_1, \ldots, w_{n-1} \xleftarrow{\$} \{0,1\}^{|sk^*|}$ and computes $w_n \leftarrow sk^* \oplus w_1 \oplus \cdots \oplus w_{n-1}$. Finally, it returns master public key $pk = pk' \| pk^*$ and master secret key $sk = sk' \| w_1 \| \ldots \| w_n$.

- KDer on input the master secret key $sk = sk' \| w_1 \| \ldots \| w_k$ and a (description of a) function $f_i \in \mathcal{F}$ first runs $\mathsf{KDer}'_{sk'}(f_i)$ to obtain sk'_{f_i}. Then, it outputs $sk_{f_i} = sk'_{f_i} \| w_i$.

- Enc on input the master public key $pk = pk' \| pk^*$ and a message $m \in \Sigma$ first computes $c' \xleftarrow{\$} \mathsf{Enc}'(pk', m)$ and $c^* \xleftarrow{\$} \mathsf{Enc}^*(pk^*, g(m))$. It returns $c' \| c^*$.

- Eval on input a secret key $sk_{f_i} = sk'_{f_i} \| w_i$ and a ciphertext $c = c' \| c^*$ computes $d \leftarrow \mathsf{Eval}'(sk'_{f_i}, c')$, and outputs d.

Theorem 3.1 If $\mathcal{AE}^*$ is IND-CPA secure and $\mathcal{FE}'$ is IND-FULL secure for $\mathcal{F} = \{f_1, \ldots, f_n\}$ as above (i.e., where g is hidden by $f_1 \| \ldots \| f_n$), then $\mathcal{FE}$ is also IND-FULL secure for $\mathcal{F}$. However, it is not SS-TNA secure.

Note that the assumptions of the theorem do not constitute any additional complexity assumptions beyond the (minimal) one of $\mathcal{FE}'$ being IND-FULL secure for $\mathcal{F}$, meaning based on the latter we can construct the other schemes and functions that are assumed.

We also remark that the separation also holds in the case of "selective-security," where the challenge messages are chosen up-front by the adversary, as considered in e.g. [BW07, KSW08]. It also holds in the case of predicate encryption [BW07, KSW08], since we can take f_i for $1 \leq i \leq n$ to output the i-th bit of a function f such that g is hidden by f (i.e., a function can always be decomposed bit-wise into predicates).

Proof: (Sketch.) To see $\mathcal{FE}$ is IND-FULL secure for $\mathcal{F}$, first consider an adversary A that does not request tokens for all of $f_1, \ldots, f_n$. Then in addition to interacting with $\mathcal{FE}'$ in the IND-FULL experiment, the adversary is just given additional random strings when it requests tokens, which in particular are independent of b, so security of $\mathcal{FE}$ follows from that of $\mathcal{FE}'$. Now consider A that requests tokens for all of $f_1, \ldots, f_n$. In this case, in addition to interacting with $\mathcal{FE}'$ the adversary obtains $g(m_b)$ where m_b is the challenge message. But by the rules of the experiment we know that $f_1(m_0) \| \ldots \| f_n(m_0) = f_1(m_1) \| \ldots \| f_n(m_1)$ and thus by assumption $g(m_0) = g(m_1)$, meaning again this information is independent of b and so IND security of $\mathcal{FE}$ follows from that of $\mathcal{FE}'$.

[8] By pointwise concatenation $f \| g$ of functions f and g on a set D we mean that $f \| g(x) = f(x) \| g(x)$ for all $x \in D$.

To show that $\mathcal{FE}$ is not SS-TNA secure, we describe an SS-TNA adversary $B = (B_1, B_2, B_3)$ for which there is no simulator with comparable probability of guessing $t = t'$. Namely, B_1 requests evaluation tokens for all of $f_1, \ldots f_n$ and passes them along as the state, and B_2 chooses a random challenge message $m \in \Sigma$, sets $t \leftarrow g(m)$, and outputs (m, t). Then, by construction B_3 can always output $t = t'$ by decrypting the part of the challenge ciphertext formed by $\mathcal{AE}^*$ (note that B_3 makes no queries itself as required). However, a simulator who outputs $t = t'$ with non-negligible probability would contradict the fact that g is hidden by f since the simulator is not given any ciphertext but just the value of $f_1(m) \| \ldots \| f_n(m)$ (also pk and the evaluation tokens for $f_1, \ldots, f_n$, but a hidden function adversary can generate these itself). ▌

4 An Equivalence under Multi-Preimage Sampleability

We show that for token non-adaptive (TNA) security the counter-example in Section 3 is essentially tight. Namely, we show an *equivalence* between indistinguishability and semantic-security under TNA security for what we call *multi-preimage sampleable* schemes. Note that TNA security seems reasonable in practical applications where what tokens a party receives does not depend on the encrypted messages.

Multi-preimage sampleability. Let $\mathcal{FE} = (\mathsf{Setup}, \mathsf{KDer}, \mathsf{Enc}, \mathsf{Eval})$ be a functional encryption scheme over message-space Σ for the class of functions $\mathcal{F}$. We call $\mathcal{FE}$ *multi-preimage sampleable* (MPS) if there is a PPT algorithm that given $(f_1, y_1 = f_1(m)), \ldots, (f_i, y_i = f_i(m))$ for any polynomial $i = i(k)$, any $f_1, \ldots, f_i \in \mathcal{F}$ and any $m \in \Sigma$ samples uniformly from the set

$$S_{y_1,\ldots,y_i} \;=\; \{m' \in \Sigma \mid f_1(m') = y_1, \ldots, f_i(m') = y_i\}$$

and moreover $|S_{y_1,\ldots,y_i}| \geq 2$. Note that multi-preimage sampleability as we have defined it is thus a property of $\mathcal{F}$.

It is sometimes more useful to consider (cf. Section 5) multi-preimage sampleability not as a property of a functionality $\mathcal{F}$ itself but rather of an SS *adversary*. Namely, call an SS adversary $A = (A_1, A_2, A_3)$ *multi-preimage samplable* if there is a PPT algorithm that given the queries $f_1, \ldots, f_q$ made by A_1 in any run of the SS-REAL-TNA experiment samples uniformly from the set

$$S_{f_1(m),\ldots,f_q(m)} \;=\; \{m' \in \Sigma \mid f_1(m') = f_1(m), \ldots, f_q(m') = f_q(m)\}$$

where m is the message output by A_2, and moreover $|S_{f_1(m),\ldots,f_q(m)}| \geq 2$.

In essence, we show that multi-preimage sampleability provides a "test" of whether equivalence between the IND and SS definitions is maintained in the case of TNA security.

Theorem 4.1 Let $\mathcal{FE}$ be an MPS functional encryption scheme. Then $\mathcal{FE}$ is SS-TNA secure if and only if it is IND-TNA secure.

An analogous theorem holds for *any* functional encryption scheme when considering only MPS adversaries in the SS-TNA case.

Proof: (Sketch.) Suppose that $\mathcal{FE}$ is not IND-TNA secure, in particular let $A = (A_1, A_2)$ be a successful IND-TNA adversary against it. Consider a SS-TNA adversary $B = (B_1, B_2, B_3)$ that works as follows. B_2 runs A_1 on pk to receive messages m_0, m_1. It then chooses $d \in \{0, 1\}$ at

random and returns (m_d, d). B_3 runs A_2 on its input and outputs the result. Note that no SS-TNA simulator can output d with probability better than $1/2$ in the SS-TNA-REAL experiment because it gets no information about it (recall that A only makes key-derivation queries whose results are independent of b, and the simulator makes no queries).

Now suppose $\mathcal{FE}$ *is* IND-TNA secure. Let $A = (A_1, A_2, A_3)$ be an SS-TNA adversary against $\mathcal{FE}$. We construct a simulator S with comparable probability of outputting $t = t'$ in the SS-IDEAL-TNA experiment to A in the SS-REAL-TNA experiment, meaing $\mathcal{FE}$ is SS-TNA secure. Simulator S works as follows: given the queries $f_1, \ldots, f_q$ made by A_1 and their values $y_1, \ldots, y_q$ on the challenge message m, S will sample uniformly a "dummy" message $m' \in \Sigma$ such that $f_1(m') = y_1, \ldots, f_q(m') = y_q$ using the sampler guaranteed by the definition of MPS. It runs A_3 on the encryption of m' and outputs the result. If A_3's success probability differs in this simulated environment, then we can use A to construct a successful IND-TNA adversary B, as follows. B runs A_1, A_2 on the appropriate inputs to receive m, t. Let $f_1, \ldots, f_q$ be the queries made by A_1. Using the sampler guaranteed by the definition of MPS, B samples uniformly a message m' such that $f_1(m') = f_1(m), \ldots, f_q(m') = f_q(m)$. It then submits m, m' as its challenge messages, runs A_3 on the result and checks whether it returns $t' = t$ or not; if so, it outputs 0, and otherwise 1. Note that it is important here that $m \neq m'$, which holds with probability at least $1/2$ by the definition of MPS. This contradicts our assumption. ▌

5 On Multi-Preimage Sampleability of Some Functionalities

We examine whether some specific functionalities (i.e., function classes) for FE considered in the literature satisfy our multi-preimage sampleability condition. In those we consider we show that either the answer is "yes" or the restriction to MPS adversaries for them under the SS-TNA notion is relatively natural.

INNER-PRODUCTS. We first show that multi-preimage sampleability is satisfied by the important class of *inner-product predicates* realized in prior work [KSW08, LOS+10, OT10]. Hence, by Theorem 4.1, schemes in the literature for this functionality proven secure relative to the IND notion also meet SS, at least under non-adaptive access to the token-derivation oracle. Namely, consider the evaluation of inner products over $\mathbb{Z}_N$ for a composite N (of which it assumed hard to find a non-trivial factor). More formally, let $n \in \mathbb{N}$ be given and let N be such a composite. Let $\Sigma = \mathbb{Z}_N^n \setminus \{0^n\}$ [9] and define the associated class of *inner-product predicates* $\mathcal{P}_{iprod} = \{p_{\mathbf{x}} \mid \mathbf{x} \in \mathbb{Z}_N^n\}$ where $p_{\mathbf{x}}(\mathbf{y}) = 1$ if $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i \cdot y_i = 0 \bmod N$, and 0 otherwise.

Proposition 5.1 The class $\mathcal{P}_{iprod}$ as defined above is multi-preimage sampleable.

Proof: (Sketch.) For any polynomial $r = r(k)$, given $y_1 = p_{\mathbf{x}_1}(m), \ldots, y_r = p_{\mathbf{x}_r}(m)$ for any $p_{\mathbf{x}_1}, \ldots, p_{\mathbf{x}_r} \in \mathcal{P}_{iprod}$ and $m \in \Sigma$, we construct an algorithm A that samples uniformly from $S_{y_1,\ldots,y_r}$. Let I_b denote the set $\{i \in [r] \mid y_i = b\}$ for $b \in \{0, 1\}$, and let B be an $|I_1| \times n$ matrix where each row is a unique element of $\{\mathbf{x}_i \mid i \in I_1\}$.

The sampling algorithm A first finds a basis $W = \{\mathbf{w}_1, \ldots, \mathbf{w}_s\}$ for $\ker(B)$ in the space $\mathbb{Z}_N^n$. This can be done by solving the homogeneous system of equations $B\mathbf{x} = \mathbf{0}$ using Gaussian elimination

[9] We need to disallow the all zeros vector as a message in order for multi-preimage sampleability to hold.

over $\mathbb{Z}_N$. (Note that while $\mathbb{Z}_N$ is not a field, if Gaussian elimination fails we have found a non-trivial factor of N.) Next, A samples a uniformly random element $\mathbf{w} \in \ker(B)$ by taking a random $\mathbb{Z}_N$-combination of the basis vectors in W. A then tests if $\langle \mathbf{w}, \mathbf{x}_i \rangle \neq 0$ for all $i \in I_0$. If so, A outputs $\mathbf{w}$ and terminates, and if not, then A continues to re-sample $\mathbf{w}$ from $\ker(B)$ until this is true (but halts after say a maximum k attempts).

First of all, A is a PPT algorithm, since a random vector from $\ker(B)$ is overwhelmingly likely to be non-orthogonal to each $\mathbf{x}_i$, $i \in I_0$, and thus by a union bound the termination condition is achieved with overwhelming probability in each attempt. Furthermore, A correctly samples $S_{y_1,\ldots,y_r}$ because $S_{y_1,\ldots,y_r}$ consists precisely of all vectors whose inner product with the elements of $\{\mathbf{x}_i \mid i \in I_1\}$ is zero and with the elements of $\{\mathbf{x}_i \mid i \in I_0\}$ is non-zero. Finally $|S_{y_1,\ldots,y_r}| \geq 2$ because, by definition of Σ, $s \geq 1$ (i.e., $\ker(B)$ has dimension at least 1). ▮

Anonymous IBE and PEKS. Observe that the (anonymous) IBE [BF03, ABC+08] and public-key encryption with keyword search (PEKS) [BCOP04, ABC+08] functionalities are *not* multi-preimage sampleable as we defined it. For example, in the case of anonymous IBE, if we know that $f_{ID}(ID' \| x) = x$ then there is only one possible preimage, namely $ID \| x$ (an analogous argument applies in the case of PEKS). However, for such schemes, asking that SS-TNA security hold only relative to MPS *adversaries* seems relatively natural. For example, in the case of anonymous IBE, it corresponds to asking that the adversary not ask for a secret key corresponding to the challenge identity; indeed, in this case, we can sample uniformly from the set of possible "messages" by sampling a random identity other than those for which the adversary has requested secret keys and a random payload. By an analogue of Theorem 4.1, we conclude that such schemes in the literature proven secure under an IND notion also meet SS-TNA under this condition.

Fuzzy IBE and attribute-based encryption. Fuzzy IBE [SW05] and more generally attribute-based encryption [GPSW06, BSW07, OT10] are also not multi-preimage sampleable as we have defined it for similar reasons to anonymous IBE and PEKS. However, we believe that by asking that the SS-TNA security hold only relative to MPS adversaries for these functionalities, we again get the relatively natural condition that the adversary not ask for a secret key that allows it to decrypt the challenge message. However, we have not checked the details.

Acknowledgements

We are very grateful to Alexandra Boldyreva for the initial conversations that led to this research, Mihir Bellare for discussions about the definitions and comments on the draft, and Nathan Chenette for discussions about multi-preimage sampleability.

References

[ABC+08] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi. Searchable encryption revisited: Consistency properties, relation to anonymous ibe, and extensions. *J. Cryptology*, 21(3):350–391, 2008.

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

[BF03] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the weil pairing. *SIAM J. Comput.*, 32(3), 2003.

[BSW07] John Bethencourt, Amit Sahai, and Brent Waters. Ciphertext-policy attribute-based encryption. In *IEEE Symposium on Security and Privacy*, pages 321–334, 2007.

[BSW10] Dan Boneh, Amit Sahai, and Brent Waters. Functional encryption: Definitions and challenges. Cryptology ePrint Archive, Report 2010/543, 2010. `http://eprint.iacr.org/`.

[BW07] Dan Boneh and Brent Waters. Conjunctive, subset, and range queries on encrypted data. In *TCC*, pages 535–554, 2007.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[GPSW06] Vipul Goyal, Omkant Pandey, Amit Sahai, and Brent Waters. Attribute-based encryption for fine-grained access control of encrypted data. In *ACM Conference on Computer and Communications Security*, pages 89–98, 2006.

[KSW08] Jonathan Katz, Amit Sahai, and Brent Waters. Predicate encryption supporting disjunctions, polynomial equations, and inner products. In *EUROCRYPT*, pages 146–162, 2008.

[LOS+10] Allison B. Lewko, Tatsuaki Okamoto, Amit Sahai, Katsuyuki Takashima, and Brent Waters. Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption. In *EUROCRYPT*, pages 62–91, 2010.

[MRS88] Silvio Micali, Charles Rackoff, and Bob Sloan. The notion of security for probabilistic cryptosystems. *SIAM J. Comput.*, 17(2), 1988.

[Nie02] Jesper Buus Nielsen. Separating random oracle proofs from complexity theoretic proofs: The non-committing encryption case. In *CRYPTO*, pages 111–126, 2002.

[OT10] Tatsuaki Okamoto and Katsuyuki Takashima. Fully secure functional encryption with general relations from the decisional linear assumption. In *CRYPTO*, pages 191–208, 2010.

[SW05] Amit Sahai and Brent Waters. Fuzzy identity-based encryption. In *EUROCRYPT*, pages 457–473, 2005.