On the Semantic Security of Functional Encryption Schemes

Manuel Barbosa[1] and Pooya Farshim[2]

[1] HASLab – INESC TEC and Universidade do Minho, Portugal
[2] Fachbereich Informatik, Technische Universität Darmstadt, Germany
mbb@di.uminho.pt, farshim@cased.de

Abstract. Functional encryption (FE) is a powerful cryptographic primitive that generalizes many asymmetric encryption systems proposed in recent years. Syntax and security definitions for FE were proposed by Boneh, Sahai, and Waters (BSW) (TCC 2011) and independently by O'Neill (ePrint 2010/556). In this paper we revisit these definitions, identify several shortcomings in them, and propose a new definitional approach that overcomes these limitations. Our definitions display good compositionality properties and allow us to obtain new feasibility and impossibility results for adaptive token-extraction attack scenarios that shed further light on the potential reach of general FE for practical applications.

Keywords. Functional encryption, Semantic security, Adaptive token extraction, Inner-product encryption, SIS problem.

1 Introduction

Functional encryption (FE) is a public-key primitive that generalizes many encryption systems, including public-key encryption (PKE), identity-based encryption (IBE), searchable encryption, attribute-based encryption (ABE), and all other variants of predicate encryption systems [5]. In such a system, each decryption key TK_f (called a *token*) is associated with a function f (which may be viewed as a circuit). When a token holder runs the decryption algorithm on a ciphertext encrypting a message m, it recovers the image $f(\mathsf{m})$. A trusted authority holding a master secret key is responsible for issuing tokens. This allows the TA to control which users can recover which images from encrypted data. Realizing such a powerful primitive for complex functionalities could revolutionize information security applications in a way that is comparable to the notable case of fully homomorphic encryption. Interestingly, very recent developments in this area indicate that this may indeed be within our reach. For example, a concrete realization of functional encryption for arbitrary functionalities has been recently proposed in [9], as well as functional encryption for regular languages in [15].

The intuitive security requirement for a functional encryption scheme is that no information should leak from a ciphertext bar that which can be recovered via legitimately obtained decryption tokens. As for other encryption primitives, there are various ways in which this intuition can be formalized. In the case of PKE, for example, the two standard formalizations are semantic security and ciphertext indistinguishability, which were shown to be equivalent in the seminal work of Goldwasser and Micali [8].

K. Kurosawa and G. Hanaoka (Eds.): PKC 2013, LNCS 7778, pp. 143–161, 2013.

Somewhat surprisingly, in independent works, Boneh, Sahai, and Waters [5] (BSW) and O'Neill [14] have shown that this is *not* the case for FE schemes supporting complex functionalities. Indeed, both works demonstrated limitations in the indistinguishability-based notion for functional encryption and proposed strictly stronger semantic security notions to overcome these problems. This highlights the importance of converging to a definition of semantic security for FE that can be adopted as a de facto standard by the community. However, the definitional approaches adopted in both works are significantly different and the relation between the two is not well understood. In particular, it is not clear whether there are fundamental differences between the two so as to determine which one of them should be favored in detriment of the other. The goal of this paper is to change this state of affairs. We analyze the positive and negative aspects of the definitions by BSW and O'Neill and find that both approaches have strengths that should be preserved, and yet they also have weaknesses that should be reconsidered. We propose a new balanced set of definitions incorporating these results.

ANALYSIS OF PREVIOUS DEFINITIONS. Boneh et al. [5] provide an elegant generalization of the syntax of FE schemes. The authors propose a natural indistinguishability-based security definition, but then present a counterexample showing that this notion of security is generally inadequate: a scheme that is intuitively insecure, but can be proven IND-CPA-secure. A notion of semantic security using black-box simulators is then proposed to address this problem. The paper concludes with a series of feasibility results. Most notably, BSW show that semantically secure schemes do not exist even for simple functionalities such as IBE. This result hinges on the adversary's capability to perform adaptive token-extraction queries. Nevertheless, in this work we show that the BSW definition is too *weak* in the sense that it also fails to exclude some intuitively insecure schemes. The problem is that the ideal-world simulator controls the generation of the global parameters for the FE scheme. We show that this renders the simulator unreasonably more powerful than the real-world adversary, as it can retain trapdoor information that permits recovering information from images $f(\mathsf{m})$ that is hidden in the real world.

Independently, O'Neill [14] proposed an alternative definitional approach to FE schemes for general functionalities. The author presents alternative syntax, correctness, and indistinguishability-based security notions that are conceptually close to those in [5], but proposes a significantly different semantic security definition. The paper then discusses the feasibility of achieving semantic security, by first presenting a separation to the indistinguishability notion, and then introducing a simple property for supported functionalities, called *preimage samplability*, under which the two notions are equivalent for non-adaptive token-extraction attacks. The fact that functionalities such as IBE and inner-product encryption [10] are shown to be preimage samplable provide positive results for semantically secure FE schemes for such functionalities.

The semantic security model proposed by O'Neill does not suffer from the same problem we identified for the BSW definition. Indeed, the ideal-world simulator in O'Neill's definition must work with honestly sampled global parameters. Nevertheless, we present other counterexamples for which the intuitive notion of security is not at all clear, but which can be proven secure under O'Neill's model. The crux of the matter here is that information is leaked via tokens, rather than by the ciphertext, which raises the question of whether such a scheme should be rejected by a semantic security

definition. However, one can also argue that a security definition for FE should reject schemes that fail to preserve the security properties of the supported functionalities. Finally, as acknowledged in [14], O'Neill's notion of semantic security does not suitably deal with adaptive token-extraction attacks.

RECENT WORK. In independent work, Bellare and O'Neill [4] proposed syntax and security definitions for FE that go in the same direction as those proposed here. Their notions of correctness and SS3 security are similar in spirit to ours, and their resamplability notion is akin to our notion of restricted preimage samplability. Gorbunov, Vaikuntanathan, and Wee [9] have also presented a new semantic security model. Similarly to our definition, their simulator does not control the generation of the global parameters. However, the simulators considered there are black-box and follow a specific simulation structure. We leave a detailed comparison to future work.

MAIN CONTRIBUTIONS. We now detail our main contributions.

Syntax. We start in Section 2 by tailoring the syntax and correctness definitions of functional encryption so as to capture the standard definitions for primitives such as IBE, ABE, PE, etc., as particular cases. This was not strictly the case with previous approaches. In particular, we identify a notion of *full correctness*, which maps to the notions adopted in [14,5], and imposes that the decryption operation explicitly returns a failure symbol when the functionality is undefined for a particular input value.

Indistinguishability. We modify the notion of *intentional leakage* [5] to the slightly different concept of *potential leakage* in Section 3. This allows us to dissociate syntactic aspects (e.g., we do not need to include a special empty token in the syntax of the primitive) from the security aspects of an FE scheme. Potential leakage captures the general restrictions that must be in place to ensure that various security models exclude attacks on functional encryption schemes based on information that the scheme is not designed to conceal, e.g., the length of messages or the identity of the receivers. Through this notion we are able to define indistinguishability-based security as a natural generalization of the equivalent notions for standard primitives, and automatically get feasibility results that do not require transforming the original schemes.

Semantic security. Having identified a number of limitations of the semantic security models proposed by BSW and O'Neill (Section 4), in Section 5 we propose a notion of semantic security that incorporates features from the definitions by BSW, and also by O'Neill. Again, our goal is to faithfully generalize the definitions of semantic security for primitives like PKE [7] and IBE [1]. We observe that full adaptive token extraction models are not typically considered in such schemes, and so we propose a *restricted adaptive* token-extraction attack model. The restriction we impose intuitively prevents an attacker from obtaining decryption tokens that would allow it to trivially corrupt an encrypted ciphertext a posteriori, in the style of non-committing encryption [13]. Put another way, our semantic security definition permits specifying the message distributions from which encrypted messages may be drawn, along with matching restrictions on the tokens that can be issued by the TA a posteriori, in order to provide FE security guarantees in a more flexible usage scenario. Using this strategy we circumvent impossibility results for unrestricted token extractions [5]. Finally, we show that our semantic security definition displays a desirable composition property: security against

single-message attacks implies security against multi-message attacks, even under restricted adaptive token-extraction attacks, thereby allowing us to present our results in the simpler single-message scenario.

Setup security. Our definition of semantic security preserves the resilience of O'Neill's definition in rejecting schemes that leak information to the adversary via the ciphertext. However, like all previous definitions, it does not provide any safeguards against leakage via decryption tokens or the master secret key. We therefore go on to introduce a new notion of *setup security* which enforces that tokens (or more strongly the setup procedure of the system) do not release any privileged information that might enable token holders or the trusted authority to obtain information which would otherwise be hidden by images values (Section 6). We show that setup security excludes all intuitively insecure schemes that we consider in the paper, while being inclusive enough to enable positive results for existing FE schemes. More precisely, we show that functionalities admitting a *conditional preimage sampling* procedure have an intrinsically secure setup procedure. We show PKE and IBE schemes, and more generally FE schemes supporting *all-or-nothing* functionalities are conditionally preimage samplable.

Adaptive equivalence. In Section 7 we present some positive feasibility results for our proposed notion of semantic security. There we extend O'Neill's results for non-adaptive token-extraction attacks and propose a variant of preimage samplability (PS) that enables us to obtain an equivalence between IND-CPA-secure and semantically secure functional encryption under *restricted* adaptive token extraction. Moreover, our requirement is *weaker* than that of O'Neill if we are only interested in the non-adaptive token extraction scenario. Finally, we show that conditional preimage samplability (as defined to establish setup security) also implies our stronger notion of preimage samplability. We immediately get that indistinguishability-based security is equivalent to semantic security under restricted adaptive token-extraction attacks for all-or-nothing functionalities. This gives a wide range of positive results for (multi-message) semantically secure functional encryption that extends previous known results.

Inner products. We conclude the paper in Section 8 by presenting negative results for inner-product encryption (IPE). These results bring a twist to our extension of O'Neill's work: it is *not* the case that all the equivalences between semantic security and indistinguishability established by O'Neill for non-adaptive token extractions carry over to our restricted adaptive scenario. Concretely, we show that although inner-product encryption is proven by O'Neill to satisfy the preimage sampling property [14], this functionality is provably *not* preimage samplable under the more restrictive PS notion that we introduce: for certain parameterizations of the inner-product functionality, a successful preimage sampler can be used to break the Small Integer Solution (SIS) problem. This leaves open the question of proving the semantic security of existing inner-product encryption schemes under restricted adaptive token-extraction attacks.

2 Functional Encryption Syntax and Correctness

NOTATION. We start by introducing notation. We denote assigning y to x by $x \leftarrow y$ and use $x \leftarrow\!\!\$\ X$ for sampling x from set X uniformly at random. If $\mathcal{A}$ is a probabilistic algorithm, $y \leftarrow\!\!\$\ \mathcal{A}(x_1, \ldots, x_n)$ denotes running $\mathcal{A}$ on $x_1, \ldots, x_n$ with random coins

chosen uniformly at random, assigning the result to y. We use ":" for appending to a list. We denote by $[X]$ the support of random variable X. We say $\nu(\lambda)$ is negligible if $|\nu(\lambda)| \in \lambda^{-\omega(1)}$. We use bold font to denote a vector, and abuse notation when applying a function to each element of a vector, writing $f(\mathbf{m})$. We use $[X]_i$ for the ith component of X, and $[X]_i^j$ for the ith to jth components. We denote the ℓ_2 norm of $\boldsymbol{x}$ by $\|\boldsymbol{x}\|_2$.

SYNTAX. We now define the syntax for a functional encryption (FE) scheme, where the function space may, in general, *depend* on the public parameters of the system; see the discussion below. Such a scheme is specified by four PPT algorithms as follows.

1. $\mathsf{Setup}(1^\lambda)$: This is the setup algorithm. On input a security parameter 1^λ, it outputs a master secret key Msk and a master public key Mpk. Implicitly included in Mpk are a function/circuit space description FunSp and a message space MsgSp. The function space FunSp consists of circuit descriptions $f : \mathsf{MsgSp} \to \mathsf{MsgSp} \cup \{\perp\}$.
2. $\mathsf{TKGen}(f, \mathsf{Msk})$: This is the token-generation algorithm. On input a function f and a master secret key Msk, it outputs a token TK for f.
3. $\mathsf{Enc}(\mathsf{m}, \mathsf{Mpk})$: This is the encryption algorithm. On input a message m and the master public key Mpk, it outputs a ciphertext c.
4. $\mathsf{Dec}(\mathsf{c}, \mathsf{TK})$: This is the deterministic decryption algorithm. On input a ciphertext c and a token TK, it outputs a message $\mathsf{m} \in \mathsf{MsgSp}$ or the special failure symbol $\perp$.

CORRECTNESS. The special symbol $\perp$ in the co-domain of functions accounts for functions that may be undefined on parts of their domain, or for which we do not expect the cryptosystem to behave correctly. We call an FE scheme *correct* if, for all $\lambda \in \mathbb{N}$, all $(\mathsf{Mpk}, \mathsf{Msk}) \in [\mathsf{Setup}(1^\lambda)]$, all $\mathsf{m} \in \mathsf{MsgSp}(\mathsf{Mpk})$, all $\mathsf{c} \in [\mathsf{Enc}(\mathsf{m}, \mathsf{Mpk})]$, all $f \in \mathsf{FunSp}(\mathsf{Mpk})$, and all $\mathsf{TK} \in [\mathsf{TKGen}(f, \mathsf{Msk})]$, we have that $f(\mathsf{m}) \neq \perp \implies \mathsf{Dec}(\mathsf{c}, \mathsf{TK}) = f(\mathsf{m})$. We call an FE scheme *fully correct* when the $f(\mathsf{m}) \neq \perp$ restriction is removed, i.e., when the decryption algorithm must return $\perp$ whenever $f(\mathsf{m}) = \perp$.

In the full version [2] we show that a number of standard cryptographic primitives can be seen as special cases of the FE syntax and correctness conditions defined above.

COMPARISON WITH THE PREVIOUS DEFINITIONS. There are two differences to the definition in [14]. First, O'Neill stipulates that the function space is indexed by the security parameter, yet it is fixed and independent of the setup algorithm. However, for a number of primitives such as inner-product encryption and attribute-based encryption, the function space may depend on the parameters generated by the setup algorithm. Second, we do not require correctness to hold when the function evaluates to $\perp$. As we will see, this weaker correctness notion allows us to see standard PKE, IBE, and other schemes as particular cases of functional encryption. Strictly speaking, this was not possible with the definition in [14]. A correct FE scheme according to [14] can be written as a fully correct scheme in our syntax, and vice versa.

One presentational difference between the definition in [5] and that of ours is that we treat functions explicitly whereas BSW define a general functionality $F(K, \cdot)$ indexed by keys K. This difference is inconsequential as the description of a function can be interpreted as a key. In both definitions various spaces can depend on the public parameters. Furthermore, we do not rely on a special empty key to model leakage of side information such as plaintext length. We will deal with this issue when defining

the security of an FE scheme later on. Finally, our definition of correctness differs from BSW in the same way as it differs from that of O'Neill.

3 Indistinguishability

We define an indistinguishability-based notion of security tailored to capture different gradings of security for the same functionality. The game is parameterized by a PPT relation R that defines the admissible set of challenge queries. This generalizes the restriction of choosing challenge messages with the same length in PKE. By requiring $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1)$ to hold in the challenge query, one acknowledges that challenge queries that violate this restriction *may* lead to a (trivial) break. This decouples security concerns from the correctness of the scheme. We refer to R as the *potential leakage relation*.

Definition 1 (IND-CPA **Security).** *Let game* $\mathsf{IND\text{-}CPA}_{\mathsf{FE},\mathsf{R},\mathcal{A}}$ *be as defined in Figure 1. The* IND-CPA *security of an FE scheme relative to potential leakage relation* R*, requires the advantage of any adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *to be negligible, when this is defined as*

$$\mathbf{Adv}^{\mathsf{ind\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A}}(\lambda) := 2 \cdot \Pr\left[\mathsf{IND\text{-}CPA}_{\mathsf{FE},\mathsf{R},\mathcal{A}}(\lambda) \Rightarrow \mathsf{T}\right] - 1.$$

Game $\mathsf{IND\text{-}CPA}_{\mathsf{FE},\mathsf{R},\mathcal{A}}(\lambda)$:	**oracle** $\mathbf{LR}(\mathsf{m}_0, \mathsf{m}_1)$:	**oracle** $\mathbf{Token}(f)$:
$b \leftarrow\!\!\$\ \{0,1\}$; $\mathsf{TKList} \leftarrow [\,]$	$\mathsf{c} \leftarrow\!\!\$\ \mathsf{Enc}(\mathsf{m}_b, \mathsf{Mpk})$	$\mathsf{TK} \leftarrow\!\!\$\ \mathsf{TKGen}(f, \mathsf{Msk})$
$(\mathsf{Msk}, \mathsf{Mpk}) \leftarrow\!\!\$\ \mathsf{Setup}(1^\lambda)$	Return c	$\mathsf{TKList} \leftarrow f : \mathsf{TKList}$
$b' \leftarrow\!\!\$\ \mathcal{A}^{\mathbf{O}}(\mathsf{Mpk})$		Return TK
Return $(b' = b)$		

Fig. 1. Game defining the IND-CPA security of an FE scheme. An adversary is legitimate if: 1) it calls $\mathbf{LR}$ once and with a pair $(\mathsf{m}_0, \mathsf{m}_1)$ such that $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1)$ holds; 2) for all $f \in \mathsf{TKList}$ have $f(\mathsf{m}_0) = f(\mathsf{m}_1)$; and 3) in the token non-adaptive model, it does not call $\mathbf{Token}$ after $\mathbf{LR}$.

SECURITY. Not only can we relate the syntax of functional encryption schemes to that of existing primitives but, under the appropriate potential leakage relations, we can also reduce IND-CPA security of an FE scheme to an existing primitive and vice versa. Our choices therefore lead to a notion of functional encryption scheme that is indeed a generalization of existing cryptographic primitives.

RELATION WITH O'NEILL'S DEFINITION. In [14] the implicit potential leakage relation is the equality of the message lengths, i.e., $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1) := (|\mathsf{m}_0| = |\mathsf{m}_1|)$. Although this is a natural choice, the resulting security definition fails to generalize those for IBE schemes (be it anonymous or non-anonymous). Our choice for the potential leakage relation deals with these issues in a cleaner way and is closer in spirit to that in [5]. It is straightforward to see that a feasibility result under O'Neill's definitional choices leads to a feasibility result in our setting with a fixed function space, full correctness, and with respect to the length equality relation. The converse also holds.

RELATION WITH BSW. Boneh et al. [5] define a special empty key (function) ϵ that is aimed at capturing information about encrypted messages that might be publicly

recoverable from ciphertexts (typically including the message length). However, this requirement implies that the standard syntax definitions for PKE, IBE, and other primitives do not naturally generalize to functional encryption and deviates from our goal. In BSW, for example, it is stated that an IBE should attach the message length and target identity to the ciphertext to strictly meet this requirement. We believe that our approach via relation R separates security issues from syntactic and correctness issues, while still maintaining the flexibility of the BSW definition. More formally, if the potential leakage relation is defined to be $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1) := (\epsilon(\mathsf{m}_0) = \epsilon(\mathsf{m}_1))$, queries that allow adversaries to exploit the empty token are excluded.

Given the discussion above, we can translate between feasibility results for the BSW definition and our definition. Any scheme that is secure under the BSW definition yields a fully correct and secure scheme under our definition, when one introduces the appropriate syntactic changes and adopts an adequate potential leakage relation. A conversion in the other direction implies transforming the scheme so that it explicitly leaks information through the empty token to match the restrictions imposed by R. In this case, full correctness and security in our model, yields a BSW-secure scheme.

4 Limitations of the Models by BSW and O'Neill

A closer look at the IND-CPA notion of security for FE schemes reveals that it is inadequate for general functionalities. The problem is as follows [5]. For some functionalities the restriction on the $\mathbf{LR}$ oracle imposing that $f(\mathsf{m}_0) = f(\mathsf{m}_1)$ can prevent the adversary from simultaneously extracting the token for f and launching a meaningful attack. For example, if the function is injective, any adversary extracting the token for this function will be prevented from querying anything other than $\mathsf{m}_0 = \mathsf{m}_1$ from the challenge oracle. However, in this case, the adversary will have no chance of winning.

Boneh et al. go on to turn this observation into a concrete functional encryption scheme supporting a one-way permutation that is intuitively insecure, but can be easily shown to satisfy the IND-CPA security definition. Roughly speaking, in this scheme one encrypts m under a standard PKE scheme. The token for the one-way permutation function f is the secret key for the PKE. Upon decryption, one first recovers m and then computes $f(\mathsf{m})$. The scheme is clearly correct. However, since $f(\mathsf{m})$ hides m computationally, the functional encryption scheme is not guaranteeing that the decryptor learns no more about the encrypted message than that which is leaked by $f(\mathsf{m})$. On the other hand, one can easily show that this FE scheme is IND-CPA-secure if the underlying PKE is itself IND-CPA-secure: if an adversary extracts the token for f, which is a permutation, then it is bound to calling the challenge oracle on $\mathsf{m}_0 = \mathsf{m}_1$; if it does not extract the token, then a simple reduction shows that it is attacking the PKE scheme. Boneh et al. also show that this scheme cannot be proven semantically secure, providing evidence that this is the correct notion of security for FE.

A Weakness in the BSW Model. We now follow the same approach to demonstrate an intuitively insecure scheme that can be proven BSW semantically secure. We restrict ourselves to the non-adaptive token extraction model so as not to fall within the range of impossibility results established in [5]. (This only strengthens our argument.)

Our argument also goes through for the weaker definition of semantic security that is used in [5, Definition 5] to present a (stronger) impossibility result.

Consider an FE scheme constructed from a PKE scheme and a one-way trapdoor permutation TDP. The scheme provides the expected functionality by encrypting an input message under a standard PKE, and evaluating the TDP upon decryption. The token corresponding to the TDP is simply the PKE secret key. The trapdoor for the TDP is *not* kept as part of the secret parameters, and to make the point clearer one should think of it as being "destroyed" upon generation. Consider also that the intentional leakage for this scheme is defined as $|\mathsf{m}|$. The scheme is clearly correct. Following the same reasoning as in the previous counterexample, this scheme leaks too much information to a decryptor holding a token for f: it will learn m, whereas only the image under the TDP should be leaked.

In the full version [2] we present a BSW simulator that always succeeds in simulating $\mathcal{A}$'s output, as long as the underlying PKE is $\mathsf{IND\text{-}CPA}$-secure. If the adversary extracts the token for f, then the simulator is able reconstruct a perfect simulation of the ciphertext in the real game using the trapdoor for the TDP that it (abusively) keeps in its state. On the other hand, if the adversary does not extract the token, then any adversary/distinguisher pair that distinguish the simulation can be used to break the $\mathsf{IND\text{-}CPA}$ security of the underlying PKE scheme. This counterexample can be extended to FE schemes where the function space is fixed and independent of the global parameters.

Potential Shortcomings in O'Neill's Model. There is a fundamental difference between O'Neill's definition of semantic security and that of BSW: the simulator is no longer in control of the generation of systems parameters. In return, a token-extraction oracle is provided in the ideal game. This means that the same strategy we presented above to argue for the inadequacy of the BSW definition does *not* directly apply. Nevertheless, other potential problems remain that we discuss next.

POTENTIALLY INSECURE SCHEME 1. We modify the BSW counterexample scheme as follows. The setup procedure is similar to before, except that the trapdoor for the randomly chosen permutation f is no longer destroyed but kept in the master secret key. The token-generation algorithm is modified so that the token for the TDP now also contains the trapdoor. The encryption and decryption routines are as before. This scheme can be proven secure under O'Neill's definition: although the simulator cannot generate the trapdoor information itself, this will become available once the adversary extracts the token for f. It is unclear if this scheme is intuitively insecure as the ciphertext does not leak any information beyond that leaked by images *and tokens*.

POTENTIALLY INSECURE SCHEME 2. Consider the following trivial construction of an FE scheme supporting its own encryption circuit. Take a PKE scheme and set the message space of the FE scheme to be (m, r) pairs. Take a PKE keys $(\mathsf{sk}, \mathsf{pk})$ and set the master secret key to be sk and the master public key to be pk. To functionally encrypt m re-encrypts under pk the ciphertext c resulting from $\mathsf{Enc}(\mathsf{m}, \mathsf{pk}; r)$. The decryption token is simply sk and decryption recovers and outputs c. This construction is correct and it can be shown to be semantically secure under the previous semantic security definitions. It is also unclear whether it should be classified as insecure. On the one hand, it is hard to argue that it is intuitively insecure. This is because the function is evaluated

on the sender's side *and* encrypted under a secure encryption scheme. Furthermore, the decryptor is the legitimate holder of the decryption key, and hence from its perspective there is no security property associated with the evaluated function. However, one can also consider things from the perspective of the encryptor, e.g., she might expect that token holders recover nothing but a ciphertext, but this is not the case.

5 The New Semantic Security Model

We now propose a new definition of semantic security that avoids the above problems while simultaneously achieving several other goals of interest. Our definition is designed to be strong enough to exclude the clearly intuitively insecure counterexample we presented for the BSW definition and capture adaptive token extractions, without being infeasible to achieve due to its excessive strength. Furthermore, the definition should be compatible with the standard definitional approaches for PKE and IBE schemes.

Definition 2 (Semantic Security). *Let games* $\mathsf{SS\text{-}Real}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{D}}$ *and* $\mathsf{SS\text{-}Ideal}_{\mathsf{FE},\mathsf{R},\mathcal{S},\mathcal{D}}$ *be as shown in Figure 2. The semantic security of an FE scheme relative to potential leakage relation* R *requires that for any PPT adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$*, there exists a legitimate PPT simulator* $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ *such that for all PPT distinguishers* $\mathcal{D}$ *the following advantage function is negligible.*

$$\mathbf{Adv}^{\mathsf{ss\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S},\mathcal{D}}(\lambda) := \Pr\left[\mathsf{SS\text{-}Real}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{D}}(\lambda) \Rightarrow \mathsf{T}\right] - \Pr\left[\mathsf{SS\text{-}Ideal}_{\mathsf{FE},\mathsf{R},\mathcal{S},\mathcal{D}}(\lambda) \Rightarrow \mathsf{T}\right]$$

Game $\mathsf{SS\text{-}Real}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{D}}(\lambda)$:
$\mathsf{FuncList} \leftarrow [\,]$
$(\mathsf{Msk}, \mathsf{Mpk}) \leftarrow\$\ \mathsf{Setup}(1^\lambda)$
$(\mathcal{M}, \mathsf{st}) \leftarrow\$\ \mathcal{A}_1^{\mathbf{Token}}(\mathsf{Mpk})$
$(\mathsf{m}, h, t) \leftarrow\$\ \mathcal{M}$
$\mathsf{c} \leftarrow\$\ \mathsf{Enc}(\mathsf{m}, \mathsf{Mpk})$
$v \leftarrow\$\ \mathcal{A}_2^{\mathbf{Token}}(\mathsf{c}, h, \mathsf{st})$
$\mathsf{trace} \leftarrow (\mathsf{Mpk}, \mathcal{M}, t, \mathsf{FuncList})$
Return $\mathcal{D}(\mathsf{trace}, v)$

Game $\mathsf{SS\text{-}Ideal}_{\mathsf{FE},\mathsf{R},\mathcal{S},\mathcal{D}}(\lambda)$:
$\mathsf{FuncList} \leftarrow [\,]; \mathsf{m} \leftarrow \perp$
$(\mathsf{Msk}, \mathsf{Mpk}) \leftarrow\$\ \mathsf{Setup}(1^\lambda)$
$(\mathcal{M}, \mathsf{st}) \leftarrow\$\ \mathcal{S}_1^{\mathbf{Eval}}(\mathsf{Mpk})$
$(\mathsf{m}, h, t) \leftarrow\$\ \mathcal{M}$
$\mathsf{ImgList} \leftarrow [f(\mathsf{m}) : f \in \mathsf{FuncList}]$
$v \leftarrow\$\ \mathcal{S}_2^{\mathbf{Eval}}(\mathsf{ImgList}, h, \mathsf{st})$
$\mathsf{trace} \leftarrow (\mathsf{Mpk}, \mathcal{M}, t, \mathsf{FuncList})$
Return $\mathcal{D}(\mathsf{trace}, v)$

oracle $\mathbf{Eval}(f)$:
$\mathsf{TK} \leftarrow\$\ \mathsf{TKGen}(f, \mathsf{Msk})$
$\mathsf{FuncList} \leftarrow f : \mathsf{FuncList}$
Return $(\mathsf{TK}, f(\mathsf{m}))$

oracle $\mathbf{Token}(f)$:
$\mathsf{TK} \leftarrow\$\ \mathsf{TKGen}(f, \mathsf{Msk})$
$\mathsf{FuncList} \leftarrow f : \mathsf{FuncList}$
Return TK

Fig. 2. Games defining the semantic security of an FE scheme. An adversary is legitimate if: 1) $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1)$ holds for every pair of messages in $[\mathcal{M}]_1$; 2) for all second-stage $\mathbf{Token}$ queries f, we have that $f(\mathsf{m}_0) = f(\mathsf{m}_1)$ for all $\mathsf{m}_0, \mathsf{m}_1 \in [\mathcal{M}]_1$; and 3) in the token non-adaptive model, $\mathcal{A}_2$ and $\mathcal{S}_2$ do not call $\mathbf{Token}$ and $\mathbf{Eval}$ respectively.

The intuition behind the definition is as in the previous definitional approaches: an adversary should learn no more about an encrypted message than that which is explicitly revealed by the functions associated to the decryption tokens that it obtains. To this end, we require the existence of a simulator that does not have access to the ciphertext, but only to the images of the encrypted message under the same set of functions. This simulator is bound to producing an output that essentially looks like that produced by the adversary in the real world, which implies the ciphertext indeed reveals no extra information. More in detail, the simulator must emulate $\mathcal{A}_1$'s output and produce an output v that matches the information recovered by $\mathcal{A}_2$ from the ciphertext. However,

the simulator is denied access to the ciphertext, and is bound to obtaining a set of images that matches those recovered by the real-world adversary via its $\mathbf{Token}$ oracle (this last restriction is imposed by including $\mathsf{FuncList}$ in trace). Like in the indistinguishability model, the potential leakage relation can be used to exclude trivial attacks whereby the real-world adversary would obtain information trivially leaked by the ciphertext (whereas this would not be available in the ideal world). Finally, we observe that the token-extraction queries performed by the adversary in the second stage are restricted to functions that are constant over the support of the message distribution. This allows us to generalize the feasibility results that are well known for particular instances of functional encryption, namely IBE schemes. For this reason, we call this model semantic security under *restricted adaptive token-extraction attacks*.

In the full version [2] we present a detailed justification of our definitional choices. Here we summarize the main features of our definition: 1) free simulators, 2) honest parameter generation (as in O'Neill), 3) use of general distinguishers (closer to BSW), 4) message generation via a message distribution (closer to O'Neill), 5) history information (closer to BSW), and 6) hint for distinguisher (present in both O'Neill and BSW).

COMPOSITION. Observe that the $\mathsf{IND\text{-}CPA}$ definition can be shown to compose from single to multiple $\mathbf{LR}$ queries (i.e., from a single-message to a multi-message attack scenario) using a standard hybrid argument [3]. One of the crucial features of our semantic security definition is that it also composes. Below we show that a multi-message variant of our definition where the message distribution outputs a *vector* of messages (see the full version [2] for the details) is equivalent to the definition above.

Theorem 1 (Composition). *Let* FE *be a functional encryption scheme that is semantically secure under the (single-message) definition in Figure 2. Suppose that there is a polynomial* poly *such that for any single-message adversary* $\mathcal{A}$*, there is a semantic security simulator* $\mathcal{S}[\mathcal{A}]$ *such that*

$$\mathsf{Time}_{\mathcal{S}[\mathcal{A}]}(\lambda) \leq \mathsf{Time}_{\mathcal{A}}(\lambda) + \mathsf{Time}_{\mathcal{M}}(\lambda) + \mathsf{poly}(\lambda)$$
$$\mathsf{Time}_{\mathcal{S}[\mathcal{M}]}(\lambda) \leq \mathsf{Time}_{\mathcal{M}}(\lambda) + \mathsf{poly}(\lambda)$$

where $\mathcal{M}$ *is the distribution output by* $\mathcal{A}$ *and* $\mathcal{S}[\mathcal{M}]$ *is the simulated message distribution. Then for every real-world multi-message PPT adversary* $\mathcal{A}'$*, there exist a real-world single-message PPT adversary* $\mathcal{A}$ *and a multi-message PPT simulator* $\mathcal{S}'$ *such that for any distinguisher* $\mathcal{D}'$*, there exists a distinguisher* $\mathcal{D}$ *for which*

$$\mathbf{Adv}^{\mathsf{m\text{-}ss\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A}',\mathcal{S}',\mathcal{D}'}(\lambda) \leq \mathbf{Q}(\lambda) \cdot \mathbf{Adv}^{\mathsf{ss\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S},\mathcal{D}}(\lambda)\,,$$

where $\mathcal{S}$ *is the simulator implied by the single-message semantic security and* $\mathbf{Q}(\lambda)$ *is an upper bound on the number of messages output by message distributions.*

Proof (Overview). We give an overview of the proof for the non-adaptive case here and leave the details to in the full version [2]. The proof is essentially a simulation-based hybrid argument. Consider the attack models where in the ith hybrid, for $i = 0, \ldots, q$, the adversary has access to $(q - i)$ ciphertexts and i image lists. In each step we change a ciphertext to the corresponding image list. We show that for any adversary in the ith

hybrid model, there is an adversary in the $(i+1)$st hybrid that does equally well. To see this, note that the adversary in the ith hybrid can be viewed as a *single-message* real-world adversary that also receives some extra auxiliary information consisting of ciphertexts and image lists. By the semantic security guarantees of the scheme we may replace this adversary by an equally good one that only gets the image list for the replaced ciphertext. This concludes the proof as the qth hybrid corresponds to the ideal-world multi-message semantic security game. Note that the running time of the final simulator in the ideal game, which recursively depends on the previous simulators, stays polynomial if the condition given in the theorem is satisfied. □

6 Setup Security

Similarly to O'Neill's model, our definition of semantic security fails to exclude the counterexamples from Section 4 (because the simulator also has access to decryption tokens). This raises the question of whether our model can be strengthened further so these schemes are also ruled out. One direct approach to achieve this would be to further restrict the simulator by denying it access to tokens (as well as the master secret key) in the ideal world. We present this model in the full version of this paper [2] and show that it is infeasible to achieve (essentially because a correct simulation of tokens would imply breaking the functional encryption scheme one is trying to prove semantically secure). We therefore take a different approach.

The first observation we make is that our definition accepts these counterexamples because indeed they do *not* leak information through the ciphertext. In fact, leakage is enabled by the combined information provided by *images* and decryption tokens (or more generally the master secret key). Intuitively, once a trapdoor for a TDP is provided to a token holder, the TDP circuit essentially becomes an efficiently invertible encoding function from messages onto images, offering no (intuitive) security whatsoever. From the point of view of the semantic security definition, where the aim is to exclude schemes where ciphertexts leak more information than that which is leaked by images, these counterexample schemes should therefore be considered secure.

However, it is still a reasonable security goal to expect that tokens do not help a token holder to extract information from images that would otherwise be hidden by the functionality. We therefore consider the stronger setting where the master secret key is required not to compromise the security properties of the supported functionalities. Informally, this means that even the trusted authority holding the master secret key should not be able to hold any "trapdoor" information on the functionalities supported by the functional encryption scheme.

A NEW NOTION OF SECURITY. Our approach to formalizing this security notion, which we call *setup security*, is as follows. We consider an attack scenario where an adversary is given the master secret key Msk and adaptively interacts with an evaluation oracle for the functionality in order to construct a trace that contains the master public key, a list of functions, a message distribution, a list of images, and a function func that models leaked information. This trace establishes the adversary's claim as to his ability to extract information from the images, which is specified by the leakage function. The FE scheme will then be setup-secure if a simulator given only the trace and the same

set of images provided to the adversary, i.e., without having access to the master secret key, can extract essentially the same information.

Definition 3 (Setup Security). *Let game* $\mathsf{SetSec}_{\mathsf{FE},\mathsf{R},\mathcal{A}}$ *be as shown in Figure 3. The setup security of an FE scheme relative to potential leakage relation* R *requires that for any adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$*, there exists a simulator* $\mathcal{S}$ *such that the following advantage function is negligible.*

$$\mathbf{Adv}^{\mathsf{setsec}}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S}}(\lambda) := 2 \cdot \Pr\left[\mathsf{SetSec}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S}}(\lambda) \Rightarrow \mathsf{T}\right] - 1\,.$$

Game $\mathsf{SetSec}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S}}(\lambda)$:	**oracle** $\mathbf{Eval}(f)$:
$\mathsf{FuncList} \leftarrow [\,]$; $b \leftarrow_{\$} \{0,1\}$	$\mathsf{FuncList} \leftarrow f : \mathsf{FuncList}$
$(\mathsf{Msk}, \mathsf{Mpk}) \leftarrow_{\$} \mathsf{Setup}(1^\lambda)$	Return $f(\mathsf{m})$
$(\mathcal{M}, \mathsf{st}) \leftarrow_{\$} \mathcal{A}_1(\mathsf{Msk}, \mathsf{Mpk})$	
$\mathsf{m} \leftarrow_{\$} \mathcal{M}$	
$(\mathsf{func}, v_0) \leftarrow_{\$} \mathcal{A}_2^{\mathbf{Eval}}(\mathsf{st})$	
$\mathsf{ImgList} \leftarrow [f(\mathsf{m}) : f \in \mathsf{FuncList}]$	
$\mathsf{trace} \leftarrow (\mathsf{Mpk}, \mathcal{M}, \mathsf{FuncList}, \mathsf{ImgList}, \mathsf{func})$	
$v_1 \leftarrow_{\$} \mathcal{S}(\mathsf{trace})$	
Return $(v_b = \mathsf{func}(\mathsf{m}))$	

Fig. 3. Game defining the setup security of an FE scheme. An adversary is legitimate if $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1)$ holds for every pair of messages in $[\mathcal{M}]$.

SETUP SECURITY VIA CONDITIONAL PREIMAGE SAMPLING. It is easy to see that all the potentially insecure TDP-based counterexamples that we have introduced are excluded by the setup security definition. We now show that the definition allows natural classes of functionalities to be proven setup secure. To this end, we introduce a notion of *conditional preimage samplability*. Roughly speaking, this asserts that given a message distribution $\mathcal{M}$ and a list of functions $[f_i]_{i=1}^n$, it is possible to efficiently sample from $\mathcal{M}$ when this is conditioned on a set of images $[f_i(\mathsf{m})]_{i=1}^n$ for some $\mathsf{m} \in [\mathcal{M}]$. The actual definition is slightly more complex, as we need to deal with possible adversarial adaptiveness as well as define indistinguishability of conditional distributions in a meaningful way. We leave the details to the full version [2], where we also prove the results in this section. The following theorem shows that conditional preimage samplability is a sufficient condition for setup security.

Theorem 2 (CPS $\Rightarrow$ Secure Setup). *Any functional encryption supporting a CPS functionality relative to potential leakage relation* R *is setup-secure with respect to* R*.*

CONCRETE SETUP-SECURE SCHEMES. We now look at concrete functionalities and show that setup security is already achieved by many existing functional encryption schemes. We begin by defining a broad class of functionalities where an image either entirely reveals the encrypted message, or nothing at all.

Definition 4 (All-or-Nothing Functionality). *We say a functional encryption scheme supports an all-or-nothing (ANOT) functionality if for all* $\lambda \in \mathbb{N}$*, all* $(\mathsf{Mpk}, \mathsf{Msk}) \in [\mathsf{Setup}(1^\lambda)]$*, all* $f \in \mathsf{FunSp}$*, and all* $\mathsf{m} \in \mathsf{MsgSp}$ *we have that* $f(\mathsf{m}) \in \{\mathsf{m}, \perp\}$*.*

As an example, consider predicate encryption systems [5] where the message space is partitioned into pairs $\mathsf{m} = (\mathsf{x}, \mathsf{idx})$. Here, x is a hidden payload and idx is extra information that determines which tokens can be used to recover x from a ciphertext encrypting m. More precisely, each secret key is associated with a predicate P, and the payload can be recovered whenever $P(\mathsf{idx}) = \mathsf{T}$. Formally,

$$f_P(\mathsf{x}, \mathsf{idx}) := \begin{cases} (\mathsf{x}, \mathsf{idx}) & \text{if } P(\mathsf{idx}) = \mathsf{T}\,; \\ \perp & \text{otherwise.} \end{cases}$$

Observe that we include idx in the output of the functionality when the predicate evaluates to T, thereby rendering the functionality all-or-nothing. It is easy to see that PKE and IBE schemes are examples of ANOT schemes. For PKE schemes this is obvious, since the functionality is the identity function and the index space is empty. For IBE, observe that whenever the output of the functionality is not $\perp$, the identity (i.e., the index) is also implicitly leaked by the functionality, thereby revealing the full message. Furthermore, this also includes variants of inner-product encryption, hidden vector encryption, etc., where a successful decryption operation explicitly reveals the encrypted index. Our first positive result for setup security is given by the following theorem.

Theorem 3 (ANOT $\Rightarrow$ CPS). *Any ANOT functionality is conditionally preimage samplable in* expected *polynomial time (for any potential leakage relation).*

The intuition behind the proof is as follows. Since the functionality is ANOT, there are two possible trace outcomes. In the first, the adversary queries a function which acts as the identity map on the message sampled from $\mathcal{M}$. Here the sampler can simply return the message. In the second, all the image values are $\perp$. Here the sampler will repeatedly sample a message from $\mathcal{M}$ until it maps to $\perp$ under all functions given to it in the trace. The number of retries, conditioned on a given trace, will depend on the probability that the image list is all $\perp$. However, the overall expected number of retries can be shown to be 1. In the full version [2] we discuss why the sampler we construct cannot be converted into a *strict* PPT black-box sampler by truncating the execution time. Combining Theorems 2 and 3 we obtain the following corollary.

Corollary 1. *Any FE scheme supporting an ANOT functionality has a secure setup procedure w.r.t. expected PPT simulators and arbitrary potential leakage relations.*

PUBLIC-INDEX PREDICATE ENCRYPTION. We now show that there exists a large class of FE schemes for which we can construct a strict PPT conditional preimage sampler. Intuitively, such schemes leak more information about encrypted messages which, when provided to the sampler, allows this stronger result to go through (in our framework, this is captured by the potential leakage relation.

Definition 5 (Jointly All-or-Nothing Functionality). *We say an FE scheme supports a jointly all-or-nothing (JNOT) functionality relative to potential leakage relation* R *if, for all* $\lambda \in \mathbb{N}$*, all* $(\mathsf{Msk}, \mathsf{Mpk}) \in [\mathsf{Setup}(1^\lambda)]$*, all subsets* $\mathcal{F} \subseteq \mathsf{FunSp}$*, all message distributions* $\mathcal{M}$ *where* $\mathsf{R}(\mathsf{m}_0, \mathsf{m}_1) = \mathsf{T}$ *for all* $\mathsf{m}_0, \mathsf{m}_1 \in [\mathcal{M}]$*, we have that*

$$\forall \mathsf{m} \in [\mathcal{M}],\ \exists f \in \mathcal{F},\ f(\mathsf{m}) = \mathsf{m} \quad \vee \quad \forall \mathsf{m} \in [\mathcal{M}],\ \forall f \in \mathcal{F},\ f(\mathsf{m}) = \perp\ .$$

In this definition the all-or-nothing property is no longer formulated over a class of admissible message distributions defined by the potential leakage relation. Concretely, R constrains the support of the message distribution $\mathcal{M}$ in such a way that, for any subset of functions $\mathcal{F}$ extracted from the function space, the list of images will be guaranteed to, either totally reveal m, or to information theoretically preserve the entropy of the message distribution. We now show that this definition is satisfied by a large class of all-or-nothing functionalities, corresponding to predicate encryption systems with *public index* [5]. For such schemes, no claim about hiding idx is made. This means that their security is analyzed with respect to the special potential leakage relation

$$\mathsf{R}^*((\mathsf{x}_0, \mathsf{idx}_0), (\mathsf{x}_1, \mathsf{idx}_1)) := (|\mathsf{x}_0| = |\mathsf{x}_1| \ \wedge \ \mathsf{idx}_0 = \mathsf{idx}_1) \ .$$

The fact that message distributions are now restricted by R^* yields the following result.

Theorem 4. *All predicate encryption systems are JNOT with respect to* R^*.

The previous result includes primitives such as PKE, (non-anonymous) IBE, non-attribute-hiding ABE, and inner-product encryption that *reveals* the index in the ciphertext. We now state the final result of this section.

Theorem 5 (JNOT $\Rightarrow$ CPS). *Take an FE scheme supporting a JNOT functionality with respect to potential leakage relation* R. *Then this scheme is conditionally preimage samplable (in strict polynomial time) with respect to* R.

The intuition behind the proof of this theorem is exactly the same as in Theorem 3. The difference to all-or-nothing functionalities is that the JNOT property guarantees that, either the sampler gets the challenge message in the image list, or sampling a message from the message distribution yields a valid result. We obtain the following corollary.

Corollary 2. *All (public-index) predicate encryption systems are setup-secure with respect to potential leakage relation* R^*.

7 Preimage Samplability

Despite the shortcomings of indistinguishability models highlighted in [14,5], O'Neill shows that, for certain classes of functionalities, indistinguishability-based security is no less adequate than his proposed notion of semantic security. Indeed, it is shown in [14] that if an FE scheme is *preimage samplable* (see the full version [2]) then, in the non-adaptive token-extraction attack scenario, indistinguishability and semantic security are equivalent. Furthermore, functionalities such as those for IBE and inner-product encryption are shown to be preimage samplable. In light of the new syntactical and definitional approach introduced above, we propose a modified definition of preimage samplability and show that a similar result holds. Our definition, however, permits extending the equivalence result to the *multi-message and restricted adaptive* token extraction model, and hence generalizes known results for, e.g., IBE schemes, in this area. This is an important extension, as (restricted) adaptive extraction of secret keys is the standard attack model for all predicate encryption systems.

Definition 6 ((Un)Restricted Preimage Samplability). *Let game* $\mathsf{PS}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathsf{Samp},\mathsf{mode}}$ *be as defined in Figure 4. We call an FE (un)restricted preimage samplable for the potential leakage relation* R *if, for any algorithm* $\mathcal{A}$ *there exists a sampling algorithm* Samp *such that the following advantage function is negligible.*

$$\mathbf{Adv}^{\text{mode-ps}}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathsf{Samp}}(\lambda) := \Pr\left[\mathsf{PS}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathsf{Samp},\mathsf{mode}}(\lambda) \Rightarrow \mathsf{F}\right] .$$

Game $\mathsf{PS}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathsf{Samp},\mathsf{mode}}(\lambda)$:
$(\mathsf{Msk},\mathsf{Mpk}) \leftarrow\!\!\$\ \mathsf{Setup}(1^\lambda)$
$(\mathcal{M},[f_j]_{j=1}^n) \leftarrow\!\!\$\ \mathcal{A}(\mathsf{Msk},\mathsf{Mpk})$
$\mathsf{m}_0 \leftarrow\!\!\$\ \mathcal{M}$
$\mathsf{m}_1 \leftarrow\!\!\$\ \mathsf{Samp}(\mathcal{M},[(f_j,f_j(\mathsf{m}_0))]_{j=1}^n,\mathsf{Mpk})$
If $(\exists j\ :\ f_j(\mathsf{m}_0) \neq f_j(\mathsf{m}_1))$ Return F
If $\neg\mathsf{R}(\mathsf{m}_0,\mathsf{m}_1)$ Return F
If $(\mathsf{mode} = \mathsf{res} \wedge \mathsf{m}_1 \notin [\mathcal{M}])$ Return F
Return T

Fig. 4. Game defining (un)restricted preimage samplability, for mode $\mathsf{mode} \in \{\mathsf{res},\mathsf{unres}\}$. $\mathcal{A}$ is legitimate if $\mathsf{R}(\mathsf{m}_0,\mathsf{m}_1)$ holds for all $\mathsf{m}_0,\mathsf{m}_1 \in [\mathcal{M}]$.

COMPARISON WITH O'NEILL PS DEFINITION. Our definition differs from that in [14] in several aspects. First, the adversary now has access to $(\mathsf{Msk},\mathsf{Mpk})$ rather than 1^λ only. Access to Mpk is consistent with our syntax of FE schemes, which permits generation of function space together with the master public key. Access to Msk in needed when arguing that the actions of some IND-CPA adversary contradict preimage samplability. More precisely, the adversary may use information dependent on the Msk (i.e., decryption tokens) to come up with a non-samplable message. This issue seems to have been overlooked in [14]. Second, the definition is parameterized by a potential leakage relation R. This ensures that the sampler obtains as much information about the challenge message as a real-world semantic security adversary. Technically, this allows the equivalence proof to go through (for the non-adaptive case) for a larger class of functional encryption schemes than those covered by O'Neill. For example, our results cover those schemes that can be captured using our syntactic conventions, but not under those in [14]. (A simple example of this is standard (non-anonymous) IBE.) Finally, the adversary now outputs a message distribution rather than a single message. The unrestricted sampler, similar to O'Neill's, is only bound to producing a message that collides with m_0 on all functions. The (stronger) restricted sampler is bound to return an m_1 that is in the support of $\mathcal{M}$. As we shall see, this is necessary to enable extending the equivalence result to the *restricted adaptive* token extraction setting. For the unrestricted case this condition is dropped, and we end up with a definition which is implied by (and hence *weaker* than) O'Neill PS definition (the $\mathcal{M}$ and $\mathcal{A}$ can be merged). Consequently, all positive feasibility results in [14] carry over to our setting.

The following theorem, proven in the full version of this paper [2], establishes equivalence between our two notions of FE security for restricted preimage samplable schemes and restricted adaptive token extraction scenarios.

Theorem 6 (Equivalence under PS). *Fix potential leakage relation* R. *For every adversary* $\mathcal{A}$ *against the* IND-CPA *security of scheme* FE*, there exist a (single-message)* SS-Real *adversary* $\mathcal{B}$ *and a distinguisher* $\mathcal{D}$ *such that for any simulator* $\mathcal{S}$

$$\mathbf{Adv}^{\mathsf{ind\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A}}(\lambda) \leq 2 \cdot \mathbf{Adv}^{\mathsf{ss\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{B},\mathcal{S},\mathcal{D}}(\lambda)\,.$$

Furthermore, for every single-message SS-Real *adversary* $\mathcal{A}$*, there is a* PS *adversary* $\mathcal{C}$ *with sampler* Samp*, and a* SS-Ideal *simulator* $\mathcal{S}$ *such that for every distinguisher* $\mathcal{D}$ *there is an* IND-CPA *adversary* $\mathcal{B}$ *with*

$$\mathbf{Adv}^{\mathsf{ss\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{A},\mathcal{S},\mathcal{D}}(\lambda) \leq \mathbf{Adv}^{\mathsf{ind\text{-}cpa}}_{\mathsf{FE},\mathsf{R},\mathcal{B}}(\lambda) + \mathbf{Adv}^{\mathsf{res\text{-}ps}}_{\mathsf{FE},\mathsf{R},\mathcal{C},\mathsf{Samp}}(\lambda)\,.$$

The running time of $\mathcal{S}$ *in the ideal world is that of running* $\mathcal{A}$ *in the real world plus the running time of* Samp.

The guarantee on the running time of the simulator allows us to obtain semantic security in the multi-message scenario from single-message indistinguishability via Theorem 1, provided that the running time of the sampler is independent of the running time of the adversary. This is indeed the case in our feasibility results below.

REMARK. The above result can be extended to a setting where samplers, adversaries and simulators may execute in expected polynomial time. More precisely, for expected PPT preimage samplers, one can prove that IND-CPA security with respect to expected PPT adversaries is equivalent to semantic security when both the real-world adversary and the ideal-world simulator may run in expected polynomial time.

FEASIBILITY. We conclude this section with a discussion of the feasibility results we obtain with the new definition of preimage samplability. On the negative side, it is easy to see that no FE scheme supporting a one-way function can be preimage samplable with respect to O'Neill's or our definition. On the positive side, and on top of all of O'Neill's feasibility results for the non-adaptive token extraction scenario, the following theorem yields feasibility results for restricted preimage samplability for a large class of functionalities, which in turn immediately yield positive feasibility results for semantically secure functional encryption under restricted adaptive token extraction scenarios.

Theorem 7 (CPS $\Rightarrow$ PS). *Any conditionally preimage samplable functional encryption scheme is also restricted preimage samplable.*

Combining this theorem (proved in the full version [2]) with the previous results we get that any IND-CPA-secure FE scheme supporting an ANOT functionality is semantically secure under restricted adaptive token-extraction attacks and also enjoys setup security, both with respect to expected PPT simulators. For the special cases where the potential leakage relation allows us to construct a strict PPT preimage sampler (e.g., PKE, IBE, and other predicate encryption schemes that explicitly leak the index) the result holds for strict PPT simulators. Furthermore, since the sampler executes $\mathcal{M}$ once, this allows us to extend the implication from single-message IND-CPA security to multi-message semantic security under restricted adaptive token-extraction attacks.

8 Inner-Product Encryption

Inner-product encryption (IPE) [10] is a form of functional encryption where the index space corresponds to vectors $\boldsymbol{x}$ in $\mathbb{Z}_q^m$, and each secret key is also associated with a vector $\boldsymbol{y}$ in $\mathbb{Z}_q^m$. The associated predicate is given by

$$P_{\boldsymbol{y}}(\boldsymbol{x}) := \begin{cases} \mathsf{T} & \text{if } \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \mod q\,; \\ \mathsf{F} & \text{otherwise.} \end{cases}$$

Without loss of generality, we will concentrate on the predicate-only version of inner-product encryption, where the payload is empty and the functionality is $f_{\boldsymbol{y}}(\mathsf{m}) = P_{\boldsymbol{y}}(\boldsymbol{x})$. Note that IPE is *not* an all-or-nothing functionality, since upon successful decryption one does not learn $\boldsymbol{x}$.

Our goal is to show that inner-product encryption is not restricted preimage samplable. To this end, we will rely on well-established intractable problems related to finding short solutions to linear equations. More precisely, we will be relying on the Small Integer Solution (SIS) problem and a decisional variant of it that we call DSIS [12,6,11] (see the full version [2] for the details). We will show that, for certain parameters q, m in the inner-product functionality, no restricted preimage sampler can be successful against the PS adversary $\mathcal{A}$ shown in Figure 5. This adversary is parameterized by four values n, m, q, and d, which we assume to be polynomial in the security parameter. This guarantees that the algorithm runs in PPT.

Algorithm $\mathcal{A}_{q,n,m,d}(\mathsf{Msk}, \mathsf{Mpk})$:
For i from 1 to n do
  $\boldsymbol{y}_i \leftarrow\!\!\$\ \mathbb{Z}_q^m$
Set $\mathcal{M}$ to uniform on $\mathrm{B}(d)^m$
Return $(\mathcal{M}, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$

Fig. 5. PS adversary for the inner-product encryption

The formal statement of our result is as follows.

Theorem 8 (IPE Is Not Restricted PS). *Let $\mathcal{A}$ be the PS adversary in Figure 5. Then for any PPT sampler* Samp*, there exist PPT adversaries $\mathcal{B}_1$ and $\mathcal{B}_2$ such that*

$$1 - \mathbf{Adv}^{\mathsf{res\text{-}ps}}_{\mathsf{FE},\mathsf{R},\mathcal{A}_{q,n,m,d},\mathsf{Samp}}(\lambda) \leq \mathbf{Adv}^{(q,m,n,d)\text{-}\mathsf{dsis}}_{\mathcal{B}_1}(\lambda) + \mathbf{Adv}^{(q',m',n',\beta)\text{-}\mathsf{sis}}_{\mathcal{B}_2}(\lambda) + \nu(\lambda)\,,$$

where $d = q^{n/m}$, $q' = q$, $m' = m$, $n' = n/q - \sqrt{n/q}\log(n/q)$, $\beta = d\sqrt{m}$, and $\nu(\lambda)$ is a negligible function depending on q and n.

We note that this is a stronger result than what we need as it establishes that the sampler will fail with *overwhelming* probability. We leave the details of the proof to the full version [2], where we also briefly discuss how to extend the theorem to large values of q, and give a high-level overview here.

The main idea behind the proof is that a successful sampler should match the zero values in the image list it receives, while being restricted to outputting solutions in the

support of the message distribution, which consists of *small* vectors. In other words, the sampler is solving a system of linear equations with a small solution. This allows us to establish a connection with the SIS problem. Despite this, in order to solve a SIS problem instance using the sampler, we need to make sure that the sampler is forced to match sufficiently many zeros. (Note it cannot be the case that sampling a random message leads to only zero image values as otherwise we can preimage sample by repeated sampling as before.) Hence enough zero and nonzero values must be present in the image list. We achieve this by making sure the adversary $\mathcal{A}$ returns more vectors $\boldsymbol{y}_i$ than the SIS dimension n'. But now there is a problem as we do not know the image values for the newly generated vectors. This is where we appeal to the DSIS problem and simply assume these values are random: any change in the sampler's success probability would translate to a DSIS break. We are now in a position where we can reduce to the SIS problem. We assign the rows of the SIS matrix to (some of) the zeros in the randomly generated values, and assign the newly generated rows to the remaining ones. A successful sampler for this set of images would also solve the SIS problem instance.

Acknowledgements. We would like to thank Daniel Cabarcas, Angelo De Caro, Gottfried Herold, Vincenzo Iovino, Vadim Lyubashevsky, and Giuseppe Persiano for helpful discussions. Pooya Farshim is supported by grant Fi 940/4-1 of the German Research Foundation (DFG). This work was partly done while Manuel Barbosa was visiting École Normale Supérieure – Paris. This work is part-financed by National Funds through the FCT - Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) within project ENIAC/2224/2009 and by ENIAC Joint Undertaking under grant agreement number 120224.

References

1. Attrapadung, N., Cui, Y., Galindo, D., Hanaoka, G., Hasuo, I., Imai, H., Matsuura, K., Yang, P., Zhang, R.: Relations Among Notions of Security for Identity Based Encryption Schemes. In: Correa, J.R., Hevia, A., Kiwi, M. (eds.) LATIN 2006. LNCS, vol. 3887, pp. 130–141. Springer, Heidelberg (2006)
2. Barbosa, M., Farshim, P.: On the semantic security of functional encryption schemes. Cryptology ePrint Archive, Report 2012/474 (2012) Full version of this paper
3. Bellare, M., Desai, A., Pointcheval, D., Rogaway, P.: Relations among Notions of Security for Public-Key Encryption Schemes. In: Krawczyk, H. (ed.) CRYPTO 1998. LNCS, vol. 1462, pp. 26–45. Springer, Heidelberg (1998)
4. Bellare, M., O'Neill, A.: Semantically-secure functional encryption: possibility results, impossibility results and the quest for a general definition. Cryptology ePrint Archive, Report 2012/515 (2012)
5. Boneh, D., Sahai, A., Waters, B.: Functional Encryption: Definitions and Challenges. In: Ishai, Y. (ed.) TCC 2011. LNCS, vol. 6597, pp. 253–273. Springer, Heidelberg (2011)
6. Gentry, C., Peikert, C., Vaikuntanathan, V.: Trapdoors for hard lattices and new cryptographic constructions. In: Dwork, C. (ed.) STOC 2008, pp. 197–206. ACM (2008)
7. Goldreich, O.: The Foundations of Cryptography. Basic Applications, vol. 2. Cambridge University Press (2004)
8. Goldwasser, S., Micali, S.: Probabilistic encryption. Journal of Computer and System Sciences 28(2), 270–299 (1984)

9. Gorbunov, S., Vaikuntanathan, V., Wee, H.: Functional Encryption with Bounded Collusions via Multi-party Computation. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 162–179. Springer, Heidelberg (2012)
10. Katz, J., Sahai, A., Waters, B.: Predicate Encryption Supporting Disjunctions, Polynomial Equations, and Inner Products. In: Smart, N.P. (ed.) EUROCRYPT 2008. LNCS, vol. 4965, pp. 146–162. Springer, Heidelberg (2008)
11. Lyubashevsky, V.: Lattice Signatures without Trapdoors. In: Pointcheval, D., Johansson, T. (eds.) EUROCRYPT 2012. LNCS, vol. 7237, pp. 738–755. Springer, Heidelberg (2012)
12. Micciancio, D., Regev, O.: Worst-case to average-case reductions based on gaussian measures. SIAM Journal on Computing 37(1), 267–302 (2007)
13. Nielsen, J.B.: Separating Random Oracle Proofs from Complexity Theoretic Proofs: The Non-committing Encryption Case. In: Yung, M. (ed.) CRYPTO 2002. LNCS, vol. 2442, pp. 111–126. Springer, Heidelberg (2002)
14. O'Neill, A.: Definitional issues in functional encryption. Cryptology ePrint Archive, Report 2010/556 (2010)
15. Waters, B.: Functional Encryption for Regular Languages. In: Safavi-Naini, R., Canetti, R. (eds.) CRYPTO 2012. LNCS, vol. 7417, pp. 218–235. Springer, Heidelberg (2012)