# Virtual Grey-Boxes Beyond Obfuscation: A Statistical Security Notion for Cryptographic Agents

Shashank Agrawal[1], Manoj Prabhakaran[2], and Ching-Hua Yu[2]

[1] University of Texas at Austin
sagrawal@cs.utexas.edu
[2] University of Illinois at Urbana-Champaign
{mmp, cyu17}@illinois.edu

**Abstract.** We extend the simulation-based definition of Virtual Grey Box (VGB) security – originally proposed for obfuscation (Bitansky and Canetti, 2010) – to a broad class of cryptographic primitives. These include functional encryption, graded encoding schemes, bi-linear maps (with über assumptions), as well as unexplored ones like homomorphic functional encryption.

Our main result is a characterization of VGB security, in all these cases, in terms of an *indistinguishability-preserving* notion of security, called $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security, formulated using an extension of the recently proposed *Cryptographic Agents* framework (Agrawal et al., 2015). We further show that this definition is equivalent to an indistinguishability based security definition that is restricted to "concentrated" distributions (wherein the outcome of any computation on encrypted data is essentially known ahead of the computation).

A result of Bitansky et al. (2014), who showed that VGB obfuscation is equivalent to strong indistinguishability obfuscation (SIO), is obtained by specializing our result to obfuscation. Our proof, while sharing various elements from the proof of Bitansky et al., is simpler and significantly more general, as it uses $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security as an intermediate notion. Our characterization also shows that the semantic security for graded encoding schemes (Pass et al. 2014), is in fact an instance of this same definition.

We also present a composition theorem for $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security. We can then recover the result of Bitansky et al. (2014) regarding the existence of VGB obfuscation for all $\mathsf{NC}^1$ circuits, simply by instantiating this composition theorem with a reduction from obfuscation of $\mathsf{NC}^1$ circuits to graded encoding schemas (Barak et al., 2014) and the assumption that there exists an $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure scheme for the graded encoding schema (Pass et al. 2014).

## 1 Introduction

Many recent advances in theoretical cryptography deal with obfuscation, multilinear maps, various forms of functional encryption and more generally, tools that enable computation on encrypted data. These tools are relatively new (compared to say, encryption, signatures and secure multi-party computation): for instance, the first formal definitions of obfuscation appeared only at the turn of the century [16,4]. As such our understanding of these tools and their security properties is relatively limited, and continues to generate steady interest within the field.

In this paper, we further push the boundaries of what we know regarding the security notions for these emerging cryptographic objects. To illustrate our findings, consider defining a new primitive, called *Homomorphic Functional Encryption* (HFE): HFE requires a private-key for encryption and decryption, but allows public homomorphic operations — for concreteness, addition — on ciphertexts, and also lets one use the private-key to generate function-keys that can be used to securely evaluate functions on ciphertexts (the function-key may reveal the function associated with it). Note that this allows a user with a collection ciphertexts $(c_1, \cdots, c_n)$ and a key for a function $f$, to evaluate $f(\sum_{i \in S} x_i)$, where $x_i$ is the plaintext of $c_i$ and $S \subseteq [n]$. We study two possible security notions for HFE, stated roughly below:

- A simulation-based security definition $s$-$\mathsf{SIM}$,[3] in which a set of ciphertexts and function-keys can be simulated by a computationally unbounded simulator which is allowed to query $f(\sum_{i \in S} x_i)$ for only polynomially many subsets $S$.
- An indistinguishability based definition $\mathsf{IND\text{-}CON}$, in which it is enough that, given a key for a function $f$, the ciphertexts for two "concentrated" distributions over plaintexts are indistinguishable. A pair of plaintext distributions $(\mathcal{D}_0, \mathcal{D}_1)$ is said to be concentrated for $f$ if there is a function $F$ such that for all $S \subseteq [n]$, $f(\sum_{i \in S} x_i) = F(S)$, with high probability over $(x_1, \cdots, x_n) \leftarrow \mathcal{D}_b$ for both $b = 0$ and $b = 1$ (i.e., the outcome is predictable just from the subset).

$\mathsf{IND\text{-}CON}$ is a fairly basic requirement: if the plaintext distribution is promised to be such that the function reveals virtually no information about the plaintexts (as the outcome of every function evaluation is known *a priori*), then the ciphertexts and function keys should hide which exact distribution the plaintexts were drawn from. On the other hand, the simulation-based definition requires the security to hold irrespective of the input distribution. The simulator needs to fool only an adversary who makes polynomially many queries, but no matter which of the *exponentially many subset queries* the adversary evaluates using the simulated ciphertexts and function keys, the outcome should match what the actual evaluation would have given. Remarkably,

*our result implies that these two definitions are equivalent to each other.*

[3] $s$ stands for statistical, indicating that the simulator is computationally unbounded.

This is a significant generalization of a similar surprising result by Bitansky et al. [9], who studied the problem of obfuscation of circuits with boolean outputs. There it was shown that virtual grey-box (VGB) obfuscation and strong-indistinguishability obfuscation (SIO) are equivalent. In this work, we abstract out the fundamental properties underlying this equivalence and show that it covers a much wider spectrum of primitives beyond obfuscation, including HFE, (function-hiding) functional encryption, graded encoding schemes (with semantic security [19]), bi-linear maps (with über assumptions similar to the ones in [2]), etc.

Our main tool for establishing this equivalence is an intermediate security definition, which we cast in the recently formulated framework of *Cryptographic Agents* [2]. The Cryptographic Agents framework unifies several disparate cryptographic *objects*, akin to how the universal composition framework [11] unifies the study of *protocols* like oblivious-transfer, commitment and zero-knowledge proofs. Perhaps more significantly, it provides a definitional framework that, unlike the universal composition framework and the constructive cryptography framework [17], is based on indistinguishability-preservation (IND-PRE). We extend this security property, as well as introduce a new "test family" (which specifies the nature of the environment in which the security property should hold) as follows:

- we introduce the notion of *statistical* indistinguishability preserving ($s$-IND-PRE) security;
- we formulate a non-interactive test family $\Gamma^*$, which provides arbitrary auxiliary information about the objects being encoded, but — being non-interactive — prevents the adversary from adaptively influencing their choice.

We show that the resulting security definition of $\Gamma^*$-$s$-IND-PRE is equivalent to both the $s$-SIM and IND-CON definitions sketched above. These two definitions are formulated to apply to all primitives in the framework: when applied to obfuscation they yield the same definitions as in the equivalence result of [9], namely VGB obfuscation and SIO, respectively, thereby recovering the main result of [9] as a corollary.

We emphasize that our result is not about a particular primitive like obfuscation or HFE, *but about the framework itself.* Thus, for any primitive which can be modeled in the Cryptographic Agents framework, this equivalence holds.[4] For example, we observe that the "semantic-security" notion for graded encoding schemes introduced by Pass et al. [19] (or more precisely, its strengthening, as used in [9]) corresponds to $\Gamma^*$-$s$-IND-PRE security, and hence is also equivalent to corresponding $s$-SIM and IND-CON security definitions.

[4] We point out that for certain primitives, like simple functional encryption and fully-homomorphic encryption, for which the number of ideal computations that a user can make — given a set of (evaluation or decryption) keys and ciphertexts — is only polynomially large, this equivalence is easier to establish. This is because, then a simulator can make *all possible ideal queries* that the user can ever make, and use plaintexts consistent with their results to generate the simulated ciphertexts.

**A New Composition Theorem for Cryptographic Agents.** Another important component in our extension of the agents framework is a composition theorem. Given that our security definition involves a computationally unbounded adversary in the ideal world, the original composition theorem of [2] breaks down. However, we present a new information-theoretic variant of the notion of reduction — statistical reduction — between two schemas, to re-establish a composition theorem. Specifically, we show that

> *a statistical reduction from a schema $\Sigma$ to another schema $\Sigma^*$ can be combined with a secure scheme for $\Sigma^*$, to obtain a secure scheme for $\Sigma$,*

where security refers to $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security.

An illustrative application of this composition theorem is to recover another result of [9] regarding the existence of VGB obfuscation for all $\mathsf{NC}^1$ circuits. Indeed, once cast in our framework, this result is natural and immediate: [5] gave (using a different terminology) a reduction from obfuscation of $\mathsf{NC}^1$ circuits to graded encoding schemas, and [19,9] put forth the assumption that there exists a $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure scheme for the graded encoding schema. Under this assumption, our composition theorem immediately yields the result that VGB obfuscation exists for all $\mathsf{NC}^1$ circuits.

**Our Contributions.** Below we summarize the contributions discussed above:

- We extend the cryptographic agents framework [2] to include the notion of statistical hiding and a new security definition called $s$-$\mathsf{IND\text{-}PRE}$. Specifically, we consider $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security, where $\Gamma^*$ is a family of computationally unbounded tests, which do not accept messages from the user. We also present two security definitions, $\mathsf{IND\text{-}CON}$ (indistinguishability for concentrated distributions) and $s$-$\mathsf{SIM}$ (statistical simulation security) for all schemas, which generalize the notions of SIO and VGB obfuscation to all schemas.
- Our main result is that all the above definitions are equivalent (for any schema). For the case of obfuscation, this result was proven in [9].
- We define a notion of *statistical reductions* and prove a *composition theorem* for $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security and statistical reductions. In particular, this can be used to reprove the existence of $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure obfuscation for all of $\mathsf{NC}^1$, assuming "strong-sampler semantically-secure" graded encoding schemes [19,9], and relying on an interpretation of a construction in [5] as a statistical reduction from the obfuscation schema to the graded encoding schema.

The above results clarify and significantly generalize the results in a small but influential collection of recent works on the foundations of security definitions for cryptographic objects [8,19,9,2]. Specifically,

> *our results generalize the notion of "Virtual Grey-Box security" beyond the realm of obfuscation.*

In particular, they help us better understand the security notions for graded encoding schemes. Also, they give concrete ways to prove VGB security for *future* constructions of homomorphic functional encryption, function-hiding functional encryption, etc.: a composition theorem that can be directly used if the construction uses VGB secure components, and an equivalence with IND-CON security, which would typically be easier to prove from scratch.

Finally, our results also enrich the nascent framework of Cryptographic Agents. We consider this an important contribution, as this framework can play a significant role in developing our understanding of the definitional aspects of emerging cryptographic primitives. Indeed, our result itself illustrates the usefulness of this framework, as it allowed us to extend a non-trivial result about obfuscation to a general result about unbounded simulation.

### 1.1 Technical Overview

We outline the definitional aspects first, and then present a high-level sketch of the proof of our main theorem ($\mathsf{IND\text{-}CON} \Leftrightarrow \Gamma^*\text{-}s\text{-}\mathsf{IND\text{-}PRE} \Leftrightarrow \Gamma^*\text{-}s\text{-}\mathsf{SIM}$), and the composition theorem.

**Security Definitions.** Cryptographic agents and $\mathsf{IND\text{-}PRE}$ security were introduced as a means to define security for a large class of modern cryptographic primitives — including obfuscation, functional encryption, fully homomorphic-encryption and graded encoding schemes — avoiding the notion of simulation [2].

A scheme $\Pi$ (consisting of two algorithms $\mathcal{O}$ and $\mathcal{E}$, analogous to the obfuscation and evaluation algorithms, in the case of obfuscation), is said to be $\mathsf{IND\text{-}PRE}$ secure for a schema $\Sigma$ (which is defined by a family of idealized "agents" to which a user will only have black-box access in an ideal world) if every *test* in the ideal world that hides a challenge bit continues to hide the challenge bit in the real world. A cryptographic primitive is fully defined by the schema $\Sigma$ as well as the test family $\Gamma$ for which the indistinguishability preservation property holds.

We extend this notion naturally to consider *statistical hiding* in the ideal world. In $s$-$\mathsf{IND\text{-}PRE}$ security, a test in $\Gamma$ is required to be hiding in the real world only if it is statistically hiding in the ideal world — i.e., hiding against computationally unbounded adversaries (who are still limited to making polynomial number of accesses to the agents uploaded by the test). Further, we introduce a sharper *quantitative notion* of $s$-$\mathsf{IND\text{-}PRE}$ security, which makes explicit the (polynomial) gap permitted between the extent of ideal world hiding and real world hiding.[5]

[5] In $\mathsf{IND\text{-}PRE}$ security as defined in [2], it is only required that a negligible distinguishing probability in the ideal world translates to a negligible distinguishing probability in the real world. The security notion here is tighter in that it requires indistinguishability to be preserved up to a polynomial loss, even if the original distinguishing probability in the ideal world is not negligible.

We also introduce a new test family denoted by $\Gamma^*$, which consists of computationally unbounded tests, which do not accept any messages from the adversary. Alternately, a test in $\Gamma^*$ can be considered as sampling a collection of agents to upload, and a string of bits to communicate to the adversary (taking only a challenge bit as input in the experiments).

Combined, the above two elements fully define $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$. Next, we turn our attention to giving two security definitions which are not of the indistinguishability-preserving genre. Firstly, $s$-$\mathsf{SIM}$ is a statistical simulation based security notion, which, on the face of it, is a stronger definition than $s$-$\mathsf{IND\text{-}PRE}$. In $s$-$\mathsf{SIM}$ security, it is required that for every real world adversary $\mathsf{Adv}$, there is an ideal world simulator $\mathcal{S}$, which has a similar distinguishing probability as $\mathsf{Adv}$ has in the real world experiment. To be a strong security guarantee, we require that the simulator cannot depend on the test (but it can depend on $\mathsf{Adv}$). We instantiate $s$-$\mathsf{SIM}$ security against the test-family $\Gamma^*$. This generalizes the notion of VGB security for obfuscation (see Proposition 1 in Section 5).

The other security definition we introduce, called $\mathsf{IND\text{-}CON}$ (for indistinguishability of concentrated distributions) generalizes the notion of SIO introduced by [9] for obfuscation, to all schemas. Here indistinguishability is required only against tests which upload agents from two distributions which are not only indistinguishable in the ideal world, but in fact "concentrated" — with high probability, the outcome of any query strategy[6] is already determined.

**Equivalence of Security Notions.** It is easy to see that $\Gamma^*$-$s$-$\mathsf{SIM} \Rightarrow \Gamma^*$-$s$-$\mathsf{IND\text{-}PRE} \Rightarrow \mathsf{IND\text{-}CON}$.[7] Our main result is a proof that the reverse implications hold as well, and hence the three notions are identical.

Our proof could be seen as a simplification and significant generalization of the proof in [9] that SIO implies VGB obfuscation. We briefly overview the proof of [9] before explaining our version. There it is shown how to construct a computationally unbounded simulator which receives access to a single circuit computing a binary function, makes only polynomially many queries to the

[6] As opposed to the case of obfuscation, for general schemas, a query can typically depend on previous queries. For example, in a graded encoding schema, it may be the case that a "zero-test" can be performed only after performing a sequence of operations on encodings provided by the test. A query-strategy is a polynomially deep (but exponentially large) tree which fully specifies a (deterministic) choice of ideal world queries based on the outcomes of the previous queries, and potentially using the agents generated by those queries.

[7] In this chain, we may insert a weaker version of $s$-$\mathsf{SIM}$, which allows the simulator to depend on the test as well as the adversary (but not on the challenge bit given to the test), between $\Gamma^*$-$s$-$\mathsf{SIM}$ and $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ security. Since all these notions turn out to be the same, in this paper we avoid defining the weaker simulation. However, for more general test families, or without the requirement of statistical security, this notion of a simulation could be of independent interest.

circuit, and learns a sufficiently accurate approximation of the circuit so that it can simulate it to the given adversary, provided that the obfuscation scheme is SIO secure. The simulator iteratively narrows down the set of possibilities for the circuit it is given access to, by making carefully chosen queries. Firstly, the simulator narrows down the possibilities to a set of circuits $R$ such that a uniform distribution over $R$ is a concentrated distribution (this is called the *concentration step* of the proof). However, the adversary may behave differently on certain circuits within this set; the computationally unbounded simulator can identify this subset $D$[8]. To determine if the circuit is from $D$ using a small number of queries, the simulator relies on SIO security: since the adversary can distinguish the obfuscation of each of the circuits in $D$ from the obfuscation of a random circuit in $R$ (with distinguishing advantage of the same sign), it follows that it can distinguish the obfuscation of a random circuit in $D$ from a random circuit in $R$. Hence, by SIO security, it must be the case that the uniform distribution over $D$ is not concentrated around the same majority outcome as $R$ is (and possibly, not concentrated at all). This is exploited to argue that a small set of queries can be found to check if the circuit is in $D$ or not (this is called the *majority-separation step*). If after making these queries, the simulator determines that the circuit is not in $D$, it can obfuscate a random circuit from $R$ and present it to the adversary. On the other hand, if it is in $D$, this allows the simulator to make significant progress, because as $D$ is not concentrated, it must be a significantly small fraction of $R$. The simulator *iterates the concentration and majority-separation steps alternately* until it determines that the circuit is not in $D$. To complete the proof, it is argued that the number of iterations (and the number of queries within each iteration) is logarithmic in the size of the space of circuits being obfuscated.

In our proofs, the simulation is required only in showing that $\Gamma^*$-$s$-IND-PRE security implies $\Gamma^*$-$s$-SIM security. Here, the simulator can rely on the "stronger" $s$-IND-PRE security guarantee, and obtain a "separating query" more directly, without relying on $R$ being concentrated: indeed, if $D$ is distinguishable from $R$ in the real world, then $s$-IND-PRE security guarantees that there is a (small-depth) query strategy that separates the two. Performing this query strategy either allows $D$ to be significantly shrunk, or allows $R$ to be significantly shrunk (since otherwise, it will not be a sufficiently separating query strategy). If $R$ shrinks, then $D$ is redefined with respect to the new $R$ (and may become as large as the new $R$). Iterating this procedure makes $D$ empty, with the number of iterations being logarithmic in the size of the space of agents.

Roughly, the above argument corresponds to the majority-separation step in the proof of [9]. An analogue of the concentration step appears in the proof that IND-CON security implies $\Gamma^*$-$s$-IND-PRE security, described below.

[8] More precisely there are two parts of $D$, corresponding to positive and negative distinguishing advantage. For simplicity, here we assume that only one such part is non-empty.

A potentially difficult part in proving IND-PRE security in general is that it requires one to show that *every* ideal-hiding test is real-hiding, and it is not clear which tests are ideal-hiding. Our proof can in fact be viewed as a characterization of tests in $\Gamma^*$ that are statistically ideal-hiding. A test in $\Gamma^*$ can be identified with a pair of distributions $\mathcal{D}_0$ and $\mathcal{D}_1$, corresponding to the collection of agents (and auxiliary information) it generates when the challenge bit is 0 and 1 respectively. For a test to be ideal hiding, the outcome of any (polynomial depth) query-strategy must have essentially the same distribution for both $\mathcal{D}_0$ and $\mathcal{D}_1$, but the distributions are not necessarily concentrated (which requires the outcome of any query strategy to be essentially deterministic). We give a simple combinatorial lemma which shows that

> *for any distribution $\mathcal{D}$ over agents and auxiliary information, there is a polynomial-depth query strategy that breaks down $\mathcal{D}$ into concentrated distributions (plus negligible mass on an unconcentrated distribution).*

The query strategy reveals which constituent concentrated distribution a collection of agents come from. Hence, if $\mathcal{D}_0$ and $\mathcal{D}_1$ are ideal-hiding, then both of them should have essentially the same distribution over concentrated distributions. Now, for each concentrated distribution, IND-CON security guarantees that the two distributions are real-hiding too.

**Simplification and Generalization.** We highlight two contributions of our result, given the prior work of [9]. Technically, it simplifies the proof by changing a nested iterative construction (used in the simulator), into two separate constructions, each with a simple iterative procedure. At a more conceptual level, seemingly technical details in the proof of [9] — namely, the concentration step and the majority-separation step — are reflected in two separate concrete concepts (namely, IND-CON $\Rightarrow$ $\Gamma^*$-$s$-IND-PRE and $\Gamma^*$-$s$-IND-PRE $\Rightarrow$ $\Gamma^*$-$s$-SIM).

But more importantly our result also ties these results to the new framework of cryptographic agents. While the development of the notions of VGB obfuscation and SIO were important contributions to our understanding of obfuscation, our result shows that their equivalence has more to do with certain structural properties of the security definition (captured in $\Gamma^*$-$s$-IND-PRE security) rather than obfuscation itself. Indeed, we show that the same security definition, applied to the graded encoding schema captures the independently developed notion of "semantic-security" for graded encoding [19].[9] More broadly, $\Gamma^*$-$s$-IND-PRE security can be used to model über assumptions for a variety of cryptographic encoding schemes (e.g., groups, groups with bi-linear pairings etc.). Our result shows that *in all these cases*, there is an equivalent simulation based security notion as well as a low-level security notion for concentrated distributions.

[9] The original notion in [19] essentially corresponds to $s$-IND-PRE security for a test family which requires the tests to be efficient. Without this requirement, the security notion is termed strong-sampler semantic-security [9].

**Composition Theorem.** In [2] a notion of reduction was defined and it was shown that IND-PRE security composes with reductions: if $\Sigma$ reduces to $\Sigma^*$, and $\Sigma^*$ has an IND-PRE secure scheme, then so does $\Sigma$. However, this composition theorem breaks down in the case of $s$-IND-PRE security, since it involves an ideal-world adversary who is computationally unbounded. However, if the reduction is a *statistical reduction* — i.e., $\Sigma$ can be information-theoretically securely constructed based on $\Sigma^*$— then we show that the composition theorem holds. Further, the composition theorem holds even if we restrict to the test family $\Gamma^*$.

A consequence of this composition theorem is that we can readily obtain the result that, if a strong-sampler semantically-secure graded encoding scheme exists, then there exists a VGB obfuscation scheme for $\mathsf{NC}^1$ circuits. We point out that in obtaining this result, we do not rely on the IND-CON security definition at all. While [9] crucially used the notion of SIO for obtaining this result, the notion of $\Gamma^*$-$s$-IND-PRE is sufficient: the proof relies on the fact that $\Gamma^*$-$s$-IND-PRE is equivalent to VGB security for obfuscation and to strong-sampler semantic security for graded encoding schemes and on the composition theorem for $\Gamma^*$-$s$-IND-PRE (as well as the existence of a statistical reduction from obfuscation for $\mathsf{NC}^1$ to graded encoding schemes).

### 1.2 Related work

A formal study of obfuscation was initiated in the works of Hada [16] and Barak et al. [4] only about a decade and a half ago. The latter proposed several notions of obfuscation: virtual black-box (VBB), differing-inputs obfuscation (diO), indistinguishability obfuscation (iO), etc., with VBB being the strongest. Further definitions appeared later [14,15,8]. In particular, Bitansky and Canetti proposed the definition of Virtual Grey-Box (VGB) obfuscation [8].

Much work has appeared on the definitional front for other primitives like functional encryption as well [10,18,7,6,3,12,1]. The recent framework of Cryptographic Agents [2] unified many of the concepts underlying the definitions of obfuscation, functional encryption and other cryptographic objects. Our results are formulated in this new framework, and hence extends to all primitives that can be expressed as cryptographic agent schemas.

Recently, Bitansky et al. [9] gave a surprising characterization of VGB obfuscation as being equivalent to a seemingly simpler definition of obfuscation, called *strong indistinguishability obfuscation* (SIO). Further, based on this, they showed that under a variant of a semantic-security assumption on graded encoding schemes (a.k.a. multi-linear maps) [19], any $\mathsf{NC}^1$ circuit can be VGB-obfuscated. Both these results can be obtained as corollaries of our result.

# 2 Preliminaries

We use $\kappa$ to denote the security parameter. For two functions $f$ and $g$, we write $f(g)$ to denote the function $f \circ g$, so that $f(g)(x) = f(g(x))$. If $X = \{X_\lambda\}_{\lambda \in \mathbb{N}}$ and $Y = \{Y_\lambda\}_{\lambda \in \mathbb{N}}$ are distribution ensembles over $\{0,1\}$, we write $X \approx Y$ if there is a negligible function $\mathsf{negl}$ such that $|\Pr[X_\lambda = 1] - \Pr[Y_\lambda = 1]| \leq \mathsf{negl}(\lambda)$.

We work with the same framework of cryptographic agents as was originally proposed by Agrawal et al. [2], except that we consider *statistical* hiding in the ideal world and focus on a new family of tests which are computationally unbounded and do not receive messages from adversaries. We summarize the salient features of the framework here, and provide further details in Appendix A for the sake of self-containment.

**Agents and Sessions.** Agents are used to model idealizations of entities like ciphertexts, keys, encodings and obfuscations. An agent is an interactive Turing Machine, derived from a family of agents all of whose programs are identical, but may have different contents in a read-only parameter tape (e.g., message in a ciphertext, the function in a functional-encryption key, or the program in an obfuscation). Agents may interact with each other (e.g., a ciphertext agent and a key agent) to produce outputs that a user can access. This is modeled by *sessions*. A session consists of a finite ordered set of agents which can interact with each other according to their programs (e.g., a ciphertext agent can send its message to a key agent), and result in updated states for the agents as well as outputs from each agent in the session. Updated state may be the same as the original state, and the outputs may be empty.

**Ideal world.** The ideal system for a schema $\Sigma = (\mathcal{P}_{\mathsf{auth}}, \mathcal{P}_{\mathsf{user}})$, where $\mathcal{P}_{\mathsf{auth}}$ and $\mathcal{P}_{\mathsf{user}}$ are agent families, consists of two parties $\mathsf{Test}$ and $\mathsf{User}$ and a fixed third party $\mathcal{B}[\Sigma]$ (for "black-box"). $\mathsf{Test}$ receives a "secret bit" $b$ as input and $\mathsf{User}$ produces an output bit $b'$. $\mathsf{Test}$ and $\mathsf{User}$ can, at any point, choose an agent and **upload** it to $\mathcal{B}[\Sigma]$. $\mathsf{Test}$ is allowed to upload agents from $\mathcal{P}_{\mathsf{test}} := \mathcal{P}_{\mathsf{auth}} \cup \mathcal{P}_{\mathsf{user}}$ and $\mathsf{User}$ agents from $\mathcal{P}_{\mathsf{user}}$. Whenever an agent is uploaded, $\mathcal{B}[\Sigma]$ sends a unique handle for that agent to $\mathsf{User}$.

A **query** is a request for session execution. At any point in time, $\mathsf{User}$ may request an execution of a session, by sending an ordered tuple of handles $(h_1, \ldots, h_t)$ along with their inputs. $\mathcal{B}[\Sigma]$ reports back the outputs from the session, and also gives new handles corresponding to the configurations of the agents when the session terminated. (Note that after a session, the old handles for the agents are not invalidated.)

We define the random variable $\text{IDEAL}\langle \mathsf{Test}(b) \mid \Sigma \mid \mathsf{User} \rangle$ to be the output of $\mathsf{User}$ in an execution of the above system, when $\mathsf{Test}$ gets $b$ as input. We write $\text{IDEAL}\langle \mathsf{Test} \mid \Sigma \mid \mathsf{User} \rangle$ in the case when the input to Test is a uniformly random bit. We also define $\text{TIME}\langle \mathsf{Test} \mid \Sigma \mid \mathsf{User} \rangle$ as the maximum number of steps taken by $\mathsf{Test}$ (with a random input), $\mathcal{B}[\Sigma]$ and $\mathsf{User}$ in total.

**Definition 1 ((Statistical) Ideal world hiding).** *A* $\mathsf{Test}$ *is* $\eta$-s-hiding w.r.t. a schema $\Sigma$ *if, for all unbounded users* $\mathsf{User}$ *who make at most* $\eta$ *queries,*

$$|\Pr[\textsc{ideal}\langle\mathsf{Test}(0) \mid \Sigma \mid \mathsf{User}\rangle = 1] - \Pr[\textsc{ideal}\langle\mathsf{Test}(1) \mid \Sigma \mid \mathsf{User}\rangle = 1]| \leq \frac{1}{\eta}.$$

**Real World.** A *cryptographic scheme* consists of programs $\mathcal{O}$ and $\mathcal{E}$, where $\mathcal{O}$ is an encoding (or objectification) procedure for agents in $\mathcal{P}_{\mathsf{test}}$ and $\mathcal{E}$ is an execution procedure. The real world execution for a scheme $(\mathcal{O}, \mathcal{E})$ consists of $\mathsf{Test}$, a user that we shall generally denote as $\mathsf{Adv}$, and the encoder $\mathcal{O}$. ($\mathcal{E}$ features as part of an honest user.) $\mathsf{Test}$ uploads agents to the encoder $\mathcal{O}$, who encodes them and sends the resulting cryptographic agents to $\mathsf{Adv}$. $(\mathcal{O}, \mathcal{E})$ are generally memory-less from one invocation to the next, except that $\mathcal{E}$ has access to a list of all objects it ever received. For certain schemes, it is important to let $\mathcal{O}$ and $\mathcal{E}$ have access to persistent keys generated during a set-up phase, which is also incorporated into the model via a public-secret key pair $(\mathsf{MPK}, \mathsf{MSK})$ (for details see Appendix A).

We define the random variable $\textsc{real}\langle\mathsf{Test}(b) \mid \mathcal{O} \mid \mathsf{Adv}\rangle$ to be the output of $\mathsf{Adv}$ in an execution of the above system, when $\mathsf{Test}$ gets $b$ as input; as before, we omit $b$ from the notation to indicate a random bit. Also, as before, $\textsc{Time}\langle\mathsf{Test} \mid \mathcal{O} \mid \mathsf{User}\rangle$ is the maximum number of steps taken by $\mathsf{Test}$ (with a random input), $\mathcal{O}$ and $\mathsf{User}$ in total.

**Definition 2 (Real world hiding).** *A* $\mathsf{Test}$ *is* $\eta$*-hiding w.r.t.* $\mathcal{O}$ *if for all adversaries* $\mathsf{Adv}$ *who run for at most* $\eta$ *time,*

$$|\Pr[\textsc{real}\langle\mathsf{Test}(0) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1] - \Pr[\textsc{real}\langle\mathsf{Test}(1) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1]| \leq \frac{1}{\eta}.$$

**Definition 3 (Admissibility of schemes).** *A cryptographic agent scheme* $\Pi = (\mathcal{O}, \mathcal{E})$ *is said to be an* admissible scheme *for a schema* $\Sigma$ *if the following conditions hold.*

- *Correctness.* $\forall$ $\mathsf{PPT}$ $\mathsf{User}$ *and* $\forall$ $\mathsf{Test}$*,*

$$\textsc{ideal}\langle\mathsf{Test} \mid \Sigma \mid \mathsf{User}\rangle \approx \textsc{real}\langle\mathsf{Test} \mid \mathcal{O} \mid \mathcal{E} \circ \mathsf{User}\rangle.$$

  *If the difference is* $0$*,* $(\mathcal{O}, \mathcal{E})$ *is said to have perfect correctness.*
- *Efficiency. There exists a polynomial* poly *such that,* $\forall$ $\mathsf{PPT}$ $\mathsf{User}$*,* $\forall$ $\mathsf{Test}$*,*

$$\textsc{Time}\langle\mathsf{Test} \mid \mathcal{O} \mid \mathcal{E} \circ \mathsf{User}\rangle \leq \mathrm{poly}(\textsc{Time}\langle\mathsf{Test} \mid \Sigma \mid \mathsf{User}\rangle, \kappa).$$

$\Gamma^*$ **test family.** This family consists of computationally unbounded tests which do not accept any messages from the user/adversary. Without loss of generality, such a test is fully characterized by a distribution over $\{0,1\}^* \times \mathcal{P}_{\mathsf{test}}^*$.[10]

[10] In proving our results, we can assume an upper-bound on the number of bits communicated by the test, as there will be a bound on the running time of an adversary that it interacts with.

The first part of a $\boldsymbol{P} \in \{0,1\}^* \times \mathcal{P}^*_{\mathsf{test}}$, which we denote as $\boldsymbol{P}_0 \in \{0,1\}^*$, is a message from test to the user/adversary; the remaining components of the vector $\boldsymbol{P}$ denote a (possibly empty) collection of agents from $\mathcal{P}_{\mathsf{test}}$.

We write $\mathcal{O}(\boldsymbol{P})$ to denote a random encoding of $\boldsymbol{P}$ which consists of $(\boldsymbol{P}_0, \mathcal{O}(\boldsymbol{P}_1), \cdots, \mathcal{O}(\boldsymbol{P}_i))$ (as well as the public-key $\mathsf{MPK}$ if $\mathcal{O}$ involves a set-up). We write $\mathsf{Adv}(\mathcal{O}(\boldsymbol{P}))$ to denote the random variable corresponding to the bit output by $\mathsf{Adv}$ when given $\mathcal{O}(\boldsymbol{P})$.

**Definition 4 (IND-PRE security).** *An admissible cryptographic agent scheme* $\Pi = (\mathcal{O}, \mathcal{E})$ *is said to be a* $p$-$\Gamma^*$-$s$-IND-PRE-secure scheme *for a schema* $\Sigma$ *if for all* $\kappa$*, all* $\mathsf{Test} \in \Gamma^*$*, and every polynomial* $\eta$*, if* $\mathsf{Test}$ *is* $p(\eta(\kappa))$*-*$s$*-hiding w.r.t.* $\Sigma$*, then it is* $\eta(\kappa)$*-hiding w.r.t.* $\mathcal{O}$*.*

*If* $\Pi = (\mathcal{O}, \mathcal{E})$ *is* $p$-$\Gamma^*$-$s$-IND-PRE*-secure for some polynomial* $p$*, then we simply refer to it as* $\Gamma^*$-$s$-IND-PRE-secure scheme*.*

We also define a simulation-based security notion in the agents framework.

**Definition 5 (Simulation-based security).** *An admissible cryptographic agent scheme* $\Pi = (\mathcal{O}, \mathcal{E})$ *is said to be a* $p$-$\Gamma^*$-$s$-SIM-secure scheme *for a schema* $\Sigma$ *if for all* $\kappa$*, all polynomials* $\ell, \eta$*, and any adversary* $\mathsf{Adv}$ *which runs in time at most* $\ell(\kappa)$*, there exists a computationally unbounded simulator* $\mathcal{S}$ *that makes at most* $p(\eta(\kappa), \ell(\kappa))$ *queries, such that for all* $\mathsf{Test} \in \Gamma^*$*,*

$$|\Pr[\text{IDEAL}\langle \mathsf{Test} \mid \Sigma \mid \mathcal{S} \rangle = 1] - \Pr[\text{REAL}\langle \mathsf{Test} \mid \mathcal{O} \mid \mathsf{Adv} \rangle = 1]| \leq \frac{1}{\eta(\kappa)}.$$

*A cryptographic agent scheme* $\Pi = (\mathcal{O}, \mathcal{E})$ *is said to be a* $\Gamma^*$-$s$-SIM-secure scheme *if it is a* $p$-$\Gamma^*$-$s$-SIM*-secure scheme for some (bivariate) polynomial* $p$*.*

We remark that one can consider a weaker notion of simulation where $\mathcal{S}$ can depend on $\mathsf{Test}$. As we shall see, for $\Gamma^*$, this weaker notion is no different from the notion defined above.

### 2.1 Concentrated distributions

Recall that in the ideal world, $\mathsf{User}$ can make queries — i.e., requests to run sessions — to $\mathcal{B}[\Sigma]$ and obtain the outcome of the session (and handles for the updated configurations of the agents involved in the session). $\mathsf{User}$ can carry this out repeatedly, and adaptively. The following definition captures this procedure (for a deterministic $\mathsf{User}$).

**Definition 6 (Query Strategy).** *A* $d$*-query-strategy is a tree of depth at most* $d$ *where each internal node* $u$ *is labeled with a query* $q_u$ *and each outgoing edge* *from $u$ is labeled with a different possible outcome of $q_u$. The* execution *of a query strategy $Q$ on a $\boldsymbol{P} \in \{0,1\}^* \times \mathcal{P}_{\mathsf{test}}^*$ is a path in this tree starting from the root node, such that an edge from node $u$, labeled with an answer* ans, *is present in the path if and only if the outcome of running a session on (the updated configurations of) $\boldsymbol{P}$ with the query $q_u$ is* ans. *The outcome of the entire execution, denoted by $\boldsymbol{P}(Q)$ is the (concatenated) outcomes of all the queries in the path. We use the convention that the first query in $Q$ is an empty query and its answer is the auxiliary information $\boldsymbol{P}_0 \in \{0,1\}^*$.*

We now define concentrated distributions over collections of agents and indistinguishability between them.

**Definition 7 (Concentrated distributions).** *A distribution ensemble $\mathcal{D}$ over $\{0,1\}^{\ell(\kappa)} \times \bigcup_{i=0}^{\ell(\kappa)} \mathcal{P}_{\mathsf{test}}^i$ is said to be $\eta$-concentrated if for all $\kappa$ there exists a function $A$ (called an answer function) which maps query strategies to answers, such that for all depth $\eta(\kappa)$ query strategy $Q$,*

$$\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}(\kappa)}[\boldsymbol{P}(Q) \neq A(Q)] \leq \frac{1}{\eta(\kappa)}.$$

*A pair of distribution ensembles $(\mathcal{D}_0, \mathcal{D}_1)$ is said to be $\eta$-concentrated if they are both $\eta$-concentrated with the same answer function.*

**Definition 8 (Indistinguishability of concentrated distributions).** *An admissible scheme $\Pi = (\mathcal{O}, \mathcal{E})$ is $q$-*IND-CON *secure for $\Sigma = (\mathcal{P}_{\mathsf{auth}}, \mathcal{P}_{\mathsf{user}})$ if for all $\kappa$, every polynomial $\eta$, and any pair of distribution ensembles $(\mathcal{D}_0, \mathcal{D}_1)$ over $\{0,1\}^{\ell(\kappa)} \times \bigcup_{i=0}^{\ell(\kappa)} \mathcal{P}_{\mathsf{test}}^i$ which are $q(\eta(\kappa))$-concentrated, we have that for any* PPT *adversary* Adv *with running time at most $\eta(\kappa)$,*

$$\left|\mathsf{Pr}_{\boldsymbol{P} \leftarrow \mathcal{D}_0(\kappa)}[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1] - \mathsf{Pr}_{\boldsymbol{P} \leftarrow \mathcal{D}_1(\kappa)}[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1]\right| \leq \frac{1}{\eta(\kappa)}.$$

*A scheme $\Pi = (\mathcal{O}, \mathcal{E})$ is* IND-CON *secure if it is $q$-*IND-CON *secure for some polynomial $q$.*

**A probability lemma.** The following is a simple lemma which can be used to relate two distributions with a small statistical difference to a single common distribution; further, the lemma allows the common distribution to avoid a subset $S$ of the sample space, provided the given distributions have low mass on it. Below, $\Delta(\cdot,\cdot)$ denotes the statistical difference between two distributions.

**Lemma 1.** *For any two probability distributions $\mathcal{A}_0$, $\mathcal{A}_1$ over the same sample space, and any subset $S$ of the sample space, there exists $\epsilon \leq \Delta(\mathcal{A}_0, \mathcal{A}_1) + \min\{\Pr_{a \leftarrow \mathcal{A}_0}[a \in S], \Pr_{a \leftarrow \mathcal{A}_1}[a \in S]\}$, a distribution $\mathcal{A}_{\overline{S}}$ over $\overline{S}$, and two distributions $\mathcal{A}'_0, \mathcal{A}'_1$ such that for each $b \in \{0,1\}$, $\mathcal{A}_b$ is equal to the distribution of $a$ in the following experiment:*

$$\alpha \sim \text{Bernoulli}(\epsilon);\ \textit{if } \alpha = 0, a \leftarrow \mathcal{A}_{\overline{S}},\ \textit{else } a \leftarrow \mathcal{A}'_b.$$

*Proof.* Given distributions $\mathcal{A}_0, \mathcal{A}_1$ over a sample space $T$ and a set $S \subseteq T$, the goal is to construct a distribution $\mathcal{A}_{\overline{S}}$ over $\overline{S}$ such that sampling according to $\mathcal{A}_0$ (resp. $\mathcal{A}_1$) is the same as sampling according to $\mathcal{A}_{\overline{S}}$ with probability $1-\epsilon$ and according to another distribution $\mathcal{A}'_0$ (resp. $\mathcal{A}'_1$) with probability $\epsilon$. Intuitively, $\mathcal{A}_{\overline{S}}$ is the "intersection" of $\mathcal{A}_0$ and $\mathcal{A}_1$ over $\overline{S}$, and $\mathcal{A}'_0$ (resp. $\mathcal{A}'_1$) is the "remaining distribution" after $\mathcal{A}_{\overline{S}}$ is cut out from $\mathcal{A}_0$ (resp. $\mathcal{A}_1$).

More formally, define weight functions $f, f_0, f_1 : T \to [0,1]$ as follows:

$$f(a) = \begin{cases} \min\{\mathcal{A}_0(a), \mathcal{A}_1(a)\} & \text{if } a \in \overline{S} \\ 0 & \text{if } a \in S \end{cases} \qquad \text{and} \qquad \begin{aligned} f_0(a) &= \mathcal{A}_0(a) - f(a) \\ f_1(a) &= \mathcal{A}_1(a) - f(a) \end{aligned}$$

where $\mathcal{A}_b(a)$ denotes the probability mass on $a$ according to the distribution $\mathcal{A}_b$. Furthermore, set $\epsilon = 1 - \sum_a f(a) = \sum_a f_0(a) = \sum_a f_1(a)$. Then, we define the distributions $\mathcal{A}_{\overline{S}}$, $\mathcal{A}'_0$, $\mathcal{A}'_1$ as follows:

$$\mathcal{A}_{\overline{S}}(a) = f(a)/(1-\epsilon), \qquad \mathcal{A}_0(a) = f_0(a)/\epsilon, \qquad \mathcal{A}_1(a) = f_1(a)/\epsilon.$$

(If $\epsilon = 1$, we let $\mathcal{A}_{\overline{S}}$ be an arbitrary probability distribution; similarly if $\epsilon = 0$, $\mathcal{A}_0, \mathcal{A}_1$ are arbitrary.) Then for $b \in \{0,1\}$, for all $a \in T$, $\mathcal{A}_b(a) = (1-\epsilon)\mathcal{A}_{\overline{S}}(a) + \epsilon\mathcal{A}'_b(a)$, as required by the lemma.

It remains to prove the claimed upper bound on $\epsilon$. Let $g(a) = \min\{\mathcal{A}_0(a), \mathcal{A}_1(a)\}$ for all $a$. Note that $\sum_a \mathcal{A}_b(a) - g(a) = \Delta(\mathcal{A}_0, \mathcal{A}_1)$ for $b \in \{0,1\}$ and $\sum_a g(a) - f(a) \leq \min\{\Pr_{a \leftarrow \mathcal{A}_0}[a \in S], \Pr_{a \leftarrow \mathcal{A}_1}[a \in S]\}$. Hence $\epsilon = \sum_a f_0(a) = \sum_a \mathcal{A}_0(a) - g(a) + \sum_a g(a) - f(a) \leq \Delta(\mathcal{A}_0, \mathcal{A}_1) + \min\{\Pr_{a \leftarrow \mathcal{A}_0}[a \in S], \Pr_{a \leftarrow \mathcal{A}_1}[a \in S]\}$. □

## 3 Equivalence of Definitions

In this section we prove our main results (Theorem 1 and Theorem 2).

**Theorem 1 (Equivalence of IND-CON and $s$-IND-PRE).** *A cryptographic agent scheme $\Pi = (\mathcal{O}, \mathcal{E})$ is a $\Gamma^*$-$s$-IND-PRE-secure scheme for a schema $\Sigma$ if and only if it is IND-CON secure for $\Sigma$.*

To prove Theorem 1, or more specifically, that IND-CON $\Rightarrow$ $\Gamma^*$-$s$-IND-PRE, we rely on the following lemma, which gives a query strategy that can be used to narrow down a distribution over agents to a concentrated distribution (except with negligible probability over the choice of the agents). As sketched in Section 1.1, this lemma gives a characterization of hiding tests in terms of concentrated distributions and is at the heart of proving Theorem 1.

Below, for a distribution $\mathcal{D}$ over agent vectors and a query strategy $Q$, $\mathcal{D}|_{Q \to \mathsf{ans}}$ denotes the distribution obtained by restricting $\mathcal{D}$ to the subset $\{\boldsymbol{P} | \boldsymbol{P}(Q) = \mathsf{ans}\}$. Below, when we say that a distribution $\mathcal{D}|_{Q \to \mathsf{ans}}$ is $\rho$-concentrated, we consider concentration against depth $\rho$ query-strategies which can optionally use the handles resulting from the query-strategy $Q$, as well as the original handles (this is relevant only for schemas with stateful agents).

**Lemma 2.** *Let $\mathcal{P}_{\mathsf{test}}$ be a set of agents with polynomially long representation. Then, for any polynomial $\rho$, there exists a polynomial $\pi$ such that for any polynomial $\eta$, any function $\varepsilon > 0$, and any distribution $\mathcal{D}$ over $\mathcal{R}^\eta = \{0,1\}^\eta \times \mathcal{P}^\eta_{\mathsf{test}}$, there is a $\pi(\eta \cdot \log \frac{1}{\varepsilon})$-query strategy $Q^\star$ such that*

$$\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}}[\mathcal{D}|_{Q^\star \to \boldsymbol{P}(Q^\star)} \textit{ not } \rho(\eta)\textit{-concentrated}] \leq \varepsilon.$$

*Proof.* The query strategy can be defined as repeatedly, conditioned on the previous queries and answers, identifying and carrying out a query strategy whose answer is not concentrated (i.e., no one answer has probability more than $1 - \rho(\eta)$) until the remaining distribution is $\rho(\eta)$-concentrated, or the budget on the number of queries (depth of the strategy) has been exhausted. We shall show that this leads to the mass in unconcentrated leaves of the query strategy tree to be at most $\varepsilon$.

More formally, consider a tree in $T$ which each node $v$ is associated with a subset $R_v \subseteq \mathcal{R}^\eta$ and (unless it is a leaf node) with a query strategy $Q_v$. The set at the root of $T$ is the entire set $\mathcal{R}^\eta$. For $R \subseteq \mathcal{R}^\eta$, let $\mathcal{D}|_R$ denote the distribution $\mathcal{D}$ restricted to the set $R$. A node $v$ in $T$ is a leaf node either if the distribution $\mathcal{D}|_{R_v}$ is $\sigma := \rho(\eta)$-concentrated or if $v$ is at a depth $\sigma$. For every internal node $v$, $Q_v$ is a query strategy of depth at most $\sigma$ such that for all $\mathsf{ans}$, $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}|_{R_v}}[\boldsymbol{P}(Q_v) = \mathsf{ans}] \leq 1 - \frac{1}{\sigma}$. Note that such a $Q_v$ exists since $\mathcal{D}|_{R_v}$ is not $\sigma$-concentrated ($v$ being an internal node). For each possible answer $\mathsf{ans}$ to $Q_v$, $v$ has a child $v_{\mathsf{ans}}$ such that $R_{v_{\mathsf{ans}}} = \{\boldsymbol{P} \in R_v \mid \boldsymbol{P}(Q_v) = \mathsf{ans}\}$.

Let $L_\ell$ be the set of all nodes at depth $\ell$ in $T$. Note that for each $v \in L_\ell$, $|R_v| \geq 1$, whereas $\sum_{v \in L_\ell} |R_v| \leq |\mathcal{R}^\eta|$. Therefore, $|L_\ell| \leq |\mathcal{R}^\eta|$. On the other hand, note that if $u$ is a child of $v$ in $T$, then $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}}[\boldsymbol{P} \in R_u \mid \boldsymbol{P} \in R_v] \leq 1 - \frac{1}{\sigma}$. Thus for all $v \in L_\ell$, $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}}[\boldsymbol{P} \in R_v] \leq (1 - \frac{1}{\sigma})^\ell$. Hence, $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}}[\boldsymbol{P} \in \bigcup_{v \in L_\ell} R_v] \leq (1 - \frac{1}{\sigma})^\ell \cdot |\mathcal{R}^\eta|$.

If we choose $\ell = \Omega(\sigma \cdot \log(|\mathcal{R}^\eta|/\varepsilon))$ then $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}}[\boldsymbol{P} \in \bigcup_{v \in L_\ell} R_v] \leq \varepsilon$. Note that $|\mathcal{R}^\eta| = \zeta^\eta$ for some polynomial $\zeta$ (determined by the size of $\mathcal{P}_{\mathsf{test}}$) and $\sigma = \mathrm{poly}(\eta)$, so that $\ell$ is polynomial in $\eta \cdot \log \frac{1}{\varepsilon}$. Our query strategy $Q^\star$ is obtained from $T$ by executing the first $\ell$ query strategies in it. The depth of $Q^\star$ is $\ell \cdot \sigma$, again a polynomial in $\eta \cdot \log \frac{1}{\varepsilon}$. □

We prove the two directions of Theorem 1 separately. Intuitively, $\mathsf{IND\text{-}CON}$ security is a "weaker" notion, and hence the first direction below is easier to see. The second direction relies on Lemma 2.

<u>$\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE} \Rightarrow \mathsf{IND\text{-}CON}$:</u> Suppose that for some polynomial $q$, $\Pi = (\mathcal{O}, \mathcal{E})$ is a $q$-$\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure scheme for a schema $\Sigma$. We shall show that $\Pi$ is $q$-$\mathsf{IND\text{-}CON}$ secure for $\Sigma$.

Let $\eta$ be a polynomial, and $(\mathcal{D}_0, \mathcal{D}_1)$ be a pair of distribution ensembles which are $q(\eta)$-concentrated. Let $A$ denote the answer function that maps depth $q(\eta)$ query strategies to answers, so that for any such query strategy $Q$, for both $b \in \{0,1\}$, we have $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_b}[\boldsymbol{P}(Q) \neq A(Q)] \leq \frac{1}{q(\eta)}$.

Consider the test $\mathsf{Test}$ which on input $b \in \{0,1\}$, uploads a sample from the distribution $\mathcal{D}_b$. Observe that $\mathsf{Test} \in \Gamma^*$. Consider any unbounded ideal-world user $\mathsf{User}$ that makes at most $q(\eta)$ queries. For each setting of the random-tape of $\mathsf{User}$, its behavior can be identified with a query strategy of depth at most $q(\eta)$. For any such strategy $Q$, irrespective of the bit $b$, with probability at least $1 - 1/q(\eta)$ $\mathsf{User}$ receives the answer $A(Q)$. Thus, for any $\mathsf{User}$ which makes at most $q(\eta)$ queries $|\Pr[\textsc{ideal}\langle \mathsf{Test}(0) \mid \Sigma \mid \mathsf{User}\rangle = 1] - \Pr[\textsc{ideal}\langle \mathsf{Test}(1) \mid \Sigma \mid \mathsf{User}\rangle = 1]| \le 1/q(\eta)$. That is, $\mathsf{Test}$ is $q(\eta)$-s-hiding w.r.t. $\Sigma$.

Then, since $\Pi$ is a $q$-$\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure scheme for $\Sigma$, we have that $\mathsf{Test}$ is $\eta$-hiding w.r.t. $\mathcal{O}$. That is, for any adversary $\mathsf{Adv}$ with running time at most $\eta$, $|\Pr[\textsc{real}\langle \mathsf{Test}(0) \mid \Sigma \mid \mathsf{User}\rangle = 1] - \Pr[\textsc{real}\langle \mathsf{Test}(1) \mid \Sigma \mid \mathsf{User}\rangle = 1]| \le 1/\eta$. But $\Pr[\textsc{real}\langle \mathsf{Test}(b) \mid \Sigma \mid \mathsf{User}\rangle = 1]$ is simply $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_b}[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1]$.

Hence, by the definition of $\mathsf{IND\text{-}CON}$ security, $\Pi$ is $q$-$\mathsf{IND\text{-}CON}$ secure for $\Sigma$.

<u>$\mathsf{IND\text{-}CON} \Rightarrow \Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$:</u> Suppose $\Pi$ is an $\mathsf{IND\text{-}CON}$ secure scheme for $\Sigma$. Then, there is a polynomial $q$ such that it is $q$-$\mathsf{IND\text{-}CON}$ secure. We shall show that $\Pi$ is $p$-$\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure, for some polynomial $p$.

Let $\mathsf{Test}$ be an arbitrary test in $\Gamma^*$, that is $\eta^*$-hiding w.r.t. $\Sigma$. We shall show that $\mathsf{Test}$ is $\eta$-hiding w.r.t. $\Pi$, where $\eta^* = p(\eta)$ (for a polynomial $p$ to be determined).

We consider the space $\mathcal{R}^\eta$ of all possible agents vector produced by tests, i.e., $\mathcal{R}^\eta = \{0,1\}^\eta \times \mathcal{P}^\eta_{\mathsf{test}}$.[11] Let $\mathcal{D}_0$ and $\mathcal{D}_1$ be the distributions over $\mathcal{R}^\eta$, produced by $\mathsf{Test}$ on input $b = 0$ and $b = 1$ respectively. Now, we apply Lemma 2 to the distribution $\mathcal{D}_0$, with $\eta$ as above, $\rho(\eta) := 2q(\eta/2)$, and (say) $\varepsilon = 2^{-\eta}$. Let $Q$ be the query strategy guaranteed by the lemma. Also, let $\mu = \rho(\eta)/2$.

Recall that each root-to-leaf path in a query strategy is labeled by a sequence of responses, $\mathsf{ans}$. We define two subsets of leaves $B$ and $C$ which correspond to answers that can potentially differentiate between $\mathcal{D}_0$ and $\mathcal{D}_1$. Let $B = \{\mathsf{ans} \mid \mathcal{D}_0|_{Q \to \mathsf{ans}}$ is not $2\mu$-concentrated$\}$. Also let $C = \{\mathsf{ans} \mid \mathcal{D}_0|_{Q \to \mathsf{ans}}$ is $2\mu$-concentrated around some answer function $A$, but $\mathcal{D}_1|_{Q \to \mathsf{ans}}$ is not $\mu$-concentrated around $A\}$. For $\mathsf{ans} \notin B \cup C$, the pair of distributions $(\mathcal{D}_0|_{Q \to \mathsf{ans}}, \mathcal{D}_1|_{Q \to \mathsf{ans}})$ is $\mu$-concentrated.

We argue, relying on the fact that $\mathsf{Test}$ is $\eta^*$-hiding, that the mass of $B \cup C$ under $\mathcal{D}_0$ is $O(\mu/\eta^*)$. Firstly, mass of $B$ under $\mathcal{D}_0$ is bounded by Lemma 2 to at most $\varepsilon$. Next, for each $\mathsf{ans} \in C$, let $A_{\mathsf{ans}}$ be the answer function that $\mathcal{D}_0|_{Q \to \mathsf{ans}}$ is $2\mu$-concentrated around. Since $\mathcal{D}_1|_{Q \to \mathsf{ans}}$ is not $\mu$-concentrated around $A_{\mathsf{ans}}$, there is some query strategy $Q_{\mathsf{ans}}$ with depth at most $\mu$, such that $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_1|_{Q \to \mathsf{ans}}}[\boldsymbol{P}(Q_{\mathsf{ans}}) \neq A_{\mathsf{ans}}(Q_{\mathsf{ans}})] > 1/\mu$. But since $\mathcal{D}_0|_{Q \to \mathsf{ans}}$ is $2\mu$-concentrated around $A_{\mathsf{ans}}$ and $Q_{\mathsf{ans}}$ has depth less than $2\mu$, $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0|_{Q \to \mathsf{ans}}}[\boldsymbol{P}(Q_{\mathsf{ans}}) \neq$

[11] Note that we truncate the auxiliary information to $\eta(\kappa)$ bits, and the number of agents uploaded by the test to $\eta(\kappa)$. This is because, to show that $\mathsf{Test}$ is $\eta$-hiding w.r.t. $\Pi$, it is enough to consider adversaries who read at most $\eta$ bits of the messages from $\mathsf{Test}$.

$A_{\mathsf{ans}}(Q_{\mathsf{ans}})] \leq 1/(2\mu)$. Now, consider a 2-phase query strategy $Q'$ that in the first phase carries out $Q$ and at the end of it, if $\mathsf{ans} \in C$ is obtained, then follows up with the query strategy $Q_{\mathsf{ans}}$. $Q'$ is of depth at most $\pi(\eta^2) + \mu$ (which we shall arrange to be less than $\eta^*$). We may write the answer $\boldsymbol{P}(Q')$ as $\mathsf{ans}_1 || \mathsf{ans}_2$, where $\mathsf{ans}_1$ and $\mathsf{ans}_2$ are the answers to the first and second phases of queries, respectively (if $\mathsf{ans}_1 \notin C$, then $\mathsf{ans}_2$ will be empty). Then,

$$\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0}[\boldsymbol{P}(Q') = \mathsf{ans}||A_{\mathsf{ans}}(Q_{\mathsf{ans}}) \text{ for } \mathsf{ans} \in C] \geq \Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0}[\boldsymbol{P}(Q) \in C] \cdot \left(1 - \frac{1}{2\mu}\right)$$

$$\begin{aligned}
&\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_1}[\boldsymbol{P}(Q') = \mathsf{ans}||A_{\mathsf{ans}}(Q_{\mathsf{ans}}) \text{ for } \mathsf{ans} \in C] \\
&\qquad\qquad < \Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_1}[\boldsymbol{P}(Q) \in C] \cdot \left(1 - \frac{1}{\mu}\right) \\
&\qquad\qquad \leq \left(\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0}[\boldsymbol{P}(Q) \in C] + 1/\eta^*\right) \cdot \left(1 - \frac{1}{\mu}\right)
\end{aligned}$$

The difference between these two probabilities is more than $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0}[\boldsymbol{P}(Q) \in C] \cdot \frac{1}{2\mu} - \frac{1}{\eta^*}$. But as the depth of $Q'$ is less than $\eta^*$ (as we ensure below), and $\mathsf{Test}$ is $\eta^*$-hiding, this difference is upper-bounded by $\frac{1}{\eta^*}$. Hence $\Pr_{\boldsymbol{P} \leftarrow \mathcal{D}_0}[\boldsymbol{P}(Q) \in C] \leq \frac{4\mu}{\eta^*}$.

Now, we view the test, on each input $b$, as sampling its agents vector $\boldsymbol{P}$ by first sampling the answer $\boldsymbol{P}(Q)$, and then sampling $\boldsymbol{P}$ conditioned on this answer. $\boldsymbol{P}(Q)$ itself is sampled from the distribution $\mathcal{A}_b = \{\boldsymbol{P}(Q)\}_{\boldsymbol{P} \leftarrow \mathcal{D}_b}$. Now, we invoke Lemma 1 on the distributions $\mathcal{A}_0$ and $\mathcal{A}_1$ with the set $S = B \cup C$. This results in $\epsilon = O(\frac{\mu}{\eta^*})$, given the above bound (and since $\Delta(\mathcal{A}_0, \mathcal{A}_1) \leq 1/\eta^*$). Thus, the test, with probability $1 - \epsilon$ samples $\mathsf{ans} \notin B \cup C$ (from a distribution independent of $b$) and then samples $\boldsymbol{P} \leftarrow \mathcal{D}_b|_{Q \to \mathsf{ans}}$. (With the remaining $\epsilon$ probability, it samples $\boldsymbol{P}$ depending on $b$ as appropriate.) Recall that, for $\mathsf{ans} \notin B \cup C$, we have that $(\mathcal{D}_0|_{Q \to \mathsf{ans}}, \mathcal{D}_1|_{Q \to \mathsf{ans}})$ is $\mu$-concentrated, where $\mu = q(\eta/2)$. Hence we can apply the $q$-$\mathsf{IND}$-$\mathsf{CON}$ security to conclude that no adversary can distinguish between $b = 0$ and $b = 1$ in the real experiment with advantage more than $\epsilon + (1-\epsilon)\eta/2$. We shall set $\epsilon < \eta/2$ so that this advantage is less than $\eta$, as we need to prove.

To finish the proof we need to ensure that $\eta^* > \pi(\eta^2) + \mu$ and $\epsilon < \eta/2$. This is satisfied by setting, say, $\eta^* > \pi(\eta^2) + q(\eta/2)$. Thus, we can set $p$ to be, say, the polynomial $p(\eta) := \pi(\eta^2) + q(\eta/2) + 1$.

**Theorem 2 (Equivalence of $s$-$\mathsf{IND}$-$\mathsf{PRE}$ and $s$-$\mathsf{SIM}$).** *A cryptographic agent scheme $\Pi = (\mathcal{O}, \mathcal{E})$ is a $\Gamma^*$-$s$-$\mathsf{IND}$-$\mathsf{PRE}$-secure scheme for a schema $\Sigma$ if and only if it is $\Gamma^*$-$s$-$\mathsf{SIM}$-secure for the same schema.*

*Proof.* Intuitively, $\Gamma^*$-$s$-$\mathsf{IND}$-$\mathsf{PRE}$ security is "weaker" than $\Gamma^*$-$s$-$\mathsf{SIM}$ security, and hence the first direction below is easier to see.

<u>$\Gamma^*$-$s$-$\mathsf{SIM} \Rightarrow \Gamma^*$-$s$-$\mathsf{IND}$-$\mathsf{PRE}$:</u> Suppose $\Sigma = (\mathcal{O}, \mathcal{E})$ is a $p$-$\Gamma^*$-$s$-$\mathsf{SIM}$ secure scheme for $\Sigma$, for some (bivariate) polynomial $p$. We shall show that $\Sigma$ is a $q$-$\Gamma^*$-$s$-$\mathsf{IND}$-$\mathsf{PRE}$ schema for a polynomial $q$ to be determined.

For a $\mathsf{Test} \in \Gamma^*$ and $\eta$, suppose there exists a $\mathsf{PPT}$ adversary $\mathsf{Adv}$ which runs in at most $\eta$ time but can distinguish between $\mathsf{Test}$ with bit 0 and 1 with probability at least $1/\eta$. That is,

$$|\Pr[\textsc{real}\langle \mathsf{Test}(0) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1] - \Pr[\textsc{real}\langle \mathsf{Test}(1) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1]| > 1/\eta.$$

We need to show that there is an ideal world user $\mathsf{User}$, which makes at most $q(\eta)$ queries and achieves a distinguishing advantage of at least $1/q(\eta)$.

Since $\Pi$ is $p$-$\Gamma^*$-$s$-$\mathsf{SIM}$ secure, given $\mathsf{Adv}$ which runs in time at most $\eta$, there exists an unbounded simulator $\mathcal{S}$ making at most $p(3\eta, \eta)$ queries, such that for all tests (and in particular, for $\mathsf{Test}$) and $b \in \{0, 1\}$:

$$|\Pr[\textsc{ideal}\langle \mathsf{Test}(b) \mid \Sigma \mid \mathcal{S}\rangle = 1] - \Pr[\textsc{real}\langle \mathsf{Test}(b) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1]| \leq \frac{1}{3\eta}.$$

And therefore,

$$|\Pr[\textsc{ideal}\langle \mathsf{Test}(0) \mid \Sigma \mid \mathcal{S}\rangle = 1] - \Pr[\textsc{ideal}\langle \mathsf{Test}(1) \mid \Sigma \mid \mathcal{S}\rangle = 1]| > \\ \frac{1}{\eta} - \frac{2}{3\eta} = \frac{1}{3\eta}.$$

We set $q$ such that $q(\eta) \geq p(3\eta, \eta)$ and $\frac{1}{3\eta} \geq \frac{1}{q(\eta)}$. For instance, we can set $q(x) = p(3x, x) + 3x$.

Note that in the above proof, we could allow $\mathcal{S}$ to depend on $\mathsf{Test}$, and therefore, even the weaker notion of simulation mentioned after Definition 5 implies $\mathsf{IND\text{-}PRE}$ security.

$\underline{\Gamma^*\text{-}s\text{-}\mathsf{IND\text{-}PRE} \Rightarrow \Gamma^*\text{-}s\text{-}\mathsf{SIM}:}$ Suppose $\Pi = (\mathcal{O}, \mathcal{E})$ is $q$-$\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure for a schema $\Sigma$. Fix a polynomial $\eta$ and a $\mathsf{PPT}$ adversary $\mathsf{Adv}$ whose running time is upper-bounded by a polynomial $\ell$. We shall construct a simulator $\mathcal{S}$ for $\mathsf{Adv}$ in the ideal world, which makes at most $p(\eta, \ell)$ queries for some polynomial $p$, and suffers a simulation error of at most $1/\eta$. Below, we write $\eta$ to mean $\max(\eta, \ell)$, so that we may assume that $\eta \geq \ell$.

In the ideal world, when a test $\mathsf{Test} \in \Gamma^*$ uploads a $\widehat{\boldsymbol{P}} \in \{0, 1\}^* \times \mathcal{P}^*_{\mathsf{test}}$, $\mathcal{S}$ attempts to learn a sufficiently accurate approximation $\boldsymbol{P}^\dagger$ using a polynomial depth query strategy, and then faithfully simulates $\mathcal{O}(\boldsymbol{P}^\dagger)$ to $\mathsf{Adv}$. Note that since $\mathsf{Adv}$'s running time is upper-bounded by the polynomial $\ell$, w.l.o.g, the simulator considers $\widehat{\boldsymbol{P}}$ to be in $\{0, 1\}^\ell \times \mathcal{P}^{\ell'}_{\mathsf{test}}$, where $\ell'$ is the lesser of $\ell$ and the actual number of agents uploaded by $\mathsf{Test}$.

$\mathcal{S}$ defines $R_i \subseteq \{0, 1\}^\ell \times \mathcal{P}^{\ell'}_{\mathsf{test}}$ and $D_i \subseteq R_i$ inductively as follows, for integers $i \geq 0$, up till $i = i^*$ such that $D_{i^*} = \emptyset$. It then samples $\boldsymbol{P}^\dagger \leftarrow R_{i^*}$ and uses it to complete the simulation.

Below, we write $\mathsf{Adv}(\mathcal{O}(R_i))$ to denote the random variable corresponding to the output of $\mathsf{Adv}$ when a random $\boldsymbol{P} \leftarrow R_i$ is encoded using $\mathcal{O}$ and given to $\mathsf{Adv}$; also, recall that $\mathsf{Adv}(\mathcal{O}(\boldsymbol{P}))$ denotes the similar random variable when the fixed agent vector $\boldsymbol{P}$ is encoded and given to $\mathsf{Adv}$.

1. Firstly, for each $i$, we define $D_i^*$ in terms of $R_i$, as follows. $D_i^* = D_{i,0}^* \cup D_{i,1}^*$, where

   $$D_{i,b}^* = \left\{ \boldsymbol{P} \in R_i \mid (-1)^b (\Pr[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1] - \Pr[\mathsf{Adv}(\mathcal{O}(R_i)) = 1]) > \frac{1}{\eta} \right\}.$$

   Below, we shall iteratively define sets $D_{i,0}$ and $D_{i,1}$, and let $D_i := D_{i,0} \cup D_{i,1}$. We shall maintain the invariant that, for all $i \geq 0$, $D_{i,\beta} \subseteq D_{i,\beta}^*$, and the uploaded agent vector $\widehat{\boldsymbol{P}} \in R_i \setminus (D_i^* \setminus D_i)$ (i.e., $\boldsymbol{P} \in R_i$, and if $\boldsymbol{P} \in D_i^*$ then $\boldsymbol{P} \in D_i$).
2. $R_0 = \{0,1\}^\ell \times \mathcal{P}_{\mathsf{test}}^{\ell'}$, $D_{0,0} = D_{0,0}^*$, and $D_{0,1} = D_{0,1}^*$.
3. If $D_i \neq \emptyset$, we define $R_{i+1}$ and $D_{i+1}$ as follows.

   Suppose $D_{i,\beta} \neq \emptyset$. Then, consider the test $\mathsf{Test}_{i,\beta} \in \Gamma^*$, which on input $b = 0$ uploads $\boldsymbol{P} \leftarrow D_{i,\beta}$, and on input $b = 1$, uploads $\boldsymbol{P} \leftarrow R_i$.[12] Since $D_{i,\beta}$ is not empty, we have

   $$\begin{aligned} &|\Pr[\textsc{real}\langle \mathsf{Test}_{i,\beta}(0) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1] - \Pr[\textsc{real}\langle \mathsf{Test}_{i,\beta}(1) \mid \mathcal{O} \mid \mathsf{Adv}\rangle = 1]| \\ &\quad = (-1)^\beta \frac{1}{|D_{i,\beta}|} \sum_{\boldsymbol{P} \in D_{i,\beta}} (\Pr[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1] - \Pr[\mathsf{Adv}(\mathcal{O}(R_i)) = 1]) > \frac{1}{\eta} \end{aligned}$$

   because for each $\boldsymbol{P} \in D_{i,\beta} \subseteq D_{i,\beta}^*$, we have $(-1)^\beta(\Pr[\mathsf{Adv}(\mathcal{O}(\boldsymbol{P})) = 1] - \Pr[\mathsf{Adv}(\mathcal{O}(R_i)) = 1]) > \frac{1}{\eta}$. That is, $\mathsf{Test}_{i,\beta}$ is not $\eta$-hiding (against $\mathsf{Adv}$, which runs for less than $\ell \leq \eta$ time). Since the scheme $\Pi = (\mathcal{O}, \mathcal{E})$ is $\Gamma^*$-$s$-$\mathsf{IND}$-$\mathsf{PRE}$-secure, there must exist an ideal world adversary, or equivalently, a query strategy $Q_{i,\beta}$ of depth at most $q(\eta)$ which has advantage of more than $\sigma := 1/q(\eta)$ in distinguishing $\mathsf{Test}_{i,\beta}(0)$ and $\mathsf{Test}_{i,\beta}(1)$.

   If $D_{i,\beta} = \emptyset$, $Q_{i,\beta}$ is taken as the empty query strategy. For each $\beta \in \{0,1\}$, $\mathcal{S}$ executes the query strategy $Q_{i,\beta}$ to obtain an answer $\mathsf{ans}_{i,\beta}$. It defines $R_i' = \{\boldsymbol{P} \in R_i \mid \boldsymbol{P}(Q_{i,0}) = \mathsf{ans}_{i,0}, \boldsymbol{P}(Q_{i,1}) = \mathsf{ans}_{i,1}\}$, and $D_{i,\beta}' = \{\boldsymbol{P} \in D_{i,\beta} \mid \boldsymbol{P}(Q_{i,\beta}) = \mathsf{ans}_{i,\beta}\}$. If $|R_i'| \leq (1-\sigma)|R_i|$, then set $R_{i+1} = R_i'$ and $D_{i+1,\beta} = D_{i+1,\beta}^*$. Otherwise, set $R_{i+1} = R_i$ (and hence $D_{i+1,\beta}^* = D_{i,\beta}^*$) and $D_{i+1,\beta} = D_{i,\beta}'$.

The above iteration terminates for the least $i$ such that $D_i = \emptyset$. Then we have the property that the uploaded agent $\widehat{\boldsymbol{P}} \in R_i \setminus D_i^*$, which means that

$$\left| \Pr[\mathsf{Adv}(\mathcal{O}(\widehat{\boldsymbol{P}})) = 1] - \Pr[\mathsf{Adv}(\mathcal{O}(R_i)) = 1] \right| \leq \frac{1}{\eta}.$$

Thus $\mathcal{S}$ completes the simulation by sampling $\boldsymbol{P}^\dagger \leftarrow R_i$ and giving $\mathcal{O}(\boldsymbol{P}^\dagger)$ to $\mathsf{Adv}$.

[12] Note that $\mathsf{Test}_{i,\beta}$ may be computationally inefficient. This is the only reason we are not able to prove analogous results for a test-family that is like $\Gamma^*$ but restricted to $\mathsf{PPT}$ tests.

Note that if $|R'_i| > (1-\sigma)|R_i|$ then $|D'_{i,\beta}| \leq (1-\sigma)|D_{i,\beta}|$, because otherwise $Q_{i,\beta}$ cannot distinguish $\mathsf{Test}_{i,\beta}$ with advantage $\sigma$ (as, for $b = 0$ and $b = 1$, it receives an answer other than $\mathsf{ans}_{i,\beta}$ with probability less than $\sigma$). Therefore, we make progress in each iteration: either $|R_{i+1}| \leq (1-\sigma)|R_i|$ (in which case $|D_{i+1}| \leq |R_{i+1}|$), or $|R_{i+1}| = |R_i|$ and $|D_{i+1,\beta}| \leq (1-\sigma)|D_{i,\beta}|$. Hence, for $i^* \leq \log^2_{1-\sigma} |R_0|$ we have $D_i = \emptyset$.

The total number of queries made by the simulator above is bounded by $q(\eta) \cdot \log^2_{1-\sigma} |R_0|$. Note that $\log_2 |R_0| \leq \ell + n_\Sigma \cdot \ell$, where $n_\Sigma$ is a (polynomial) upper-bound on the number of bits required to represent an agent in the schema $\Sigma$. Also, $\left|\frac{1}{\log_2(1-\sigma)}\right| = O(q(\eta))$, so that $\log^2_{1-\sigma} |R_0| = O((n_\Sigma \cdot \ell \cdot q(\eta))^2)$. Hence, we can set $p(\eta, \ell)$ to be $q(\eta)$ times this polynomial. □

### 3.1 Extensions: Limited Agent-Space and Resettable Tests

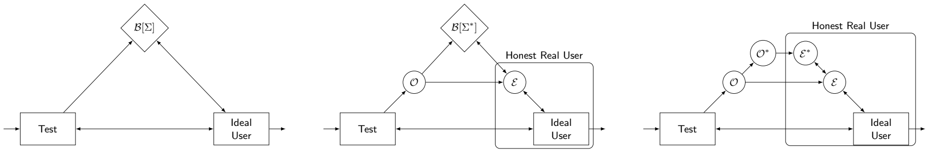
Firstly, in the above results we can use a test-family which is a subset of $\Gamma^*$ as follows. Note that the tests in $\Gamma^*$ may upload any number of agents and send messages of any length (i.e., we considered agents in $\{0,1\}^* \times \mathcal{P}^*_{\mathsf{test}}$). But our proofs go through unchanged if we restrict to a subset of $\Gamma^*$ which uses an arbitrary subset of $\{0,1\}^* \times \mathcal{P}^*_{\mathsf{test}}$. (In this case, $\mathsf{IND\text{-}CON}$ is suitably modified to use the same subset.) In particular, we may restrict to the test-family $\Gamma^*_1 \subseteq \Gamma^*$ which uploads a single agent and does not give any auxiliary information. Thus, every test in this family is fully characterized by a distribution over $\mathcal{P}^*_{\mathsf{test}}$. A variant of $\mathsf{IND\text{-}CON}$, say $\mathsf{IND\text{-}CON}_1$, can be defined where distribution ensembles only over $\mathcal{P}_{\mathsf{test}}$ are considered.

Secondly, we consider the possibility of using a test-family that is larger than $\Gamma^*$. Above, the restriction to $\Gamma^*$ was crucial in allowing the construction of a composite query strategy by grafting a query strategy onto the leaves of another query strategy. However, if the test allowed itself to be treated as an agent — i.e., allowing a $\mathsf{User}$ to access $\mathsf{Test}$ from any state in its history — then the above equivalences would carry over. Thus, we may define a test-family $\Gamma_{\mathsf{reset}}$ consisting of tests which are allowed to accept messages from the user/adversary and react to them, but also allows the user/adversary to reset it to the beginning (without changing its random tape). Then the above proofs extend to show that $\mathsf{IND\text{-}CON}$ $\Leftrightarrow$ $\Gamma_{\mathsf{reset}}$-$s$-$\mathsf{IND\text{-}PRE}$ $\Leftrightarrow$ $\Gamma_{\mathsf{reset}}$-$s$-$\mathsf{SIM}$, for all schemas. Note that tests in $\Gamma^*$ are effectively resettable and hence $\Gamma_{\mathsf{reset}} \supseteq \Gamma^*$. We defer a formal definition of $\Gamma_{\mathsf{reset}}$ to the final version.

## 4 Reductions and Compositions

A *hybrid scheme* $(\mathcal{O}, \mathcal{E})^{\Sigma^*}$ is a cryptographic agent scheme in which $\mathcal{O}$ and $\mathcal{E}$ have access to $\mathcal{B}[\Sigma^*]$, as shown in Figure 1 (in the middle), where $\Sigma^* = (\mathcal{P}^*_{\mathsf{auth}}, \mathcal{P}^*_{\mathsf{user}})$.[13]

[13] If $\mathcal{O}$ has a setup phase (see Appendix A), we require that $\mathcal{O}_{\mathsf{user}}$ uploads agents only in $\mathcal{P}^*_{\mathsf{user}}$ but $\mathcal{O}_{\mathsf{auth}}$ can upload any agent in $\mathcal{P}^*_{\mathsf{auth}} \cup \mathcal{P}^*_{\mathsf{user}}$.



(a) (b) (c)

Fig. 1

In general, the honest user would be replaced by an adversarial user $\mathsf{Adv}$. Note that the output bit of $\mathsf{Adv}$ in such a system is given by the random variable $\text{IDEAL}\langle \mathsf{Test} \circ \mathcal{O} \mid \Sigma^* \mid \mathsf{Adv}\rangle$, where $\mathsf{Test} \circ \mathcal{O}$ denotes the combination of $\mathsf{Test}$ and $\mathcal{O}$.

We introduce a new *information-theoretic* notion of reduction between schemata which would allow for composition of $\Gamma^*$-$s$-IND-PRE secure schemes. When compared to [2], the main difference is that we require the hybrid world to be secure against *unbounded* adversaries (who make a polynomial number of queries). Further, the simulator is allowed to depend on the adversary.

**Definition 9 (Statistical Reduction).** *We say that a (hybrid) cryptographic agent scheme $\Pi = (\mathcal{O}, \mathcal{E})$* statistically reduces *$\Sigma$ to $\Sigma^*$ with respect to $\widetilde{\Gamma}$, if there exists a polynomial $p$ such that for all* unbounded $\mathsf{User}$ *who make at most $\eta(\kappa)$ queries for some polynomial $\eta$,*

1. *Correctness:* $\forall$ $\mathsf{Test}$, $\text{IDEAL}\langle \mathsf{Test} \mid \Sigma \mid \mathsf{User}\rangle \approx \text{IDEAL}\langle \mathsf{Test} \circ \mathcal{O} \mid \Sigma^* \mid \mathcal{E} \circ \mathsf{User}\rangle$.
2. *Simulation: $\exists$ a simulator $\mathcal{S}_{\mathsf{User}}$ which makes at most $p(\eta(\kappa))$ queries s.t.* $\forall$ $\mathsf{Test} \in \widetilde{\Gamma}$, $\text{IDEAL}\langle \mathsf{Test} \mid \Sigma \mid \mathcal{S}_{\mathsf{User}}\rangle \approx \text{IDEAL}\langle \mathsf{Test} \circ \mathcal{O} \mid \Sigma^* \mid \mathsf{User}\rangle$.

*If there exists a scheme that reduces $\Sigma$ to $\Sigma^*$, then we say $\Sigma$ reduces to $\Sigma^*$. (Note that correctness is required for all tests, not just those in $\widetilde{\Gamma}$.)*

Figure 1 illustrates a reduction. It also shows how such a reduction can be composed with an IND-PRE-secure scheme for $\Sigma^*$. We now prove the main result of this section, in a manner very similar to that of Agrawal et al. [2].

**Theorem 3 (Composition).** *For any two schemata, $\Sigma$ and $\Sigma^*$, if $(\mathcal{O}, \mathcal{E})$ reduces $\Sigma$ to $\Sigma^*$ with respect to $\Gamma^*$ and $(\mathcal{O}^*, \mathcal{E}^*)$ is a $\Gamma^*$-$s$-*IND-PRE *secure scheme for $\Sigma^*$, then $(\mathcal{O} \circ \mathcal{O}^*, \mathcal{E}^* \circ \mathcal{E})$ is a $\Gamma^*$-$s$-*IND-PRE *secure scheme for $\Sigma$.*

*Proof.* Let $(\mathcal{O}', \mathcal{E}') = (\mathcal{O} \circ \mathcal{O}^*, \mathcal{E}^* \circ \mathcal{E})$. Also, let $\mathsf{Test}' = \mathsf{Test} \circ \mathcal{O}$ and $\mathsf{User}' = \mathcal{E} \circ \mathsf{User}$. We first show that for all $\mathsf{Test}$ and PPT $\mathsf{User}$, $(\mathcal{O}', \mathcal{E}')$ is a correct agent scheme for $\Sigma$. We have

$$\begin{aligned}\textsc{real}\langle \mathsf{Test} \mid \mathcal{O}' \mid \mathcal{E}' \circ \mathsf{User}\rangle &= \textsc{real}\langle \mathsf{Test}' \mid \mathcal{O}^* \mid \mathcal{E}^* \circ \mathsf{User}'\rangle \\ &\overset{(a)}{\approx} \textsc{ideal}\langle \mathsf{Test}' \mid \Sigma^* \mid \mathsf{User}'\rangle \\ &= \textsc{ideal}\langle \mathsf{Test} \circ \mathcal{O} \mid \Sigma^* \mid \mathcal{E} \circ \mathsf{User}\rangle \\ &\overset{(b)}{\approx} \textsc{ideal}\langle \mathsf{Test} \mid \Sigma \mid \mathsf{User}\rangle\end{aligned}$$

where $(a)$ follows from the correctness guarantee of IND-PRE security of $(\mathcal{O}^*, \mathcal{E}^*)$ (Definition 3), and $(b)$ follows from the correctness guarantee of $(\mathcal{O}, \mathcal{E})$ being a reduction of $\Sigma$ to $\Sigma^*$ (Definition 9). (Both $(a)$ and $(b)$ hold for all tests.) The other equalities are by regrouping the components in the system.

It remains to prove that there exists a polynomial $p$ such that for all large enough $\kappa$, all $\mathsf{Test} \in \Gamma^*$, and every polynomial $\eta$, if $\mathsf{Test}$ is $p(\eta(\kappa))$-s-hiding w.r.t. $\Sigma$ then $\mathsf{Test}$ is $\eta(\kappa)$-hiding w.r.t. $\mathcal{O}'$.

Suppose that for some polynomial $p'$, $(\mathcal{O}^*, \mathcal{E}^*)$ is a $p'$-$\Gamma^*$-$s$-IND-PRE secure scheme for $\Sigma^*$. We know that since $(\mathcal{O}, \mathcal{E})$ is a statistical reduction of $\Sigma$ to $\Sigma^*$ w.r.t. $\Gamma^*$, there exists a polynomial $p^*$ such that for all unbounded $\mathsf{User}$ who make at most $\mu(\kappa)$ queries (for some polynomial $\mu$), there exists a simulator $\mathcal{S}_{\mathsf{User}}$ which makes at most $p^*(\mu(\kappa))$ queries such that for all $\mathsf{Test} \in \Gamma^*$,

$$\begin{aligned}|\Pr[\textsc{ideal}\langle \mathsf{Test} \mid \Sigma \mid \mathcal{S}_{\mathsf{User}}\rangle = 1] - {}&\\ \Pr[\textsc{ideal}\langle \mathsf{Test} \circ \mathcal{O} \mid \Sigma^* \mid \mathsf{User}\rangle = 1]| &\leq \mathsf{negl}(\kappa). \qquad (1)\end{aligned}$$

So let $p$ be a polynomial such that $1/p(x) \leq \max\{1/p'(x) - 2 \cdot \mathsf{negl}(x), 1/p^*(p'(x))\}$ for all $x \geq 0$.

Let $\mathsf{Test}$ be an arbitrary test in $\Gamma^*$, $\eta$ be any polynomial, and $\overline{\mathsf{User}}$ be any unbounded user who makes at most $p'(\eta(\kappa))$ queries. We can apply Equation 1 on $\mathsf{Test}(b)$ and $\overline{\mathsf{User}}$ to get

$$\begin{aligned}|\Pr[\textsc{ideal}\langle \mathsf{Test}(b) \mid \Sigma \mid \mathcal{S}_{\overline{\mathsf{User}}}\rangle = 1] - {}&\\ \Pr[\textsc{ideal}\langle \mathsf{Test}(b) \circ \mathcal{O} \mid \Sigma^* \mid \overline{\mathsf{User}}\rangle = 1]| &\leq \mathsf{negl}(\kappa) \qquad (2)\end{aligned}$$

for $b \in \{0, 1\}$. Here the simulator $\mathcal{S}_{\overline{\mathsf{User}}}$ makes at most $p^*(p'(\eta(\kappa))) \leq p(\eta(\kappa))$ queries.

If $\mathsf{Test}$ is $p(\eta(\kappa))$-s-hiding w.r.t. $\Sigma$, then for all unbounded $\mathsf{User}'$ who make at most $p(\eta(\kappa))$ queries,

$$\begin{aligned}|\Pr[\textsc{ideal}\langle \mathsf{Test}(0) \mid \Sigma \mid \mathsf{User}'\rangle = 1] - {}&\\ \Pr[\textsc{ideal}\langle \mathsf{Test}(1) \mid \Sigma \mid \mathsf{User}'\rangle = 1]| &\leq \frac{1}{p(\eta(\kappa))}. \qquad (3)\end{aligned}$$

Recall that $\mathsf{Test}' = \mathsf{Test} \circ \mathcal{O}$ and if $\mathsf{Test} \in \Gamma^*$ then $\mathsf{Test}' \in \Gamma^*$ too. Now by using Equation 2 and Equation 3 with $\mathsf{User}'$ set to $\mathcal{S}_{\overline{\mathsf{User}}}$, we get

$$|\Pr[\text{IDEAL}\langle \mathsf{Test}'(0) \mid \Sigma^* \mid \overline{\mathsf{User}}\rangle = 1] - \Pr[\text{IDEAL}\langle \mathsf{Test}'(1) \mid \Sigma^* \mid \overline{\mathsf{User}}\rangle = 1]| \\ \leq \frac{1}{p(\eta(\kappa))} + 2 \cdot \mathsf{negl}(\kappa) \leq \frac{1}{p'(\eta(\kappa))}.$$

Thus $\mathsf{Test}'$ is $p'(\eta(\kappa))$-s-hiding w.r.t. $\Sigma^*$. This implies that $\mathsf{Test}'$ is $\eta(\kappa)$-hiding w.r.t. $\mathcal{O}^*$, and by regrouping the components, we have that $\mathsf{Test}$ is $\eta(\kappa)$-hiding w.r.t. $\mathcal{O}'$. □

We also have the following result regarding transitivity of reduction.

**Theorem 4 (Transitivity of Reduction).** *For any three schemata, $\Sigma_1, \Sigma_2, \Sigma_3$, if $\Sigma_1$ statistically reduces to $\Sigma_2$ and $\Sigma_2$ statistically reduces to $\Sigma_3$, then $\Sigma_1$ statistically reduces to $\Sigma_3$.*

*Proof.* If $\Pi_1 = (\mathcal{O}_1, \mathcal{E}_1)$ and $\Pi_2 = (\mathcal{O}_2, \mathcal{E}_2)$ are schemes that carry out the statistical reduction of $\Sigma_1$ to $\Sigma_2$ and that of $\Sigma_2$ to $\Sigma_3$, respectively, we claim that the scheme $\Pi = (\mathcal{O}_1 \circ \mathcal{O}_2, \mathcal{E}_2 \circ \mathcal{E}_1)$ is a statistical reduction of $\Sigma_1$ to $\Sigma_3$. The correctness of this reduction follows from the correctness of the given reductions. Further, if $\mathcal{S}_1$ and $\mathcal{S}_2$ are the simulators associated with the two reductions, we can define a simulator $\mathcal{S}$ for the composed reduction as $\mathcal{S}_2 \circ \mathcal{S}_1$. □

## 5 Applications

In this section we briefly summarize how the above results can be instantiated to rederive the main results of [9]. We start off by defining the obfuscation schema.

**Obfuscation Schema.** If $\mathcal{F}$ is a family of circuits, we define

$$\Sigma_{\text{OBF}(\mathcal{F})} := (\emptyset, \mathcal{F}).$$

That is, in the ideal execution $\mathsf{User}$ obtains handles for agents which simple compute $\mathcal{F}$ on their inputs and write the result on to their output tapes. We shall consider setup-free, $\mathsf{IND\text{-}PRE}$ secure implementations $(\mathcal{O}, \mathcal{E})$ of $\Sigma_{\text{OBF}(\mathcal{F})}$.

The following propositions which easily follow from the definitions. Below we refer to the test-family $\Gamma_1^*$ from Section 3.1.

**Proposition 1.** *For a function family $\mathcal{F}$, a $\Gamma_1^*$-s-$\mathsf{SIM}$ secure scheme for $\Sigma_{\text{OBF}(\mathcal{F})}$ is a VGB obfuscation scheme for $\mathcal{F}$, and vice-versa.*

With the modification to $\mathsf{IND\text{-}CON}$ also to distributions over a single agent (circuit), which we called $\mathsf{IND\text{-}CON}_1$ in Section 3.1, we have the following proposition.

**Proposition 2.** *For a function family $\mathcal{F}$, an* $\mathsf{IND\text{-}CON}_1$ *secure scheme for* $\Sigma_{\text{OBF}(\mathcal{F})}$ *is an SIO scheme for $\mathcal{F}$ and vice versa.*

These propositions, combined with Theorem 1, Theorem 2 (as extended in Section 3.1), yields the following result of [9] as a corollary.

**Corollary 5** *An obfuscation scheme is a VGB obfuscation for a function family $\mathcal{F}$ if and only if it is an SIO for $\mathcal{F}$.*

Next we describe how the security of the VGB obfuscation construction given in [9] follows as a corollary of our composition theorem.

**Graded Encoding Schema.** Following "set-based" graded encoding [9,13,5,19], we define the graded encoding schema $\Sigma_{GE} = (\emptyset, \mathcal{P}^{GE}_{\mathsf{user}})$, where $\mathcal{P}^{GE}_{\mathsf{user}}$ contains a single type of agent. The schema is specified by a ring $\mathcal{R}(+, \times)$ and a subset $\mathfrak{S}$ of $2^{[k]}$ for a level $k \in \mathbb{N}$ (where $[k] = \{1, 2, \ldots, n\}$). The persistent state of an agent $P \in \mathcal{P}^{GE}_{\mathsf{user}}$ is a pair $(x, S)$ where $x \in \mathcal{R}$ and $S \in \mathfrak{S}$, which it maintains on its work-tape (initially copied from its parameter tape). When invoked without an input, it sends $(x, S)$ to a peer agent in the session. When invoked with an input *Oper* on its input tape, it operates as follows (before entering a blocking state):

- *Oper* $= +$ (resp. $-$): It reads a message $(x', S')$ from its incoming communication tape. If $S = S'$, it updates its work-tape with $(x + x', S)$ (resp. $(x - x', S)$); otherwise, it writes $\perp$ on its output tape.
- *Oper* $= \times$: It reads a message $(x', S')$ from its incoming communication tape. If $S' \in \mathfrak{S}$ and $S \cap S' = \emptyset$, it updates its work-tape with $(x \times x', S \cup S')$; otherwise, it writes $\perp$ on its output tape.
- *Oper* $=$ $\mathsf{Zero\text{-}Test}$: It first checks whether $S$ is the universe set $[k]$. If not, it writes $\perp$ on its output tape. Otherwise, if $x = 0$ it writes 1; otherwise, 0.

The following proposition is an immediate consequence of the definition of strong-sampler semantic security [9].

**Proposition 3.** *A graded encoding scheme is strong-sampler semantically secure if and only if it is a* $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ *secure scheme for the schema $\Sigma_{GE}$.*

The following is a restatement of a result in [5] (that [9] relies on), formalized as a statistical reduction.

**Proposition 4.** *For any function family $\mathcal{F} \in \mathsf{NC}^1$, there exists a statistical reduction from $\Sigma_{\text{OBF}(\mathcal{F})}$ to $\Sigma_{GE}$.*

The following result of [9] is then an immediate corollary of the above two propositions and the composition theorem (Theorem 3) as well as the fact that a $\Gamma^*$-$s$-$\mathsf{IND\text{-}PRE}$ secure scheme for $\Sigma_{\text{OBF}(\mathcal{F})}$ is a VGB obfuscation (from Theorem 2 and Proposition 1).

**Corollary 6** *If there exists a strong-sampler semantically-secure graded encoding scheme, then there exists a VGB obfuscation scheme for any function family* $\mathcal{F} \in \mathsf{NC}^1$.

## Acknowledgments

This work was supported in part by NSF grant 12-28856. Part of this work was carried out while the authors were visiting the Simons Institute for Theoretical Computer Science, supported by the Simons Foundation and by the DIMACS/Simons Collaboration in Cryptography through NSF grant CNS 15-23467.

Part of this work was done when the first author was at the University of Illinois at Urbana-Champaign. At University of Texas at Austin, he is supported by NSF CNS-1228599,CNS-1414082 and DARPA SafeWare.

## References

1. Agrawal, S., Agrawal, S., Badrinarayanan, S., Kumarasubramanian, A., Prabhakaran, M., Sahai, A.: On the practical security of inner product functional encryption. In: Public-Key Cryptography - PKC 2015 - 18th IACR International Conference on Practice and Theory in Public-Key Cryptography, Gaithersburg, MD, USA, March 30 - April 1, 2015, Proceedings. pp. 777–798 (2015)
2. Agrawal, S., Agrawal, S., Prabhakaran, M.: Cryptographic agents: Towards a unified theory of computing on encrypted data. In: Oswald, E., Fischlin, M. (eds.) Advances in Cryptology - EUROCRYPT 2015, Lecture Notes in Computer Science, vol. 9057, pp. 501–531. Springer Berlin Heidelberg (2015), http://dx.doi.org/10.1007/978-3-662-46803-6_17
3. Agrawal, S., Gurbanov, S., Vaikuntanathan, V., Wee, H.: Functional encryption: New perspectives and lower bounds. In: CRYPTO (2013)
4. Barak, B., Goldreich, O., Impagliazzo, R., Rudich, S., Sahai, A., Vadhan, S., Yang, K.: On the (im)possibility of obfuscating programs. In: CRYPTO (2001)
5. Barak, B., Garg, S., Kalai, Y.T., Paneth, O., Sahai, A.: Protecting obfuscation against algebraic attacks. In: Advances in Cryptology - EUROCRYPT 2014 - 33rd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Copenhagen, Denmark, May 11-15, 2014. Proceedings. pp. 221–238 (2014)
6. Barbosa, M., Farshim, P.: On the semantic security of functional encryption schemes. In: Kurosawa, K., Hanaoka, G. (eds.) Public-Key Cryptography, PKC 2013. Lecture Notes in Computer Science, vol. 7778, pp. 143–161. Springer Berlin Heidelberg (2013)
7. Bellare, M., O'Neill, A.: Semantically-secure functional encryption: Possibility results, impossibility results and the quest for a general definition. In: CANS. pp. 218–234 (2013)
8. Bitansky, N., Canetti, R.: On strong simulation and composable point obfuscation. In: Rabin, T. (ed.) Advances in Cryptology - CRYPTO 2010, Lecture Notes in Computer Science, vol. 6223, pp. 520–537. Springer Berlin Heidelberg (2010), http://dx.doi.org/10.1007/978-3-642-14623-7_28

9. Bitansky, N., Canetti, R., Kalai, Y., Paneth, O.: On virtual grey box obfuscation for general circuits. In: Garay, J., Gennaro, R. (eds.) Advances in Cryptology - CRYPTO 2014, Lecture Notes in Computer Science, vol. 8617, pp. 108–125. Springer Berlin Heidelberg (2014), http://dx.doi.org/10.1007/978-3-662-44381-1_7
10. Boneh, D., Sahai, A., Waters, B.: Functional encryption: Definitions and challenges. In: TCC. pp. 253–273 (2011)
11. Canetti, R.: Universally composable security: A new paradigm for cryptographic protocols. In: Proceedings of the 42Nd IEEE Symposium on Foundations of Computer Science. FOCS '01 (2001)
12. Caro, A.D., Iovino, V., Jain, A., O'Neill, A., Paneth, O., Persiano, G.: On the achievability of simulation-based security for functional encryption. In: CRYPTO (2013)
13. Garg, S., Gentry, C., Halevi, S., Raykova, M., Sahai, A., Waters, B.: Candidate indistinguishability obfuscation and functional encryption for all circuits. In: FOCS (2013), http://eprint.iacr.org/
14. Goldwasser, S., Kalai, Y.T.: On the impossibility of obfuscation with auxiliary input. In: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science. FOCS '05 (2005)
15. Goldwasser, S., Rothblum, G.N.: On best-possible obfuscation. In: Proceedings of the 4th Conference on Theory of Cryptography. TCC'07 (2007)
16. Hada, S.: Zero-knowledge and code obfuscation. In: Advances in Cryptology - ASIACRYPT 2000, 6th International Conference on the Theory and Application of Cryptology and Information Security, Kyoto, Japan, December 3-7, 2000, Proceedings. pp. 443–457 (2000)
17. Maurer, U.: Constructive cryptography - A new paradigm for security definitions and proofs. In: Theory of Security and Applications - Joint Workshop, TOSCA 2011, Saarbrücken, Germany, March 31 - April 1, 2011, Revised Selected Papers. pp. 33–56 (2011), http://dx.doi.org/10.1007/978-3-642-27375-9_3
18. O'Neill, A.: Definitional issues in functional encryption. Cryptology ePrint Archive, Report 2010/556 (2010), http://eprint.iacr.org/
19. Pass, R., Seth, K., Telang, S.: Indistinguishability obfuscation from semantically-secure multilinear encodings. In: Garay, J., Gennaro, R. (eds.) Advances in Cryptology - CRYPTO 2014, Lecture Notes in Computer Science, vol. 8616, pp. 500–517. Springer Berlin Heidelberg (2014), http://dx.doi.org/10.1007/978-3-662-44371-2_28

## A Preliminaries

The following description of the Cryptographic Agents model is adapted from [2], and follows it closely.

### A.1 Agents

**Definition 10 (Agents and Family of Agents).** *An agent is an interactive Turing Machine, with the following modifications:*

- *There is a special read-only parameter tape, which always consists of a security parameter $\kappa$, and possibly other parameters.*
- *There is an a priori restriction on the size of all the tapes other than the randomness tape (including input, communication and work tapes), as a function of the security parameter.*
- *There is a special* blocking state *such that if the machine enters such a state, it remains there if the input tape is empty. Similarly, there are blocking states which let the machine block if any combination of the communication tape and the input tape is empty.*

*An* agent family *is a maximal set of agents with the same program (i.e., state space and transition functions), but possibly different contents in their parameter tapes. We also allow an agent family to be the empty set $\emptyset$.*

Note that an agent who enters a blocking state can move out of it if its configuration is changed by adding a message to its input tape and/or communication tape. However, if the agent enters a halting state, it will not move out of that state. An agent who never enters a blocking state is called a *non-reactive agent*. An agent who never reads or writes from a communication tape is called a *non-interactive agent*.

**Definition 11 (Session).** *A session maps a finite ordered set of agents, their configurations and inputs, to outputs and (updated) configurations of the same agents, as follows. The agents are initialized with the given inputs on their input tapes, and then executed together until they are deadlocked.*[14] *The result of applying the session is defined as the collection of outputs and configurations of the agents when the session terminates (if it terminates; if not, the result is left undefined).*

We shall be restricting ourselves to collections of agents such that sessions involving them are guaranteed to terminate. Note that we have defined a session to have only an initial set of inputs, so that the outcome of a session is well-defined (without the need to specify how further inputs would be chosen).

**Definition 12 (Ideal Agent Schema).** *A (well-behaved) ideal agent schema $\Sigma = (\mathcal{P}_{\mathsf{auth}}, \mathcal{P}_{\mathsf{user}})$, or simply* schema*, is a pair of* agent families*, such that there is a polynomial* poly *such that for any session of agents belonging to $\mathcal{P}_{\mathsf{auth}} \cup \mathcal{P}_{\mathsf{user}}$ (with any inputs and any configurations, with the same security parameter $\kappa$), the session terminates within* $\mathrm{poly}(\kappa, t)$ *steps, where $t$ is the number of agents in the session.*

[14] More precisely, the first agent is executed till it enters a blocking or halting state, and then the second and so forth, in a round-robin fashion, until all the agents remain in blocking or halting states for a full round. After each execution of an agent, the contents of its outgoing communication tape are interpreted as an ordered sequence of messages to each of the other agents in the session (some or all of them possibly being empty messages), and copied over to the respective agents' incoming communication tapes.

### A.2 Security Definitions

We define what it means for a cryptographic agent scheme to securely implement a given ideal agent schema. Intuitively, the security notion is of *indistinguishability preservation*: if two executions using an ideal schema are indistinguishable, we require them to remain indistinguishable when implemented using a cryptographic agent scheme.

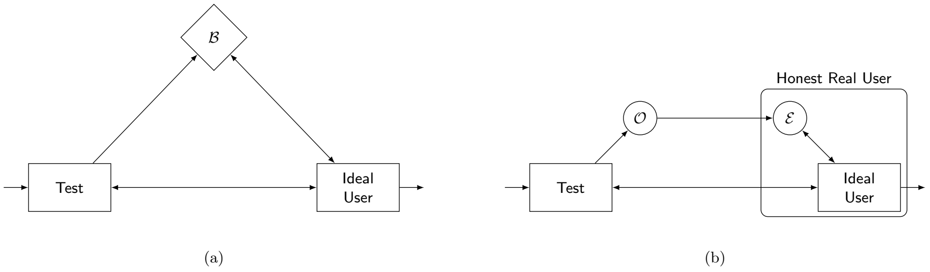


Fig. 2: The ideal world (on the left) and the real world with an honest user.

**Ideal World.** The ideal system for a schema $\Sigma$ consists of two parties $\mathsf{Test}$ and $\mathsf{User}$ and a fixed third party $\mathcal{B}[\Sigma]$ (for "black-box"). All three parties have a security parameter $\kappa$ built-in. We shall explicitly refer to their random-tapes as $r, s$ and $t$. $\mathsf{Test}$ receives a "secret bit" $b$ as input and $\mathsf{User}$ produces an output bit $b'$. The interaction between $\mathsf{User}$, $\mathsf{Test}$ and $\mathcal{B}[\Sigma]$ can be summarized as follows:

- **Uploading agents.** Let $\Sigma = (\mathcal{P}_{\mathsf{auth}}, \mathcal{P}_{\mathsf{user}})$ where we associate $\mathcal{P}_{\mathsf{test}} := \mathcal{P}_{\mathsf{auth}} \cup \mathcal{P}_{\mathsf{user}}$ with $\mathsf{Test}$ and $\mathcal{P}_{\mathsf{user}}$ with $\mathsf{User}$. $\mathsf{Test}$ and $\mathsf{User}$ can, at any point, choose an agent from its agent family and send it to $\mathcal{B}[\Sigma]$. More precisely, $\mathsf{User}$ can send a string to $\mathcal{B}[\Sigma]$, and $\mathcal{B}[\Sigma]$ will instantiate an agent $\mathcal{P}_{\mathsf{user}}$, with the given string (along with its own security parameter) as the contents of the parameter tape, and all other tapes being empty. Similarly, $\mathsf{Test}$ can send a string and a bit indicating whether it is a parameter for $\mathcal{P}_{\mathsf{auth}}$ or $\mathcal{P}_{\mathsf{user}}$, and it is used to instantiate an agent $\mathcal{P}_{\mathsf{auth}}$ or $\mathcal{P}_{\mathsf{user}}$, accordingly.[15] Whenever an agent is instantiated, $\mathcal{B}[\Sigma]$ sends a unique handle (a serial number) for that agent to $\mathsf{User}$; the handle also indicates whether the agent belongs to $\mathcal{P}_{\mathsf{auth}}$ or $\mathcal{P}_{\mathsf{user}}$.
- **Query.** A query is a request for session execution. At any point in time, $\mathsf{User}$ may request an execution of a session, by sending an ordered tuple of handles $(h_1, \ldots, h_t)$ (from among all the handles obtained thus far from $\mathcal{B}[\Sigma]$) to specify the configurations of the agents in the session, along with their inputs. $\mathcal{B}[\Sigma]$ reports back the outputs from the session, and also gives new

[15] In fact, for convenience, we allow $\mathsf{Test}$ and $\mathsf{User}$ to specify multiple agents in a single message to $\mathcal{B}[\Sigma]$.

handles corresponding to the configurations of the agents when the session terminated.[16] If an agent halts in a session, no new handle is given for that agent.

Observe that only $\mathsf{User}$ receives any output from $\mathcal{B}[\Sigma]$; the communication between $\mathsf{Test}$ and $\mathcal{B}[\Sigma]$ is one-way. (See Figure 2.)

**Real World.** A *cryptographic scheme* (or simply scheme) consists of a pair of (possibly stateful and randomized) programs $(\mathcal{O}, \mathcal{E})$, where $\mathcal{O}$ is an encoding procedure for agents in $\mathcal{P}_{\mathsf{test}}$ and $\mathcal{E}$ is an execution procedure. The real world execution for a scheme $(\mathcal{O}, \mathcal{E})$ consists of $\mathsf{Test}$, a user that we shall generally denote as $\mathsf{Adv}$ and the encoder $\mathcal{O}$. ($\mathcal{E}$ features as part of an honest user in the real world execution: see Figure 2.) $\mathsf{Test}$ remains the same as in the ideal world, except that instead of sending an agent to $\mathcal{B}[\Sigma]$, it sends it to the encoder $\mathcal{O}$. In turn, $\mathcal{O}$ encodes this agent and sends the resulting cryptographic agent to $\mathsf{Adv}$.

**Syntactic Requirements on $(\mathcal{O}, \mathcal{E})$.** $(\mathcal{O}, \mathcal{E})$ may or may not use a "setup" phase. In the latter case we call it a *setup-free cryptographic agent scheme*, and $\mathcal{O}$ is required to be a memory-less program that takes an agent $P \in \mathcal{P}_{\mathsf{test}}$ as input and outputs a cryptographic agent that is sent to $\mathsf{Adv}$. If the scheme has a setup phase, $\mathcal{O}$ consists of a triplet of memory-less programs $(\mathcal{O}_{\mathsf{setup}}, \mathcal{O}_{\mathsf{auth}}, \mathcal{O}_{\mathsf{user}})$: in the real world execution, first $\mathcal{O}_{\mathsf{setup}}$ is run to generate a secret-public key pair $(\mathsf{MSK}, \mathsf{MPK})$;$\mathsf{MPK}$ is sent to $\mathsf{Adv}$. Subsequently, when $\mathcal{O}$ receives an agent $P \in \mathcal{P}_{\mathsf{auth}}$ it will invoke $\mathcal{O}_{\mathsf{auth}}(P, \mathsf{MSK})$, and when it receives an agent $P \in \mathcal{P}_{\mathsf{user}}$, it will invoke $\mathcal{O}_{\mathsf{user}}(P, \mathsf{MPK})$, to obtain a cryptographic agent that is then sent to $\mathsf{Adv}$.

$\mathcal{E}$ is required to be memoryless as well, except that when it gives a handle to a $\mathsf{User}$, it can record a string against that handle, and later when $\mathsf{User}$ requests a session execution, $\mathcal{E}$ can access the string recorded for each handle in the session. There is a *compactness requirement* that the size of this string is *a priori* bounded (note that the state space of the ideal agents are also *a priori* bounded). If there is a setup phase, $\mathcal{E}$ can also access $\mathsf{MPK}$ each time it is invoked.

[16] Note that if the same handle appears more than once in the tuple $(h_1, \ldots, h_t)$, it is interpreted as multiple agents with the same configuration (but possibly different inputs). Also note that after a session, the old handles for the agents are not invalidated; so a $\mathsf{User}$ can access a configuration of an agent any number of times, by using the same handle.