A Note on Black-Box Separations for Indistinguishability Obfuscation

Mohammad Mahmoody* Ameer Mohammed† Soheil Nematihaji‡ Rafael Pass§ abhi shelat¶

May 24, 2016

Abstract

Mahmoody et al. (TCC 2016-A) showed that basing indistinguishability obfuscation (IO) on a wide range of primitives in a black-box way is *as hard as* basing public-key cryptography on one-way functions. The list included any primitive $\mathcal{P}$ that could be realized relative to random trapdoor permutation or degree-$O(1)$ graded encoding oracle models in a secure way against computationally unbounded polynomial-query attackers.

In this work, relying on the recent result of Brakerski, Brzuska, and Fleischhacker (ePrint 2016/226) in which they ruled out statistically secure approximately correct IO, we show that there is no fully black-box constructions of IO from any of the primitives listed above, assuming the existence of one-way functions and $\mathbf{NP} \not\subseteq \mathbf{coAM}$.

At a technical level, we provide an alternative lemma to the Borel-Cantelli lemma that is useful for deriving black-box separations. In particular, using this lemma we show that attacks in idealized models that succeed with only a *constant* advantage over the trivial bound are indeed sufficient for deriving fully black-box separations from primitives that exist in such idealized models unconditionally.

Keywords: Indistinguishability Obfuscation, Black-Box Separations.

*University of Virginia, `mohammad@cs.virginia.edu`. Supported by NSF CAREER award CCF-1350939. The work was done in part while the author was visiting the Simons Institute for the Theory of Computing, supported by the Simons Foundation and by the DIMACS-Simons Collaboration in Cryptography through NSF grant CNS-1523467.

†University of Virginia, `am8zv@virginia.edu`. Supported by University of Kuwait.

‡University of Virginia, `sn8fb@virginia.edu`. Supported by NSF award CCF-1350939.

§Cornell University, `rafael@cs.cornell.edu`. Work supported in part by a Microsoft Faculty Fellowship, Google Faculty Award, NSF Award CNS-1217821, NSF Award CCF-1214844, AFOSR Award FA9550-15-1-0262 and DARPA and AFRL under contract FA8750-11-2-0211.

¶University of Virginia, `shelat@cs.virginia.edu`. Work supported by NSF CAREER Award 0845811, NSF TC Awards 1111781 and 0939718, DARPA and AFRL under contract FA8750-11-C-0080, Microsoft New Faculty Fellowship, SAIC Scholars Research Award, and Google Research Award.

1 Introduction

The study of reductions between cryptographic primitives as computational building blocks has long occupied a central role in the theory of cryptography. In this paper, we apply this lens to *indistinguishability obfuscation* (IO), a primitive which has attracted special interest during the last few years. IO was proposed nearly 15 years ago by Barak et al. [5, 6], recently constructed by Gentry et al. [23] for general circuits based on multi-linear assumptions [21], and shown to be a "central hub" [45] for cryptographic tasks/primitives (see [23, 24, 22, 16, 11, 26, 20, 46, 9] to name a few).

Assumptions behind IO. Due to its applicability, it is important to identify the assumptions that are necessary to construct IO. The first candidate construction of IO [23] and many subsequent alternative constructions rely on polynomial-degree "multi-linear maps" or their idealized form of "graded encoding schemes" [23, 14, 4, 43, 28, 2, 40, 47, 37, 25].[1] Such assumptions are still considered to be extremely strong, as there has been multiple attacks on various forms of multi-linear maps [27, 19, 18].[2] Thus, a fascinating question is to study whether we can base IO on more standard assumptions such as trapdoor permutations, collision-resistant hash functions, DDH, bilinear maps, etc. Goldwasser and Rothblum [31, 32] take the first step towards answering this question and completely rule out the possibility of *statistically* secure IO if $\mathbf{NP} \not\subseteq \mathbf{coAM}$. Their result, however, leaves open whether (computational) IO can be based on standard computational assumptions. In a recent beautiful work, Brakerski, Brzuska, and Fleischhacker [13] extend the result of [32] to IO schemes that are only required to be *approximately-correct* assuming one-way functions exist and that $\mathbf{NP} \not\subseteq \mathbf{coAM}$.[3]

Black-box lower bounds. The most widely used framework to study the *impossibility* of basing cryptographic tasks on other (more basic) assumptions is the black-box framework of Impagliazzo and Rudich [36] and its subsequent formalization by Reingold, Trevisan, and Vadhan [44]. Considering the versatility of IO, it seems one should be able to prove that IO is indeed "too complex" to be constructed in a black-box way from well-studied standard assumptions such as OWFs, CRHFs, etc. Note that until we resolve the $\mathbf{P}$ vs $\mathbf{NP}$ question, any black-box separation result for the assumptions behind IO will depend on some computational hardness assumption, because if $\mathbf{P} = \mathbf{NP}$, then *statistically* secure IO exists.[4]

Relying on the work of [15, 42, 38] which studies *virtual black-box* obfuscation in idealized models of computations, Mahmoody et al. [39] show the first barrier towards obtaining black-box constructions of IO from certain powerful cryptographic assumptions. In particular, they show that if $\mathbf{NP} \neq \mathbf{co\text{-}NP}$, IO with perfect completeness cannot be based on collision-resistant hash functions (CRHFs) in a black-box way, and that basing IO on a large set of other stronger primitives such as trapdoor permutations, bilinear maps, etc. is *as hard* as constructing public-key encryption (PKE) from one-way functions. Impagliazzo and Rudich [36] rule out black-box methods for the latter question; however, finding a non-black-box approach remains a major open

[1] Interestingly, the work of [37] shows how to get IO from *constant-degree* multi-linear maps in a non-black-box way, using some extra assumptions.

[2] The work of [25] gives a new IO scheme that is resilient to these vulnerabilities.

[3] As we will describe, this result plays a major role in the proof of our main result.

[4] A statistically secure construction could be interpreted as a black-box construction from any primitive $\mathbf{P}$ that simply ignores the oracle providing $\mathbf{P}$!

question in cryptography. Indeed, the authors do not believe that a construction of PKE from OWFs is impossible; in particular, assuming IO, such constructions [45] already exist!

Thus, Mahmoody et al. [39] leave open whether their hardness of black-box constructions for IO can be extended to fully black-box separations.[5]

Our main result. In this short paper we extend the *hardness* results of [39] into the following black-box separation: Let $\mathcal{P}$ be any primitive that can be realized relative to the random trapdoor permutation oracle or the degree-$O(1)$ graded encoding model in a way that is secure against polynomial-query attackers. Examples of $\mathcal{P}$ include CCA-secure public-key encryption [41, 7], hierarchical identity based encryption [29, 35], non-interactive zero-knowledge proofs for **NP** [10, 8, 30], etc. We rule out fully-black-box constructions of indistinguishability obfuscators (IO) from any such $\mathcal{P}$ under the widely believed assumption that one-way functions exist and that $\mathbf{NP} \not\subseteq \mathbf{coAM}$.

1.1 Technical Overview

Recall that a fully black-box construction [44] of a primitive $\mathcal{Q}$ from another primitive $\mathcal{P}$ consists of two oracle PPT algorithms (Q, S) such that Q^P implements $\mathcal{Q}$ given access to any oracle P that implements $\mathcal{P}$, and $S^{P,A}$ turns any oracle attacker A against Q^P into an attack against P itself (see Definition 2.1).

Big picture of the argument. Similar to previous black-box separations (e.g., [36]), our proof that $\mathcal{P} \not\Rightarrow_{\mathsf{BB}}$ IO presents a polynomial-query attacker A that breaks the security of any IO construction iO in an idealized model $\mathcal{I}$ that provides an "unquestionably secure" instantiation $\mathcal{P}$ (against computationally unbounded polynomial-query attackers).[6] Intuitively, the existence of such an A rules out the possibility of a fully black-box construction construction (iO, S) of IO from $\mathcal{P}$ by simple *composition*. First, the construction $iO^{P^{\mathcal{I}}} = (iO^P)^{\mathcal{I}}$ yields an implementation of IO in the same idealized model $\mathcal{I}$. But attacker A breaks every such construction of IO and therefore the security reduction $S^{P^{\mathcal{I}},A^{\mathcal{I}}}$ implies the existence of a new attacker $(S^A)^{\mathcal{I}}$ that calls the idealized oracle $\mathcal{I}$ a polynomial number of times and breaks the implementation $P^{\mathcal{I}}$ of $\mathcal{P}$. But this leads to a contradiction because $P^{\mathcal{I}}$ is an "unquestionably secure" construction of $\mathcal{P}$ in $\mathcal{I}$.

Attacks on IO in idealized models. The recent elegant work of Brakerski, Brzuska, and Fleischhacker [13] when combined with previous works of [15, 42, 38] show an attacker that can break any IO scheme in either of the idealized model $\mathcal{I}$ of random trapdoor permutations and degree-$O(1)$ graded encoding models by asking a polynomial number of oracle queries. In particular, the previous works of [15, 42, 38] show how to "compile out" the idealized oracle $\mathcal{I}$ from the IO scheme and achieve an *approximately-correct* IO scheme in the plain model that is correct on, say, $99/100$ of the input points. Brakerksi et al. then show that any such approximately-correct IO scheme can be

[5]As we pointed our earlier, such separations will be necessarily based on computational assumptions unless we manage to prove that $\mathbf{P} \neq \mathbf{NP}$. However, proving separations based on assumptions like $\mathbf{P} \neq \mathbf{NP}$ are qualitatively different than just proving that such constructions are possible but hard to achieve.

[6]For example in the context of OWF $\not\Rightarrow_{\mathsf{BB}}$ Key-Agreement, the idealized oracle that provides the OWF is a random oracle, and [36] shows how to break any key-agreement protocol in the random oracle by asking only a polynomial number of oracle queries to derive the separation.

broken by a computationally unbounded attacker.[7] As pointed out in [13] this means that any IO scheme will be broken in the idealized model $\mathcal{I}$, and in particular the computationally unbounded attacker B can be modified into a computationally unbounded, yet *polynomial-query* attacker A against the original IO in the idealized model $\mathcal{I}$.

The challenge: fixing $\mathcal{I}$ while keeping the attack successful. At a first glance, it seems that the attacker A of [13] against IO in an idealized model $\mathcal{I}$ would immediately imply the desired black-box separation between IO and primitives that exist in model $\mathcal{I}$. However, the challenge, roughly speaking, is that the attacker of [13] does not succeed in breaking IO with probability close to 1, and doing so is left as an open question. In order to see the challenge more clearly, we need to further discuss the *big picture* argument above and see how an attack in the idealized model $\mathcal{I}$ exactly implies the black-box separation.

A crucial point is that to apply the security reduction $S^{P^{\mathcal{I}},A^{\mathcal{I}}}$ and get the desired attack against $P^{\mathcal{I}}$, we must *fix* the oracles $P^{\mathcal{I}}$ and $A^{\mathcal{I}}$ into deterministic functions (which requires us to sample and fix $\mathcal{I}$) because only then is S guaranteed to generate an attack. However, while fixing $\mathcal{I}$, we want to keep the promise that $A^{\mathcal{I}}$ is still a "successful" attack. Handling both tasks simultaneously may raise an issue because all attacks in idealized models (e.g., the attack of [36] against key-agreement in random oracle model) and in particular the attack against IO in idealized models that is implied by [13] are successful with probability taken over the randomness of the idealized oracle $\mathcal{I}$.

Borel-Cantelli for highly successful attacks. Here is where the Borel-Cantelli lemma (Lemma 2.8) usually comes to help, but only if the attack succeeds with high probability. In particular, if the demonstrated attacker A wins the security game for security parameter n with probability e.g., $1-1/n^4$, then by an averaging argument, with probability at least $1-1/n^2$ over the sampled oracle $\mathcal{I}$, A successfully attacks the game on security parameter n. Therefore, since the probability of the "fail" event is $\sum_{n=1}^{\infty} 1/n^2 = O(1)$, Borel-Cantelli lemma implies that with measure one over[8] the sampled oracle $\mathcal{I}$ it holds that A is a successful attack for all but finitely many security parameters.

An alternative to Borel-Cantelli lemma for mildly-successful attacks. By the above discussion on how to use Borel-Cantelli, we would be done if the attacker of [13] succeeds with probability $1 - 1/\operatorname{poly}(n)$. However, their attack works against (ϵ, δ) statistical approximate[9] IO when $2\epsilon + 3\delta < 1$; thus, by making optimal parameter choices, their attacker only succeeds (in guessing the obfuscated circuit) with probability $\approx 1/2 + 1/6$ which is not arbitrarily close to 1. As a result, when combined with the results of [15, 42, 38] we would only get an attack against IO in idealized models that succeeds with *some constant* advantage over 1/2 (and thus fails with some constant probabilty). Thus, we can no longer apply the Borel-Cantelli lemma as we did before because the summation of the probability of failure becomes unbounded. Thus, we can no longer conclude that this attack would remain successful for an infinite sequence of security parameters n [10] when we sample and fix the idealized oracle $\mathcal{I}$. In fact, there are examples of protocols in idealized models with attacks against them with $1/\operatorname{poly}(n)$ advantage over the trivial bound, but

[7] The attack of [13] assumes the existence of OWFs and that $\mathbf{NP} \not\subseteq \mathbf{coAM}$, and that is where we get these assumptions for our separation as well.

[8] Since the probability distribution here is over infinite-size oracles, we cannot assign probabilities to arbitrary events, but we can alternatively work with measurable sets.

[9] Here ε refers to the correctness error, and δ refers to the statistical closeness.

[10] Here n, the security parameter, is equal to the circuit size.

once the randomized oracle is sampled and fixed, they do *not* remain successful over an infinite sequence of security parameters (see Remark 2.5).

To overcome this issue, we provide a variant of the Borel-Cantelli lemma (see Lemma 2.9) which allows us to make sufficiently strong conclusions about the attacker as long as the attacker A succeeds with a *constant* advantage over the trivial bound. Note that Borel-Cantelli (when applicable) would imply a stronger result, because it shows that the attack will remain successful for *all but finitely* many security parameters, while our lemma shows that it only succeeds for an infinite sequence of security parameters. However, even this weaker conclusion is still enough for the security reduction $S^{P^{\mathcal{I}},A^{\mathcal{I}}}$ to be able to use A and give a polynomial-query attack against $P^{\mathcal{I}}$.

The scope of this argument does not seem to be at all limited to proving separations for IO, and we believe that it could potentially be applied to other primitives as well. Namely, it shows that to derive a black-box separation $\mathcal{P} \not\Rightarrow_{\mathsf{BB}} \mathcal{Q}$ it is enough to break $\mathcal{Q}$ in an idealized model that gives $\mathcal{P}$ by asking a polynomial number of queries and a *constant* advantage over the trivial bound.

Organization. In the next section, we provide the necessary definitions, the borrowed results of [42, 38] and [13] as well as the new measure theoretic alternative lemma to Borel-Cantelli. In Section 3 we formally prove the main result.

2 Preliminaries

2.1 Definitions

Definition 2.1 (Fully black-box constructions [44])**.** A *fully-black-box* construction of a primitive $\mathcal{Q}$ from a primitive $\mathcal{P}$ consists of two PPT algorithms (Q, S) as follows:

- **Implementation:** if oracle P implements $\mathcal{P}$, then Q^P implements $\mathcal{Q}$.
- **Security reduction:** for any oracle P implementing $\mathcal{P}$ and for any (computationally unbounded) oracle adversary A breaking the security of Q^P, $S^{P,A}$ breaks the security of P.

Reingold, Trevisan and Vadhan [44] also defined other (more relaxed) notions of black-box constructions, and Baecher, Brzuska, and Fischlin [3] further studied those notions in more details. We refer the readers to [44, 3] for those extensions. We will, however, assume one general property about the primitives that we deal with in this work: function P implementing $\mathcal{P}$ will be partitioned into sub-domains indexed by "security parameter" n and any adversary A who successfully breaks P would have to "win" over an infinite number of security parameters for a "noticeable" advantage.

We skip defining IO and approximate IO and directly define the generalized notion of approximate computational IO. We first recall a statistical variant of this notion defined by [13].

Definition 2.2 ([13] Approximate Statistical Correlation IO)**.** A PPT O is an (ε, δ)-*approximate statistical correlation IO* (CIO for short) if:

- **Approximate correctness:** $\Pr[O(C)(x) \neq C(x)] \leq \varepsilon(|C|)$ where the probability is over the randomness of the obfuscator and the input x.
- **Statistical correlation:** For every pair of circuits $C_1 \equiv C_2$ of the same size n, the statistical distance between $O(C_1)$ and $O(C_2)$ (both defined over the randomness of O) is at most $\delta(n)$.

A computational variant of Definition 2.2 can be defined analogously:

Definition 2.3 (Approximate Computational Correlation IO)**.** A PPT O is an (ε, δ)-approximate *computational* CIO if it satisfies the same correctness condition as approximate statistical CIO and:

- **Computational correlation:** For every poly-time adversary A and for every pair of circuits $C_1 \equiv C_2$ of equal size n, it holds that $\Pr[A(O(C_1)) = 1] - \Pr[A(O(C_2))] \leq \delta(n)$.

Fully-black-box constructions of IO. A fully-black-box construction of approximate computational CIO from primitive $\mathcal{P}$ could be defined through a combination of Definitions 2.1 and 2.3. Here we emphasize that the input circuits do not have any oracle gates while the obfuscation algorithm and the final circuits could use the oracle implementing $\mathcal{P}$. This seemingly restricted model is in fact sufficient for all known applications (see [39] for more discussions).

Idealized Models. An idealized model $\mathcal{I}$ is a randomized oracle; examples include the random oracle, random trapdoor permutation oracle, generic group model, graded encoding model, etc. An $I \leftarrow \mathcal{I}$ can (usually) be represented as a sequence $(I_1, I_2, \dots)$ where I_n is the part of I that is defined for "security parameter" n. The distribution over the infinite object $I \leftarrow \mathcal{I}$ could naturally be defined through finite distributions $\mathcal{D}_i$ over the finite space of I_i. Caratheodory's extension theorem shows that such finite probability distributions could always be extended consistently to a measure space over the full infinite space of $I \leftarrow \mathcal{I}$ (see Theorem 4.6 of [34] for a proof).

Definition 2.4 (Oracle-fixed Constructions in Idealized Models [39])**.** We say a primitive $\mathcal{P}$ has an *oracle-fixed* black-box construction in the idealized model $\mathcal{I}$ if there is an oracle-aided algorithm P such that:

- **Completeness:** P^I implements $\mathcal{P}$ correctly for every $I \leftarrow \mathcal{I}$.
- **Black-box security:** Let A be an oracle-aided adversary $A^{\mathcal{I}}$ where the *query complexity* of A is bounded by the specified complexity of the attacks for primitive $\mathcal{P}$. For example if $\mathcal{P}$ is polynomially secure (resp., quasi-polynomially secure), then A only asks a polynomial (resp., quasi-polynomial) number of queries but is computationally unbounded otherwise. Then, for any such A, with measure one over the choice of $I \xleftarrow{\$} \mathcal{I}$, it holds that A does *not* break P^I.

Remark 2.5 (Oracle-fixed vs. Oracle-mixed Constructions)**.** We called the constructions of Definition 2.4 "oracle-fixed" because many constructions in idealized models use an "oracle-mixed" security definition. In an oracle-mixed construction P of a primitive $\mathcal{P}$ in an idealized model $\mathcal{I}$, the completeness is defined similarly to Definition 2.4, but when it comes to security, the advantage of A in breaking the scheme is calculated *also over the randomness of* $\mathcal{I}$. Even though oracle-fixed constructions seem to enjoy a stronger security guarantee than oracle-mixed ones, it can be shown that the oracle-fixed security does not imply oracle-mixed security when the advantage of the attack is only $1/\operatorname{poly}(n)$. For example consider a trivial primitive in the Boolean random oracle model $\mathcal{B}$ in which a trivial attacker A succeeds in its attack over security parameter n if $\mathcal{B}$ is equal to 0 over the first $\log(n)$ queries. Then the only oracle for which A succeeds in its attack for an infinite sequence of security parameters is the constant zero oracle, which has a measure zero of being sampled.[11] However, looking ahead, the proof of our main theorem shows that when the attacker

[11] In [39] oracle-fixed and orale-mixed constructions are, in order, called strong and weak constructions. However, exactly because of such cases where oracle-fixed$\not\Rightarrow$oracle-mixed we did not use the same terminology as strong vs. weak might might be very insightful.

achieves constant $\Omega(1)$ advantage over the trivial bound, an oracle-fixed black-box construction is also an oracle-mixed black-box construction.

In what follows, unless specified otherwise, by constructions in idealized models we refer to oracle-fixed black-box constructions.

2.2 Borrowed Results

Theorem 2.6 ([13]). *Suppose one-way functions exist,* $\mathbf{NP} \not\subseteq \mathbf{coAM}$*, and* $\delta, \varepsilon \colon \mathbb{N} \mapsto [0,1]$ *are such that* $2\varepsilon(n) + 3\delta(n) < 1 - 1/\operatorname{poly}(n)$*, then there is no* (ε, δ)*-approximate statistical CIO for all poly-size circuits.*

Theorem 2.7 ([42, 38]). *Suppose* O' *is an approximately correct obfuscation algorithm with error at most* ε' *in idealized model* $\mathcal{I}$ *where* $\mathcal{I}$ *is random trapdoor permutation oracle or the degree-*$O(1)$ *graded encoding model for finite rings. Suppose* $\varepsilon'' \geq 1/\operatorname{poly}(n)$*. Then there is another obfuscation algorithm* O *in the plain model such that:*

- *The running time of* O *is* $\operatorname{poly}(n/\varepsilon''(n))$ *where* n *is the size of the input circuit and it is approximately correct with error at most* $\varepsilon = \varepsilon' + \varepsilon''$*.*
- *There is a simulator* Sim *in the idealized model that runs in time* $\operatorname{poly}(n/\varepsilon''(n))$*, and for any circuit* C*, the distributions* $\mathsf{Sim}^{\mathcal{I}}(O'^{\mathcal{I}}(C))$ *and* $O(C)$ *have statistical distance* $\operatorname{negl}(|C|)$*.*

2.3 Measure Theoretic Tools

By a *probability space* we mean a *measure space* with total measure equal to one, and by $\Pr[E]$ we denote the measure of the measurable set E. For a sequence of measurable sets $\mathcal{E} = (E_1, E_2, \dots)$ defined over some measure space, the limit supremum of $\mathcal{E}$ is defined as $\operatorname{limSup}(\mathcal{E}) = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m$. It can be shown that $\operatorname{limSup}(\mathcal{E})$ is measurable if E_i is so for all i.

Lemma 2.8 (Borel–Cantelli [12, 17]). *Let* $\mathcal{E} = (E_1, E_2, \dots)$ *be a sequence of measurable sets over some probability space, and* $\sum_{n=1}^{\infty} \Pr[E_i] = O(1)$*. Then* $\operatorname{limSup}(\mathcal{E})$ *has measure zero.*

The following lemma follows from Exercise 2 of Section 7.3 of [33]. For completeness we give a proof using continuity of probability.

Lemma 2.9. *If* $\mathcal{E} = (E_1, E_2, \dots)$ *is a sequence of measurable sets over some probability space, and* $\Pr[E_i] \geq \delta$ *for all* $i \in \mathbb{N}$*, then* $\Pr[\operatorname{limSup}(\mathcal{E})] \geq \delta$*.*

Proof. We use the following well-known lemma whose proof could be found in [1] Proposition 37, Part (iii).

Lemma 2.10 (Continuity of Probability). *Let* $B_1 \supseteq B_2 \supseteq \dots$ *be a sequence of measurable sets over some measure space, and* $\Pr[B_1] < \infty$*. Then* $\Pr\left[\bigcap_{n=1}^{\infty} B_n\right] = \lim_{n\to\infty} \Pr[B_n]$*.*

Now let $B_n = \bigcup_{m=n}^{\infty} E_m$, and so $\operatorname{limSup}(\mathcal{E}) = \bigcap_{n=1}^{\infty} B_n$. Since the measure space is a probability space, thus we have $\Pr[B_1] \leq 1$, and we can apply the above lemma to conclude that

$$\lim_{n\to\infty} \Pr[B_n] = \Pr\left[\bigcap_{n=1}^{\infty} B_n\right] = \Pr[\operatorname{limSup}(\mathcal{E})].$$

Finally, because $\Pr[B_n] \geq \Pr[E_i] \geq \delta$ for every n, we get $\delta \leq \lim_{n\to\infty} \Pr[B_n] = \Pr[\operatorname{limSup}(\mathcal{E})]$. $\square$

3 Proving the Main Separation

In this section we formally prove our main result. First we formalize the statement by specifying the way $\mathcal{P}$ is constructed in the idealized models.

Theorem 3.1 (Main Result)**.** *Assuming the existence of one-way functions and* $\mathbf{NP} \not\subseteq \mathbf{coAM}$*, there is no fully-black-box construction of IO from any primitive* $\mathcal{P}$ *that has a oracle-fixed black-box construction in the random trapdoor permutation oracle or the degree-*$O(1)$ *graded encoding model for any finite ring.*

In fact, we prove a stronger separation that holds for approximate computational CIO as well.

Theorem 3.2. *Assuming there is no* (ε, δ)*-approximate statistical CIO, there is no fully-black-box construction of* (ε', δ')*-approximate computational CIO for any* $\varepsilon' \leq \varepsilon - n^{-\Omega(1)}, \delta' \leq \delta - \Omega(1)$ *from any of the primitives listed in Theorem 3.1.*

Proving Theorem 3.1 using Theorems 2.6 and 3.2. Theorem 2.6 rules out (ε, δ)-approximate statistical CIO (assuming OWFs and $\mathbf{NP} \not\subseteq \mathbf{coAM}$) for some $\varepsilon = 1/\operatorname{poly}(n)$ and $\delta = 0.3$. Thus, if we choose $\varepsilon' = \varepsilon/2$ and $\delta' = \delta/2$, then Theorem 3.1 follows from Theorems 3.2 and 2.6.

In the following we will focus on proving Theorem 3.2.

Remark 3.3 (The need for constant δ.)**.** Our proof of Theorem 3.2 crucially relies on the fact that $\delta - \delta' \geq \Omega(1)$ which in turn requires $\delta \geq \Omega(1)$. Thus, the separation holds because the attacker of [13] could achieve $\delta \approx 1/3$ (as opposed to just $1/\operatorname{poly}(n)$). More technically, our proof will make use of Lemma 2.9 rather than the Borel-Cantelli lemma, and that is the source of our need for $\delta \geq \Omega(1)$. However, in case one can improve the result of [13] to cover the setting of $\varepsilon = 1/\operatorname{poly}(n)$ and $\delta = 1 - \alpha$ for arbitrary small $\alpha = 1/\operatorname{poly}(n)$, then our Theorem 3.2 could be improved to any $\delta' = \delta - 1/\operatorname{poly}(n)$. In fact the proof will be simple and will not use our Lemma 2.9 and could be based on the Borel-Cantelli lemma (see the end of this section for a sketch).

Remark 3.4 (Ruling out relativizing constructions)**.** In Theorem 3.1 we focus on ruling out fully-black-box constructions. However, the proof can be extended to rule out relativizing constructions (of IO from the set of listed primitives) using standard techniques and the fact that an optimal statistical distinguisher can be implemented in **PSPACE**. In particular, the separating oracle would be a random sample from the idealized oracle $I \leftarrow \mathcal{I}$ and an oracle for a **PSPACE**-complete oracle. However, interestingly, in our case the sampled $I \leftarrow \mathcal{I}$ would only work with *constant* measure (which is enough since it is still a positive measure) due to using Lemma 2.9 as opposed to measure one, which is typically the case in black-box separations.

of Theorem 3.2. In the following, let $\mathcal{Q}$ denote the primitive of (ε', δ')-approximate computational CIO. Also let $\mathcal{P}$ be any primitive that can be constructed in the idealized models listed in Theorem 3.1 (according to Definition 2.4), and let P be the implementation of $\mathcal{P}$ relative to $\mathcal{I}$.

For sake of contradiction, in the following we let Q be the fully-black-box construction of $\mathcal{Q}$ from $\mathcal{P}$. First we recall a composition lemma from [39] showing that $\mathcal{Q}$ could also be implemented relative to $\mathcal{I}$ as well.[12] Then we rule out the existence of black-box constructions of $\mathcal{Q}$ from $\mathcal{I}$ to conclude that Q could not exist.

[12][39] proved a variant of Lemma 3.5 for semi-black-box constructions, and sketched the proof for fully-black-box case. For sake of completeness here we recall the proof for fully-black-box constructions.

Lemma 3.5 (Composition lemma [39]). *Suppose Q is a fully-black-box construction of $\mathcal{Q}$ from $\mathcal{P}$, and suppose P is an (oracle-fixed black-box) implementation of $\mathcal{P}$ relative to $\mathcal{I}$. Then Q^P is an (oracle-fixed black-box) implementation of $\mathcal{Q}$ relative to the same idealized model $\mathcal{I}$.*

Proof. It is easy to see that Q^P is an implementation of $\mathcal{Q}$ relative to $\mathcal{I}$ (by completeness of the constructions P and Q), and so the completeness holds. The proof of security follows. For sake of contradiction, let $A^{\mathcal{I}}$ be any efficient query successful attacker against the implementation Q^P (of $\mathcal{Q}$) in the idealized model $\mathcal{I}$ which rules out its oracle-fixed black-box property. Namely, there is a non-zero measure fraction of $I \xleftarrow{\$} \mathcal{I}$ for which it holds that A^I breaks the security of Q^{P^I}. For any such fixed I, the security reduction $S^{A^I,I}$ (of the fully-black-box construction Q of P) would break the security of P^I. By combining the algorithms S and A we get that the efficient query attacker $(S^A)^I = B^I$ breaks the security of P^I with non-zero measure over the sampled oracle $I \xleftarrow{\$} \mathcal{I}$. But this contradicts the assumption that $\mathcal{P}$ is securely realized in $\mathcal{I}$ in an oracle-fixed black-box way. Therefore Q^P is also an *oracle-fixed black-box* construction of $\mathcal{Q}$ relative to $\mathcal{I}$. □

In the following we will use Theorems 2.7 and 2.6 to rule out the possibility of any *oracle-fixed black-box* construction of $\mathcal{Q}$ relative to $\mathcal{I}$ which (with Lemma 3.5) shows that Q could not exist.

Let $\varepsilon'' = \varepsilon - \varepsilon' \geq 1/\operatorname{poly}(n)$ and $\delta'' = \delta - \delta' \geq \Omega(1)$. Since P is a construction of $\mathcal{P}$ relative to $\mathcal{I}$, we have that $O'^{\mathcal{I}} = (Q^P)^{\mathcal{I}}$ is an ε'-approximate obfuscation mechanism relative to $\mathcal{I}$. Let O be the ε-approximate obfuscator in the plain model that exists due to Theorem 2.7. The assumption in Theorem 3.2 is that O cannot be an (ε, δ)-approximate statistical CIO. Therefore, there exists a computationally unbounded adversary A and an infinite sequence of circuit pairs $(C_0^1, C_1^1), \dots, (C_0^i, C_1^i), \dots$ such that for all i: $|C_0^i| = |C_1^i|$, $C_0^i \equiv C_1^i$, and $\Pr_{b\leftarrow\{0,1\}}[A(O(C_b^i)) = b] \geq 1/2 + \delta(n)/2$.

Now consider another attacker A' in the idealized model $\mathcal{I}$ which, given a circuit B' as input, runs the simulator of Theorem 2.7 to get the circuit $B = \mathsf{Sim}^{\mathcal{I}}(B')$ and then runs A over B to output whatever A does. By the property of the simulator Sim we conclude that A' is an efficient query (computationally unbounded) attacker in the idealized model $\mathcal{I}$ that achieves

$$\Pr_{b\leftarrow\{0,1\}, I\leftarrow\mathcal{I}}[A'^I(O'^I(C_b^i)) = b] \geq 1/2 + \delta(n)/2 - \operatorname{negl}(n)$$

where $|C_0^i| = |C_1^i| = n$.

A crucial point is that the above probability is *also over the randomness of the oracle* $I \leftarrow \mathcal{I}$ for every i, while we are interested in *fixing* $I \leftarrow \mathcal{I}$ and getting a successful attack for infinitely many pairs of circuits at the same time. By a simple averaging argument we can get:

$$\Pr_{I\leftarrow\mathcal{I}}\left[\Pr_{b\leftarrow\{0,1\}}[A'^I(O'^I(C_b^i)) = b] \geq 1/2 + \delta'(n)/2\right] \geq \delta''(n)/2 - \operatorname{negl}(n).$$

Thus, if we define the event E_i over the sampled oracle $I \leftarrow \mathcal{I}$ as:

$$E_i \text{ holds if: } \Pr_{b\leftarrow\{0,1\}}[A'^I(O'^I(C_b^i)) = b] \geq 1/2 + \delta'/2$$

then we get $\Pr[E_i] \geq \delta''(n)/2 - \operatorname{negl}(n) \geq \delta''/3$ for every $i \in \mathbb{N}$. Now we can apply Lemma 2.9 to conclude that, with probability at least $\delta''/3$ over the choice of $I \leftarrow \mathcal{I}$, an infinite number of the events E_i's would happen at the same time for I. We call $I \leftarrow \mathcal{I}$ a good oracle if it is indeed the case that infinitely many of the events E_i's happen over I. By definition, for any good oracle I, the attacker A' successfully breaks $(Q^P)^I$ (as an implementation of $\mathcal{Q}$ in model $\mathcal{I}$) over infinitely many pairs of circuits while asking only an efficient number of oracle queries to I. The existence of such A' who breaks $(Q^P)^I$ for non-zero (in fact $\geq \delta''/3$) measure of the choice of the oracles $I \leftarrow \mathcal{I}$ prevents Q^P from being a oracle-fixed black-box construction of $\mathcal{Q}$ relative to $\mathcal{I}$. $\square$

Case of $\delta' \approx 1 - 1/\operatorname{poly}(n)$. Theorem 3.2 was sufficient for us to derive Theorem 3.1, however that is not the strongest separation one can imagine for approximate computational CIO as it does not cover the case of $1 - 1/\operatorname{poly}(n)$. The work of [13] shows that whenever $2\varepsilon + \delta > 1$ then there is in fact a way to achieve (ε, δ)-approximate statistical CIO. Thus one can imagine the possibility that the result of [13] could ultimately be improved to rule out (ε, δ)-approximate statistical CIO for $O(\varepsilon) + \delta < 1 - 1/\operatorname{poly}(n)$. Below, we show that such a result, if proved, could be used to derive lower bounds on the complexity of (ε', δ')-approximate computational CIO for $\delta' \approx 1 - 1/\operatorname{poly}(n)$.

Theorem 3.6. *If there is no (ε, δ)-approximate statistical CIO for $\delta = 1 - \rho$ for sufficiently small $\rho = 1/\operatorname{poly}(n)$ (e.g., $\rho = 1/n^4$ suffices), then there is no fully-black-box construction of $(\varepsilon', \delta' = 1 - \sqrt{\rho})$-approximate computational CIO for any $\varepsilon' \leq \varepsilon - n^{-\Omega(1)}$ from the primitives of Theorem 3.1.*

Thus, the main difference between Theorem 3.2 and Theorem 3.6 is that in Theorem 3.6 we cover the case of $\delta' = 1 - 1/\operatorname{poly}(n)$, but we also rely on stronger assumption that $\delta = 1 - 1/\operatorname{poly}(n)$.

of Theorem 3.6. The proof is identical to that of Theorem 3.2 except for the following. Since the attackers A and A' will succeed in guessing the correct circuit with probability $1 - 1/\operatorname{poly}(1) \approx 1$ we can do a better averaging argument to get a better attack after fixing the oracle. Namely, define the event E_i as:

$$E_i \text{ holds if: } \Pr_{b \leftarrow \{0,1\}}[A'^I(O'^I(C_b^i)) = b] \geq 1 - \sqrt{\rho(n)/2}$$

where n is the size of the circuits C_0^i, C_1^i. Then we can conclude that $\Pr[E_i] \geq 1 - 10\sqrt{\rho(n)}$. Now, since the events E_i happen with large probability and that $\sum_n 10\sqrt{\rho(n)} < \infty$ we can apply the Borel-Cantelli lemma (Lemma 2.8) to conclude that with measure *one* over the choice of the oracle $I \leftarrow \mathcal{I}$ all but finitely many of E_i's would happen. The rest of the proof remains unchanged. $\square$

Acknowledgement. W thank Kasra Alishahi and Erfan Salavati for the reference of Lemma 2.9, and we thank Leonid Alexandrovich Petrov for pointing out the name of Lemma 2.10 to us.

References

[1] Michelle Alexopoulos. Notes on set theory and probability theory, 2003. `http://www.biostat.umn.edu/~dipankar/pubh7440/ProbSets.pdf`. 7

[2] Prabhanjan Ananth, Divya Gupta, Yuval Ishai, and Amit Sahai. Optimizing obfuscation: Avoiding barrington's theorem. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, CCS '14, pages 646–658, New York, NY, USA, 2014. ACM. 2

[3] Paul Baecher, Christina Brzuska, and Marc Fischlin. Notions of black-box reductions, revisited. In *Advances in Cryptology-ASIACRYPT 2013*, pages 296–315. Springer, 2013. 5

[4] Boaz Barak, Sanjam Garg, Yael Tauman Kalai, Omer Paneth, and Amit Sahai. Protecting obfuscation against algebraic attacks. In *Advances in Cryptology–EUROCRYPT 2014*, pages 221–238. Springer, 2014. 2

[5] Boaz Barak, Oded Goldreich, Rusell Impagliazzo, Steven Rudich, Amit Sahai, Salil Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs. In *CRYPTO 2001*, pages 1–18. Springer, 2001. 2

[6] Boaz Barak, Oded Goldreich, Russell Impagliazzo, Steven Rudich, Amit Sahai, Salil Vadhan, and Ke Yang. On the (im) possibility of obfuscating programs. *Journal of the ACM (JACM)*, 59(2):6, 2012. 2

[7] Mihir Bellare and Phillip Rogaway. Random Oracles are Practical: A Paradigm for Designing Efficient Protocols. In *ACM Conference on Computer and Communications Security*, pages 62–73, 1993. 3

[8] Mihir Bellare and Moti Yung. Certifying permutations: Noninteractive zero-knowledge based on any trapdoor permutation. *J. Cryptology*, 9(3):149–166, 1996. 3

[9] Nir Bitansky, Omer Paneth, and Daniel Wichs. *Theory of Cryptography: 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, chapter Perfect Structure on the Edge of Chaos, pages 474–502. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. 2

[10] Manuel Blum, Alfredo De Santis, Silvio Micali, and Giuseppe Persiano. Noninteractive zero-knowledge. *SIAM Journal on Computing*, 20(6):1084–1118, 1991. 3

[11] Dan Boneh, Divya Gupta, Ilya Mironov, and Amit Sahai. *Advances in Cryptology - EUROCRYPT 2015: 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part II*, chapter Hosting Services on an Untrusted Cloud, pages 404–436. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[12] Émile Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, 27(1):247–271, 1909. 7

[13] Zvika Brakerski, Christina Brzuska, and Nils Fleischhacker. On statistically secure obfuscation with approximate correctness. Cryptology ePrint Archive, Report 2016/226, 2016. `http://eprint.iacr.org/`. 2, 3, 4, 5, 7, 8, 10

[14] Zvika Brakerski and Guy N Rothblum. Virtual black-box obfuscation for all circuits via generic graded encoding. In *Theory of Cryptography Conference, TCC*, pages 1–25. Springer, 2014. 2

[15] Ran Canetti, Yael Tauman Kalai, and Omer Paneth. On obfuscation with random oracles. Cryptology ePrint Archive, Report 2015/048, 2015. `http://eprint.iacr.org/`. 2, 3, 4

[16] Ran Canetti, Huijia Lin, Stefano Tessaro, and Vinod Vaikuntanathan. *Theory of Cryptography: 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part II*, chapter Obfuscation of Probabilistic Circuits and Applications, pages 468–497. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[17] Francesco Paolo Cantelli. Sulla probabilita come limite della frequenza. *Atti Accad. Naz. Lincei*, 26(1):39–45, 1917. 7

[18] Jung Hee Cheon, Kyoohyung Han, Changmin Lee, Hansol Ryu, and Damien Stehlé. *Advances in Cryptology – EUROCRYPT 2015: 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part I*, chapter Cryptanalysis of the Multilinear Map over the Integers, pages 3–12. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[19] Jean-Sebastien Coron, Tancrede Lepoint, and Mehdi Tibouchi. Cryptanalysis of two candidate fixes of multilinear maps over the integers. Cryptology ePrint Archive, Report 2014/975, 2014. `http://eprint.iacr.org/`. 2

[20] Dana Dachman-Soled, Jonathan Katz, and Vanishree Rao. *Theory of Cryptography: 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part II*, chapter Adaptively Secure, Universally Composable, Multiparty Computation in Constant Rounds, pages 586–613. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[21] Sanjam Garg, Craig Gentry, and Shai Halevi. Candidate multilinear maps from ideal lattices. In *Eurocrypt*, volume 7881, pages 1–17. Springer, 2013. 2

[22] Sanjam Garg, Craig Gentry, Shai Halevi, and Mariana Raykova. *Theory of Cryptography: 11th Theory of Cryptography Conference, TCC 2014, San Diego, CA, USA, February 24-26, 2014. Proceedings*, chapter Two-Round Secure MPC from Indistinguishability Obfuscation, pages 74–94. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. 2

[23] Sanjam Garg, Craig Gentry, Shai Halevi, Mariana Raykova, Anant Sahai, and Brent Waters. Candidate indistinguishability obfuscation and functional encryption for all circuits. In *IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 40–49. IEEE, 2013. 2

[24] Sanjam Garg, Craig Gentry, Amit Sahai, and Brent Waters. Witness encryption and its applications. In *Proceedings of the Forty-fifth Annual ACM Symposium on Theory of Computing*, STOC '13, pages 467–476, New York, NY, USA, 2013. ACM. 2

[25] Sanjam Garg, Pratyay Mukherjee, and Akshayaram Srinivasan. Obfuscation without the vulnerabilities of multilinear maps. Cryptology ePrint Archive, Report 2016/390, 2016. `http://eprint.iacr.org/`. 2

[26] Sanjam Garg and Antigoni Polychroniadou. *Theory of Cryptography: 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part II*, chapter Two-Round Adaptively Secure MPC from Indistinguishability Obfuscation, pages 614–637. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[27] Craig Gentry, Shai Halevi, Hemanta K. Maji, and Amit Sahai. Zeroizing without zeroes: Cryptanalyzing multilinear maps without encodings of zero. Cryptology ePrint Archive, Report 2014/929, 2014. `http://eprint.iacr.org/`. 2

[28] Craig Gentry, Allison B Lewko, Amit Sahai, and Brent Waters. Indistinguishability obfuscation from the multilinear subgroup elimination assumption. *IACR Cryptology ePrint Archive*, 2014:309, 2014. 2

[29] Craig Gentry and Alice Silverberg. Hierarchical id-based cryptography. In *Proceedings of the 8th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology*, ASIACRYPT '02, pages 548–566, London, UK, UK, 2002. Springer-Verlag. 3

[30] Oded Goldreich. Basing non-interactive zero-knowledge on (enhanced) trapdoor permutations: The state of the art. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 406–421. Springer, 2011. 3

[31] Shafi Goldwasser and Guy N Rothblum. On best-possible obfuscation. In *Theory of Cryptography*, pages 194–213. Springer, 2007. 2

[32] Shafi Goldwasser and Guy N Rothblum. On best-possible obfuscation. *Journal of Cryptology*, 27(3):480–505, 2014. 2

[33] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*. Oxford university press, 2001. 7

[34] Thomas Holenstein. Complexity theory, 2015. `http://www.complexity.ethz.ch/education/Lectures/ComplexityFS15/skript_printable.pdf`. 6

[35] Jeremy Horwitz and Ben Lynn. Toward hierarchical identity-based encryption. In LarsR. Knudsen, editor, *Advances in Cryptology EUROCRYPT 2002*, volume 2332 of *Lecture Notes in Computer Science*, pages 466–481. Springer Berlin Heidelberg, 2002. 3

[36] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *ACM Symposium on Theory of Computing (STOC)*, pages 44–61. ACM Press, 1989. 2, 3, 4

[37] Huijia Lin. Indistinguishability obfuscation from constant-degree graded encoding schemes. Cryptology ePrint Archive, Report 2016/257, 2016. `http://eprint.iacr.org/`. 2

[38] Mohammad Mahmoody, Ameer Mohammed, and Soheil Nematihaji. More on impossibility of virtual black-box obfuscation in idealized models. Cryptology ePrint Archive, Report 2015/632, 2015. `http://eprint.iacr.org/`. 2, 3, 4, 5, 7

[39] Mohammad Mahmoody, Ameer Mohammed, Soheil Nematihaji, Rafael Pass, and Abhi Shelat. Lower bounds on assumptions behind indistinguishability obfuscation. In *Theory of Cryptography*, pages 49–66. Springer, 2016. 2, 3, 6, 8, 9

[40] Eric Miles, Amit Sahai, and Mor Weiss. Protecting obfuscation against arithmetic attacks. Cryptology ePrint Archive, Report 2014/878, 2014. `http://eprint.iacr.org/`. 2

[41] Moni Naor and Moti Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *In Proc. of the 22nd STOC*, pages 427–437. ACM Press, 1990. 3

[42] Rafael Pass and abhi shelat. Impossibility of vbb obfuscation with ideal constant-degree graded encodings. Cryptology ePrint Archive, Report 2015/383, 2015. `http://eprint.iacr.org/`. 2, 3, 4, 5, 7

[43] Rafael Pass, Karn Seth, and Sidharth Telang. Indistinguishability obfuscation from semantically-secure multilinear encodings. In *CRYPTO 2014*, pages 500–517. Springer, 2014. 2

[44] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In *Theory of Cryptography Conference, TCC 2004*, volume 2951 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2004. 2, 3, 5

[45] Amit Sahai and Brent Waters. How to use indistinguishability obfuscation: Deniable encryption, and more. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, pages 475–484. ACM, 2014. 2, 3

[46] Brent Waters. *Advances in Cryptology – CRYPTO 2015: 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part II*, chapter A Punctured Programming Approach to Adaptively Secure Functional Encryption, pages 678–697. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2

[47] Joe Zimmerman. *Advances in Cryptology - EUROCRYPT 2015: 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Sofia, Bulgaria, April 26-30, 2015, Proceedings, Part II*, chapter How to Obfuscate Programs Directly, pages 439–467. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. 2