Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 91 (2016)

Structure vs. Hardness through the Obfuscation Lens*

Nir Bitansky Akshay Degwekar Vinod Vaikuntanathan

Abstract

Much of modern cryptography, starting from public-key encryption and going beyond, is based on the hardness of structured (mostly algebraic) problems like factoring, discrete log, or finding short lattice vectors. While structure is perhaps what enables advanced applications, it also puts the hardness of these problems in question. In particular, this structure often puts them in low (and so called structured) complexity classes such as $\mathsf{NP} \cap \mathsf{coNP}$ or statistical zero-knowledge (SZK).

Is this structure really necessary? For some cryptographic primitives, such as one-way permutations and homomorphic encryption, we know that the answer is *yes* — they imply hard problems in $\mathsf{NP} \cap \mathsf{coNP}$ and SZK, respectively. In contrast, one-way functions do *not* imply such hard problems, at least not by *black-box reductions*. Yet, for many basic primitives such as public-key encryption, oblivious transfer, and functional encryption, we do not have any answer.

We show that the above primitives, and many others, do *not* imply hard problems in $\mathsf{NP} \cap \mathsf{coNP}$ or SZK via black-box reductions. In fact, we first show that even the very powerful notion of Indistinguishability Obfuscation (IO) does not imply such hard problems, and then deduce the same for a large class of primitives that can be constructed from IO.

Keywords: Indistinguishability Obfuscation, Statistical Zero-knowledge, $\mathsf{NP} \cap \mathsf{coNP}$, Structured Hardness, Collision-Resistant Hashing.

*MIT. E-mail: {nirbitan,akshayd,vinodv}@csail.mit.edu. Research supported in part by NSF Grants CNS-1350619 and CNS-1414119, Alfred P. Sloan Research Fellowship, Microsoft Faculty Fellowship, the NEC Corporation, a Steven and Renee Finn Career Development Chair from MIT. This work was also sponsored in part by the Defense Advanced Research Projects Agency (DARPA) and the U.S. Army Research Office under contracts W911NF-15-C-0226.

ISSN 1433-8092

Contents

1 Introduction

The last four decades of research in cryptography has produced a host of fantastic objects, starting from one-way functions and permutations to public-key encryption [DH76, RSA78, GM82] and zero-knowledge proofs [GMR85] in the 1980s, all the way to fully homomorphic encryption [RAD78, Gen09, BV11] and indistinguishability obfuscation [BGI+01, GGH+13a] in the modern day.

The existence of all these objects requires at the very minimum that $\mathsf{NP} \not\subseteq \mathsf{BPP}$, but that is hardly ever enough. While one-way functions (OWFs), the most basic cryptographic object, does not seem to require much structure, as we advance up the ranks, we seem to require that certain *structured problems are hard*. For example, conjectured hard problems commonly used in cryptography (especially the public-key kind), such as factoring, discrete logarithms, and shortest (or closest) vectors on lattices all have considerable algebraic structure. On one hand, it is this structure that enables strong applications such as public-key and homomorphic encryption. On the other hand, this structure is also what puts their hardness in question, and is exactly what algorithms may try to exploit in order to solve these problems. There is of course the fear that this structure will (eventually, if not today) deem these problems *easy*. Or, as Barak says more eloquently [Bar13]:

> [...] *based on the currently well studied schemes, structure is strongly associated with (and perhaps even implied by) public key cryptography. This is troubling news, since it makes public key crypto somewhat of an "endangered species" that could be wiped out by a surprising algorithmic advance. Therefore the question of whether structure is inherently necessary for public key crypto is not only of mathematical interest but also of practical importance as well.*

Thus, a fundamental question in cryptography is *what type of structure is necessary for different primitives?* Indeed, the answer to this question may be crucial to our understanding of what are the minimal assumptions required to construct these primitives. While there may be different ways of approaching this question, one main approach, which is also taken in this work, has been through the eyes of complexity theory. That is, we wish to understand which cryptographic primitives require hardness in low (and so called structured) complexity classes such as $\mathsf{NP} \cap \mathsf{coNP}$, TFNP (the class of total NP search problems), or SZK (the class of problems with statistical zero-knowledge proofs).

Aiming to answer this question, one line of research demonstrates that, for some cryptographic primitives, hardness in structured complexity classes is indeed necessary. The existence of one-way permutations (OWPs) requires a hard problem in $\mathsf{NP} \cap \mathsf{coNP}$ [Bra79]; the same holds for restricted cases of public-key encryption schemes satisfying specific structural properties (e.g ciphertext certification) [Bra79, GG98]; homomorphic encryption schemes and non-interactive computational private-information retrieval schemes imply hard problems in SZK [BL13, LV16]; and indistinguishability obfuscation schemes imply a hard problem in $\mathsf{PPAD} \subseteq \mathsf{TFNP}$ (assuming $\mathsf{NP} \not\subseteq \mathsf{ioBPP}$) [BPR15].

Yet, for many primitives such hardness is not known to be inherent. While this is perhaps expected for OWFs, it is also the case for seemingly structured primitives such as collision-resistant hash functions, oblivious transfer, and general public-key encryption schemes. *Do these primitives require hardness in structured complexity classes? Can we prove that they do or that they don't?*

Black-Box Separations. Formalizing this question in a meaningful way requires care. Indeed, it may be easy to formalize a statement of the form "the existence of crypto primitive $\mathcal{P}$ *implies*

hardness in a complexity class $\mathcal{C}$": one just needs to show a reduction from breaking $\mathcal{P}$ to solving problems in $\mathcal{C}$. However, it is not clear how to prove statements of the form "the existence of crypto primitive $\mathcal{P}$ does *not* imply hardness in a complexity class $\mathcal{C}$". For example, it is commonly believed that $\mathsf{NP} \cap \mathsf{coNP}$ *does* contain hard problems. So in a trivial logical sense the existence of such problems is implied by any primitive $\mathcal{P}$. Instead, we follow the methodology of black-box separations, whose study in cryptography was pioneered in a remarkable work by Impagliazzo and Rudich [IR89]. Faced with a similar problem of how to show that a primitive $\mathcal{P}$ (OWFs) cannot be used to construct another primitive $\mathcal{P}'$ (public-key encryption), they prove this cannot be shown through *black-box reductions* — cryptography's de facto technique for showing such implications.

A bit more elaborately, a *fully black-box reduction* [RTV04] of a primitive (or, in our case, a problem) $\mathcal{P}'$ to a primitive $\mathcal{P}$ consists of a black-box *construction* and a black-box *security reduction*. The construction of $\mathcal{P}'$ from $\mathcal{P}$ does not exploit the actual implementation of primitive $\mathcal{P}$, but rather just its input-output interface. The security reduction can use any adversary that breaks (or, in our case, solves) $\mathcal{P}'$ to break $\mathcal{P}$, and is oblivious to the implementation of the adversary (as well as of that of $\mathcal{P}$).

Following [IR89], there has been a rich study of black-box separations in cryptography (see, e.g., [Rud91, Sim98, KST99, GKM+00, GT00, GMR01, BT03, RTV04, HR04, GGKT05, Pas06, GMM07, BM09, HH09, KSS11, BKSY11, DLMM11, GKLM12, DHT12, BBF13, Fis12, Pas13, BB15, HHRS15a] and many others). Most of this study has been devoted to establishing separations between different cryptographic primitives. (In particular, the most relevant to us are the recent works of Asharov and Segev [AS15, AS16] that study black-box separations for indistinguishability obfuscation, which we elaborate on below.) Some of this study puts limitations on basing cryptographic primitives on NP-hardness [GG98, AGGM06, BL13, BB15, LV16].

Going back to our main question of which primitives require structured hardness, we know the following.

- As described above, OWPs imply a hard problem in $\mathsf{NP} \cap \mathsf{coNP}$ [Bra79], homomorphic encryption and PIR imply hard problems in SZK [BL13, LV16] and IO(with OWFs) implies a hard problem in PPAD [BPR15] via *black-box reductions*.

- On the flip side, we know that there are no black-box reductions from hard problems in $\mathsf{NP} \cap \mathsf{coNP}$ to OWFs [BI87, Rud88], and from hard-on-average problems in SZK to OWPs (corollary from [OW93a, OV08, HHRS15a]).

For more advanced primitives, most notably (general) public-key encryption, we do not have results in either direction. In fact, many existing constructions are based on problems in $\mathsf{NP} \cap \mathsf{coNP}$ or SZK. We are thus left with (quite basic) primitives at an unclear state; as far as we know, they may very well imply hard problems in structured complexity classes, even by black-box reductions.

1.1 Our Results

We revisit the relationship between two structured complexity classes, statistical zero-knowledge (SZK) and $\mathsf{NP} \cap \mathsf{coNP}$, and cryptographic primitives. In broad strokes, we show that there are no fully black-box reductions of hard problems in these classes to any one of a variety of cryptographic primitives, including (general) public-key encryption, oblivious transfer, deniable encryption, and functional encryption. More generally, we separate SZK and $\mathsf{NP} \cap \mathsf{coNP}$ from indistinguishability obfuscation (IO). Then, leveraging on the fact that IO can be used to construct a wide variety of

cryptographic primitives in a black-box way, we derive corresponding separations for these primitives.[1] One complexity-theoretic corollary of this result is a separation between SZK and $\mathsf{NP} \cap \mathsf{coNP}$ from the class PPAD [MP91] that captures the complexity of computing Nash Equilibria.

On the positive side, we construct collision-resistant hash functions from a strong form of SZK-hardness and IO. It was previously known [AS15] that IO by itself does not imply collision-resistant hashing in a black-box way; we show that it does if one adds SZK-hardness as a "catalyst".

We now go into more detail on each of the results.

Statistical Zero-Knowledge and Cryptography. The notion of statistical zero-knowledge proofs was introduced in the seminal work of Goldwasser, Micali and Rackoff [GMR85]. The class of *promise problems* with statistical zero-knowledge proofs (SZK) can be characterized by several complete problems, such as *statistical difference* [SV03] and *entropy difference* [GV99]. SZK hardness is known to follow from various number-theoretic problems that are commonly used in cryptography, such as Discrete Logarithms [GK93], Quadratic Residuosity [GMR85], Lattice Problems [GG98, MV03] as well as problems like Graph Isomorphism [GMW91]. As mentioned, we also know that a handful of cryptographic primitives such as homomorphic encryption [BL13], private information retrieval [LV16] and rerandomizable encryption imply SZK-hardness. (On the other hand, $\mathsf{SZK} \subseteq \mathsf{AM} \cap \mathsf{coAM}$ [For89, AH91], and thus, SZK cannot contain NP-hard problems, unless the polynomial hierarchy collapses [BHZ87].)

We ask more generally which cryptographic primitives can be shown to imply such hardness, with the intuition that such primitives are *structured* in a certain way. In particular, whereas one may not expect a seemingly unstructured object like OWFs to imply such hardness, what can we say for instance about OWPs, public-key encryption, or even IO (which has proven to be powerful enough to yield almost any known cryptographic goal)?

We prove that none of these primitives imply such hardness through black-box reductions.

Theorem 1.1 (Informal)**.** *There is no fully black-box reduction of any (even worst-case) hard problem in* SZK *to IO and OWPs.*

Corollary 1.2 (from [SW14, Wat15], Informal)**.** *There is no such reduction to (general) public-key encryption, oblivious transfer, deniable encryption, functional encryption, or any other object that has a black-box reduction to IO and OWPs.*

We would like to elaborate a bit more on what a black-box construction of a hard problem in SZK means. We shall focus on the characterization of SZK by the *statistical difference* promise problem [SV03]. In this problem, an instance is a pair of circuit samplers $C_0, C_1 : \{0,1\}^n \to \{0,1\}^m$ which induce distributions $\boldsymbol{C}_0$ and $\boldsymbol{C}_1$ where the distribution $\boldsymbol{C}_b$ obtained by evaluating the circuit C_b on a uniformly random input. The promise is that the statistical distance $s = \Delta(\boldsymbol{C}_0, \boldsymbol{C}_1)$ of the corresponding distributions is either large (say, $s \geq 2/3$) or small (say, $s \leq 1/3$). The problem, named $\mathbf{SD}^{1/3,2/3}$ (or just $\mathbf{SD}$), is to decide which is the case.

Let us look at a specific example of the construction of such a problem from *rerandomizable encryption*. In a (say, symmetric-key) rerandomizable encryption scheme, on top of the usual encryption and decryption algorithms $(\mathsf{Enc}, \mathsf{Dec})$ there is a ciphertext rerandomization algorithm

[1] More accurately, these primitives follow from IO and OWFs (OWFs), and accordingly our separation addresses IO and OWFs in conjunction. The concept of a black-box reduction from IO and OWF requires clarification and discussion. Here we will follow the framework of Asharov and Segev [AS15]. We elaborate below.

ReRand that can statistically refresh ciphertexts. Namely, for any ciphertext CT encrypting a bit b, $\mathsf{ReRand}(\mathsf{CT})$ produces a ciphertext that is statistically close to a fresh encryption $\mathsf{Enc}_{\mathsf{sk}}(b)$. This immediately gives rise to a hard statistical difference problem [BL13]: given a pair of ciphertexts $(\mathsf{CT}_0, \mathsf{CT}_1)$, decide whether the corresponding rerandomized distributions given by the circuits $(C_0(\cdot), C_1(\cdot)) := (\mathsf{ReRand}(\mathsf{CT}_0; \cdot), \mathsf{ReRand}(\mathsf{CT}_1; \cdot))$ are statistically far or close. Indeed, this corresponds to whether they encrypt the same bit or not, which is hard to decide by the security of the encryption scheme.

A feature of this reduction of hard statistical difference instances to rerandomizable encryption is that, similarly to most reductions in cryptography, it is *fully black-box* [RTV04] in the sense that the circuits C_0, C_1 only make black-box use of the encryption scheme's algorithms, and can in fact be represented as oracle-aided circuits $(C_0^{\mathsf{ReRand}(\cdot)}, C_1^{\mathsf{ReRand}(\cdot)})$. Furthermore, "hardness" can be shown by a black-box security proof that can use any decider for the problem in a black-box way to break the underlying encryption scheme. More generally, one can consider the statistical difference problem relative to different oracles implementing different cryptographic primitives and ask when can hardness be shown based on a black-box reduction. Theorem 1.1 rules out such reductions relative to IO and OWPs (and everything that follows from these in a fully black-box way). For more details, see Section 1.2 and Section 3.

$\mathsf{NP} \cap \mathsf{coNP}$ and Cryptography. Hard (on average) problems in $\mathsf{NP} \cap \mathsf{coNP}$ are known to follow based on several number-theoretic problems in cryptography, such as Discrete Log, Factoring and Lattice Problems [Has88, LLJS90, AR04]. As in the previous section for SZK, we are interested in understanding which cryptographic primitives would imply such hardness, again with the intuition that this implies structure. For instance, it is known [Bra79] that any OWP $f : \{0,1\}^n \to \{0,1\}^n$ implies a hard problem in $\mathsf{NP} \cap \mathsf{coNP}$, e.g. given an index $i \in [n]$ and an image $f(x)$ find the ith preimage bit x_i. In contrast, Blum and Impagliazzo [BI87] and Rudich [Rud88] proved that seemingly unstructured objects like OWFs do not imply hardness in $\mathsf{NP} \cap \mathsf{coNP}$ by fully black-box reductions. In this context, a fully black-box reduction essentially means that the non-deterministic verifiers only make black-box use of the OWF (or OWP in the previous example) and the reduction establishing the hardness is also black-box (in both the decider and the OWF).[2]

But what about more structured primitives such as public-key encryption, oblivious transfer, or even IO? We rule out fully black-box reductions from OWFs (or even *injective OWFs*) and IO to hard problems in $\mathsf{NP} \cap \mathsf{coNP}$. Hence, also for the other primitives, which can be constructed from IO (with OWFs) in a fully black-box way.

Theorem 1.3 (Informal)**.** *There is no fully black-box reduction of any (even worst-case) hard problem in $\mathsf{NP} \cap \mathsf{coNP}$ to IO and OWFs.*

Corollary 1.4 (from [SW14, Wat15], Informal)**.** *There is no such reduction to (general) public-key encryption, oblivious transfer, deniable encryption, functional encryption, or any other object that has a black-box reduction to IO and OWFs.*

Our approach also gives a new (rather different) proof to the original separation between OWFs and $\mathsf{NP} \cap \mathsf{coNP}$ [BI87, Rud88]. For more details, see Section 1.2 and Section 4.

[2] Roughly speaking, [BI87] rule out *perfectly correct constructions*, where the $\mathsf{NP} \cap \mathsf{coNP}$ structure is guaranteed for any implementation of the OWF oracle. In [Rud88], this is generalized also to *almost perfectly correct constructions* that only work for an overwhelming fraction of OWF oracles. We also rule out constructions that are perfectly correct.

We remark that unlike our result for SZK (which ruled out hard *promise problems*), the above result only rules out hard *languages* in NP ∩ coNP. Indeed, Even, Selman, and Yacobi [ESY84] demonstrated promise problems in NP ∩ coNP that are NP-hard. Hence even the assumption P ≠ NP (let alone OWFs) gives us hard promise problems in NP ∩ coNP. (See [Gol06] for further reading.)

Relation to the Work of Asharov and Segev. The flood of IO applications following, starting from [GGH+13b, SW14], has lead many to conjecture that IO may be "complete for cryptography" (assuming also OWFs, or just NP ⊈ ioBBP [KMN+14]). Nevertheless, some cryptographic goals could not be constructed based on IO.

Asharov and Segev [AS15, AS16] were the first to initiate a formal study to understand *the limits of IO*. Our separations for IO are based on their framework [AS15]. We aim to draw the complexity-theoretic boundaries of IO. Indeed, black-box separations from IO require some care, given that the typical use of IO makes non-black-box use of the circuits it obfuscates and thus any associated cryptographic primitive such as OWFs. The Asharov-Segev framework considers obfuscators that take as input circuits with OWF (or OWP) gates. They observe, most known IO-based constructions fall into this category. Thus, a separation in this model allows deriving the corresponding separations between SZK or NP ∩ coNP and a wide variety of cryptographic primitives. See Section 1.2 for more details.

In terms of results, they show that collision-resistant hashing and (domain invariant) OWPs do not have black-box reductions to IO (and OWFs). Our separation of IO and NP ∩ coNP is more general and implies their previous result for OWPs (and gives a rather different proof for this fact). Their result for collision-resistant hashing is not captured by our results (indeed collision-resistance is not known to imply hardness in either SZK or NP ∩ coNP). We also stress that our separation of SZK from IO and OWPs does not follow from their results; indeed, SZK-hardness is not known to imply collision-resistance.[3]

Indistinguishability Obfuscation: Perspective. Since the breakthrough of [GGH+13b], the notion of IO has been extensively studied. While we already understand that IO has far reaching implications, our understanding of how it can be constructed and under what assumptions is still at an early stage. Indeed, basing IO on solid foundations is one of cryptography's greatest challenges today. In this context, we stress that the results presented in this work hold regardless of the state of existing candidates. In fact, even if it turned out that there is no secure realization of IO, the separation of SZK and NP ∩ coNP from primitives such as public-key encryption, which follow from IO, still holds. The expressiveness of IO (established in [GGH+13b, SW14] and onwards) allows us to prove many separations in one shot. (Indeed, three years ago we would have probably addressed each primitive separately.)

As for the search for candidates itself, while at this point candidates are based on lattice-related problems that do break in SZK, our work suggests the theoretical possibility that IO candidates may not require such structure. A similar conclusion is true of course for the much more basic and long-studied question of public-key encryption. Almost all known public-key encryption candidates

[3] We note that previous work [OW93a, OV08] does imply that constant-round statistically-hiding commitments have a black-box reduction to any *hard-on-average* SZK problem. However, [AS15] do not rule these out (but only collision-resistant hashing). We also note that in any case, our result also rules out constructions of worst-case hard SZK problems (rather than average-case hard problems).

rely on very algebraic assumptions (that do break in SZK or NP ∩ coNP). Constructing public key encryption from less structured assumptions remains a fascinating open question. While there has been initial steps trying to diverge from such structure[Ale03, ABW10], there is yet a long way to go.

On TFNP vs. NP ∩ coNP. One of the corollaries of our result is a separation between SZK and NP ∩ coNP from the complexity class PPAD. PPAD, a subclass of total NP search problems called TFNP [MP91], was defined by Papadimitriou [Pap94] and has been shown to capture the complexity of computing Nash equilibria [DGP06, CDT09]. It was recently shown [BPR15] that IO and injective OWFs can be used (in a black-box way) to construct hard problems in PPAD. Put together with our separation, we get that there is no black-box construction of an SZK (resp. NP ∩ coNP) hard problem from PPAD-hardness.[4]

Given that TFNP, which contains PPAD, is commonly thought of as a search version of NP ∩ coNP, it is interesting to note that the result shows that hardness in NP ∩ coNP (of decisional problems) does not follow from hardness in TFNP (aka, hardness of search problems) in a black-box way. Namely, there is no black-box "search-to-decision reduction" between these classes.

The Positive Result: Collision-Resistant Hashing from Strong SZK-Hardness. We end our paper with a positive result. While most of our focus has been on showing that hardness in SZK and NP ∩ coNP does *not* follow from cryptography, here we ask the "inverse question", namely whether certain cryptographic primitives can be built from other cryptographic primitives together with hardness in certain structured complexity classes. Little is known in this direction with the exception of the beautiful work of Ostrovsky and Wigderson [OW93a] who construct a OWF from average-case SZK-hardness, and the recent work of Applebaum and Raykov [AR16] who showed that average-case hardness in the subclass PRE ⊆ SRE ⊆ SZK of languages with a perfect randomized encoding gives us collision-resistant hashing.

We construct collision-resistant hashing from a strong form of SZK-hardness and IO. It was previously known [AS15] that IO by itself does not imply collision-resistant hashing in a black-box way; we show that it does if one adds SZK-hardness as a "catalyst". Slightly more precisely, in the SZK-complete problem $\mathbf{SD}^{1/3,2/3}$ is is required to distinguish between distributions that are 1/3-close from ones that are 2/3-far. We show that IO together with average-case hardness of $\mathbf{SD}^{0,1}$ (a stronger assumption) implies collision-resistant hashing.

Theorem 1.5 (Informal)**.** *Assuming average-case hardness of* $\mathbf{SD}^{0,1}$ *and the existence of IO, there is a collision-resistant hashing scheme.*

Organization. We give an overview of the methodology and techniques used in the following Section 1.2. Section 2 provides required preliminaries. The black-box separation between SZK and IO (plus OWPs) is given in Section 3. The separation between NP ∩ coNP and IO (plus injective OWFs) is given in Section 4. Our construction of collision-resistant hashing functions from IO and SZK hardness is given in Section 5.

[4] We note that in concurrent and independent work, Rosen, Shahaf, and Segev [RSS16] show that one-way functions do not have black-box reductions to PPAD-hardness, which combined with [OW93a], also yields a separation between SZK and PPAD.

1.2 Overview of Techniques

We now give an overview of our approach and main ideas. We start by discussing how to capture fully black-box constructions in the context of indistinguishability obfuscation following [AS15]. We then recall the common methodology for ruling out black-box constructions [IR89, RTV04, BBF13], and explain the main ideas behind our impossibility results for SZK and $\mathsf{NP} \cap \mathsf{coNP}$. In the last part of this section, we outline the construction of collision-resistant hashing from indistinguishability obfuscation and SZK-hardness and the main ideas behind it.

Indistinguishability Obfuscation and Black-Box Constructions. Traditionally, when thinking about a *black-box construction* of one cryptographic primitive $\mathcal{P}'$ (e.g., a pseudo-random generator) from a primitive $\mathcal{P}$ (e.g., a one-way function), we mean that all algorithms in the construction of $\mathcal{P}'$ invoke $\mathcal{P}$ as a black-box, oblivious of its actual implementation. This is hardly the case in constructions based on indistinguishability obfuscation where circuits that explicitly invoke the primitive $\mathcal{P}$ may be obfuscated.

Nonetheless, as observed by Asharov and Segev [AS15], in almost all existing constructions, the code implementing $\mathcal{P}$ is used in a very restricted manner. Typically, obfuscated circuits can be implemented as oracle aided circuits $C^{\mathcal{P}}$ that are completely black-box in $\mathcal{P}$, where $\mathcal{P}$ is some low-level primitive, such as a one-way function. Indeed, in most cases the circuits obfuscated are symmetric-key primitives, such as puncturable pseudo-random functions [SW14], which can be constructed in a black-box way from one-way functions (in some constructions more structured low-level primitives may be used, like injective one-way functions, or one-way permutations). Furthermore, in these constructions, the obfuscator $i\mathcal{O}$ itself is also treated as a black-box.

Accordingly, almost all existing constructions based on indistinguishability obfuscation can be cast into a model in which indistinguishability obfuscation exists for oracle-aided circuits $C^{\mathcal{P}}$, where $\mathcal{P}$ is say a one-way function, and both $\mathcal{P}$ and the obfuscator $i\mathcal{O}$ can only be accessed as black-boxes. On top of that, they can be proven secure in this model by a *black-box reduction* that makes black-box use of $(\mathcal{P}, i\mathcal{O})$ and any attacker against the constructed primitive $\mathcal{P}'$. Such constructions where both the construction itself and the reduction are black-box are called *fully black-box constructions* [RTV04]. Following Asharov and Segev [AS15, AS16], we shall prove our results in this model, ruling out black-box constructions of hard problems in SZK and $\mathsf{NP} \cap \mathsf{coNP}$ based on indistinguishability obfuscation for oracle-aided circuits. Further details follow.

Ruling out Black-Box Reductions. We prove our results in the model described above following the methodology of oracle separations (see e.g. [IR89, Sim98, RTV04, HR04]). Concretely, to prove that there is no fully black-box construction of a primitive $\mathcal{P}'$ from primitive $\mathcal{P}$, we demonstrate oracles $(\Psi, \mathcal{A})$ such that:

- relative to Ψ, there exists a construction $C^{\Psi}_{\mathcal{P}}$ realizing $\mathcal{P}$ that is secure in the presence of $\mathcal{A}$,
- but *any* construction $C^{\Psi}_{\mathcal{P}'}$ realizing $\mathcal{P}'$ can be broken in the presence of $\mathcal{A}$.

Indeed, if such oracles $(\Psi, \mathcal{A})$ exist, then no efficient reduction will be able to use (as a black-box) the attacker $\mathcal{A}$ against $\mathcal{P}'$ to break $\mathcal{P}$ (as the construction of $\mathcal{P}$ is secure in the presence of $\mathcal{A}$). In our case, we would like to apply this paradigm rule out black-box constructions of hard instances in either SZK or $\mathsf{NP} \cap \mathsf{coNP}$ from a low-level primitive (e.g. a one-way function) indistinguishability obfuscation for oracle-aided circuits. We next outline the main ideas behind the construction and analysis of the oracles $(\Psi, \mathcal{A})$ in each of the two cases.

Ruling out Black-Box Constructions of Hard SZK Problems. As explained in the previous section, we focus on the characterization of SZK by its complete problem: the statistical difference problem $\mathbf{SD}$ [SV03]. We demonstrate oracles $(\Psi, \mathcal{A})$ such that relative to Ψ there exist constructions of one-way permutations (OWPs) and IO for circuits with OWP gates, and these constructions are secure in the presence of $\mathcal{A}$. At the same time, $\mathcal{A}$ will decide (in the worst-case) $\mathbf{SD}^{\Psi}$. Since $\mathbf{SD}$ is complete for SZK in a relativizing manner, deciding $\mathbf{SD}^{\Psi}$ suffices to break SZK^{Ψ}. That is, $\mathcal{A}$ will decide *all* instances (C_0^{Ψ}, C_1^{Ψ}) of circuit samplers that only use the IO and OWPs realized by Ψ in a black-box manner. We next explain how each of the two are constructed.

The construction of Ψ follows a general recipe suggested in [AS15, AS16]. The oracle consists of three parts $(f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$ where:

1. f is a random permutation, realizing the one-way permutation primitive.
2. $\mathcal{O}$ is a random injective function, realizing the obfuscation algorithm. It takes as input an oracle-aided circuit $C^{(\cdot)}$ along with randomness r and outputs an obfuscation $\widehat{C} = \mathcal{O}(C, r)$.
3. $\mathsf{Eval}^{\mathcal{O},f}$ realizes evaluation of obfuscated circuits. On input $(\widehat{C}, x)$, it inverts $\mathcal{O}$ to find (C, r), and outputs $C^f(x)$. If $\widehat{C}$ is not in the image of $\mathcal{O}$, it returns $\perp$.

The above construction readily satisfies the syntactic (or "functionality") requirements of one-way permutations and indistinguishability obfuscation. Furthermore, using standard techniques, it is not hard to show that relative to Ψ, the function f is one-way and $\mathcal{O}$ satisfies IO indistinguishability requirement. The challenge is to now come up with an oracle $\mathcal{A}$ that, on one hand, will decide $\mathbf{SD}^{\Psi}$, but on the other, will not compromise the security of the latter primitives.

Recall that deciding $\mathbf{SD}^{\Psi}$ means that given two oracle-aided circuit samplers (C_0, C_1) such that the statistical distance of the corresponding distributions $(\boldsymbol{C}_0^{\Psi}, \boldsymbol{C}_1^{\Psi})$ is $s = \Delta(\boldsymbol{C}_0^{\Psi}, \boldsymbol{C}_1^{\Psi}) \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$, the oracle $\mathcal{A}$ must decide in which of the two intervals s lies, whereas if the promise is not satisfied and $s \in (\frac{1}{3}, \frac{2}{3})$, there is no requirement whatsoever. With this in mind, a first naive attempt would be the following. $\mathcal{A}$ will have unbounded access to Ψ, give a query (C_0, C_1), it would compute compute $s = \Delta(\boldsymbol{C}_0, \boldsymbol{C}_1)$, and simply say whether $s < \frac{1}{2}$ or $s \geq \frac{1}{2}$. While such an oracle would definitely decide $\mathbf{SD}^{\Psi}$, it is not too hard to show that it is simply too powerful, and would not only break IO and OWPs, but would, in fact, allow solving any problem in NP^{Ψ} (or even in PP^{Ψ}). Other naive attempts such as refusing to answer outside the promise intervals, encounter a similar problem.

At high-level, the problem with such oracles is that solutions to hard problems can be easily correlated with "tiny" differences in the statistical distance of the two input circuits, whereas the above oracle may reflect tiny changes when the statistical distance is close to some threshold (1/2 in the above example) on which the oracle changes its behaviour. This motivates our actual definition of $\mathcal{A}$ as a *noisy oracle* that produces its answer, not according to some fixed threshold, but according to a random threshold, chosen afresh for each and every query. Concretely, the oracle, which we call StaDif^{Ψ}, for any query (C_0, C_1), chooses a uniformly random threshold $t \leftarrow (\frac{1}{3}, \frac{1}{3})$, and answers accordingly:

$$\mathsf{StaDif}^{\Psi}(C_0, C_1) = \begin{cases} Y & \text{if } s \geq t \text{ (far distributions)} \\ N & \text{if } s < t \text{ (similar distributions)} \end{cases}.$$

The main challenge in proving that the security of the IO and OWPs realized by $\mathcal{A}$ is not compromised by this oracle is that StaDif^{Ψ} has the power to query Ψ on exponentially many points in

order to compute s. For instance, it may query Ψ on the preimage of a OWP challenge $f(x)$ or of a given obfuscation $\mathcal{O}(C, r)$. The key observation behind the proof is that the oracle's final answer still does not reflect how Ψ behaves locally on random points.

Intuitively, choosing the threshold t at random, for each query (C_0, C_1), guarantees that with high probability t is "far" from the corresponding statistical distance $s = \Delta(C_0^\Psi, C_1^\Psi)$. Thus, changing the oracle Ψ on, say, a single input x, such as the preimage of an OWP challenge $f(x)$, should not significantly change s and will not affect the oracle's answer; that is, unless the circuits query Ψ on x with high probability to begin with. We give a reduction showing that we can always assume that (C_0, C_1) are "smooth", in the sense that they do not make any specific query to Ψ with too high probability.

Following this intuition, we are able to show that through such local changes that go undetected by StaDif^Ψ, we can move to an ideal world where inverting the OWP or breaking IO can be easily shown to be impossible. We refer the reader to Section 3 for further details.

Ruling out Black-Box Constructions of Hard $\mathsf{NP} \cap \mathsf{coNP}$ Problems. As mentioned earlier, a fully black-box construction of hard problems in $\mathsf{NP} \cap \mathsf{coNP}$ is actually known assuming one-way permutations (OWPs), and cannot be ruled out as in the case of SZK. Instead, we rule out constructions from (non-surjective) injective one-way functions (IOWFs) and IO for circuits with IOWF gates. This generalizes several previous results by Blum and Impagliazzo [BI87] and Rudich [Rud88], showing that OWFs do not give hardness in $\mathsf{NP} \cap \mathsf{coNP}$, by Matsuda and Matsuura [MM11], showing that IOWFs do not give OWPs (which are a special case of hardness $\mathsf{NP} \cap \mathsf{coNP}$), and by Asharov and Segev [AS16], showing that OWFs and IO for circuits with OWF gates do not give OWPs. In fact, our approach yields a new (and rather different) proof for each one of these results.

We follow a similar methodology to one we used for the case of SZK. That is, we would like to come up with oracles $(\Psi, \mathcal{A})$ such that Ψ realizes IOWFs and IO for circuits with IOWFs gates, which are both secure in the presence of $\mathcal{A}$, whereas black-box constructions of problems in $\mathsf{NP} \cap \mathsf{coNP}$ from these primitives can be easily solved by $\mathcal{A}$. By black-box constructions here we mean a pair of efficient oracle-aided non-deterministic verifiers $V_0^{(\cdot)}, V_1^{(\cdot)}$ that for every oracle Ψ implementing IOWFs and IO, yield co-languages $\overline{L}^\Psi, L^\Psi$ in $\mathsf{NP}^\Psi \cap \mathsf{coNP}^\Psi$.

The requirement that V_0, V_1 give a language in $\mathsf{NP} \cap \mathsf{coNP}$ for *every* oracle implementing IOWFs and IO follows previous modeling [BI87],[5] and aligns with how we usually think about *correctness* of black-box constructions of cryptographic primitives. For instance, the construction of public-key encryption from trapdoor permutations is promised to be correct, for all oracles implementing the trapdoor permutation. Similarly, the construction of hard $\mathsf{NP} \cap \mathsf{coNP}$ languages from one-way permutations, give an $\mathsf{NP} \cap \mathsf{coNP}$ language for any oracle implementing a permutation.[6]

We stress that a construction where correctness is only guaranteed for particular (even if natural) oracles may definitely exist. This is for example the case if we only consider implementations of IO similar to those presented above in the context of SZK. Indeed, in that construction the implementation of IO has an additional property — it allows identifying *invalid obfuscations* (the

[5] Rudich [Rud88] also considered a slight relaxation of constructions that are correct for an overwhelming fraction of oracles rather than all.

[6] We note that this issue does not come up for black-box constructions of SZK *promise* problems, because the construction is allowed to yield instances that do not obey the promise; there correctness is always guaranteed, and the only question is whether the instances that do satisfy the promise are hard to decide.

Eval oracle would simply return $\perp$ on such obfuscations). This "verifiability" property coupled with the injectivity of obfuscators actually imply a hard problem in $\mathsf{NP} \cap \mathsf{coNP}$ in a black-box way.[7] Our separation thus leverages the fact that IO need not necessarily be verifiable, and rules out constructions that are required to be correct for any implementation of IO, even a non-verifiable one.

Accordingly, the oracles $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$ that we consider are a tweaked version of the oracles considered in the SZK case. Now f is a random injective function that is expanding, rather than a permutation, the oracle $\mathcal{O}$ is defined as before, and the oracle $\mathsf{Eval}^{f,\mathcal{O}}$ is defined as before for valid obfuscations $\widehat{C} \in \text{Image}(\mathcal{O})$ *but is allowed to act arbitrarily for invalid obfuscations.* As for $\mathcal{A}$, this time it is trivially implemented by an oracle Decide^{Ψ} that, given input x, simply returns the unique bit b such that $V_b(x) = 1$, namely it just decides the corresponding language L^{Ψ}.[8]

In the results mentioned above [Rud88, MM11, AS16], it is actually shown that any query to such an oracle can be completely simulated with a small number of queries to Ψ.[9] We do not show such a simulation process. Instead, we take a different approach inspired by our proof for the SZK setting described above. Roughly speaking, we show that somewhat similarly to our statistical difference oracle StaDif^{Ψ}, the oracle Decide^{Ψ} is also rather robust to random local changes. The main observation here is that for any fixed yes-instance $x \in L^{\Psi}$, tweaking Ψ at a random input into a new oracle Ψ', it is likely that x will still be a yes-instance in $L^{\Psi'}$, as long as Ψ' is in our allowed family of oracles and $L^{\Psi'}$ is indeed in $\mathsf{NP}^{\Psi'} \cap \mathsf{coNP}^{\Psi'}$ (and the same is true for no-instances).

In slightly more detail, fixing a witness w such that $V_1^{\Psi}(x, w) = 1$, we can show that since V_1 makes a small number of oracle calls, with high probability tweaking the oracle Ψ at a random place will not affect these oracle calls and thus $V_1^{\Psi'}(x, w) = V_1^{\Psi}(x, w) = 1$. Then, assuming $L^{\Psi'}$ is guaranteed to be in $\mathsf{NP} \cap \mathsf{coNP}$, we can deduce that x must still a yes-instance (other witnesses for this fact may be added or disappear, but this does not change the oracle's answer). In the body, we argue that indeed $L^{\Psi'} \in \mathsf{NP}^{\Psi'} \cap \mathsf{coNP}^{\Psi'}$, where we strongly rely on the fact that arbitrary behavior of Eval is permitted on invalid obfuscations.

Once again, we show that through local changes that go undetected by Decide^{Ψ}, we can move to an ideal world where inverting the IOWF or breaking IO can be easily shown to be impossible. We refer the reader to Section 4 for further details.

Implied Separations. As a result of the two separations discussed above, we can rule out black-box constructions of hard problems in SZK or $\mathsf{NP} \cap \mathsf{coNP}$ from various cryptographic primitives or complexity classes. This essentially includes all primitives that have fully black-box constructions from OWPs (or IOWFs) and IO for circuits with OWP (or IWOF) gates. This includes public-key encryption, oblivious transfer, deniable encryption [SW14], functional encryption [Wat15], delegation, [BGL+15, CHJV15, KLW15], hard (on-average) PPAD instances [BPR15], and more.

We note that there a few applications of IO that do not fall under this characterization. For instance, the construction of IO for Turing machines from IO-based succinct randomized encodings [BGL+15, CHJV15, KLW15] involves obfuscating a circuit that itself outputs (smaller) obfuscated circuits. To capture this, we would need to extend the above model to IO for circuits that can also make IO oracle calls (on smaller circuits). Another example is the construction of non-interactive

[7] E.g. the language of all valid obfuscations and indices i, such that the ith bit of the obfuscated circuit is 1

[8] In the body, we further allow it to answer relative to other languages L' provided that they are indeed in $\mathsf{NP} \cap \mathsf{coNP}$. This allows us later to prove a more general oracle separation. See details in Section 4.6.

[9] More accurately, this is the case for Rudich's result for $\mathsf{NP} \cap \mathsf{coNP}$, whereas for the other results that rule out constructions of one-way permutations, one can simulate an analog of Decide that inverts the permutation.

witness indistinguishable proofs from IO [BP15]. There an obfuscated circuit may get as input another obfuscated circuit and would have to internally run it; furthermore, in this application, the code of the obfuscator is used in a (non-black-box) ZAP. Extending the above model to account for this type of IO applications is an interesting question that we leave for future exploration.

Full Oracle Separations. As explained, the methodology we rely on rules out fully black-box constructions by exhibiting two oracles $(\Psi, \mathcal{A})$, the first which may be used by the construction of a primitive $\mathcal{P}'$ from $\mathcal{P}$, and the second which breaks $\mathcal{P}'$. In the literature (e.g., in [IR89, Sim98]), a stronger type of separation is often shown where a single oracle Γ is exhibited and can be fully accessed, not only by the adversary, but also by the construction (whereas above the construction can only access Ψ, but not $\mathcal{A}$). This rules out an even weaker type of reductions called *relativizing reductions* [RTV04], which guarantee that $\mathcal{P}'$ can be securely realized in any oracle world where $\mathcal{P}$ can. In the body, we show how to extend our result for $\mathsf{NP} \cap \mathsf{coNP}$ to also imply this stronger type of separation.

The Positive Result: Collision-Resistance from IO and SZK-Hardness. We now described the main ideas behind our construction of collision-resistant hash functions. The starting point for the construction is the work of Ishai, Kushilevitz, and Ostrovsky [IKO05] that shows how to construct collision-resistant hash functions from commitments that are additively homomorphic (for simplicity, say over $\mathbb{F}_2$). The idea is simple: we can hash ℓ bits to m bits, where m is the size of a single bit commitment and ℓ can be arbitrarily longer, as follows. The hash key is a commitment $\gamma := (\mathsf{com}(\beta_1), \ldots, \mathsf{com}(\beta_\ell))$ to a random vector $\beta \in \mathbb{F}_2^\ell$, and hashing $x \in \mathbb{F}_2^\ell$, is done by homomorphically computing a commitment to the inner product $\mathsf{CRH}_\gamma(x) = \mathsf{com}(\langle \beta, x \rangle)$. Intuitively, the reason this works is that any collision in CRH_γ reveals a vector that is orthogonal to β and thus leaks information about it and violating the hiding of the commitment.

At a high-level, we aim to mimic the above construction based on obfuscation. As a key for the collision-resistant hash we can obfuscate a program Π_β associated with a random vector β that given x outputs a commitment $\mathsf{com}(\langle \beta, x \rangle)$, where the commitment is derandomized using a PRF.[10] The obfuscation $i\mathcal{O}(\Pi_\beta)$ can be thought of as the commitment to β, and evaluating this program at x, corresponds to homomorphic evaluation. Despite the clear intuition behind this construction, it is not clear how to prove its security based on IO. In fact, by the work of Asharov and Segev [AS15], it cannot be proven based on a black-box reduction as long as plain statistically-binding commitments are used, as these can be constructed from OWPs in a fully black-box manner, and [AS15] rule out black-box constructions of collision-resistant hashing from OWPs and IO for circuits with OWP gates.

We show, however, that relying on a relaxed notion of perfectly-hiding commitments, as well as subexponential hardness of IO and puncturable PRFs, the construction can be proven secure. The perfect hiding of the commitment is leveraged in a probabilistic IO argument [CLTV15] that involves a number of hybrids larger than the overall number of commitments. We then observe that these relaxed commitments follow from average-case hardness of the polar statistical difference problem $\mathbf{SD}^{0,1}$.[11]

[10] In the body, we describe a slightly more abstract construction where inner product is replaced by an arbitrary 2-universal hash function.

[11] Similar SZK-hardness is known to imply statistically-hiding commitments against malicious receivers, but with a larger (constant) number of rounds [OV08].

2 Preliminaries

In this section, we introduce the basic definitions and notation used throughout the paper.

2.1 Conventions

For a distribution D, we denote the process of sampling from D by $x \leftarrow D$. A function $\mathsf{negl} : \mathbb{N} \to \mathbb{R}^+$ is negligible if for every constant c, there exists a constant n_c such that for all $n > n_c$ $\mathsf{negl}(n) < n^{-c}$. We refer to uniform probabilistic polynomial-time algorithms as PPT algorithms.

Randomized Algorithms. As usual, for a random algorithm A, we denote by $A(x)$ the corresponding output distribution. When we want to be explicit about the algorithm using randomness r, we shall denote the corresponding output by $A(x; r)$.

Oracles. We consider *oracle-aided algorithms (or circuits)* that make repeated calls to an oracle Γ. Throughout, we will consider deterministic oracles Γ that are a-priori sampled from a distribution Γ on oracles. More generally, we consider infinite oracle ensembles $\Gamma = \{\Gamma_n\}_{n\in\mathbb{N}}$, one distribution Γ_n for each security parameter $n \in \mathbb{N}$ (each defined over a finite support). For example, we may consider an ensemble $f = \{f_n\}$ where each $f_n : \{0,1\}^n \to \{0,1\}^n$ is a random function. For such an ensemble Γ and an oracle aided algorithm (or circuit) A with finite running time, we will often abuse notation and denote by $A^{\Gamma}(x)$ and execution of A on input x where each of (finite number of) oracle calls that A makes is associated with a security parameter n and is answered by the corresponding oracle Γ_n. When we write $A_1^{\Gamma}, \ldots, A_k^{\Gamma}$ for k algorithms, we mean that they all access the same realization of Γ.

2.2 Indistinguishability Obfuscation for Oracle-Aided Circuits

The notion of *indistinguishability obfuscation* (IO) was introduced by Barak et al. [BGI+01] and the first candidate construction was demonstrated in the work of Garg et al. [GGH+13a]. Since then, IO has given rise to a plethora of applications in cryptography and beyond. Nevertheless, Asharov and Segev [AS15, AS16] demonstrated that IO is insufficient to achieve some cryptographic tasks, most notably (domain-invariant) one-way permutations, collision-resistant hashing, and as a corollary, private information retrieval and (even additively) homomorphic encryption. To formally show such a statement, they introduced the framework of indistinguishability obfuscation for oracle-aided circuits. We follow their framework.

We begin by recalling the notion of two oracle-aided circuits being equivalent, and move on to defining IO relative to oracles.

Definition 2.1. Let C_0 and C_1 be two oracle-aided circuits and let f be a function. C_0 and C_1 are said to be *functionally equivalent relative to* f, denoted as $C_0^f \equiv C_1^f$, if for every input x, $C_0^f(x) = C_1^f(x)$.

Definition 2.2. Let $\mathcal{C} = \{\mathcal{C}_n\}_{n\in\mathbb{N}}$ be a class of oracle aided circuits, where each $C \in \mathcal{C}_n$ is of size n.[12] A PPT algorithm $i\mathcal{O}$ is an *indistinguishability obfuscator* for $\mathcal{C}$ relative to an oracle distribution

[12] As in [AS15], we assume throughout that the size of the obfuscated circuits equals the security parameter. This is only for simplicity of notation, and is without loss of generality as the circuits can be padded up if they are too small, and the security parameter can be polynomially increased if the circuits are too large.

ensemble $\Gamma = \{\Gamma_n\}_{n\in\mathbb{N}}$ if the following conditions are met:

1. **Functionality.** For all $n \in \mathbb{N}$ and for all $C \in \mathcal{C}_n$ it holds that

$$\Pr_{\Gamma, i\mathcal{O}}\left[C^\Gamma \equiv \widehat{C}^\Gamma \mid \widehat{C} \leftarrow i\mathcal{O}^\Gamma(1^n, C)\right] = 1 \ .$$

2. **Indistinguishability.** For any non-uniform PPT distinguisher $D = (D_1, D_2)$ there exists a negligible function negl such that for all $n \in \mathbb{N}$

$$\mathsf{Adv}^{\mathsf{iO}}_{\Gamma, i\mathcal{O}, \mathcal{C}, D}(n) = \left|\Pr\left[\mathsf{Exp}^{\mathsf{iO}}_{\Gamma, i\mathcal{O}, \mathcal{C}, D}(n) = 1\right] - \frac{1}{2}\right| \leq \mathsf{negl}(n)$$

where the random variable $\mathsf{Exp}^{\mathsf{iO}}_{\Gamma, i\mathcal{O}, \mathcal{C}, D}(n)$ is defined via the following experiment:

(a) $b \leftarrow \{0,1\}$.
(b) $(C_0, C_1, \mathsf{state}) \leftarrow D_1^\Gamma(1^n)$ where $C_0, C_1 \in \mathcal{C}_n$ and $C_0^\Gamma \equiv C_1^\Gamma$.
(c) $\widehat{C} \leftarrow i\mathcal{O}^\Gamma(1^n, C_b)$.
(d) $b' = D_2^\Gamma(\mathsf{state}, \widehat{C})$.
(e) If $b = b'$ output 1 else output 0.

We further say that $i\mathcal{O}$ satisfies δ-indistinguishability if the above negligible advantage is at most δ.

3 One-Way Permutations, Indistinguishability Obfuscation, and Hardness in Statistical Zero Knowledge

In this section, we ask which cryptographic primitives imply hardness in the class statistical zero-knowledge (SZK). Roughly speaking, we show that one-way permutations (OWPs) and indistinguishability obfuscation (IO), for circuits with OWP-gates, do not give rise to a black-box construction of hard problems in SZK. This, in turn implies that many cryptographic primitives (e.g., public-key encryption, functional encryption, and delegation), and hardness in certain low-level complexity classes (e.g. PPAD), also do not yield black-box constructions of hard problems in SZK.

We first motivate and define a framework of SZK relative to oracles, define fully black-box constructions of hard SZK problems, and then move on to the actual separation.

3.1 SZK and Statistical Difference

The notion of statistical zero-knowledge proofs was introduced in the seminal work of Goldwasser, Micali and Rackoff [GMR85]. The class of promise problems with statistical zero-knowledge proofs (SZK) can be characterized by several complete problems, such as *statistical difference* [SV03] and *entropy difference* [GV99] (see also [Vad99] and references within). We shall focus on the characterization of SZK by the statistical difference problem. Here an instance is a pair of circuit samplers $C_0, C_1 : \{0,1\}^n \to \{0,1\}^m$ with the promise that the statistical distance $s = \Delta(\boldsymbol{C}_0, \boldsymbol{C}_1)$ of the corresponding distributions is either large (say, $s \geq 2/3$) or small (say, $s \leq 1/3$). The problem is to decide which is the case.

Hard Statistical Difference Problems from Cryptography: Motivation. SZK hardness, and in particular hard statistical difference problems, are known to follow from various number-theoretic and lattice problems that are commonly used in cryptography, such as Decision Diffie-Hellman, Quadratic Residuosity, and Learning with Errors. We ask more generally which cryptographic primitives can be shown to imply such hardness, with the intuition that such primitives are *structured* in a certain way. In particular, whereas one would not expect a completely unstructured object like one-way functions to imply such hardness, what can we say for instance about public-key encryption, or even indistinguishability obfuscation (which has proven to be structured enough to yield almost any known cryptographic goal).

We prove that none of these primitives imply such hardness through the natural class of black-box constructions and security reductions. To understand what a black-box construction of a hard statistical difference problem means, let us look at a specific example of the construction of such a problem from *rerandomizable encryption*. In a (say, symmetric-key) rerandomizable encryption scheme, on top of the usual encryption and decryption algorithms $(\mathsf{Enc}, \mathsf{Dec})$ there is a ciphertext rerandomization algorithm ReRand that can statistically refresh ciphertexts. Namely, for any ciphertext CT encrypting a bit b, $\mathsf{ReRand}(\mathsf{CT})$ produces a ciphertext that is statistically close to a fresh encryption $\mathsf{Enc}(b)$. Note that this immediately gives rise to a hard statistical difference problem: given a pair of ciphertexts $(\mathsf{CT}, \mathsf{CT}')$, decide whether the corresponding rerandomized distributions given by the circuits $(C_0(\cdot), C_1(\cdot)) := (\mathsf{ReRand}(\mathsf{CT}; \cdot), \mathsf{ReRand}(\mathsf{CT}'; \cdot))$ are statistically far or close. Indeed, this corresponds to whether they encrypt the same bit or not, which is hard to decide by the security of the encryption scheme.

A feature of this construction of hard statistical difference instances is that, similarly to most constructions in cryptography, it is *fully black-box* [RTV04] in the sense that the circuits C_0, C_1 only make black-box use of the encryption scheme's algorithms, and can in fact be represented as oracle-aided circuits $(C_0^{\mathsf{ReRand}(\cdot)}, C_1^{\mathsf{ReRand}(\cdot)})$. Furthermore, "hardness" can be shown by a black-box reduction that can use any decider for the problem in a black-box way to break the underlying encryption scheme. More generally, one can consider the statistical difference problem relative to different oracles implementing different cryptographic primitives and ask when can hardness be shown based on a black-box reduction. We will rule out such reductions relative to IO and OWPs (and everything that follows from these in a fully black-box way).

3.2 Fully Black-Box Constructions of Hard SD Problems from IO and OWPs

We start by defining statistical difference problem relative to oracles. In the following definition, for an oracle-aided (sampler) circuit $C^{(\cdot)}$ with n-bit input and an oracle Ψ, we denote by $\boldsymbol{C}^{\Psi}$ the output distribution $C^{\Psi}(r)$ where $r \leftarrow \{0,1\}^n$. For two distributions $\mathbf{X}$ and $\mathbf{Y}$ we denote their statistical distance by $\Delta(\mathbf{X}, \mathbf{Y})$.

Definition 3.1 (Statistical difference relative to oracles)**.** For an oracle Ψ, the statistical difference promise problem relative to Ψ, denoted as $\mathbf{SD}^{\Psi} = (\mathbf{SD}_Y^{\Psi}, \mathbf{SD}_N^{\Psi})$, is given by

$$\mathbf{SD}_Y^{\Psi} = \left\{ (C_0, C_1) \;\middle|\; \Delta(\boldsymbol{C}_0^{\Psi}, \boldsymbol{C}_1^{\Psi}) \geq \frac{2}{3} \right\} ,$$

$$\mathbf{SD}_N^{\Psi} = \left\{ (C_0, C_1) \;\middle|\; \Delta(\boldsymbol{C}_0^{\Psi}, \boldsymbol{C}_1^{\Psi}) \leq \frac{1}{3} \right\} .$$

We now formally define the class of constructions and reductions ruled out. That is, *fully black-box* constructions of hard statistical distance problems from OWPs and IO for OWP-aided circuits. The definition is similar in spirit to those in [AS15, AS16], adapted to our context of SZK-hardness.

Definition 3.2. A fully black-box construction of a hard statistical distance problem from OWPs and IO for the class $\mathcal{C}$ of circuits with OWP-gates consists of a collection of oracle-aided circuit pairs $\Pi^{(\cdot)} = \left\{\Pi_n^{(\cdot)} = \left\{(C_0^{(\cdot)}, C_1^{(\cdot)}) \in \{0,1\}^{n\times 2}\right\}\right\}_{n\in\mathbb{N}}$ and a probabilistic oracle-aided reduction $\mathcal{R}$ that satisfy:

- **Black-box security proof:** There exist functions $q_{\mathcal{R}}(\cdot), \varepsilon_{\mathcal{R}}(\cdot)$ such that the following holds. Let f be any distribution on permutations and let $i\mathcal{O}$ be any distribution on functions such that $\widehat{C}^f \equiv C^f$ for any $C^{(\cdot)}$ and r, where $\widehat{C}^{(\cdot)} := i\mathcal{O}(C^{(\cdot)}, r)$. Then for any probabilistic oracle-aided $\mathcal{A}$ that *decides* Π *in the worst-case*, namely, for all $n \in \mathbb{N}$

$$\Pr_{f,i\mathcal{O},\mathcal{A}}\left[\mathcal{A}^{f,i\mathcal{O}}(C_0, C_1) = B \quad \text{for all} \quad \begin{array}{l}(C_0, C_1) \in \Pi_n, B \in \{Y, N\}\\ \text{such that } (C_0, C_1) \in \mathbf{SD}_B^{f,i\mathcal{O}}\end{array}\right] = 1$$

  the reduction breaks either f or $i\mathcal{O}$, namely, for infinitely many $n \in \mathbb{N}$ either

$$\Pr_{\substack{x\leftarrow\{0,1\}^n\\ f,i\mathcal{O},\mathcal{A}}}\left[\mathcal{R}^{\mathcal{A},f,i\mathcal{O}}(f(x)) = x\right] \geq \varepsilon_{\mathcal{R}}(n) \ ,$$

  or

$$\left|\Pr\left[\mathsf{Exp}^{\mathsf{iO}}_{(f,i\mathcal{O}),i\mathcal{O},\mathcal{C},\mathcal{R}^{\mathcal{A}}}(n) = 1\right] - \frac{1}{2}\right| \geq \varepsilon_{\mathcal{R}}(n) \ ,$$

  where in both $\mathcal{R}$ makes at most $q_{\mathcal{R}}(n)$ queries to any of its oracles $(\mathcal{A}, f, i\mathcal{O})$, and any query $(C_0^{(\cdot)}, C_1^{(\cdot)})$ it makes to $\mathcal{A}$ consists of circuits that also make at most $q_{\mathcal{R}}(n)$ queries to their oracles $(f, i\mathcal{O})$. The random variable $\mathsf{Exp}^{\mathsf{iO}}_{(f,i\mathcal{O}),i\mathcal{O},\mathcal{C},\mathcal{R}^{\mathcal{A}}}(n)$ represents the reductions winning probability in the IO security game (Definition 2.2) relative to $(f, i\mathcal{O})$.

We make several remarks about the definition:

- **Correctness.** Typically, we also require certain *correctness* from the black-box construction. For instance, in the next section, we shall require that the construction always satisfies the $\mathsf{NP} \cap \mathsf{coNP}$ structure. In the above definition, the construction is allowed to yield instances $(C_0^{f,i\mathcal{O}}, C_1^{f,i\mathcal{O}})$ that do not satisfy the SZK promise; namely $(C_0^{f,i\mathcal{O}}, C_1^{f,i\mathcal{O}}) \notin \mathbf{SD}_Y^{f,i\mathcal{O}} \cup \mathbf{SD}_N^{f,i\mathcal{O}}$. It is natural to think of more stringent definitions that require that the corresponding problem $\Pi^{f,i\mathcal{O}}$ is non-trivial, in the sense that $\Pi^{f,i\mathcal{O}} \cap \mathbf{SD}_Y^{f,i\mathcal{O}} \neq \emptyset$ and $\Pi^{f,i\mathcal{O}} \cap \mathbf{SD}_N^{f,i\mathcal{O}} \neq \emptyset$ (which is the case for known constructions of SZK hardness from cryptographic primitives). Our impossibility is more general and would, in particular, rule out such definitions as well.

- **Worst-Case vs. Average-Case Hardness.** In the above, we address *worst-case hardness*, in the sense that the reduction $\mathcal{R}$ has to break the underlying primitives only given a decider $\mathcal{A}$ that is always correct. One could further ask whether IO and OWPs even imply average-case hardness in SZK (as do many of the algebraic hardness assumptions in cryptography). Ruling out worst-case hardness (as we will do shortly) in particular rules out such average-case hardness as well.

- **IO for Oracle-Aided Circuits.** Following [AS15, AS16], we consider indistinguishability obfuscation for oracle-aided circuits C^f that can make calls to the one-way permutation oracle. This model captures constructions where IO is applied to circuits that use pseudo-random generators, puncturable pseudo-random functions, or injective one-way functions as all of those have fully black-box constructions from one-way permutations (see further discussion in [AS15]). This includes almost all known constructions from IO, including public-key encryption, deniable encryption [SW14], functional encryption [Wat15], delegation [BGL+15, CHJV15, KLW15], and hard (on-average) PPAD instances [BPR15]. Accordingly, separating SZK from IO and OWPs in this model, results in a similar separation between SZK and any one of these primitives.

  We note that there a few applications though that do not fall under this model. The first is in applications where the obfuscated circuit might itself output (smaller) obfuscated circuit, for instance in the construction of IO for Turing machines from IO-based succinct randomized encodings [BGL+15, CHJV15, KLW15]. To capture such applications, one would have to extend the model to also account for circuits with IO gates (and not only OWP gates). A second example is the construction of non-interactive witness indistinguishable proofs from IO [BP15]. There an obfuscated circuit may get as input another obfuscated circuit and would have to internally run it; furthermore, in this application, the code of the obfuscator is used in a (non-black-box) ZAP. Extending our results (and those of [AS15, AS16]) to these models is an interesting question, left for future work.

- **Security Loss.** In the above definition the functions $q_{\mathcal{R}}$ and $\varepsilon_{\mathcal{R}}$ capture the *security loss* of the reduction. Most commonly in cryptography, the query complexity is polynomial $q_{\mathcal{R}}(n) = n^{O(1)}$ and the probability of breaking the underlying primitive is inverse polynomial $\varepsilon_{\mathcal{R}}(n) = n^{-O(1)}$. Our lower-bounds will in-fact apply for *exponential* $q_{\mathcal{R}}, \varepsilon_{\mathcal{R}}^{-1}$. This allows capturing also constructions that rely on subexponentially secure primitives (e.g., [BGL+15, CHJV15, KLW15, BPR15, BPW16]).

Ruling Out Fully Black-Box Constructions: A Road Map. Our main result in this section is that a fully black-box construction of a hard statistical difference problem from IO and OWPs does not exist. Furthermore, this holds even if the latter primitives are exponentially secure.

Theorem 3.3. *Any fully black-box construction of a statistical difference problem* Π *from OWPs and IO for circuits with OWP gates has an exponential security loss:* $\max(q_{\mathcal{R}}(n), \varepsilon_{\mathcal{R}}^{-1}(n)) \geq \Omega(2^{n/12})$.

The proof of the theorem follows a common methodology (applied for instance in [HR04, HHRS15b, AS15]). We exhibit two (distributions on) oracles $(\Psi, \mathsf{StaDif}^{\Psi})$, where Ψ realizes OWPs and IO for circuits with OWP gates, and StaDif^{Ψ} that decides $\mathbf{SD}^{\Psi}$, the statistical difference problem relative to Ψ, in the worst case. Since $\mathbf{SD}$ is complete for SZK in a relativizing manner, solving $\mathbf{SD}^{\Psi}$ suffices to break SZK^{Ψ}. We then show that the primitives realized by Ψ are (exponentially) secure even in the presence of StaDif^{Ψ}. Then viewing StaDif as a worst-case decider $\mathcal{A}$ (as per Definition 3.2) directly implies Theorem 3.3, ruling out fully black-box constructions with a subexponential security loss.

The rest of this section is organized according to the above plan. First, in Section 3.3, we describe the oracle StaDif^{Ψ} (which is independent of the specific way that Ψ realizes IO and OWPs). Then,

in Sections 3.4 and 3.5, we describe the oracle Ψ realizing OWPs and IO and prove its (exponential) security in the presence of StaDif^{Ψ}.

3.3 A Noisy Statistical-Distance Oracle

We now define the oracle StaDif^{Ψ} that will solve the statistical difference problem $\mathbf{SD}^{\Psi}$ in all the separations proved in this section. Our goal is to design StaDif^{Ψ} in a way that will not break the security of the cryptographic primitives realized by Ψ (OWPs in the warmups, and then OWPs and IO for circuits with OWP-gates). For this purpose, in our definition of the oracle StaDif^{Ψ}, we will try to exploit the fact that statistical distance is insensitive to *local changes* in the input distributions. Then, we will show that breaking the relevant cryptographic primitives, captured by Ψ, is impossible without detecting such local changes.

The concrete way of capturing the spoken insensitivity will be to define a "noisy oracle" that would be correct on distribution pairs whose distance is within the promise range $\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]$, but would behave randomly within $\left(\frac{1}{3},\frac{2}{3}\right)$.

Definition 3.4 (Oracle StaDif^{Ψ})**.** The oracle consists of $\mathbf{t}=\{\mathbf{t}_n\}_{n\in\mathbb{N}}$ where $\mathbf{t}_n:\{0,1\}^{2n}\to\left(\frac{1}{3},\frac{2}{3}\right)$ is a uniformly random function. Given n-bit descriptions of oracle-aided circuits $C_0,C_1\in\{0,1\}^n$, let $t=\mathbf{t}_n(C_0,C_1)$, and let $s=\Delta(\boldsymbol{C}_0^{\Psi},\boldsymbol{C}_1^{\Psi})$, return

$$\mathsf{StaDif}^{\Psi}(C_1,C_2;t):=\begin{cases}N & \text{If } s<t\\ Y & \text{If } s\geq t\end{cases}$$

It is immediate to see that StaDif^{Ψ} decides $\mathbf{SD}^{\Psi}$ in the worst-case.

Claim 3.5. *For any oracle Ψ,*

$$\mathbf{SD}^{\Psi}\in\mathsf{P}^{\Psi,\mathsf{StaDif}^{\Psi}}\ .$$

The main challenge is in showing that Ψ can implement OWPs and IO (for OWP-aided circuits) that will be secure in the presence of StaDif^{Ψ}. We next develop the terminology and establish several useful properties of StaDif that will allow us to carry out the above plan.

Capturing Insensitivity to Local Changes. We introduce two general notions of *farness* and *smoothness* that aim to capture the sense in which the statistical difference oracle StaDif^{Ψ} defined above is insensitive to local changes.

Roughly speaking *farness* says that the random threshold t used for a query (C_0,C_1) to StaDif^{Ψ} is "far" from the actual statistical distance. We will show that with high probability over the choice of random threshold $\mathbf{t}$, farness holds for all queries (C_0,C_1) made to StaDif^{Ψ} by any (relatively) efficient adversary. This intuitively means that changing the distributions $(\boldsymbol{C}_0^{\Psi},\boldsymbol{C}_1^{\Psi})$, on sets of small density, will not change the oracle's answer.

Definition 3.6 (Farness)**.** The oracles $(\Psi,\mathsf{StaDif}^{\Psi})$ satisfy δ-*farness* with respect to oracle-aided circuits $(C_0,C_1)\in\{0,1\}^n$ if the statistical difference $s=\Delta(\boldsymbol{C}_0^{\Psi},\boldsymbol{C}_1^{\Psi})$ and the threshold $t=\mathbf{t}_n(C_0,C_1)$ sampled by StaDif are δ-far:

$$|s-t|\geq\delta\ .$$

For an adversary $\mathcal{A}$, we denote by $\mathbf{Far}(\mathcal{A},\Psi,\delta)$ the event that $\Gamma=(\Psi,\mathsf{StaDif}^{\Psi})$ satisfies δ-farness for all queries (C_0,C_1) made by $\mathcal{A}$ to StaDif^{Ψ}.

Claim 3.7. *Fix any Ψ and any oracle-aided adversary $\mathcal{A}$ such that $\mathcal{A}^{\Psi,\mathsf{StaDif}^\Psi}$ makes at most q queries to StaDif^Ψ. Then*

$$\Pr_{\mathbf{t}}[\mathbf{Far}(\mathcal{A},\Psi,\delta)] \geq 1 - 6\delta q \ ,$$

where the probability is over the choice $\mathbf{t}$ of random thresholds by StaDif.

Proof. This follows from the fact that, for any query (C_0, C_1) to StaDif^Ψ with $s = \Delta(\boldsymbol{C}_0^\Psi, \boldsymbol{C}_1^\Psi)$, δ-farness does not hold only if the threshold $t = \mathbf{t}(C_0, C_1)$, chosen at random for this query, happens to be in the interval $(s-\delta, s+\delta)$, which occurs with probability at most $|(s-\delta, s+\delta)| \,/\, \left|\left(\frac{1}{3}, \frac{2}{3}\right)\right| = 6\delta$. The lemma then follows by a union bound over at most q queries. □

We now turn to define the notion of *smoothness*. Roughly speaking we will say that an oracle-aided circuit C is smooth with respect to some oracle Ψ if any specific oracle query is only made with small probability. In particular, for a pair of smooth circuits (C_0, C_1), local changes to the oracle Ψ should not change significantly the statistical distance $s = \Delta(\boldsymbol{C}_0^\Psi, \boldsymbol{C}_1^\Psi)$.

Definition 3.8 ((Ψ,δ)-Smoothness)**.** An oracle-aided circuit $C^{(\cdot)} : \{0,1\}^n \to \{0,1\}^m$ is said to be (Ψ,δ)-smooth if for all $x \in \{0,1\}^\star$,

$$\Pr_{r \leftarrow \{0,1\}^n}\left[C^\Psi(r) \text{ queries } \Psi \text{ at } x\right] \leq \delta \ .$$

For an adversary $\mathcal{A}$, we denote by $\mathbf{Smo}(\mathcal{A},\Psi,\delta)$ the event that all queries (C_0, C_1) made by $\mathcal{A}$ to StaDif^Ψ are (Ψ,δ)-smooth.

Claim 3.9. *Let Ψ, Ψ' be oracles that differ on at most c values in the domain. Let (C_0, C_1) be (Ψ,δ)-smooth. Let $s = \Delta(C_0^\Psi, C_1^\Psi)$ and $s' = \Delta(C_0^{\Psi'}, C_1^{\Psi'})$ then $|s - s'| \leq 2c\delta$.*

Proof. For either $b \in \{0,1\}$,

$$\begin{aligned}
&\Delta(C_b^\Psi, C_b^{\Psi'}) \leq \\
&\Pr_r\left[C_b^\Psi(r) \neq C_b^{\Psi'}(r)\right] \leq \\
&\Pr_r\left[C_b^\Psi(r) \text{ queries } \Psi \text{ at } x \text{ where } \Psi(x) \neq \Psi'(x)\right] \leq \\
&\sum_{x:\Psi(x)\neq\Psi'(x)} \Pr_r\left[C_b^\Psi(r) \text{ queries } \Psi \text{ at } x\right] \leq c \cdot \delta \ .
\end{aligned}$$

The claim then follows by the fact that

$$\left|s - s'\right| := \left|\Delta(C_0^\Psi, C_1^\Psi) - \Delta(C_0^{\Psi'}, C_1^{\Psi'})\right| \leq \Delta(C_0^\Psi, C_0^{\Psi'}) + \Delta(C_1^\Psi, C_1^{\Psi'}) \leq 2c\delta \ .$$

□

The above roughly means that (under the likely event that farness holds) making smooth queries should not help the adversary detect local changes in the oracle Ψ. We will next show that, in fact, we can always "smoothen" the adversary's circuit at the expense of making (a few) more queries to Ψ, which intuitively deems the statistical difference oracle StaDif^Ψ useless altogether for detecting local changes in Ψ. Looking ahead, we will later show that breaking certain cryptographic

primitives (OWPs and IO) is impossible without detecting such local changes, and then deduce that they do not break in the presence of StaDif^{Ψ}.

In what follows, we say that an adversary $\mathcal{A}$ is q-query if $\mathcal{A}^{\Psi,\mathsf{StaDif}^{\Psi}}$ makes at most q queries to Ψ and q queries to StaDif^{Ψ}, and any query made to StaDif^{Ψ} consist of oracle-aided circuits (C_0, C_1) that make at most q queries to Ψ, on any specific input. (We do not restrict the size of these circuits, but only the number of queries they make.)

Lemma 3.10 (Smoothing Lemma)**.** *For any q-query algorithm $\mathcal{A}$ and $\beta \in \mathbb{N}$, there exists a $(q + 2\beta q)$-query algorithm $\mathcal{S}$ such that for any input $z \in \{0,1\}^{\star}$ and oracles $\Psi, \mathsf{StaDif}^{\Psi}$:*

1. *$\mathcal{S}^{\Psi,\mathsf{StaDif}^{\Psi}}(z)$ perfectly simulates the view of $\mathcal{A}^{\Psi,\mathsf{StaDif}^{\Psi}}(z)$,*
2. *$\mathcal{S}^{\Psi,\mathsf{StaDif}^{\Psi}}(z)$ only makes (Ψ,δ)-smooth queries to StaDif^{Ψ} with probability:*

$$\Pr_{\mathcal{S}}\left[\mathbf{Smo}(\mathcal{S},\Psi,\delta)\right] \geq 1 - 2^{-\delta\beta+\log(2q^2/\delta)} \ ,$$

over its own random coin tosses.

Proof. The simulator $\mathcal{S}$ emulates $\mathcal{A}$ and whenever $\mathcal{A}$ makes a query (C_0, C_1) to Ψ, $\mathcal{A}$ first evaluates each of the two circuits C_0^{Ψ}, C_1^{Ψ} on β random inputs and stores all the queries they make to Ψ along with their answers in a table T. It then generates a new query consisting of circuits (C_0', C_1') that have the table T hardwired in them. Each C_b' emulates C_b, but whenever the emulated C_b makes an oracle query to Ψ, C_b' first tries to answer using the table T, and only if the answer is not there turns to the oracle Ψ.

By construction, $\mathcal{S}$ perfectly emulates the view of $\mathcal{A}$. We now bound the probability that $\mathcal{S}$ generates a circuit that is not (Ψ,δ)-smooth. Fix any query (C_0, C_1) and let x be a *heavy query* in the sense that it is queried with probability larger than δ by one of the two circuits. Then the query x will be put in the table T except with probability $(1-\delta)^{\beta} \leq 2^{-\delta\beta}$. Furthermore, each one of the two circuits makes at most q oracle queries and thus each has at most q/δ inputs x as above. The claim now follows by a union bound over at most q queries (C_0, C_1) and at most q/δ heavy inputs that each of the two has. □

3.4 Warmup: One-Way Permutations in the Presence of StaDif

In this section, we show that a random permutation f is hard to invert even given access to the noisy statistical difference oracle StaDif^{f}. We start by defining the oracle. In what follows, $\mathbf{P}_n$ denotes the set of permutations of $\{0,1\}^n$.

Definition 3.11 (The Oracle f)**.** $f = \{f_n\}_{n\in\mathbb{N}}$ on input $x \in \{0,1\}^n$ answers with $f_n(x)$ where f_n is a random permutation $f_n \leftarrow \mathbf{P}_n$.

Our main theorem states that f cannot be inverted, except with exponentially small probability, even given an exponential number of oracle queries to f and StaDif^{f}. Here, consistently with the previous subsection, we say that an adversary $\mathcal{A}$ is q-query if $\mathcal{A}^{\Psi,\mathsf{StaDif}^{\Psi}}$ makes at most q queries to f and q queries to StaDif^{f}, and any query made to StaDif^{f} consists of oracle-aided circuits (C_0, C_1) that make at most q queries to f, on any specific input.

Theorem 3.12. *Let $q \leq O(2^{n/6})$. Then for any q-query adversary $\mathcal{A}$*

$$\Pr_{f,\mathsf{StaDif},x}\left[\mathcal{A}^{f,\mathsf{StaDif}^f}(f(x)) = x\right] \leq O(2^{-n/6}) \ ,$$

where the probability is over the random choices of f, StaDif and $x \leftarrow \{0,1\}^n$.

At a very high level, the proof of the theorem follows the plan outlined above, showing that in order to invert a random permutation the adversary must be able to detect certain local changes to the permutation, which the noisy statistical difference oracle is insensitive to.

Proof. We, in fact, prove a stronger statement: the above holds when fixing the oracles $f_{-n} := \{f_k\}_{k \neq n}$. Fix a q-query adversary $\mathcal{A}$ and let $\mathcal{S}$ be its smooth $(q + 2\beta q)$-query simulator given by Lemma 3.10, where β will be specified later on. Since $\mathcal{S}$ perfectly emulates $\mathcal{A}$, it is enough to bound the probability that $\mathcal{S}$ successfully inverts. To bound $\mathcal{S}$'s inversion probability, we consider four hybrid experiments $\{\mathbf{H_i}\}_{i\in[4]}$ given in Table 1. Throughout, for a permutation $f \in \mathbf{P}_n$ and $x, y \in \{0,1\}^n$, we denote by $f_{x\mapsto y}$ the function that maps x to y and is identical to f on all other inputs (in particular, $f_{x\mapsto y}$ is no longer a permutation when $x \neq f^{-1}(y)$).

Hybrid	**$\mathbf{H_1}$ (Real)**	**$\mathbf{H_2}$**	**$\mathbf{H_3}$**	**$\mathbf{H_4}$ (Ideal)**
Permutation	$f_n \leftarrow \mathbf{P}_n$			
Preimage	$x \leftarrow \{0,1\}^n$			
2nd Preimage	$z \leftarrow \{0,1\}^n$			
Planted Image	$y \leftarrow \{0,1\}^n$			
Challenge	$f(x)$		y	
Oracle	f, StaDif^f	$f_{z\mapsto f(x)}, \mathsf{StaDif}^{f_{z\mapsto f(x)}}$	$f_{x\mapsto y}, \mathsf{StaDif}^{f_{x\mapsto y}}$	f, StaDif^f
Winning Condition	Find x			

Table 1: The hybrid experiments.

Hybrid $\mathbf{H_1}$ is identical to the real world where $\mathcal{S}$ wins if it successfully inverts the permutation at a random output. We show that the probability that the simulator wins in any of the experiments is roughly the same, and that in hybrid $\mathbf{H_4}$ the probability that $\mathcal{S}$ wins is tiny.

Claim 3.13. $|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_1}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_2}]| \leq O(2^{-n/6})$.

Proof. The difference between the two hybrids is in the oracle that $\mathcal{S}$ is given: simply f in the first, and its slightly tweaked version $f_{z\mapsto f(x)}$ in the second. We can bound the difference between the winning probabilities in $\mathbf{H_1}$ and $\mathbf{H_2}$ as follows:

$$|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_1}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_2}]| \leq$$
$$\Pr_{\substack{\mathcal{S},x,z \\ f,\mathsf{StaDif}}}\left[\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x)) \neq \mathcal{S}^{f_{z\mapsto f(x)},\mathsf{StaDif}^{f_{z\mapsto f(x)}}}(f(x))\right] \ ,$$

where the probability is over the coins of $\mathcal{S}$ and StaDif and the choice of $x, z \leftarrow \{0,1\}^n, f_n \leftarrow \mathbf{P}_n$.

In what follows, we denote by $\mathbf{Hit} = \mathbf{Hit}(\mathcal{S}, f, x, z)$ the event that $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$ queries f on z. Also, let $\mathbf{Far} = \mathbf{Far}(\mathcal{S}(f(x)), f, 2\delta)$ be the event that 2δ-farness holds for all StaDif-queries made by $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$ (Definition 3.8), and $\mathbf{Smo} = \mathbf{Smo}(\mathcal{S}(f(x)), f, \delta)$ is the event that all StaDif-queries made by $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$ are (f,δ)-smooth (Definition 3.8).

We now claim

Claim 3.14. *For any $\delta < 1$,*

$$\Pr_{\substack{\mathcal{S},x,z \\ f,\mathsf{StaDif}}}\left[\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x)) \neq \mathcal{S}^{f_{z\mapsto f(x)},\mathsf{StaDif}^{f_{z\mapsto f(x)}}}(f(x))\right] \leq \Pr_{\substack{\mathcal{S},x,z \\ f,\mathsf{StaDif}}}\left[\mathbf{Hit} \vee \overline{\mathbf{Far}} \vee \overline{\mathbf{Smo}}\right] \ .$$

Proof. We argue that whenever the complement $\overline{\mathbf{Hit}} \wedge \mathbf{Far} \wedge \mathbf{Smo}$ occurs then

$$\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x)) = \mathcal{S}^{f_{z\mapsto f(x)},\mathsf{StaDif}^{f_{z\mapsto f(x)}}}(f(x)) \ .$$

Indeed, for any StaDif-query (C_0, C_1) made by $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$, we know by (f,δ)-smoothness that changing f at one point does not affect the statistical distance by much. Concretely, by Claim 3.9:

$$\left|\Delta(\boldsymbol{C}_0^f, \boldsymbol{C}_1^f) - \Delta(\boldsymbol{C}_0^{f_{z\mapsto f(x)}}, \boldsymbol{C}_1^{f_{z\mapsto f(x)}})\right| \leq 2\delta \ .$$

Furthermore, if 2δ-farness also holds for any such query (for some threshold $\mathbf{t}$ sampled by StaDif), then

$$\mathsf{StaDif}^f(C_0, C_1; \mathbf{t}) = \mathsf{StaDif}^{f_{z\mapsto f(x)}}(C_0, C_1; \mathbf{t}) \ .$$

If in addition $\mathbf{Hit}$ does not occur, then for any f-query w made by $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$,

$$f(w) = f_{z\mapsto f(x)}(w) \ .$$

It follows that the views of $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$ and $\mathcal{S}^{f_{z\mapsto f(x)},\mathsf{StaDif}^{f_{z\mapsto f(x)}}}(f(x))$ are identical. □

It is left to bound the probability of each of the events $\mathbf{Hit}, \overline{\mathbf{Far}}, \overline{\mathbf{Smo}}$. First, noting that the view of $\mathcal{S}^{f,\mathsf{StaDif}^f}(f(x))$ is independent of the random z, we can bound

$$\Pr[\mathbf{Hit}] \leq 2^{-n} \cdot \#\{f\text{-queries made by } \mathcal{S}\} \leq 2^{-n} \cdot (q + 2\beta q) \ .$$

Furthermore, by the farness Claim 3.7 and smoothing Lemma 3.10

$$\begin{aligned} \Pr\left[\overline{\mathbf{Far}}\right] &\leq 12q\delta \ . \\ \Pr\left[\overline{\mathbf{Smo}}\right] &\leq 2^{-\delta\beta + \log(2q^2/\delta)} \ , \end{aligned}$$

Overall we can bound the difference between $\mathbf{H_1}$ and $\mathbf{H_2}$ by

$$2^{-\delta\beta+\log(2q^2/\delta)} + 2^{-n} \cdot (q + 2\beta q) + 12q\delta \leq O(2^{-n/6}) \ ,$$

when setting $\delta = 2^{-n/3}, \beta = 2^{n/3} \cdot n$, and recalling that $q \leq O(2^{n/6})$. □

Claim 3.15. $\Pr[\mathcal{S} \text{ wins in } \mathbf{H_2}] = \Pr[\mathcal{S} \text{ wins in } \mathbf{H_3}]$.

Proof. The difference between $\mathbf{H_2}$ and $\mathbf{H_3}$ is in the input of $\mathcal{S}$, $f(x)$ in the first and a random y in the second, and in the oracle $\mathcal{S}$ is given, $f_{z\mapsto f(x)}$ in the first and $f_{x\mapsto y}$ in the second. We argue, however, that the distribution $\left\{(f(x), f_{z\mapsto f(x)}, x) \;\middle|\; f \leftarrow \mathbf{P}_n, x, z \leftarrow \{0,1\}^n\right\}$ in $\mathbf{H_1}$ is identical to that of $\{(y, f_{x\mapsto y}, x) \mid f \leftarrow \mathbf{P}_n, x, z \leftarrow \{0,1\}^n\}$ are in $\mathbf{H_2}$. Indeed, in $\mathbf{H_1}$, $(f(x), x)$ are distributed uniformly and independently just as (y, x) in $\mathbf{H_2}$. Then, conditioned on any (y, x), the oracle in both distribution can be sampled as a random permutation f conditioned on $y = f(x)$ and diverting a random z from $f(z)$ to y. □

Claim 3.16. $|\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_3}] - \Pr[\mathcal{S} \textit{ wins in } \mathbf{H_4}]| \leq O(2^{-n/6})$.

The difference between the two hybrids is in the oracle that $\mathcal{S}$ is given: simply f in the second and its slightly tweaked version $f_{x\mapsto y}$ in the first. The proof of their indistinguishability is essentially identical to that of Claim 3.13 and is omitted.

To conclude the proof of Theorem 3.12, we observe that

Claim 3.17. $\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_4}] \leq 2^{-n}$.

Proof. The view of $\mathcal{S}$ in this hybrid is completely independent of the random choice of x. □

□

3.5 Indistinguishability Obfuscation (and OWPs) in the Presence of StaDif

In this section, we consider an oracle Ψ that realizes both indistinguishability obfuscation (IO) and one-way permutations (OWPs) and show that neither break in the presence of the noisy statistical difference oracle StaDif^{Ψ}. We start by defining the oracle Ψ. In a nutshell, the oracle realizes OWPs through a random permutation oracle. IO for circuits with OWP-gates is captured in a similar way to [AS15] by a random injective mapping coupled with a corresponding evaluation algorithm.

In what follows, $\mathbf{P}_n$ denotes the set of permutations of $\{0,1\}^n$, $\mathbf{F}_n^m$ denotes the set of functions mapping $\{0,1\}^n$ to $\{0,1\}^m$, and $\mathbf{I}_n^m$ denotes the set of injective functions mapping $\{0,1\}^n$ to $\{0,1\}^m$.

Definition 3.18 (The Oracle Ψ)**.** The oracle $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$ consists of three parts:

- $f = \{f_n\}_{n\in\mathbb{N}}$ on input $x \in \{0,1\}^n$ answers with $f_n(x)$, where f_n is a random permutation $f_n \leftarrow \mathbf{P}_n$.
- $\mathcal{O} = \{\mathcal{O}_n\}_{n\in\mathbb{N}}$ on input $(C, r) \in \{0,1\}^n \times \{0,1\}^n$ answers with $\widehat{C} := \mathcal{O}_n(C, r)$ where $\mathcal{O}_n$ is a random injective function $\mathcal{O}_n \leftarrow \mathbf{I}_{2n}^{5n}$ into $\{0,1\}^{5n}$.
- $\mathsf{Eval}^{f,\mathcal{O}}$ given $\widehat{C} \in \{0,1\}^{5n\times 2}, x \in \{0,1\}^\star$ computes $(C, r) = \mathcal{O}_n^{-1}(\widehat{C})$, interprets C as an oracle-aided circuit, and returns $C^f(x)$. If $\widehat{C}$ does not have a unique preimage, or the input size of C is inconsistent with $|x|$, the oracle returns $\perp$.

In the next two subsections, we show that the oracle Ψ securely realizes OWPs and IO in the presence of the noisy statistical difference oracle StaDif^{Ψ}. Throughout, we address adversaries with oracles $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{\mathcal{O},f})$ and StaDif^{Ψ}. We will say that such an adversary is q-query if they

1. make only q queries to f,

2. make only q queries to either $\mathcal{O}$ or Eval, and any query $\widehat{C}$ to Eval is of size at most $5q$, and in particular, any oracle aided circuit C that is mapped to $\widehat{C}$ by $\mathcal{O}$ is of size at most q, and makes at most q queries to f,

3. make only q queries to StaDif^{Ψ}, and for any query (C_0, C_1) made to StaDif^{Ψ}, (C_0, C_1) are Ψ-aided and each of them is q-query (according to the two conditions above).

3.5.1 One-Way Permutations

We show that f cannot be inverted, except with exponentially small probability even given an exponential number of oracle queries to $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{\mathcal{O},f})$ and StaDif^{Ψ}.

Theorem 3.19. *Let $q(n) \leq O(2^{n/12})$. Then for any q-query adversary $\mathcal{A}$*

$$\Pr_{\substack{\Psi=(f,\mathcal{O},\mathsf{Eval}) \\ \mathsf{StaDif},x}}\left[\mathcal{A}^{\Psi,\mathsf{StaDif}^{\Psi}}(f(x)) = x\right] \leq O(2^{-n/6}) \ ,$$

where the probability is over the random choice of Ψ, StaDif and $x \leftarrow \{0,1\}^n$.

Proof. We will, in fact, prove a stronger statement: the above holds when fixing the oracles $f_{-n} := \{f_k\}_{k\neq n}$, $\mathcal{O} = \{\mathcal{O}_n\}_{n\in\mathbb{N}}$. We prove the theorem by a reduction to the case that Ψ only consists of the permutation f (and does not include $\mathcal{O}, \mathsf{Eval}$). Concretely, fix any q-query adversary $\mathcal{A}$ that inverts the random permutation f_n given access to $\Psi = (f, \mathcal{O}, \mathsf{Eval})$ and StaDif^{Ψ}, we show how to reduce it to a q^2-query adversary $\mathcal{B}^f(f_n(x))$ that inverts f_n for a random $x \leftarrow \{0,1\}^n$ with the same probability as $\mathcal{A}$. The proof then follows from Theorem 3.12.

The new adversary $\mathcal{B}^{f,\mathsf{StaDif}^f}(f_n(x))$ emulates $\mathcal{A}^{\Psi,\mathsf{StaDif}^{\Psi}}(f_n(x))$ answering Ψ-queries as follows:

- f **queries:** answered according to $\mathcal{B}$'s oracle f. This translates to at most q queries to f.
- $\mathcal{O}$ **queries:** answered according to the fixed oracle $\mathcal{O}$. This does not add any calls to f.
- $\mathsf{Eval}^{f,\mathcal{O}}$ **queries:** given query $(\widehat{C}, x)$ to Eval, invert the fixed oracle $\mathcal{O}$ to find $(C, r) = \mathcal{O}^{-1}(\widehat{C})$. If no such preimage exists, return $\perp$. If a preimage does exist, using the f-oracle, compute $C^f(x)$ and return the result. This translates to at most q^2 queries to f: q queries by C, for each of the q queries $\widehat{C}$ to Eval.
- StaDif^{Ψ} **queries:** given query (C_0, C_1), where C_b makes Ψ-queries translate to D_0, D_1 that only make f-queries, where each query to $\Psi = (f, \mathcal{O}, \mathsf{Eval})$ is translated to a query to f according to the previous three items. The resulting oracle-aided (D_0, D_1) may thus make up to $q + q^2$ queries f: q corresponding to the first item, and q^2 corresponding to the third.[13]

Overall $\mathcal{B}^f$ is $O(q^2)$-query and perfectly emulates the view of $\mathcal{A}^{\Psi}$. The theorem now follows from Theorem 3.12.

□

[13] We note that while there is a bound on the number of queries that they make, we do not put any restrictions on their size, which allows to hardwire the fixed $\mathcal{O}$ and f_{-n} as required in the previous three items. Indeed, Theorem 3.12 does not put any restriction on the size of these circuits.

3.5.2 Indistinguishability Obfuscation

We now turn to show that Ψ also realizes an indistinguishability obfuscator that does not break in the presence of StaDif^{Ψ}. We start by describing the construction, which is similar to the one in [AS15].

Construction 3.20 (The Obfuscator $i\mathcal{O}^{\Psi}$)**.** *Let $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$. Given an oracle-aided circuit $C \in \{0,1\}^n$, $i\mathcal{O}^{\Psi}(1^n, C)$ samples a random $r \leftarrow \{0,1\}^n$, computes $\widehat{C} = \mathcal{O}(C, r)$, and returns an oracle aided circuit $E_{\widehat{C}}$ that given input x, computes $\mathsf{Eval}^{f,\mathcal{O}}(\widehat{C}, x)$.*

It is easy to see that $i\mathcal{O}^{f,\mathcal{O},\mathsf{Eval}}$ satisfies the functionality requirement of Definition 2.2 for the class $\mathcal{C}$ of f-aided circuits; indeed, this follows by the fact that $\mathcal{O}$ is injective, and by the definition of $i\mathcal{O}$ and the oracles $\mathcal{O}, \mathsf{Eval}$. We now show that it also satisfies indistinguishability, with an exponentially small distinguishing gap, even given an exponential number of oracle queries to $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{\mathcal{O},f})$ and the statistical difference oracle StaDif^{Ψ}.

Theorem 3.21. *Let $q(n) \leq O(2^{n/6})$. Then for any q-query adversary $\mathcal{A}$*

$$\left| \Pr\left[\mathsf{Exp}^{\mathsf{iO}}_{\Psi,\mathsf{StaDif},i\mathcal{O},\mathcal{C},\mathcal{A}}(n) = 1 \right] - \frac{1}{2} \right| \leq O(2^{-n/6})$$

where the random variable $\mathsf{Exp}^{\mathsf{iO}}_{\Gamma,i\mathcal{O},\mathcal{C},\mathcal{A}}(n)$ denotes the adversary's winning probability in the IO security game (Definition 2.2) relative to $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$ and StaDif^{Ψ}.

At a very high-level, the proof of the theorem follows a similar rationale to the proof of Theorem 3.12 showing that one-way permutations do not break in the presence of the noisy statistical difference oracle. Roughly speaking, we show that in order to break the above construction of IO, the adversary must be able to detect local changes in the oracles realizing it, whereas the noisy statistical difference oracle is insensitive of these changes. At a technical level, the case of IO requires somewhat more care than the case of one-way permutations. For once, it has a more elaborate interface consisting not only of a hard to invert mapping $\mathcal{O}$, but also of the evaluation oracle $\mathsf{Eval}^{f,\mathcal{O}}$. In particular, a single change to $\mathcal{O}$ may introduce many changes to $\mathsf{Eval}^{f,\mathcal{O}}$, which could potentially be detected by the statistical difference oracle. Another aspect that complicates the proof is that the IO game is more interactive in its nature. In particular, we need to deal with the fact that the actual circuits of the IO challenge are chosen adaptively, after the adversary had already interacted with all the oracles. We now turn to the actual proof.

Proof. We prove a stronger statement: the above holds when fixing the oracles f and $\mathcal{O}_{-n} = \{\mathcal{O}_k\}_{k \neq n}$. Fix a q-query adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ and let $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ be its smooth $(q + 2\beta q)$-query simulator given by Lemma 3.10, where β will be specified later on. Since $\mathcal{S}$ perfectly emulates $\mathcal{A}$, it suffices to prove the theorem for $\mathcal{S}$. To bound $\mathcal{S}$'s advantage in breaking $i\mathcal{O}$, we consider six hybrid experiments $\{\mathbf{H_i}\}_{i \in [6]}$ given in Table 2.

We introduce some notation that will be useful to describe the hybrids:

- For a function $\mathcal{O} = \left\{ \mathcal{O}_k : \{0,1\}^{2k} \to \{0,1\}^{5k} \right\}_{k \in \mathbb{N}}$, a pair $(C, r) \in \{0,1\}^{n \times 2}$, and $\widehat{C} \in \{0,1\}^{5n}$, we denote by $\mathcal{O}_{(C,r) \mapsto \widehat{C}}$ the function that maps (C, r) to $\widehat{C}$ and is otherwise identical to $\mathcal{O}$.

- For a function $\mathcal{O} = \left\{\mathcal{O}_k : \{0,1\}^{2k} \to \{0,1\}^{5k}\right\}_{k\in\mathbb{N}}$, we denote by $\Gamma(f,\mathcal{O})$ the oracle

$$\Gamma(f,\mathcal{O}) := f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}}, \mathsf{StaDif}^{f,\mathcal{O},\mathsf{Eval}^{f,\mathcal{O}}} \quad .$$

- For a function $\mathcal{O} = \left\{\mathcal{O}_k : \{0,1\}^{2k} \to \{0,1\}^{5k}\right\}_{k\in\mathbb{N}}$, a string $\widehat{C} \in \{0,1\}^{5n}$, and a circuit C, we denote by $\Gamma(f,\mathcal{O},\widehat{C},C)$ the oracle

$$\Gamma(f,\mathcal{O},\widehat{C},C) := f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}}_{\widehat{C},C}, \mathsf{StaDif}^{f,\mathcal{O},\mathsf{Eval}^{f,\mathcal{O}}_{\widehat{C},C}} \quad ,$$

  where $\mathsf{Eval}^{f,\mathcal{O}}_{\widehat{C},C}$ is an oracle that

  - Given $(\widehat{D},x)$ where $\widehat{D} \neq \widehat{C}$, acts like $\mathsf{Eval}^{f,\mathcal{O}}(\widehat{D},x)$. Namely, it computes $(D,r) = \mathcal{O}^{-1}(\widehat{D})$, and returns $D(x)$, or $\perp$ in case there is no unique preimage or the size of x does not match the input size of D.
  - Given $(\widehat{C},x)$ returns $C(x)$, or $\perp$ in case $C = \perp$, or the size of x does not match the input size of C.

Hybrid	$\mathbf{H_1}$ **(Real)**	$\mathbf{H_2}$	$\mathbf{H_3}$	$\mathbf{H_4}$	$\mathbf{H_5}$	$\mathbf{H_6}$ **(Ideal)**
Obfuscator Function	$\mathcal{O}_n \leftarrow \mathbf{I}^{5n}_{2n}$		$\mathcal{O}_n \leftarrow \mathbf{F}^{5n}_{2n}$		$\mathcal{O}_n \leftarrow \mathbf{I}^{5n}_{2n}$	
Challenger Randomness	$b \leftarrow \{0,1\}$, $r \leftarrow \{0,1\}^n$					
Chosen Circuits	$(C_0,C_1) \leftarrow \mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$ where $C_0^f \equiv C_1^f$ (relative to the fixed f)					
Planted Obfuscation	$\widehat{C} \leftarrow \{0,1\}^{5n}$					
Prechallenge Oracle	$\Gamma(f,\mathcal{O})$		$\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)$			$\Gamma(f,\mathcal{O})$
Challenge Obfuscation	$\mathcal{O}(C_b,r)$			$\widehat{C}$		
Postchallenge Oracle	$\Gamma(f,\mathcal{O})$			$\Gamma(\mathcal{O}_{(C_b,r)\mapsto\widehat{C}},f)$		$\Gamma(f,\mathcal{O},\widehat{C},C_0)$
Winning Condition	Guess b					

Table 2: The hybrid experiments.

Hybrid $\mathbf{H_1}$ is identical to the real world where $\mathcal{S}$ wins if it produces functionally equivalent C_0, C_1, and it successfully guesses the bit b. We show that the probability that the simulator wins in any of the experiments is roughly the same, and that in hybrid $\mathbf{H_6}$ the probability that $\mathcal{S}$ wins is $1/2$.

Claim 3.22. $|\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_1}] - \Pr[\mathcal{S} \textit{ wins in } \mathbf{H_2}]| \leq O(2^{-n/6})$

Proof. The difference between the two hybrids is in the oracle that $\mathcal{S}_1$ is given before the challenge phase: $\Gamma(f,\mathcal{O})$ in the first, and its tweaked version $\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)$ in the second. We

stress that in $\mathbf{H_2}$, the circuit C_b is defined according to the circuits (C_0, C_1) that $\mathcal{S}_1$ would have chosen given the non-tweaked oracle $\Gamma(f,\mathcal{O})$ (so there is no circularity).[14]

We can bound the difference between the winning probabilities in $\mathbf{H_1}$ and $\mathbf{H_2}$ as follows:

$$|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_1}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_2}]| \leq$$
$$\Pr_{\substack{\mathcal{S}_1,\mathcal{O}\\ r,b,\Gamma}}\left[\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n) \neq \mathcal{S}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)}(1^n)\right] \ ,$$

where $\mathcal{S}_1$ is the part of $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ that participates in the post challenge phase, and the probability is over the coins of $\mathcal{S}_1$ and Γ (specifically, StaDif) and the choice of $r \leftarrow \{0,1\}^n$, and $\mathcal{O} \leftarrow \mathbf{I}_{2n}^{5n}$, and $b \leftarrow \{0,1\}$. We will, in fact, show that the above is bounded for any fixed $b \in \{0,1\}$. Indeed, for the rest of the claim, fix $b \in \{0,1\}$.

In what follows, we denote by $\mathbf{ZHit} = \mathbf{ZHit}(\mathcal{O})$ the event that 0^{5n} is in the image of $\mathcal{O}$, and by $\mathbf{Hit} = \mathbf{Hit}(\mathcal{S}_1, \mathcal{O}, r)$ the event that $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$ queries $\mathcal{O}$ on (C_b, r). Also, let $\mathbf{Far} = \mathbf{Far}(\mathcal{S}_1, \mathcal{O}, 2\delta)$ be the event that 2δ-farness holds for all StaDif-queries made by $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$ (Definition 3.8), and $\mathbf{Smo} = \mathbf{Smo}(\mathcal{S}_1, \Gamma(f,\mathcal{O}), \delta)$ be the event that all StaDif-queries made by $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$ are (Ψ,δ)-smooth (Definition 3.8), where $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$.

We now claim

Claim 3.23. *For any $\delta < 1$,*

$$\Pr_{\substack{\mathcal{S}_1,\mathcal{O}\\ r,\Gamma}}\left[\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n) \neq \mathcal{S}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)}(1^n)\right] \leq \Pr_{\substack{\mathcal{S}_1,\mathcal{O}\\ r,\Gamma}}\left[\mathbf{ZHit} \vee \mathbf{Hit} \vee \overline{\mathbf{Far}} \vee \overline{\mathbf{Smo}}\right] \ .$$

Proof. We argue that whenever the complement $\overline{\mathbf{ZHit}} \wedge \overline{\mathbf{Hit}} \wedge \mathbf{Far} \wedge \mathbf{Smo}$ occurs then

$$\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n) = \mathcal{S}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)}(1^n) \ .$$

We first note that when $\mathbf{ZHit}$ does not occur, the tweaked evaluation function $\mathsf{Eval}_{0^{5n},\perp}^{f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}}}$ in $\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)$ behaves exactly as the non-tweaked function $\mathsf{Eval}^{f,\mathcal{O}}$. Indeed, the only potential change in Eval is on inputs of the form $(0^{5n}, x)$, however, since 0^{5n} is not in the image of $\mathcal{O}$, $\mathsf{Eval}^{f,\mathcal{O}}$ returns $\perp$ on such inputs just like its tweaked version. Accordingly, the function $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}})$ changes on the single input (C_b, r) for $\mathcal{O}$.

Also, for any StaDif-query (C_0, C_1) made by $\mathcal{S}^{\Gamma(f,\mathcal{O})}(1^n)$, we know by (Ψ,δ)-smoothness that changing Ψ at one point does not affect the statistical distance by much. Concretely, by Claim 3.9:

$$\left|\Delta(\boldsymbol{C}_0^{\Psi}, \boldsymbol{C}_1^{\Psi}) - \Delta(\boldsymbol{C}_0^{\Psi'}, \boldsymbol{C}_1^{\Psi'})\right| \leq 2\delta \ ,$$

where Ψ' is the tweaked version of Ψ in $\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)$.

Furthermore, if 2δ-farness also holds for any such query (for some threshold $\mathbf{t}$ sampled by StaDif), then

$$\mathsf{StaDif}^{\Psi}(C_0, C_1; \mathbf{t}) = \mathsf{StaDif}^{\Psi'}(C_0, C_1; \mathbf{t}) \ .$$

[14] In more detail, we first look at an execution of $\mathcal{S}_1$ with $\Gamma(f,\mathcal{O})$, as in $\mathbf{H_1}$, with respect to the sampled $\mathcal{O}, b, r$ (and coins of $\mathcal{S}_1$). This defines circuits (C_0, C_1), one of which is the challenge circuit C_b. Then we consider an execution with exactly the same samples $\mathcal{O}, b, r$, but with a pre-challenge oracle $\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\perp)$.

If in addition **Hit** does not occur, then for any $\mathcal{O}$-query (C, s) made by $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$,

$$\mathcal{O}(C, s) = \mathcal{O}_{(C_b,r)\mapsto 0^{5n}}(C, s) \ .$$

It follows that the views of $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)$ and $\mathcal{S}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}})}(1^n)$ are identical. □

It is left to bound the probability of each of the events $\mathbf{ZHit}, \mathbf{Hit}, \overline{\mathbf{Far}}, \overline{\mathbf{Smo}}$.
First, by counting

$$\Pr[\mathbf{ZHit}] = 2^{2n}/2^{5n} = 2^{-3n} \ .$$

Second, noting that the view of $\mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n))$ is independent of the random r, we can bound

$$\Pr[\mathbf{Hit}] \leq 2^{-n} \cdot \#\{\mathcal{O}\text{-queries made by } \mathcal{S}_1\} \leq 2^{-n} \cdot (q + 2\beta q) \ .$$

Further more by the farness Claim 3.7 and smoothing Lemma 3.10

$$\Pr\left[\overline{\mathbf{Far}}\right] \leq 12q\delta \ .$$
$$\Pr\left[\overline{\mathbf{Smo}}\right] \leq 2^{-\delta\beta+\log(2q^2/\delta)} \ ,$$

Overall we can bound the difference between $\mathbf{H_1}$ and $\mathbf{H_2}$ by

$$2^{-3n} + 2^{-\delta\beta+\log(2q^2/\delta)} + 2^{-n} \cdot (q + 2\beta q) + 12q\delta \leq O(2^{-n/6}) \ ,$$

when setting $\delta = 2^{-n/3}, \beta = 2^{n/3} \cdot n$, and recalling that $q \leq O(2^{n/6})$. □

Claim 3.24. $|\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_2}] - \Pr[\mathcal{S} \textit{ wins in } \mathbf{H_3}]| \leq 2^{-n}$

Proof. The difference between the two hybrids is in the choice of the oracle $\mathcal{O}$: a random injective function in the first, and a random function in the second. Thus,

$$|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_2}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_3}]| \leq$$
$$\Pr_{\mathcal{O}\leftarrow\mathbf{F}_{2n}^{5n}}[\mathcal{O} \text{ is not injective}] \leq 2^{-n} \ .$$

□

Claim 3.25. $\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_3}] = \Pr[\mathcal{S} \textit{ wins in } \mathbf{H_4}]$.

Proof. The difference between $\mathbf{H_3}$ and $\mathbf{H_4}$ is that in $\mathbf{H_4}$, in the challenge and post challenge phases, the value $\mathcal{O}(C_b, r)$ is re-sampled uniformly at random, i.e. it is replaced everywhere by $\widehat{C} \leftarrow \{0,1\}^{5n}$. We claim that this induces exactly the same distribution on $\mathcal{S}$'s view as in $\mathbf{H_3}$. Indeed, in $\mathbf{H_3}$, the view of $\mathcal{S}$ in prechallenge phase is completely independent of $\mathcal{O}(C_b, r)$ because $\mathcal{O}$ is a random function and $\mathcal{O}_{(C_b,r)\mapsto 0^{5n}}$ is completely independent $\mathcal{O}(C_b, r)$. □

Claim 3.26. $|\Pr[\mathcal{S} \textit{ wins in } \mathbf{H_4}] - \Pr[\mathcal{S} \textit{ wins in } \mathbf{H_5}]| \leq 2^{-n}$

Proof. The difference between the two hybrids is in the choice of the oracle $\mathcal{O}$: a random injective function in the first, and a random function in the second. The proof is thus identical to the proof Claim 3.24. □

Claim 3.27. $|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_5}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_6}]| \leq O(2^{-n/6})$.

Proof. There are two differences between the hybrids. The first is in the oracle that $\mathcal{S}_1$ is given before the challenge phase: $\Gamma(f, \mathcal{O})$ in $\mathbf{H_6}$, and its tweaked version $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto 0^{5n}}, 0^{5n}, \bot)$ in $\mathbf{H_5}$. The second is in the oracle that $\mathcal{S}_2$ is given after the challenge phase: $\Gamma(f, \mathcal{O}, \widehat{C}, C_0)$ in $\mathbf{H_6}$, and $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto \widehat{C}}, \widehat{C}, C)$ in $\mathbf{H_5}$. We can thus bound the difference between the winning probabilities in $\mathbf{H_5}$ and $\mathbf{H_6}$ as follows:

$$
\begin{aligned}
&|\Pr[\mathcal{S} \text{ wins in } \mathbf{H_5}] - \Pr[\mathcal{S} \text{ wins in } \mathbf{H_6}]| \leq \\
&\Pr_{\substack{\mathcal{S}_1,\mathcal{O}\\ r,\Gamma}}\left[\mathsf{state} := \mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n) \neq \mathcal{S}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}},0^{5n},\bot)}(1^n)\right] + \\
&\Pr_{\substack{\mathcal{S},\mathcal{O}\\ r,\Gamma}}\left[\mathcal{S}_2^{\Gamma(f,\mathcal{O},\widehat{C},C_0)}(\mathsf{state},\widehat{C}) \neq \mathcal{S}_2^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto \widehat{C}},\widehat{C},C)}(\mathsf{state},\widehat{C}) \;\middle|\; \mathsf{state} = \mathcal{S}_1^{\Gamma(f,\mathcal{O})}(1^n)\right] \quad ,
\end{aligned}
$$

where the probabilities are over the coins of $\mathcal{S} = (\mathcal{S}_1, \mathcal{S}_2)$ and Γ (specifically, StaDif) and the choice of $r \leftarrow \{0,1\}^n$, and $\mathcal{O} \leftarrow \mathbf{I}_{2n}^{5n}$, and $b \leftarrow \{0,1\}$.

As proved in Claim 3.13, the first summand is bounded by $O(2^{-n/6})$. We argue that a similar bound holds for the second summand as well. The proof is essentially identical to that of Claim 3.13 with one exception: in Claim 3.13, we argued that $\mathsf{Eval}_{0^{5n},\bot}^{f,\mathcal{O}_{(C_b,r)\mapsto 0^{5n}}}$ in $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto 0^{5n}}, 0^{5n}, \bot)$ behaves exactly as $\mathsf{Eval}^{f,\mathcal{O}}$. Here, we need to argue that $\mathsf{Eval}_{\widehat{C},C}^{f,\mathcal{O}_{(C_b,r)\mapsto \widehat{C}}}$ behaves exactly as $\mathsf{Eval}_{\widehat{C},C_0}^{f,\mathcal{O}}$. Indeed, the two can only differ on inputs of the form $(\widehat{C}, x)$, where the first would return $C_b(x)$ and the second $C_0(x)$. However, by the functional equivalence of (C_0, C_1), the two are identical. □

To conclude the proof of Theorem 3.12, we observe that

Claim 3.28. $\Pr[\mathcal{S} \text{ wins in } \mathbf{H_6}] = \frac{1}{2}$.

Proof. The view of $\mathcal{S}$ in this hybrid is completely independent of the random choice of b. □

□

4 One-Way Functions, Indistinguishability Obfuscation, and Hardness in NP ∩ coNP

In this section, we show that injective one-way functions (IOWFs) and indistinguishability obfuscation (IO), for circuits with IOWF-gates, do not give rise to a black-box construction of hard problems in NP ∩ coNP. This can be seen as a generalization of previous separations by Rudich [Rud88], showing that OWFs do not give hardness in NP ∩ coNP, by Matsuda and Matsuura [MM11], showing that IOWFs do not give one-way permutations (which are a special case of hardness NP∩coNP), and by Asharov and Segev [AS16], showing that OWFs and IO do not give one-way permutations. As in the previous section, the result implies that many cryptographic primitives and hardness in PPAD, also do not yield black-box constructions of hard problems in NP ∩ coNP.

We first define the framework of NP ∩ coNP relative to oracles, define fully black-box constructions of hard NP ∩ coNP problems, and then move on to the actual separation.

4.1 NP ∩ coNP

Throughout, we shall canonically represent languages $L \in \mathsf{NP} \cap \mathsf{coNP}$ by their corresponding non-deterministic poly-time verifiers V_1, V_0, where

$$L = \{x \in \{0,1\}^\star \mid \exists w : V_1(x,w) = 1\} \ ,$$
$$\overline{L} = \{x \in \{0,1\}^\star \mid \exists w : V_0(x,w) = 1\} = \{0,1\}^\star \setminus L \ .$$

Hardness in NP ∩ coNP from Cryptography - Motivation. Hard (on average) problems in NP∩coNP are known to follow based on certain number-theoretic problems in cryptography, such as Discrete Log and Factoring. As in the previous section for SZK, we are interested in understanding which cryptographic primitives would imply such hardness, again with the intuition that these should be appropriately structured. For instance, it is known [Bra79] that any one-way permutation $f : \{0,1\}^n \to \{0,1\}^n$ implies a hard problem in NP ∩ coNP, e.g. given an index $i \in [n]$ and an image $f(x)$ find the i-th pre-image bit x_i. In contrast, in his seminal work, Rudich [Rud88] proved that completely unstructured objects like one-way functions cannot construct even worst-case hard instances by fully black-box constructions. Here a fully black-box construction essentially means that the non-deterministic verifiers only make black-box use of the OWF (or OWP in the previous example) and the reduction establishing the hardness is also black-box (in both the adversary and the OWF).

But what about more structured primitives such as public-key encryption, oblivious transfer, or even indistinguishability obfuscation. Indeed, IO (plus OWFs) has-been shown to imply hardness in PPAD and more generally in the class TFNP of total search problems, which is often viewed as the search analog of NP ∩ coNP [MP91]. We will show, however, that fully black-box constructions do not give rise to a hard problem in NP ∩ coNP from OWFs (or even injective OWFs) and IO for circuits with OWF gates.

4.2 Fully Black-Box Constructions of Hardness in NP∩coNP from IO and IOWFs

We start by defining NP ∩ coNP relative to oracles [Rud88]. This, in particular, captures black-box constructions of such languages from cryptographic primitives, such as one-way functions in [Rud88] or indistinguishability obfuscation, which we will consider in this work.

Definition 4.1 (NP ∩ coNP relative to oracles)**.** Let $\mathfrak{S}$ be a family of oracles and let $V_1^{(\cdot)}, V_0^{(\cdot)}$ be a pair of oracle-aided non-deterministic polynomial-time verifiers. We say that V_1, V_0 define a collection of languages $L^{\mathfrak{S}} = \left\{L^\Gamma \;\middle|\; \Gamma \in \mathfrak{S}\right\}$ in NP ∩ coNP relative to $\mathfrak{S}$ if for any $\Gamma \in \mathfrak{S}$, the machines V_1^Γ, V_0^Γ define a language $L^\Gamma \in \mathsf{NP}^\Gamma \cap \mathsf{coNP}^\Gamma$. That is

$$L^\Gamma = \left\{x \in \{0,1\}^\star \;\middle|\; \exists w : V_1^\Gamma(x,w) = 1\right\} \ ,$$
$$\overline{L}^\Gamma = \left\{x \in \{0,1\}^\star \;\middle|\; \exists w : V_0^\Gamma(x,w) = 1\right\} = \{0,1\}^\star \setminus L \ .$$

We now formally define the class of constructions and reductions ruled out. That is, *fully black-box* constructions of hard problems in NP ∩ coNP from injective one-way functions (IOWFs) and IO for IOWF-aided circuits. The definition is similar in spirit to those in [AS15, AS16] and in the Section 3, adapted to the context of NP ∩ coNP hardness.

Definition 4.2. A fully black-box construction of a hard $\mathsf{NP} \cap \mathsf{coNP}$ problem L from IOWFs and IO for the class $\mathcal{C}$ of circuits with IOWF-gates is given by two oracle aided poly-time machines (V_0, V_1) and a probabilistic oracle-aided reduction $\mathcal{R}$ that satisfy:

1. **Structure:** Let $\mathfrak{S}$ be the family of all oracles $(f, i\mathcal{O})$ such that f is injective and $i\mathcal{O}$ is a function such that $\widehat{C^f} \equiv C^f$ for any $C^{(\cdot)} \in \mathcal{C}$, r, and $\widehat{C}^{(\cdot)} := i\mathcal{O}(C, r)$. Then (V_0, V_1) define a language $L^{f,i\mathcal{O}} \in \mathsf{NP}^{f,i\mathcal{O}} \cap \mathsf{coNP}^{f,i\mathcal{O}}$ relative to any oracle $(f, i\mathcal{O}) \in \mathfrak{S}$ (as per Definition 4.1).

2. **Black-box security proof:** There exist functions $q_{\mathcal{R}}(\cdot), \varepsilon_{\mathcal{R}}(\cdot)$ such that the following holds. Let $(f, i\mathcal{O})$ be any distribution supported on the family $\mathfrak{S}$ defined above. Then for any probabilistic oracle-aided $\mathcal{A}$ that *decides* $L^{f,i\mathcal{O}}$ *in the worst-case*, namely, for all $n \in \mathbb{N}$

$$\Pr_{f,i\mathcal{O},\mathcal{A}}\left[\mathcal{A}^{f,i\mathcal{O}}(x) = b \quad \text{for all} \quad \begin{array}{c} x \in \{0,1\}^n, b \in \{0,1\} \\ \text{such that } V_b(x) = 1 \end{array}\right] = 1$$

   the reduction breaks either f *or* $i\mathcal{O}$, namely, for infinitely many $n \in \mathbb{N}$ either

$$\Pr_{\substack{x \leftarrow \{0,1\}^n \\ f,i\mathcal{O},\mathcal{A}}}\left[\mathcal{R}^{\mathcal{A},f,i\mathcal{O}}(f(x)) = x\right] \geq \varepsilon_{\mathcal{R}}(n) \ ,$$

   or

$$\left|\Pr\left[\mathsf{Exp}^{\mathsf{iO}}_{(f,i\mathcal{O}),i\mathcal{O},\mathcal{C},\mathcal{R}^{\mathcal{A}}}(n) = 1\right] - \frac{1}{2}\right| \geq \varepsilon_{\mathcal{R}}(n) \ ,$$

   where in both $\mathcal{R}$ makes at most $q_{\mathcal{R}}(n)$ queries to any of its oracles $(\mathcal{A}, f, i\mathcal{O})$, and for any query x made to $\mathcal{A}$, the non-deterministic verifiers $V_0^{f,i\mathcal{O}}(x), V_1^{f,i\mathcal{O}}(x)$ make at most $q_{\mathcal{R}}(n)$ queries to their oracles (for any non-deterministic choice of a witness w). The random variable $\mathsf{Exp}^{\mathsf{iO}}_{(f,i\mathcal{O}),i\mathcal{O},\mathcal{C},\mathcal{R}^{\mathcal{A}}}(n)$ represents the reductions winning probability in the IO security game (Definition 2.2) relative to $(f, i\mathcal{O})$.

Remark about Correct Structure. We note that here we explicitly do put a *correctness* requirement, which we refer to as *structure*; namely, that the construction yields a language in $\mathsf{NP} \cap \mathsf{coNP}$ for any implementation of OWPs and IO. This is different from the setting from Definition 3.2 where we considered *promise problems* and allowed the construction not to satisfy the promise occasionally.

Concretely, we require that V_0, V_1 give a language in $\mathsf{NP} \cap \mathsf{coNP}$ for *every* oracle implementing IOWFs and IO. This follows the modeling of [BI87],[15] and aligns with how we usually think about *correctness* of black-box constructions of cryptographic primitives. For instance, the construction of public-key encryption from trapdoor permutations is promised to be correct, for all oracles implementing the trapdoor permutation. Similarly, the construction of hard $\mathsf{NP} \cap \mathsf{coNP}$ languages from one-way permutations, give an $\mathsf{NP} \cap \mathsf{coNP}$ language for any oracle implementing a permutation.

We also note that as in Section 3, our definition addresses *worst-case hardness*, which makes our impossibility result stronger. See further discussion after Definition 3.2 in Section 3.

[15] Rudich [Rud88] also considered a slight relaxation of constructions that are correct for an overwhelming fraction of oracles rather than all.

Ruling out Fully Black-Box Constructions: A Road Map. Our main result in this section is that fully black-box constructions of a hard $\mathsf{NP} \cap \mathsf{coNP}$ problem from IO and IOWFs do not exist. Furthermore, this holds even if the latter primitives are exponentially secure.

Theorem 4.3. *Any fully black-box construction of an* $\mathsf{NP} \cap \mathsf{coNP}$ *problem* L *from IOWFs and IO for circuits with IOWF gates has an exponential security loss:* $\max(q_{\mathcal{R}}(n), \varepsilon_{\mathcal{R}}^{-1}(n)) \geq \Omega(2^{n/6})$.

The proof of the theorem follows a similar methodology to that in Section 3. We exhibit two (distributions on) oracles $(\Psi, \mathsf{Decide}^{\Psi})$, where Ψ realizes IOWFs and IO for circuits with IOWF gates, and Decide^{Ψ} that decides $L^{\Psi} \in \mathsf{NP}^{\Psi} \cap \mathsf{coNP}^{\Psi}$ in the worst case. We then show that the primitives realized by Ψ are (exponentially) secure even in the presence of Decide^{Ψ}. Then viewing Decide as a worst-case decider $\mathcal{A}$ (as per Definition 4.2) directly implies Theorem 4.3, ruling out fully black-box constructions with a subexponential security loss.

The rest of this section is organized according to the above plan. First, in Section 4.3, we describe the oracle Decide^{Ψ}. As a warm-up, in Section 4.4 we show that injective one-way functions cannot construct hard languages in $\mathsf{NP} \cap \mathsf{coNP}$ in a black-box manner. Then in Section 4.5, we describe the oracle Ψ such that even in the presence of Decide^{Ψ}, (exponentially) secure OWFs and indistinguishability obfuscation exist. This rules out *fully black-box* constructions of even *worst-case*-hard problems in $\mathsf{NP} \cap \mathsf{coNP}$. In Section 4.6, we generalize our impossibility result to rule out what are called *relativizing reductions*.

4.3 The Decision Oracle

In this section, we construct an oracle $\mathsf{Decide}_{\mathfrak{S}}$ that is defined with respect to a family $\mathfrak{S}$ of oracles (e.g., all oracles implementing IOWF and IO), and which given access to $\Psi \in \mathfrak{S}$ decides any language in $\mathsf{NP}^{\Psi} \cap \mathsf{coNP}^{\Psi}$.

Definition 4.4 (Oracle $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$)**.** For a family of oracles $\mathfrak{S}$, we define the $\mathsf{Decide}_{\mathfrak{S}}$ oracle as follows:

- $\mathsf{Decide}_{\mathfrak{S}}$ is given oracle access to some Ψ. (In our setting, it will always be the case that $\Psi \in \mathfrak{S}$).
- Decide takes as input a pair of oracle-aided circuits (V_0, V_1) along with an input z where the circuits V_0, V_1 (allegedly) define a language in $\mathsf{NP} \cap \mathsf{coNP}$ relative to $\mathfrak{S}$.
- $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}(V_0, V_1, z)$ does the following:
  1. Checks that $V_0^{\Psi'}, V_1^{\Psi'} \in \mathsf{NP}^{\Psi'} \cap \mathsf{coNP}^{\Psi'}$ for all $\Psi' \in \mathfrak{S}$. If not, output $\perp$.
  2. For the input z, it outputs the unique b such that there exists a witness w satisfying $V_b^{\Psi}(z, w) = 1$. (Since V_0^{Ψ}, V_1^{Ψ} define an $\mathsf{NP} \cap \mathsf{coNP}$ language such b indeed exists and is unique.)

A few remarks about the $\mathsf{Decide}_{\mathfrak{S}}$ oracle.

1. We will use the $\mathsf{Decide}_{\mathfrak{S}}$ oracle in a similar way to the StaDif oracle in Section 3. We will be interested in the family of oracles $\mathfrak{S}$ that implements a required primitive $\mathcal{P}$ (eventually IOWFs and IO). We will show a distribution Ψ supported on $\mathfrak{S}$ that securely implements $\mathcal{P}$ in the presence of $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$, whereas at the same time, $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$ will enable to decide any language in $\mathsf{NP}^{\Psi} \cap \mathsf{coNP}^{\Psi}$ given by verifiers that define n $\mathsf{NP} \cap \mathsf{coNP}$ language relative to any oracle in $\mathfrak{S}$.

2. Queries to the oracle are represented as circuit verifiers V_0, V_1. We will consider adversaries that only produce V_0, V_1 that make some bounded number of oracle queries to Ψ.

3. The behavior of the oracle $\mathsf{Decide}^{\Psi}_{\mathfrak{S}}$ may be undefined for oracle Ψ outside $\mathfrak{S}$. In our analysis, all oracles considered will be taken from the family $\mathfrak{S}$.

To rule out fully black-box constructions of hard languages in $\mathsf{NP} \cap \mathsf{coNP}$ we have to show two things. First, that $\mathsf{Decide}^{\Psi}_{\mathfrak{S}}$ is sufficient to decide any $\mathsf{NP}^{\Psi} \cap \mathsf{coNP}^{\Psi}$ language given by verifiers that define an $\mathsf{NP} \cap \mathsf{coNP}$ language relative to any oracle in $\mathfrak{S}$. Second, it is not helpful in breaking IOWFs and indistinguishability obfuscation.

The first part follows directly from the definition of this oracle.

Claim 4.5. *Let $\mathfrak{S}$ be any family and let (V_0, V_1) be any pair of polynomial-time verifiers that define a collection $L^{\mathfrak{S}} = \left\{L^{\Psi}\right\}_{\Psi \in \mathfrak{S}}$ in $\mathsf{NP} \cap \mathsf{coNP}$, then for any oracle $\Psi \in \mathfrak{S}$,*

$$L^{\Psi} \in \mathsf{P}^{\Psi, \mathsf{Decide}^{\Psi}_{\mathfrak{S}}} \ .$$

The second part is the more challenging one. Our proof strategy is somewhat inspired by the proof of Theorem 3.19 for the case of SZK. Roughly speaking, we will aim to show that the oracle $\mathsf{Decide}^{\Psi}_{\mathfrak{S}}$ is in some sense insensitive to *random local changes*, whereas breaking the latter cryptographic primitives does require the ability to detect such changes.

Towards fulfilling this proof strategy, we now prove a general claim that roughly says that the answers of $\mathsf{Decide}^{\Psi}_{\mathfrak{S}}$ to any specific query are always determined by the behavior of Ψ on a relatively small "critical" set. Intuitively, this means that random changes that "evade" this critical set will go undetected by the oracle.

In what follows, we call a verifier circuit $V : \{0,1\}^n \times \{0,1\}^m \to \{0,1\}$ q-query if for any $z \in \{0,1\}^n$, and any potential witness $w \in \{0,1\}^m$, the circuit $V^{\Psi}(z, w)$ makes at most q queries to Ψ. Similarly, we call a query (V_0, V_1, z) to the oracle $\mathsf{Decide}_{\mathfrak{S}}$ q-bounded if both the verifiers V_0 and V_1 are q-query verifiers.

Claim 4.6. *Let $\mathfrak{S}$ be any family of oracles. Consider an oracle Ψ from $\mathfrak{S}$. Consider any q-bounded query (V_0, V_1, z) to $\mathsf{Decide}^{\Psi}_{\mathfrak{S}}$. Then there exists a set of queries $\mathbf{C} = \mathbf{C}(\Psi, V_0, V_1, z)$, which we call* a critical set*, such that*

1. *The critical set $\mathbf{C}$ is small: $|\mathbf{C}| \leq q$.*

2. *Consider another oracle $\Psi' \in \mathfrak{S}$. If the two oracles agree on the set $\mathbf{C}$, then the corresponding $\mathsf{Decide}_{\mathfrak{S}}$ oracles also agree. That is for every $\Psi' \in \mathfrak{S}$ such that $\Psi|_{\mathbf{C}} = \Psi'|_{\mathbf{C}}$,*

$$\mathsf{Decide}^{\Psi}_{\mathfrak{S}}(V_0, V_1, z) = \mathsf{Decide}^{\Psi'}_{\mathfrak{S}}(V_0, V_1, z) \ .$$

Proof. At high-level, the proof exploits the $\mathsf{NP} \cap \mathsf{coNP}$ structure; namely, for (V_0, V_1) corresponding to a language $L \in \mathsf{NP}^{\Psi} \cap \mathsf{coNP}^{\Psi}$, and any input z, if $z \in L$, then all the accepting witnesses w certify that $V_1^{\Psi}(z, w) = 1$ and no witness exists that certifies $V_0^{\Psi}(z, w) = 1$ (and vice versa, for $z \notin L$). So, as long as one witness is consistent across the oracles Ψ, Ψ', the answer of $\mathsf{Decide}_{\mathfrak{S}}$ remains invariant. The critical set $\mathbf{C}(\Psi, V_0, V_1, z)$ would simply correspond to the queries made by the verifiers for some specific witness.

Formally, consider any query (V_0, V_1, z). If (V_0, V_1) do not define a language in $\mathsf{NP} \cap \mathsf{coNP}$ relative to some oracle in $\mathfrak{S}$, then by definition $\mathsf{Decide}_{\mathfrak{S}}$ always returns $\perp$, and the claim trivially

follows ($\mathbf{C}$ can be set to be the empty set). Hence, from hereon, we assume that (V_0, V_1) do define a collection of languages $L^{\mathfrak{S}} = \{L_\Psi\}_{\Psi \in \mathfrak{S}}$ in $\mathsf{NP} \cap \mathsf{coNP}$.

Let $b := \mathsf{Decide}^{\Psi}_{\mathfrak{S}}(V_0, V_1, z)$. Consider the lexicographically first witness w which certifies this fact; namely, the first witness for which $V_b^{\Psi}(z, w) = 1$. We define $\mathbf{C} = \mathbf{C}(\Psi, V_0, V_1, z)$ to be the queries $V_b^{\Psi}(z, w)$ makes to Ψ to verify that $V_b^{\Psi}(z, w) = 1$. The bound on the size of $\mathbf{C}$ follows from the fact that V_b is a q-query verifier.

Now, we consider any $\Psi' \in \mathfrak{S}$ that is consistent with Ψ on $\mathbf{C}$:

$$\Psi|_{\mathbf{C}} = \Psi'|_{\mathbf{C}} \ .$$

Then by definition $V_b^{\Psi'}(z, w) = 1$. Since $\Psi' \in \mathfrak{S}$, the language $L^{\Psi'}$ defined by $V_0^{\Psi'}, V_1^{\Psi'}$ is in $\mathsf{NP}^{\Psi'} \cap \mathsf{coNP}^{\Psi'}$. This fixes the answer $\mathsf{Decide}^{\Psi'}_{\mathfrak{S}}(V_0, V_1, z)$ to b as required. □

4.4 Warmup: Injective One-Way Functions in the Presence of $\mathsf{Decide}_{\mathfrak{S}}$

As a warmup, we consider the case where an oracle family that only implements injective one-way functions (IOWFs), and show there is no fully black-box construction of a hard $\mathsf{NP} \cap \mathsf{coNP}$ problem from such oracles. This generalizes a result of [MM11] which shows that injective one-way functions cannot be used to construct one-way permutations in a black-box manner.[16]

Let $\mathfrak{S}$ be the family of injective one-bit expanding functions. As an implementation for the IOWF we will consider an oracle f that is sampled uniformly at random from $\mathfrak{S}$.

Definition 4.7 (Oracle f)**.** Let $\mathbf{I}_n^m$ denote the distribution on all injective functions from $\{0,1\}^n$ to $\{0,1\}^m$. The IOWF oracle is defined as $f = \{f_n\}_{n \in \mathbb{N}}$ where $f_n \leftarrow \mathbf{I}_n^{n+1}$ for all $n \in \mathbb{N}$.

As already discussed above, the oracle $\mathsf{Decide}^{f}_{\mathfrak{S}}$ allows deciding any language in $\mathsf{NP}^f \cap \mathsf{coNP}^f$ given by verifiers V_0, V_1 that define an $\mathsf{NP} \cap \mathsf{coNP}$ language relative to any oracle in $\mathfrak{S}$. We will show that f is one-way, even in the presence of the oracle $\mathsf{Decide}_{\mathfrak{S}}$. We will show that this is the case, even given an exponential number of queries to f and $\mathsf{Decide}^{f}_{\mathfrak{S}}$, and even if the queries (V_0, V_1, z) consist of verifiers that make an exponential number of queries.

In what follows, we call an adversary q-query if on any input y, the adversary makes at most q queries to either f or $\mathsf{Decide}^{f}_{\mathfrak{S}}$. Furthermore, each query (V_0, V_1, z) to $\mathsf{Decide}^{f}_{\mathfrak{S}}$ is q-bounded (as previously defined — the verifiers are circuits that make at most q queries to f).

Theorem 4.8. *Let $q = O(2^{n/3})$. Then any q-query adverasry cannot invert f_n except with exponential small probability:*

$$\Pr_{x \leftarrow \{0,1\}^n, f}\left[\mathcal{A}^{f, \mathsf{Decide}^{f}_{\mathfrak{S}}}(f_n(x)) = x\right] \leq O(2^{-n/3}) \ .$$

Proof. We need to show that even given access to the Decide oracle, an adversary cannot invert f. We show this is via a coupling argument. We want to look at the adversary's view in two worlds — the *real* world where the adversary gets a challenge $f(x)$ for a random x and the *ideal* world where the adversary gets a random element in the co-image $y \leftarrow \{0,1\}^{n+1} \setminus \mathrm{Image}(f)$ as the challenge that is completely independent of x. We will show that with very high probability, the adversary's view in both the worlds is identical. To this end, we consider three hybrids.

A description of the hybrids is given below and in Table 3.

[16] [MM11] show a slightly different statement — they consider injective functions that are adaptively one-way. That is, even given the ability to invert the function at all values except the challenge, it is still hard to invert. Our proof works unchanged for this stronger definition. We omit it for simplicity of exposition.

$\mathbf{H_1}$ This is the OWF security game. We pick a random injective function f and a random input x. The adversary gets $f(x)$ as the challenge to invert.

$\mathbf{H_2}$ This is also the OWF security game, but sampling is done differently. We sample the OWF by first sampling f along with a random pair $x \in \{0,1\}^n, y \in \{0,1\}^{n+1}$ where $y \notin \text{Image}(f)$ and the oracle is $f_{x \mapsto y}$, which behaves identically to f, except on x that is mapped to y.

$\mathbf{H_3}$ This is the ideal world. Here we sample f at random, and $y \in \{0,1\}^{n+1} \setminus \text{Image}(f)$ and set y to be the challenge.

Hybrid	**$\mathbf{H_1}$ (Real)**	**$\mathbf{H_2}$**	**$\mathbf{H_3}$ (Ideal)**
Injective OWF	$f = \left\{ f_k \leftarrow \mathbf{I}_k^{k+1} \right\}_{k \in \mathbb{N}}$		
Preimage	$x \leftarrow \{0,1\}^n$		
Planted Image	$y \leftarrow \{0,1\}^{n+1} \setminus \text{Image(f)}$		
Challenge	$f(x)$	y	y
Oracle	f, Decide^f	$f_{x \mapsto y}, \mathsf{Decide}^{f_{x \mapsto y}}$	f, Decide^f
Winning Condition	Find x		

Table 3: The hybrid experiments.

We will now show that the adversary cannot distinguish between the hybrids and hence cannot invert.

Claim 4.9. $\Pr_{f,x,y}[\mathcal{A} \textit{ wins in } \mathbf{H_1}] = \Pr_{f,x,y}[\mathcal{A} \textit{ wins in } \mathbf{H_2}]$.

Proof. We observe that the view of the adversary is distributed identically in the two hybrids. We are picking a random f and a random y outside the range and planting it at a random $x \in \{0,1\}^n$. The new oracle $f_{x \mapsto y}$ is also uniformly distributed in $\mathbf{I}_n^{n+1}$. Also, in both cases, conditioned on the function, y is distributed uniformly at random in $\text{Image}(f) \cap \{0,1\}^{n+1}$. Overall, the views are identically distributed:

$$(f, f(x)) \equiv (f_{x \mapsto y}, y) \ .$$

□

We next show that the hybrids $\mathbf{H_2}$ and $\mathbf{H_3}$ are indistinguishable.

Claim 4.10. $|\Pr_{f,x,y}[\mathcal{A} \textit{ wins in } \mathbf{H_2}] - \Pr_{f,x,y}[\mathcal{A} \textit{ wins in } \mathbf{H_3}]| \leq 2^{-n/3}$.

At high-level, to show this, we note that $f_{x \mapsto y}$ and f differ in exactly one location — x. Furthermore, we know that in the ideal world ($\mathbf{H_3}$), x is completely independent of the adversary's view. It immediately follows that the probability that queries made to f coincide with x is exponentially small, and thus the answers to these queries wouldn't change in $\mathbf{H_2}$. We would then like to show that the answers given by $\mathsf{Decide}^f_{\mathfrak{S}}$ are also invariant with overwhelming probability. Here we shall crucially use the $\mathsf{NP} \cap \mathsf{coNP}$ structure of queries given by Claim 4.6, from which we can deduce that it suffices to show that x does not coincide some small critical set. We now turn to the formal proof.

Proof. We show that, with overwhelming probability, the adversary has the same view (and thus the same output) in both $\mathbf{H_2}$ and $\mathbf{H_3}$:

$$\Pr_{\substack{x\leftarrow\{0,1\}^n, f \\ y\leftarrow\{0,1\}^{n+1}\setminus\mathrm{Image}(f)}}\left[\mathcal{A}^{f_{x\mapsto y},\mathsf{Decide}^{f_{x\mapsto y}}_{\mathfrak{S}}}(y)\neq\mathcal{A}^{f,\mathsf{Decide}^{f}_{\mathfrak{S}}}(y)\right]\leq 2^{-n/3}\ .$$

To show this we prove the following claim:

Claim 4.11. *Fix any $f\in\mathfrak{S}$ and $y\in\{0,1\}^{n+1}\setminus\mathrm{Image}(f)$. Then*

1. *For any query (V_0,V_1,z) that $\mathcal{A}^{f,\mathsf{Decide}^{f}_{\mathfrak{S}}}(y)$ makes to $\mathsf{Decide}^{f}_{\mathfrak{S}}$,*

$$\Pr_{x\leftarrow\{0,1\}^n}\left[\mathsf{Decide}^{f}(V_0,V_1,z)\neq\mathsf{Decide}^{f_{x\mapsto y}}(V_0,V_1,z)\right]\leq 2^{-2n/3}\ .$$

2. *For any query z that $\mathcal{A}^{f,\mathsf{Decide}^{f}_{\mathfrak{S}}}(y)$ makes to f,*

$$\Pr_{x}\left[f(z)\neq f_{x\mapsto y}(z)\right]\leq 2^{-n}\ .$$

Proof. To prove the first part of the claim, we crucially rely on Claim 4.6 (with respect to our family $\mathfrak{S}$ of injective functions). Recall that the adversary $\mathcal{A}$ is a q-query adversary and thus the query (V_0,V_1,z) is q-bounded. Accordingly, by Claim 4.6, there exists a critical queries $\mathbf{C}=\mathbf{C}(f,V_0,V_1,z)$, such that for any other $f'\in\mathfrak{S}$ that agrees with f on $\mathbf{C}$,

$$\mathsf{Decide}^{f}(V_0,V_1,z)=\mathsf{Decide}^{f'}(V_0,V_1,z)\ .$$

Thus all that we need to show is that overwhelming probability x is such that $f_{x\mapsto y}$ is injective (namely, in $\mathfrak{S}$), and agrees with f on $\mathbf{C}$. Indeed, $f_{x\mapsto y}$ is always injective since $y\notin\mathrm{Image}(f)$. Second, $f_{x\mapsto y}|_{\mathbf{C}}=f|_{\mathbf{C}}$ unless $x\in\mathbf{C}$. Since x is sampled independently of $\mathcal{A}$'s view in $\mathbf{H_3}$, and in particular independently of $\mathbf{C}$,

$$\Pr_{x\leftarrow\{0,1\}^n}\left[x\in\mathbf{C}\right]\leq|\mathbf{C}|\cdot 2^{-n}\leq q\cdot 2^{-n}\leq 2^{n/3}\cdot 2^{-n}=2^{-2n/3}\ .$$

For the second part of the claim, note that $f(z)\neq f_{x\mapsto y}(z)$, unless $z=x$. As before, since x is sampled independently of $\mathcal{A}$'s view in $\mathbf{H_3}$, and in particular independently of z , this probability is at most 2^{-n}.

□

Given Claim 4.11, we can take a union bound over all queries that $\mathcal{A}^{f,\mathsf{Decide}^{f}_{\mathfrak{S}}}(y)$ makes to the deduce that the answers to all remains invariant when considering the oracles $f_{x\mapsto y},\mathsf{Decide}^{f_{x\mapsto y}}_{\mathfrak{S}}$ except with probability

$$q\cdot\max(2^{-2n/3},2^{-n})\leq 2^{n/3}\cdot 2^{-2n/3}=2^{-n/3}\ .$$

This completes the proof of the Claim 4.10. □

To complete the proof of Theorem 4.8, it is left to note that in the ideal world, the adversary cannot invert.

Claim 4.12. *The adversary cannot win in the Ideal world. Concretely, for every fixed f,*

$$\Pr_{x,y}[\mathcal{A} \text{ wins in } \mathbf{H_3}] = 2^{-n} \ .$$

Proof. In the third hybrid $\mathbf{H_3}$, the challenge y is independent of the answer x, which is chosen uniformly at random. So, with probability 2^{-n}, the adversary's response will be x. □

Putting all of the above claims together, the adversary inverts in the real world ($\mathbf{H_1}$) with probability at most

$$2^{-2n/3} + 2^{-n} \leq O(2^{-n/3}) \ .$$

□

4.5 Indistinguishability Obsfuscation (and IOWFs) in the Presence of Decide

In this section, we generalize Theorem 4.8 to show that injective one-way functions (IOWFs) *and indistinguishability obfuscation (IO)* cannot be used to construct worst-case hard $\mathsf{NP} \cap \mathsf{coNP}$ instances in a fully black-box way. We start by discussing an aspect of IO that turns out to be crucial for this separation — *verifiability*.

Verifiability of IO. Looking back at our separation for the SZK case in Section 3, we observe that it, in fact, holds also for a stronger definition of IO that is verifiable and *unambiguous*; namely, it is possible to efficiently determine whether a given string is a valid obfuscation of *some* circuit, and this circuit is uniquely determined. Indeed, looking at the oracle $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{\mathcal{O},f})$, implementing OWFs and IO there, it induces *valid* obfuscation which are strings $\widehat{C} = \mathcal{O}(C, r)$ in the image of the *injective* $\mathcal{O}$, and *invalid* ones, which are strings outside the image of $\mathcal{O}$. Furthermore, it is possible to efficiently identify which is the case, since the oracle Eval would return $\perp$ on invalid obfuscations.

Going back to the case of $\mathsf{NP} \cap \mathsf{coNP}$, we observe that verifiable and unambiguous IO actually does imply hardness in $\mathsf{NP} \cap \mathsf{coNP}$ (in a fully black-box way). Indeed, consider the language including all $(\widehat{C}, i, b)$ such that $\widehat{C}$ is a valid obfuscation and b the the ith bit of the unique circuit C it determines. Indeed, due to verifiablility and unambiguity, this language is in $\mathsf{NP} \cap \mathsf{coNP}$, and clearly any decider for this language completely breaks IO. This means that we cannot hope to rule out fully black-box constructions of $\mathsf{NP} \cap \mathsf{coNP}$ hardness from a family of oracles $\mathfrak{S}$, if this family only includes verifiable and unambiguous IO constructions. Indeed, our Definition 4.2 of black-box constructions of hard $\mathsf{NP} \cap \mathsf{coNP}$ problems considers constructions that should work for the family $\mathfrak{S}$ of all IO constructions, and we will crucially (and necessarily) rely on this. (In fact, our separation would also work for the restricted family of IO constructions that are not verifiable, but still unambiguous.)

Capturing Non-Verifiable IO. We augment our previous definition of the oracle $\Psi = (f, \mathcal{O}, \mathsf{Eval}^{\mathcal{O},f})$ in a way that allows the Eval oracle to answer arbitrarily on invalid obfuscations, which would capture non-verifable IO constructions. To this end, we consider an augmented Eval_φ parameterized by a "backup map" $\varphi : \left\{\varphi_n : \{0,1\}^{5n} \to \{0,1\}^n\right\}_n$ from obfuscations $\widehat{C}$ to circuits C. Given a query $(\widehat{C}, x)$, if the obfuscation $\widehat{C}$ is valid, Eval_φ answers it faithfully as the previously defined Eval; otherwise, Eval_φ obtains some circuit $C = \varphi(\widehat{C})$ from φ and uses it to answer the query. Indeed, this new oracle still implements indistinguishability obfuscation and does so in a non-verifiable way. This is formally defined below.

Definition 4.13 (Oracle Ψ_φ)**.** The oracle $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}})$ consists of three parts:

- $f = \{f_n\}_{n\in\mathbb{N}}$ on input $x \in \{0,1\}^n$ answers with $f_n(x)$, where f_n is a random injective one-bit expanding function $f_n \leftarrow \mathbf{I}_n^{n+1}$.
- $\mathcal{O} = \{\mathcal{O}_n\}_{n\in\mathbb{N}}$ on input $(C, r) \in \{0,1\}^n \times \{0,1\}^n$ answers with $\widehat{C} := \mathcal{O}_n(C, r)$ where $\mathcal{O}_n$ is a random injective function $\mathcal{O}_n \leftarrow \mathbf{I}_{2n}^{5n}$ into $\{0,1\}^{5n}$.
- $\mathsf{Eval}_\varphi^{f,\mathcal{O}}(\widehat{C}, x)$ checks if $\widehat{C}$ is in the image of $\mathcal{O}_n$. If it is, it finds $(C, r) = \mathcal{O}_n^{-1}(\widehat{C})$ and returns the answer $C^f(x)$. If $\widehat{C}$ is not in the image, it uses φ to answer. That is

$$\mathsf{Eval}_\varphi^{f,\mathcal{O}}(\widehat{C}, x) = \begin{cases} C^f(x) & \text{If } \widehat{C} \in \text{Image}(\mathcal{O}_n) \text{ and } \mathcal{O}_n(C, r) = \widehat{C} \\ C_\varphi^f(x) & \text{If } \widehat{C} \notin \text{Image}(\mathcal{O}_n) \text{ and } C_\varphi = \varphi(\widehat{C}) \end{cases} .$$

For any choice of φ, and realization of Ψ_φ, we obtain a construction of an obfuscator similarly to Construction 3.20.

Construction 4.14 (Obfuscator $i\mathcal{O}^{\Psi_\varphi}$)**.** *Let $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}})$. Given an oracle-aided circuit $C \in \{0,1\}^n$, $i\mathcal{O}^{\Psi}(1^n, C)$ samples a random $r \leftarrow \{0,1\}^n$, computes $\widehat{C} = \mathcal{O}(C, r)$, and returns an oracle aided circuit $E_{\widehat{C}}$ that given input x, computes $\mathsf{Eval}_\varphi^{f,\mathcal{O}}(\widehat{C}, x)$.*

As in Section 3, $i\mathcal{O}^{\Psi_\varphi}$ satisfies the functionality requirement of Definition 2.2 for f-aided circuits, and this is the case for any choice of mapping φ. Indeed, functionality puts no restriction on how evaluation behaves for invalid obfuscations. Accordingly, the family $\mathfrak{S}$, considered in our Definition 4.2 of black-box constructions of hard problems in $\mathsf{NP} \cap \mathsf{coNP}$, includes $i\mathcal{O}^{\Psi_\varphi}$ for all φ. From hereon, we shall often abuse notation and write $\Psi_\varphi \in \mathfrak{S}$ rather than $i\mathcal{O}^{\Psi_\varphi} \in \mathfrak{S}$.

To rule out fully black-box constructions relative to the family $\mathfrak{S}$, we consider again the oracle $\mathsf{Decide}_{\mathfrak{S}}$. We show that for any specific choice of φ, in the presence of $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$,

1. Any language L^{Ψ_φ} defined by $(V_0^{\Psi_\varphi}, V_1^{\Psi_\varphi})$ is easy to decide provided that (V_0, V_1) define a language in $\mathsf{NP} \cap \mathsf{coNP}$ relative to any oracle in $\mathfrak{S}$.
2. f is a one-way function.
3. $i\mathcal{O}^{\Psi_\varphi}$ is a secure indistinguishability obfuscation.

The first item above indeed follows from Definition 4.4 of the oracle $\mathsf{Decide}_{\mathfrak{S}}$ as claimed in Section 4.3. We stress the crucial reliance on the fact that V_0, V_1 define a language in $\mathsf{NP} \cap \mathsf{coNP}$ for all oracles in $\mathfrak{S}$. Indeed, the oracle $\mathsf{Decide}_{\mathfrak{S}}$ only responds on such V_0, V_1. The fact that $\mathsf{Decide}_{\mathfrak{S}}$ only responds on such queries, is crucially used to prove one-wayness (the second item) and indistinguishability obfuscation (the third item), where we shall use the fact $\{\Psi_\varphi\} \in \mathfrak{S}$ for all φ.

In the next two subsections, we prove the last two items. Throughout, we address adversaries with oracles $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{\mathcal{O},f})$ and $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$. We say that such an adversary is q-query if they

1. make only q queries to f,
2. make only q queries to either $\mathcal{O}$ or Eval, and any query $\widehat{C}$ to Eval is of size at most $5q$, and in particular, any oracle aided circuit C that is mapped to $\widehat{C}$ by $\mathcal{O}$ is of size at most q, and makes at most q queries to f,

3. make only q queries to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$, and for any query (V_0, V_1, z) made to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$, the verification circuits V_0, V_1 are Ψ_φ-aided and each of them is q-query (according to the two conditions above).

4.5.1 One-Wayness

We show that f is a one-way function in the presence of the $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$ oracle.

Theorem 4.15. *Let $q(n) \leq O(2^{n/6})$. Fix any φ. Then for any q-query adversary $\mathcal{A}$,*

$$\Pr_{\substack{\Psi_\varphi=(f,\mathcal{O},\mathsf{Eval}_\varphi^{f,\mathcal{O}}) \\ x}}\left[\mathcal{A}^{\Psi,\mathsf{Decide}_{\mathfrak{S}}^{\Psi}}(f(x)) = x\right] \leq O(2^{-n/6}) \ ,$$

where the probability is over the randomness of Ψ_φ and $x \leftarrow \{0,1\}^n$.

Proof. We will, in fact, prove a stronger statement: the above holds when fixing the oracles $\mathcal{O}$ and $f_{-n} := \{f_k\}_{k\neq n}$. We prove the theorem by a reduction to the case that Ψ only consists of the injective function f (and does not include $\mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}}$), proven in Theorem 4.8. Concretely, fix any q-query adversary $\mathcal{A}$ that inverts the random injective function f_n given access to $\Psi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}})$ and $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$, we show how to reduce it to an $O(q^2)$-query adversary $\mathcal{B}^f(f_n(x))$ that inverts f_n for a random $x \leftarrow \{0,1\}^n$ with the same probability as $\mathcal{A}$. (This is done similarly to the proof of Theorem 3.12).

The new adversary $\mathcal{B}^{f,\mathsf{Decide}_{\mathfrak{S}}^{f}}(f_n(x))$ emulates $\mathcal{A}^{\Psi,\mathsf{Decide}_{\mathfrak{S}}^{\Psi}}(f_n(x))$ answering Ψ-queries as follows:

- f **queries:** answered according to $\mathcal{B}$'s oracle f. This translates to at most q queries to f.
- $\mathcal{O}$ **queries:** answered according to the fixed oracle $\mathcal{O}$. This does not add any calls to f.
- $\mathsf{Eval}_\varphi^{f,\mathcal{O}}$ **queries:** given query $(\widehat{C}, x)$ to Eval, invert the fixed oracle $\mathcal{O}$ to find $(C, r) = \mathcal{O}^{-1}(\widehat{C})$. If no such preimage exists, set $C = \varphi(\widehat{C})$. Using the f-oracle, compute $C^f(x)$ and return the result. This translates to at most q^2 queries to f: q queries by C, for each of at most q queries to $\mathsf{Eval}_\varphi^{f,\mathcal{O}}$.
- $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$ **queries:** given query (V_0, V_1, z), where V_b makes Ψ_φ-queries, translate the query to D_0, D_1, z that only make f-queries, where each query to $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}})$ is translated to a query to f according to the previous three items. The resulting oracle-aided query (D_0, D_1, z) may thus make up to $q + q^2$ queries to f: q corresponding to the first item, and q^2 corresponding to the third.[17]

Overall, $\mathcal{B}^f$ is $O(q^2)$-query and perfectly emulates the view of $\mathcal{A}^\Psi$. The theorem now follows from Theorem 4.8. □

[17] We note that while there is a bound on the number of queries that they make, we do not put any restrictions on their size, which allows to hardwire the fixed $\mathcal{O}$ and f_{-n} as required in the previous three items. Indeed, Theorem 4.8 does not put any restriction on the size of these circuits.

4.5.2 Indistinguishability Obfuscation

We now show that Construction 4.14 also satisfies indistinguishability, with an exponentially small distinguishing gap, even given an exponential number of oracle queries to $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{\mathcal{O},f})$ and the decide oracle $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$.

Theorem 4.16. *Let $q(n) \leq O(2^{n/3})$. Fix any φ. Then for any q-query adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$*

$$\left| \Pr\left[\mathsf{Exp}^{\mathsf{iO}}_{\Gamma, i\mathcal{O}, \mathcal{A}}(n) = 1 \right] - \frac{1}{2} \right| \leq O(2^{-n/3}) \ ,$$

where the random variable $\mathsf{Exp}^{\mathsf{iO}}_{\Gamma, i\mathcal{O}, \mathcal{A}}(n)$ the adversary's winning probability in the IO security game (Definition 2.2) relative to $\Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}})$ and $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$.

Proof. We prove a stronger statement: the above holds when fixing the oracles f and $\mathcal{O}_{-n} = \{\mathcal{O}_k\}_{k \neq n}$. For simplicity, we often suppress oracle access to the fixed $\mathcal{O}_{-n}, f$ in our notation and only denote the oracle $\mathcal{O}_n$. Fix a q-query (w.l.o.g deterministic) adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$. To bound $\mathcal{A}$'s advantage in breaking $i\mathcal{O}$, we rely on a similar proof strategy to the one in Theorem 4.8. We will consider an *ideal* world where the given challenge is uncorrelated to the bit b that the adversary will be required to guess. We will then show, through a sequence of hybrid experiments, that this world is indistinguishable from the *real* world, where the adversary get an obfuscation of the circuit C_b among two circuits C_0, C_1, which it chose.

We introduce some notation that will be useful to describe the hybrids:

- For a function $\mathcal{O} = \left\{ \mathcal{O}_k : \{0,1\}^{2k} \to \{0,1\}^{5k} \right\}_{k \in \mathbb{N}}$, a pair $(C, r) \in \{0,1\}^{n \times 2}$, and $\widehat{C} \in \{0,1\}^{5n}$, we denote by $\mathcal{O}_{(C,r) \mapsto \widehat{C}}$ the function that maps (C, r) to $\widehat{C}$ and is otherwise identical to $\mathcal{O}$.
- For a function $\varphi = \left\{ \varphi_k : \{0,1\}^{2k} \to \{0,1\}^{5k} \right\}_{k \in \mathbb{N}}$, $C \in \{0,1\}^n$, and $\widehat{C} \in \{0,1\}^{5n}$, we denote by $\varphi_{\widehat{C} \mapsto C}$ the function that maps $\widehat{C}$ to C and is otherwise identical to φ.
- For functions $\mathcal{O}, \varphi = \left\{ \mathcal{O}_k, \varphi_k : \{0,1\}^{2k} \to \{0,1\}^{5k} \right\}_{k \in \mathbb{N}}$, we denote by $\Gamma(f, \mathcal{O}, \varphi)$ the oracle

  $$\Gamma(f, \mathcal{O}, \varphi) := f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}}, \mathsf{Decide}_{\mathfrak{S}}^{f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}}} \ ,$$

  where we extend $\mathsf{Eval}_\varphi^{f,\mathcal{O}}$ to also be defined for non-injective $\mathcal{O}$: given $\widehat{C} \in \{0,1\}^{5n}$ with more than a single preimage in $\{0,1\}^{2n}$, it returns $\perp$.

The hybrid experiments are formally described in Table 4, followed by a less formal description in words.

$\mathbf{H_1}$ This is the **Real World** security game for IO. The adversary gives the challenger a pair of functionally equivalent circuits (C_0, C_1), gets back the obfuscation $\mathcal{O}(C_b, r)$ for a random b, and has to guess b.

Hybrid	$\mathbf{H_1}$ (Real)	$\mathbf{H_2}$	$\mathbf{H_3}$	$\mathbf{H_4}$	$\mathbf{H_5}$	$\mathbf{H_6}$	$\mathbf{H_7}$	$\mathbf{H_8}$ (Ideal)
Obfuscator Function	$\mathcal{O}_n \leftarrow \mathbf{I}_{2n}^{5n}$			$\mathcal{O}_n \leftarrow \mathbf{F}_{2n}^{5n}$		$\mathcal{O}_n \leftarrow \mathbf{I}_{2n}^{5n}$		
Backup Map	φ							
Challenger Randomness	$b \leftarrow \{0,1\}$, $r \leftarrow \{0,1\}^n$							
Planted Obfuscation	$\widehat{D} \leftarrow \{0,1\}^{5n} \setminus \text{Image}(\mathcal{O})$		$\widehat{D} \leftarrow \{0,1\}^{5n}$				$\widehat{D} \leftarrow \{0,1\}^{5n} \setminus \text{Image}(\mathcal{O})$	
Planted Challenge	$\widehat{C} \leftarrow \{0,1\}^{5n}$						$\widehat{C} \leftarrow \{0,1\}^{5n} \setminus \text{Image}(\mathcal{O})$	
Prechallenge Oracle	$\Gamma(f, \mathcal{O}, \varphi)$	$\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$						$\Gamma(f, \mathcal{O}, \varphi)$
Chosen Circuits	$(C_0, C_1) \leftarrow \mathcal{A}_1^{\Gamma(f,\mathcal{O})}(1^n)$ where $C_0^f \equiv C_1^f$ (relative to the fixed f)							
Challenge Obfuscation	$\mathcal{O}(C_b, r)$				$\widehat{C}$			
Postchallenge Oracle	$\Gamma(f, \mathcal{O}, \varphi)$				$\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \varphi)$			$\Gamma(f, \mathcal{O}, \varphi_{\widehat{C}\mapsto C_0})$
Winning Condition	Guess b							

Table 4: The hybrid experiments

$\mathbf{H_2}$ The pre-challenge oracle is changed to $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$ where $\widehat{D}$ is a uniformly random element outside the image of $\mathcal{O}_n$. As in Section 3, we note that in $\mathbf{H_2}$, the circuit C_b is defined according to the circuits (C_0, C_1) that $\mathcal{A}_1$ would have chosen given the non-tweaked oracle $\Gamma(f, \mathcal{O}, \varphi)$ (so there is no circularity).[18]

$\mathbf{H_3}$ In this hybrid, $\widehat{D}$ is picked uniformly at random from $\{0,1\}^{5n}$ rather than $\{0,1\}^{5n} \setminus \text{Image}(\mathcal{O})$.

$\mathbf{H_4}$ The obfuscator function $\mathcal{O}_n$ is sampled from $\mathbf{F}_n^{5n}$. That is as a completely random function rather than an injective one.

$\mathbf{H_5}$ We switch the challenge the adversary gets from $\mathcal{O}(C_b, r)$ to $\widehat{C}$. We accordingly change the post-challenge oracle to $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \varphi)$.

$\mathbf{H_6}$ We switch the obfuscator function back to an injective function from a from a random function.

$\mathbf{H_7}$ $\widehat{D}$ and $\widehat{C}$ are now sampled from $\{0,1\}^{5n} \setminus \text{Image}(\mathcal{O})$ rather than at random.

$\mathbf{H_8}$ This is the **Ideal** world. Here the oracle given to the adversary is $\Gamma(\mathcal{O}, \varphi)$ before the challenge. For the challenge, we give $\widehat{C}$ that is not in the image and change the post-challenge oracle from $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \varphi)$ to $\Gamma(f, \mathcal{O}, \varphi_{\widehat{C}\mapsto C_0})$.

We next show that the winning probability of the adversary is roughly the same in all of the above hybrids.

[18] In more detail, we first look at an execution of $\mathcal{A}_1$ with $\Gamma(f, \mathcal{O}, \varphi)$, as in $\mathbf{H_1}$, with respect to the sampled $\mathcal{O}, b, r$ (and coins of $\mathcal{A}_1$). This defines circuits (C_0, C_1), one of which is the challenge circuit C_b. Then we consider an execution with exactly the same samples $\mathcal{O}, b, r$, but with a pre-challenge oracle $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$.

Claim 4.17. $|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_1}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_2}]| \leq O(2^{-n/3})$.

Proof. The hybrids $\mathbf{H_1}$ and $\mathbf{H_2}$ differ in the pre-challenge oracle, which is changed from $\Gamma(f, \mathcal{O}, \varphi)$ to $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$ where $\widehat{D}$ is a uniformly random element in the co-image of $\mathcal{O}_n$.

We bound the difference between the above two probabilities by a coupling argument. Concretely, we can bound the difference as follows

$$\left|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_1}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_2}]\right| \leq$$
$$\Pr_{\mathcal{O},r,b,\widehat{D}}\left[\mathcal{A}_1^{\Gamma(f,\mathcal{O},\varphi)}(1^n) \neq \mathcal{A}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto\widehat{D}},\varphi)}(1^n)\right] \ ,$$

where the probability is over the choice of $r \leftarrow \{0,1\}^n$, $b \leftarrow \{0,1\}$, $\mathcal{O} \leftarrow \mathbf{I}_{2n}^{5n}$, and $\widehat{D} \leftarrow \{0,1\}^{5n} \setminus \text{Image}(\mathcal{O}_n)$. We will, in fact, show that the above is bounded for any fixed $b \in \{0,1\}, \mathcal{O}$. Indeed, for the rest of the claim, fix $b \in \{0,1\}$ and $\mathcal{O} \in \mathbf{I}_{2n}^{5n}$.

In what follows, let $\mathbf{Q}$ be the set of queries made by $\mathcal{A}_1^{\Gamma(f,\mathcal{O},\varphi)}(1^n)$ to its oracle

$$\Gamma(f, \mathcal{O}, \varphi) = (\Psi_\varphi, \mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}) \quad \text{where} \quad \Psi_\varphi = (f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}}) \ .$$

For any query $Q = (V_0, V_1, z) \in \mathbf{Q}$ made to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$, let $\mathbf{C}_Q$ denote the set of critical queries corresponding to Q, given by Claim 4.6. Note that indeed $\Psi_\varphi \in \mathfrak{S}$; namely, it is a valid oracle in the family $\mathfrak{S}$, so the claim can be applied. We denote by $\mathbf{C_Q}$ the union of all such critical sets. We now define the event $\mathbf{Hit}$, aimed at capturing the cases when the adversary's views in $\mathbf{H_1}$ and $\mathbf{H_2}$ may differ. Concretely, let $\mathbf{Hit} = \mathbf{Hit}(r, \widehat{D})$ be the event that any of the following occurs:

1. $(C_b, r) \in \mathbf{Q} \cup \mathbf{C_Q}$: the query (C_b, r) is made to $\mathcal{O}$.
2. $(\widehat{D}, x) \in \mathbf{Q} \cup \mathbf{C_Q}$ for some $x \in \{0,1\}^n$: the query $(\widehat{D}, x)$ is made to Eval_φ.
3. $(\mathcal{O}(C_b, r), x) \in \mathbf{Q} \cup \mathbf{C_Q}$ for some $x \in \{0,1\}^n$: the query $(\mathcal{O}(C_b, r), x)$ is made to Eval_φ.

Note that any query Q as above is either made directly to the corresponding oracle, i.e. $Q \in \mathbf{Q}$, or is within the critical set of queries $\mathbf{C_Q}$ (meaning that there is a query (V_0, V_1, z) to $\mathsf{Decide}_{\mathfrak{S}}$ where the one of the verifiers might make the query Q on some canonical witness, see Claim 4.6).

Claim 4.18. $\Pr_{r,\widehat{D}}\left[\mathcal{A}_1^{\Gamma(f,\mathcal{O},\varphi)}(1^n) \neq \mathcal{A}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto\widehat{D}},\varphi)}(1^n)\right] \leq \Pr_{r,\widehat{D}}[\mathbf{Hit}]$.

Proof. We observe that if $\mathbf{Hit}$ does not occur, then the view of $\mathcal{A}_1$ is the same for both oracles. Indeed, in $\mathbf{H_2}$, the oracle $\mathcal{O}$ is changed to $\mathcal{O}_{(C_b,r)\mapsto\widehat{D}}$. In particular:

- for any query $Q \neq (C_b, r)$ to $\mathcal{O}$
$$\mathcal{O}_{(C_b,r)\mapsto\widehat{D}}(Q) = \mathcal{O}(Q) \ ,$$
- for any query $Q \notin \left\{(\widehat{D}, x), (\mathcal{O}(C_b, r), x)\right\}_{x\in\{0,1\}^n}$ to $\mathsf{Eval}_\varphi^{f,\mathcal{O}}$,
$$\mathsf{Eval}_\varphi^{f,\mathcal{O}_{(C_b,r)\mapsto\widehat{D}}}(Q) = \mathsf{Eval}_\varphi^{f,\mathcal{O}}(Q) \ ,$$
since $\left(\mathcal{O}_{(C_b,r)\mapsto\widehat{D}}\right)^{-1}(Q) = \mathcal{O}^{-1}(Q)$.

It follows that for all queries $Q \in \mathbf{Q} \cup \mathbf{C_Q}$, $\Psi_\varphi(Q) = \Psi'_\varphi(Q)$, where

$$\Psi_\varphi = f, \mathcal{O}, \mathsf{Eval}_\varphi^{f,\mathcal{O}} \quad \text{and} \quad \Psi'_\varphi = f, \mathcal{O}_{(C_b,r)\mapsto \widehat{D}}, \mathsf{Eval}_\varphi^{f,\mathcal{O}_{(C_b,r)\mapsto \widehat{D}}} \ .$$

This implies that all queries made by $\mathcal{A}_1$ directly to its oracle Ψ_φ in $\mathbf{H_1}$, are answered in the same way in $\mathbf{H_2}$. It is left to show that this is also the case for queries made to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}$. For this purpose, note that $\Psi'_\varphi \in \mathfrak{S}$, namely is a valid oracle (indeed, $\widehat{D}$ is chosen outside the image of $\mathcal{O}$, so injectivity of the obfuscation oracle is guaranteed as required). Furthermore, the oracles Ψ_φ and Ψ'_φ agree on all critical sets $\mathbf{C}_Q \subseteq \mathbf{C_Q}$. It follows, by Claim 4.6, that for any $Q \in \mathbf{Q}$

$$\mathsf{Decide}_{\mathfrak{S}}^{\Psi_\varphi}(Q) = \mathsf{Decide}_{\mathfrak{S}}^{\Psi'_\varphi}(Q) \ ,$$

which completes the proof of the claim. □

It is left to bound the probability that **Hit** occurs. First, since r is chosen at random from $\{0,1\}^n$, and $\widehat{D}$ is sampled at random from $\{0,1\}^{5n} \setminus \mathrm{Image}(\mathcal{O})$, which is of size $\Omega(2^{5n})$, for any fixed query $Q \in \mathbf{Q} \cup \mathbf{C_Q}$, the probability that any of the three cases defining **Hit** occurs is at most $O(2^{-n})$. Thus, by a union bound we have

$$\Pr[\mathbf{Hit}] \leq |\mathbf{Q} \cup \mathbf{C_Q}| \cdot O(2^{-n}) \leq \left(|\mathbf{Q}| + \sum_{Q \in \mathbf{Q}} |\mathbf{C}_Q| \right) \cdot O(2^{-n}) \leq O(q^2 \cdot 2^{-n}) \leq O(2^{-n/3}) \ .$$

This completes the proof of Claim 4.17. □

Claim 4.19. $|\Pr[\mathcal{A} \textit{ wins in } \mathbf{H_2}] - \Pr[\mathcal{A} \textit{ wins in } \mathbf{H_3}]| \leq 2^{-3n}$.

Proof. The only difference between the two hybrids is the way $\widehat{D}$ is sampled. In $\mathbf{H_2}$, $\widehat{D} \notin \mathrm{Image}(\mathcal{O}_n)$ and in $\mathbf{H_3}$ the restriction is removed, and $\widehat{D}$ is uniformly random in $\{0,1\}^{5n}$. Hence the difference in advantage can be bounded by

$$\Pr_{\widehat{D} \leftarrow \{0,1\}^{5n}} \left[\widehat{D} \in \mathrm{Image}(\mathcal{O}_n) \right] = 2^{2n}/2^{5n} = 2^{-3n} \ .$$

□

Claim 4.20. $|\Pr[\mathcal{A} \textit{ wins in } \mathbf{H_3}] - \Pr[\mathcal{A} \textit{ wins in } \mathbf{H_4}]| \leq 2^{-n}$.

Proof. The only difference between the two hybrids is the way $\mathcal{O}_n$ is sampled. In $\mathbf{H_3}$, it is a random injective function, and in $\mathbf{H_4}$ it is a truly random function. Hence the difference in advantage can be bounded by

$$\Pr_{\mathcal{O}_n \leftarrow \mathbf{F}_{2n}^{5n}} [\mathcal{O}_n \text{ is not injective}] \leq 2^{-n} \ .$$

□

Claim 4.21. $\Pr[\mathcal{A} \textit{ wins in } \mathbf{H_4}] = \Pr[\mathcal{A} \textit{ wins in } \mathbf{H_5}]$.

Proof. The difference between $\mathbf{H_4}$ and $\mathbf{H_5}$ is that in $\mathbf{H_4}$, in the challenge and post challenge phases, the value $\mathcal{O}(C_b, r)$ is re-sampled uniformly at random, i.e. it is replaced everywhere by $\widehat{C} \leftarrow \{0,1\}^{5n}$. We claim that this induces exactly the same distribution on $\mathcal{A}$'s view as in $\mathbf{H_4}$. Indeed, in $\mathbf{H_3}$, the view of $\mathcal{A}$ in prechallenge phase is completely independent of $\mathcal{O}(C_b, r)$ because $\mathcal{O}$ is a random function and $\mathcal{O}_{(C_b,r)\mapsto\widehat{D}}$ is independent of $\mathcal{O}(C_b, r)$. □

Claim 4.22. $|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_5}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_6}]| \leq 2^{-n}$.

Proof. The only difference between the two hybrids is the way $\mathcal{O}_n$ is sampled. In $\mathbf{H_6}$, it is a random injective function, and in $\mathbf{H_5}$ it is a truly random function. The proof is thus identical to that of Claim 4.20. □

Claim 4.23. $|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_6}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_7}]| \leq O(2^{-3n})$.

Proof. The only difference between the two hybrids is the way $\widehat{D}, \widehat{C}$ are sampled, truly at random in $\mathbf{H_6}$ and from the co-image of $\mathcal{O}_n$ in $\mathbf{H_7}$. Hence the difference in advantage can be bounded by

$$2 \Pr_{\widehat{D}\leftarrow\{0,1\}^{5n}}\left[\widehat{D} \in \text{Image}(\mathcal{O}_n)\right] = 2 \cdot 2^{2n}/2^{5n} \leq O(2^{-3n}) \ .$$

□

Finally we show that the adversary's advantage does not change much as we shift from hybrid $\mathbf{H_7}$ to $\mathbf{H_8}$. This proof is almost identical to the proof of Claim 4.17. Here we use the fact that modifying φ does not alter the fact that Ψ_φ is a valid implementation of IO.

Claim 4.24. $|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_7}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_8}]| \leq O(2^{-n/3})$.

Proof. There are two differences between the hybrids. The first is in the oracle that $\mathcal{A}_1$ is given before the challenge phase: $\Gamma(f, \mathcal{O}, \varphi)$ in $\mathbf{H_8}$, and its tweaked version $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$ in $\mathbf{H_7}$. The second is in the oracle that $\mathcal{A}_2$ is given after the challenge phase: $\Gamma(f, \mathcal{O}, \varphi_{\widehat{C}\mapsto C_0})$ in $\mathbf{H_8}$, and $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \varphi)$ in $\mathbf{H_7}$. We can thus bound the difference between the winning probabilities in $\mathbf{H_7}$ and $\mathbf{H_8}$ as follows:

$$\begin{aligned}
&|\Pr[\mathcal{A} \text{ wins in } \mathbf{H_7}] - \Pr[\mathcal{A} \text{ wins in } \mathbf{H_8}]| \leq \\
&\Pr_{\mathcal{O},r,b,\widehat{D}}\left[\mathsf{state} := \mathcal{A}_1^{\Gamma(f,\mathcal{O},\varphi)}(1^n) \neq \mathcal{A}_1^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto\widehat{D}},\varphi)}(1^n)\right] + \\
&\Pr_{\mathcal{O},r,b,\widehat{C}}\left[\mathcal{A}_2^{\Gamma(f,\mathcal{O},\varphi_{\widehat{C}\mapsto C_0})}(\mathsf{state}, \widehat{C}) \neq \mathcal{A}_2^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto\widehat{C}},\varphi)}(\mathsf{state}, \widehat{C}) \;\middle|\; \mathsf{state} = \mathcal{A}_1^{\Gamma(f,\mathcal{O},\varphi)}(1^n)\right] \ ,
\end{aligned}$$

where the probabilities are over the choice of $r \leftarrow \{0,1\}^n$, $b \leftarrow \{0,1\}$, $\mathcal{O} \leftarrow \mathbf{I}_{2n}^{5n}$, and $\widehat{C}, \widehat{D} \leftarrow \{0,1\}^{5n} \setminus \text{Image}(\mathcal{O}_n)$.

As proved in Claim 4.17, the first summand is bounded by $O(2^{-n/3})$. We argue that a similar bound holds for the second summand as well. The proof follows the same outline as that of Claim 4.17 with the following exception: in Claim 4.17, we argued that $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{D}}, \varphi)$ agrees with $\Gamma(f, \mathcal{O}, \varphi)$ on the set of queries $\mathbf{Q}$ made by the adversary to the last oracle. However, this argument crucially relied on on $\widehat{D}$ being chosen at random independently of the adversary's view. Now, this is no longer the case, the oracle in $\mathbf{H_7}$ is $\Gamma(f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \varphi)$, where $\widehat{C}$ is the challenge, which is known to

the adversary. Instead, we will be able to show that the last oracle agrees with $\Gamma(f, \mathcal{O}, \varphi_{\widehat{C}\mapsto C_0})$ on the adversary's queries.

We will, in fact, show that the above second summand is bounded for any fixed $b, \mathcal{O}, \widehat{C}, \mathsf{state}$. Indeed, for the rest of the claim, fix all of the above. In what follows, let $\mathbf{Q}$ be the set of queries made by $\mathcal{A}_2^{\Gamma(f,\mathcal{O},\varphi_{\widehat{C}\mapsto C_0})}(\mathsf{state}, \widehat{C})$ to its oracle

$$\Gamma(f, \mathcal{O}, \varphi_{\widehat{C}\mapsto C_0}) = (\Psi_{\varphi_{\widehat{C}\mapsto C_0}}, \mathsf{Decide}_{\mathfrak{S}}^{\Psi_{\varphi_{\widehat{C}\mapsto C_0}}}) \quad \text{where} \quad \Psi_{\varphi_{\widehat{C}\mapsto C_0}} = (f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}}) \ .$$

For any query $Q = (V_0, V_1, z) \in \mathbf{Q}$ made to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_{\varphi_{\widehat{C}\mapsto C_0}}}$, let $\mathbf{C}_Q$ denote the set of critical queries corresponding to Q, given by Claim 4.6. We stress that $\Psi_{\varphi_{\widehat{C}\mapsto C_0}} \in \mathfrak{S}$; namely, it is a valid oracle in the family $\mathfrak{S}$, so the claim can be applied. (This is where we rely on the fact that $\mathfrak{S}$ includes *non-verifiable* indistinguishability obfuscationconstructions, where Eval may produce arbitrary answers on invalid obfuscations such as $\widehat{C} \notin \mathrm{Image}(\mathcal{O})$.)

We denote by $\mathbf{C_Q}$ the union of all such critical sets. We now define the event $\mathbf{Hit}$, aimed at capturing the cases when the adversary's views in $\mathbf{H_7}$ and $\mathbf{H_8}$ may differ. Concretely, let $\mathbf{Hit} = \mathbf{Hit}(r)$ be the event that one of the following occurs:

1. $(C_b, r) \in \mathbf{Q} \cup \mathbf{C_Q}$: the query (C_b, r) is made to $\mathcal{O}$.
2. $(\mathcal{O}(C_b, r), x) \in \mathbf{Q} \cup \mathbf{C_Q}$ for some $x \in \{0,1\}^n$: the query $(\mathcal{O}(C_b, r), x)$ is made to $\mathsf{Eval}_{\varphi_{\widehat{C}\mapsto C_0}}$.

Claim 4.25. $\Pr_r\left[\mathcal{A}_2^{\Gamma(f,\mathcal{O},\varphi_{\widehat{C}\mapsto C_0})}(\mathsf{state}, \widehat{C}) \neq \mathcal{A}_2^{\Gamma(f,\mathcal{O}_{(C_b,r)\mapsto\widehat{C}},\varphi)}(\mathsf{state}, \widehat{C})\right] \leq \Pr_r[\mathbf{Hit}]$.

Proof. We observe that if $\mathbf{Hit}$ does not occur, then the view of $\mathcal{A}_2$ is the same for both oracles. Indeed, in $\mathbf{H_8}$, the oracle $\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}$ is changed to $\mathcal{O}$ and φ is changed to $\varphi_{\widehat{C}\mapsto C_0}$. In particular:

- for any query $Q \neq (C_b, r)$ to $\mathcal{O}$

$$\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}(Q) = \mathcal{O}(Q) \ ,$$

- for any query $Q = (\widehat{D}, x) \notin \left\{(\widehat{C}, x), (\mathcal{O}(C_b, r), x)\right\}_{x\in\{0,1\}^n}$ to $\mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}}$,

$$\mathsf{Eval}^{f,\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}}_{\varphi}(Q) = \mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}}(Q) \ ,$$

  since $\left(\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}\right)^{-1}(Q) = \mathcal{O}^{-1}(Q)$ and $\varphi(\widehat{D}) = \varphi_{\widehat{C}\mapsto C_0}(\widehat{D})$,

- for any $x \in \{0,1\}^n$ and query $Q = (\widehat{C}, x)$ to $\mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}}$,

$$\mathsf{Eval}^{f,\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}}_{\varphi}(\widehat{C}, x) = C_b^f(x) = C_0^f(x) = \left(\varphi_{\widehat{C}\mapsto C_0}(\widehat{C})\right)^f(x) = \mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}}(\widehat{C}, x) \ ,$$

  since the circuits are functionally equivalent: $C_0^f \equiv C_1^f$.

It follows that for all queries $Q \in \mathbf{Q} \cup \mathbf{C_Q}$, $\Psi_{\varphi_{\widehat{C}\mapsto C_0}}(Q) = \Psi'_\varphi(Q)$, where

$$\Psi_{\varphi_{\widehat{C}\mapsto C_0}} = f, \mathcal{O}, \mathsf{Eval}^{f,\mathcal{O}}_{\varphi_{\widehat{C}\mapsto C_0}} \quad \text{and} \quad \Psi'_\varphi = f, \mathcal{O}_{(C_b,r)\mapsto\widehat{C}}, \mathsf{Eval}^{f,\mathcal{O}_{(C_b,r)\mapsto\widehat{C}}}_{\varphi} \ .$$

This implies that all queries made by $\mathcal{A}_2$ directly to its oracle $\Psi_{\varphi_{\widehat{C}\mapsto C_0}}$ in $\mathbf{H_8}$, are answered in the same way in $\mathbf{H_7}$. It is left to show that this is also the case for queries made to $\mathsf{Decide}_{\mathfrak{S}}^{\Psi_{\varphi_{\widehat{C}\mapsto C_0}}}$. For this purpose, note that $\Psi'_\varphi \in \mathfrak{S}$, namely is a valid oracle (indeed, $\widehat{C}$ is chosen outside the image of $\mathcal{O}$, so injectivity of the obfuscation oracle is guaranteed as required). Furthermore, the oracles $\Psi_{\varphi_{\widehat{C}\mapsto C_0}}$ and Ψ'_φ agree on all critical sets $\mathbf{C}_Q \subseteq \mathbf{C_Q}$. It follows, by Claim 4.6, that for any $Q \in \mathbf{Q}$

$$\mathsf{Decide}_{\mathfrak{S}}^{\Psi_{\varphi_{\widehat{C}\mapsto C_0}}}(Q) = \mathsf{Decide}_{\mathfrak{S}}^{\Psi'_\varphi}(Q) \ ,$$

which completes the proof of the claim. □

It is left to bound the probability that **Hit** occurs. Since r is chosen at random from $\{0,1\}^n$, for any fixed query $Q \in \mathbf{Q} \cup \mathbf{C_Q}$, the probability that any of the two cases defining **Hit** occurs is at most 2^{-n}. Thus, by a union bound we have

$$\Pr[\mathbf{Hit}] \leq |\mathbf{Q} \cup \mathbf{C_Q}| \cdot 2^{-n} \leq \left(|\mathbf{Q}| + \sum_{Q \in \mathbf{Q}} |\mathbf{C}_Q| \right) \cdot 2^{-n} \leq O(q^2 \cdot 2^{-n}) \leq O(2^{-n/3}) \ .$$

This completes the proof of Claim 4.24. □

To conclude the proof of Theorem 3.12, we observe that

Claim 4.26. *The adversary has no advantage in* $\mathbf{H_8}$*. That is,*

$$\Pr\left[\mathcal{A} \textit{ wins in } \mathbf{H_8}\right] = \frac{1}{2}$$

Proof. The view of $\mathcal{A}$ in this hybrid is completely independent of the random choice of b. □

This completes the proof of Theorem 4.16. □

4.6 Extension to Relativizing Separations

In Section 4, we showed that, for any pair of verifiers V_0, V_1 that define a language $L^\Psi \in \mathsf{NP}^\Psi \cap \mathsf{coNP}^\Psi$ relative to any oracle Ψ in the family $\mathfrak{S}$ (of correct IO and IOWFs), there exists an oracle $\Psi \in \mathfrak{S}$ and oracle $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$ such that

$$L^\Psi \in \mathsf{P}^{\Psi, \mathsf{Decide}_{\mathfrak{S}}^{\Psi}} \ ,$$

even though Ψ securely implements IO and IOWFs in the presence of $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$.

We now strengthen the statement to show that the above holds also for constructions that may call both Ψ and $\mathsf{Decide}_{\mathfrak{S}}^{\Psi}$, rater than just the first. In the parlance of black-box reductions [Fis12, HR04, RTV04], the original statement is sufficient for ruling out *fully black-box* reductions while the current statement would rule out *relativizing reductions*.[19]

[19] At high level, a relativizing reduction from a primitive $\mathcal{P}$ to a primitive $\mathcal{Q}$ says that $\mathcal{P}$ exists in any oracle world where $\mathcal{Q}$ does. In particular, when think about a world relative to the oracle $\Gamma = (\Psi, \mathsf{Decide}_{\mathfrak{S}}^{\Psi})$ the construction of $\mathcal{Q}$ (in our case, a problem in $\mathsf{NP} \cap \mathsf{coNP}$) is allowed to rely on the full functionality of Γ.

In what follows, $\mathfrak{S}$ is the same family of oracles considered in Section 4 and PSPACE denotes an oracle for a PSPACE-complete problem.

Theorem 4.27. *Let (V_0, V_1) be a pair of non-deterministic polynomial-time verifiers such that for every oracle $\Psi \in \mathfrak{S}$, and letting $\Gamma_\Psi := (\Psi, \mathsf{Decide}^\Psi_{\mathfrak{S}}, \mathsf{PSPACE})$, the verifiers $(V_0^{\Gamma_\Psi}, V_1^{\Gamma_\Psi})$ define a language $L^{\Gamma_\Psi} \in \mathsf{NP}^{\Gamma_\Psi} \cap \mathsf{coNP}^{\Gamma_\Psi}$.*

1. *For $\Psi \in \mathfrak{S}$, and Γ_Ψ, the corresponding language is trivial: $L^\Gamma \in \mathsf{P}^\Gamma$.*
2. *The oracle Ψ_φ from Definition 4.13 securely realizes IO and IOWFs in the presence of Γ_{Ψ_φ}.*

Proof sketch. The second item in the theorem follows directly from the proof of Theorem 4.16 in Section 4. Indeed, the only difference between the two is that security here has to be shown also in the presence of a PSPACE oracle, and not just $\Psi_\varphi, \mathsf{Decide}^{\Psi_\varphi}_{\mathfrak{S}}$. However, in Theorem 4.16, we prove security against a computationally unbounded adversary that is only bounded in the number of queries it is allowed to make to its oracle, in particular, it may perform PSPACE computations.

We thus focus on the first part. Here we will show that a pair of verifiers as given in the condition of the theorem can be reduced to a new pair of polynomial-time verifiers (S_0, S_1) which can simulate V_0, V_1, while not making any queries to Decide^Ψ. More formally, for every oracle $\Psi \in \mathfrak{S}$, letting $\Phi_\Psi := (\Psi, \mathsf{PSPACE})$, $(S_0^{\Phi_\Psi}, S_1^{\Phi_\Psi})$ define a language $L^{\Phi_\Psi} \in \mathsf{NP}^\Phi \cap \mathsf{coNP}^\Phi$, such that deciding L^{Γ_Ψ} can be reduced to deciding L^{Φ_Ψ}. The result will then follow from the fact that $L^{\Phi_\Psi} \in \mathsf{P}^{\Gamma_\Psi}$.[20]

To generate the simulator verifiers S_0, S_1, we exploit the fact that that the output of the oracle $\mathsf{Decide}_{\mathfrak{S}}$ is certifiable. Concretely, let $q = q(n)$ be a bound on the number of queries that V_0, V_1 make to the oracle $\mathsf{Decide}_{\mathfrak{S}}$. The simulator verifier $S_b^{\Phi_\Psi}$, given input z, non-deterministically guesses a witness $W = (w, (b_1, w_1), \ldots, (b_q, w_q))$. It then emulates $V_b^{\Gamma_\Psi}(z, w)$ as follows:

- Queries made to either Ψ, PSPACE are forwarded by S_b to the corresponding oracles.
- For the i-th query (C_0, C_1, x) made to $\mathsf{Decide}^\Psi_{\mathfrak{S}}$:
  - S_b first uses its PSPACE oracle to test whether C_0, C_1 define an $\mathsf{NP} \cap \mathsf{coNP}$ collection relative to $\mathfrak{S}$.[21]
  - If not, the answer is simulated as $\perp$.
  - Otherwise, S_b tests whether $C_{b_i}^\Psi(x, w_i)$ accepts. If this is the case, the answer is simulated as b_i, and otherwise S_b aborts the entire emulation process and rejects.

First, note that the running time of S_b is polynomial in that of V_b. By definition, S_b non-deterministically accepts x iff V_b accepts x, as required. □

5 Collision-Resistance from IO and SZK-Hardness

Asharov and Segev [AS15] showed that collision-resistant hashing cannot be constructed from (even subexponentially hard) indistinguishability obfuscation (IO) and one-way permutations (OWPs)

[20] Formally, we need to slightly extend $\mathsf{Decide}_{\mathfrak{S}}$ to account for queries (C_0, C_1, z), where C_0, C_1 are circuit verifiers that may make also PSPACE queries. Theorem 4.16 also accounts for computationally unbounded (C_0, C_1, z) (as long as the number of queries they make to Ψ is bounded), thus the second part of Theorem 4.27 still holds.

[21] We note that, as in [IR89, AS16], for this purpose, it is enough to consider "partial oracles" Ψ that are only defined for the polynomial set of queries that C_0, C_1. Thus, this can be done in PSPACE.

relying on common IO techniques. Slightly more accurately, they rule out fully black-box constructions where (as in previous sections) IO is defined with respect to circuits with OWP oracle gates. In this section, we show that, assuming IO and a strong form of SZK-hardness, there is indeed a construction of collision-resistant hashing (CRH).

The High-Level Idea Behind the Construction. The starting point for our construction is the work of Ishai, Kushilevitz, and Ostrovsky [IKO05] that shows how to construct collision-resistant hash functions from commitments that are additively homomorphic (for simplicity, say over $\mathbb{F}_2$). The idea is simple: we can hash ℓ bits to m bits, where m is the size of a single bit commitment and ℓ can be arbitrarily longer, as follows. The hash key is a commitment $\gamma := (\mathsf{com}(\beta_1), \ldots, \mathsf{com}(\beta_\ell))$ to a random vector $\beta \in \mathbb{F}_2^\ell$, and hashing $x \in \mathbb{F}_2^\ell$, is done by homomorphically computing a commitment to the inner product $\mathsf{CRH}_\gamma(x) = \mathsf{com}(\langle \beta, x \rangle)$.

This idea can, in fact, be abstracted to work with any commitment scheme wherein given a commitment $\mathsf{com}(\beta)$ for a random key for a 2-universal hash allows to homomorphically compute a commitment $\mathsf{com}(\mathsf{2UH}_\beta(x))$ to the hash at any point x, so that the resulting commitment is compact in the sense that it depends only on the size of $\mathsf{2UH}_\beta(x)$ and not on the size of x. Intuitively, the reason this works is that any collision in CRH_γ implies a collision in the underlying 2-universal hash $\mathsf{2UH}_\beta$, which leaks information about the hash key β (concretely, any fixed x, x' form a collision in a random hash function with small probability) thereby violating the hiding of the commitment.

At a high-level, we aim to mimic the above construction based on obfuscation. As a key for the collision-resistant hash we can obfuscate a program Π_β associated with a secret hash key β that given x outputs a commitment $\mathsf{com}(\mathsf{2UH}_\beta(x))$, where the commitment is derandomized using a PRF. The obfuscation $i\mathcal{O}(\Pi_\beta)$ can be thought of as the commitment to β, and evaluating this program at x, corresponds to homomorphic evaluation. Despite the clear intuition behind this construction, it is not clear how to prove its security based on IO. In fact, by [AS15], it cannot be proven based on a black-box reduction as long as plain statistically-binding commitments are used, as these can be constructed from OWPs in a fully black-box manner.

We show, however, that relying on a relaxed notion of perfectly-hiding commitments, as well as subexponential hardness of IO and puncturable PRFs, the construction can be proven secure. The perfect hiding of the commitment is leveraged in a probabilistic IO argument [CLTV15] that involves a number of hybrids larger than the overall number of commitments. We then observe that these relaxed commitments follow from appropriate average-case hardness of SZK.[22]

5.1 Definitions and Tools

We define our notion of relaxed perfectly-hiding commitments and the SZK-hardness they follow from. We also define 2-universal hashing and puncturable pseudorandom functions, which are used in our construction.

Relaxed Perfectly-Hiding Commitments. We consider two message bit commitment schemes $(\mathcal{R}, \mathcal{S})$, where the receiver $\mathcal{R}$ samples a first message σ, and $\mathcal{S}(\sigma, b)$ samples a commitment ξ to a bit b. We require that the commitment is computationally binding and that there exists a

[22] Similar SZK-hardness is known to imply statistically-hiding commitments against malicious receivers, but with a larger (constant) number of rounds [OV08].

distribution $\widetilde{\mathcal{R}}$ (not necessarily efficiently samplable) that is computationally indistinguishable from that generated by $\mathcal{R}$ and under which the commitment is perfectly hiding.

Definition 5.1 (Relaxed Statistically-Hiding Commitments)**.** $(\mathcal{R}, \mathcal{S})$ is a relaxed statistically-hiding commitment scheme if it satisfies:

1. **Computational Binding:** for any non-uniform PPT sender $\mathcal{S}^*$,

$$\Pr_{\sigma \leftarrow \mathcal{R}(1^n)} \left[\begin{array}{l} r_0, r_1 \leftarrow \mathcal{S}^*(\sigma) \\ \mathcal{S}(\sigma, 0; r_0) = \mathcal{S}(\sigma, 1; r_1) \end{array} \right] \leq \mathsf{negl}(n) \ .$$

2. **Relaxed Perfect-Hiding:** there exists a (possibly inefficient) sampler $\widetilde{\mathcal{R}}$ such that
   - $\mathcal{R}(1^n)$ and $\widetilde{\mathcal{R}}(1^n)$ are computationally indistinguishable,
   - for any $\widetilde{\sigma}$ in the support of $\widetilde{\mathcal{R}}(1^n)$, $\mathcal{S}(\widetilde{\sigma}, 0)$ and $\mathcal{S}(\widetilde{\sigma}, 1)$ are identically distributed.

Relaxed perfectly-hiding commitments are implied by the standard definition of 2-message perfectly hiding commitments. The standard definition is stronger in the sense that perfect-hiding holds for *any* (even maliciously chosen) σ. We next show that this definition is implied by appropriate average-hardness of SZK.

Average-Case Hardness of $\mathbf{SD}^{0,1}$. Roughly speaking, we require average-case hardness of an extreme case of the statistical-distance problem, referred to as $\mathbf{SD}^{0,1}$ in [Vad99]. Here YES-instances consist of pairs of samplers with disjoint support, whereas NO-instances consist of samplers with identical distribution.

Definition 5.2 (Average-Case Hardness of $\mathbf{SD}^{0,1}$)**.** The promise problem $\mathbf{SD}^{0,1} = (\mathbf{SD}^{0,1}_Y, \mathbf{SD}^{0,1}_N)$ is given by

$$\mathbf{SD}^{0,1}_Y = \{(C_0, C_1) \mid \Delta(\boldsymbol{C}_0, \boldsymbol{C}_1) = 1\} \ ,$$
$$\mathbf{SD}^{0,1}_N = \{(C_0, C_1) \mid \Delta(\boldsymbol{C}_0, \boldsymbol{C}_1) = 0\} \ .$$

We say that the problem is hard on average if there exists a PPT sampler S with support $\mathbf{SD}^{0,1}$ such that for any non-uniform PPT decider D,

$$\Pr_{(C_0, C_1) \leftarrow \mathsf{S}(1^n)} \left[\begin{array}{l} B \leftarrow D(C_0, C_1) \\ (C_0, C_1) \in \mathbf{SD}^{0,1}_B \end{array} \right] \leq \frac{1}{2} + \mathsf{negl}(n) \ .$$

The above definition should be contrasted with the standard definition of the statistical differenceproblem $\mathbf{SD} = \mathbf{SD}^{\frac{1}{3}, \frac{2}{3}}$ (Definition 3.1) where the notions of statistical farness and closeness are not absolute (but given by the constants $\frac{1}{3}, \frac{2}{3}$). We note that the ploarization lemma in [SV03] gives an efficient reduction from deciding $\mathbf{SD}^{\frac{1}{3}, \frac{2}{3}}$ to deciding $\mathbf{SD}^{2^{-n}, 1-2^{-n}}$, but such a reduction is not known if we replace $\mathbf{SD}^{2^{-n}, 1-2^{-n}}$ by $\mathbf{SD}^{0,1}$. Average-case hardness of $\mathbf{SD}^{0,1}$ is known under number-theoretic assumptions such as Decision-Diffie-Hellman and Quadratic Residuosity [GMR85], which are already known to imply collision-resistance directly. However, it may also follow from problems that are not known to imply collision-resistance, and may be of a non-algebraic nature. For instance, average-case hardness of $\mathbf{SD}^{0,1}$ would follow from the average-case hardness of Graph Non-Isomorphism [GMW91].

Claim 5.3. *Average-case hardness of* $\mathbf{SD}^{0,1}$ *implies relaxed perfectly-hiding commitments.*

Proof sketch. We define the receiver and sender $(\mathcal{R}, \mathcal{S})$. The receiver $\mathcal{R}$ is simply the instance sampler S, which outputs first messages of the form $\sigma = (C_0, C_1) \in \mathbf{SD}^{0,1}$. The sender $\mathcal{S}(C_0, C_1, b)$ outputs a random sample from $\boldsymbol{C}_b$.

We next prove binding and hiding. Denote by $\mathsf{S}_B(1^n)$ the distribution given by sampling (C_0, C_1) from $\mathsf{S}(1^n)$ conditioned on $(C_0, C_1) \in \mathbf{SD}^{0,1}_B$. We note that $\mathsf{S}_Y(1^n)$, $\mathsf{S}_N(1^n)$, and $\mathsf{S}(1^n)$ are computationally indistinguishable by the average-case hardness of S. Computational binding now holds by the fact that $\mathsf{S}_Y(1^n)$ and $\mathsf{S}(1^n)$ are computationally indistinguishable and $\mathsf{S}_Y(1^n)$ samples C_0, C_1 with disjoint supports. The fake receiver sampler $\widetilde{\mathcal{R}}$ is S_N, which is as required indistinguishable from $\mathcal{R} = \mathsf{S}$, and in which commitments are samples from $\boldsymbol{C}_0$ or $\boldsymbol{C}_1$, which are identically distributed. □

Remark 5.4. In our definition of *relaxed* perfectly-hiding commitments. The commitment scheme itself (corresponding to the honest $\mathcal{R}$) is neither perfectly hiding nor perfectly hiding according to the common definition. We note that assuming stronger notions of hardness in $\mathbf{SD}^{0,1}$ we they can be made such. Specifically, if S_Y is efficient, then using it in the actual scheme would make it perfectly binding. Alternatively, if S_N is efficient, then using it in the actual scheme would make it perfectly hiding.

Puncturable Pseudo-Random Functions. We consider a simple case of the puncturable pseudo-random functions (PRFs) where any PRF may be punctured at a single point. The definition is formulated as in [SW14], and is satisfied by the GGM [GGM86] PRF [BW13, KPTZ13, BGI14] and can be constructed from any one-way function. One-way functions are, in turn, implied by the average case hardness of $\mathbf{SD}^{0,1}$ [OW93b]. (Here we will need sub-exponential security of the PRF and thus sub-exponentially-hard one-way functions.)

Definition 5.5 (Puncturable PRFs)**.** Let k, ℓ, m be polynomially bounded functions. An efficiently computable family of functions

$$\mathcal{PRF} = \left\{ \mathsf{PRF}_\alpha : \{0,1\}^{\ell(n)} \to \{0,1\}^{m(n)} \;\middle|\; \alpha \in \{0,1\}^{k(n)}, n \in \mathbb{N} \right\} \ ,$$

is a puncturable PRF if there exists a poly-time puncturing algorithm Punc that takes as input a key α, and a point x^*, and outputs a punctured key $\alpha\{x^*\}$, so that the following conditions are satisfied:

1. **Functionality is preserved under puncturing:** For every $x^* \in \{0,1\}^{\ell(n)}$,

$$\Pr_{\alpha \leftarrow \{0,1\}^{k(n)}} \left[\forall x \neq x^* : \mathsf{PRF}_\alpha(x) = \mathsf{PRF}_{\alpha\{x^*\}}(x) \;\middle|\; \alpha\{x^*\} = \mathsf{Punc}(\alpha, x^*) \right] = 1 \ .$$

2. **Indistinguishability at punctured points:** for any nonuniform PPT distinguisher D there exists a negligible function $\mathsf{negl}(\cdot)$, such that for all $n \in \mathbb{N}$, and any $x^* \in \{0,1\}^{\ell(n)}$,

$$|\Pr[D(x^*, \alpha\{x^*\}, \mathsf{PRF}_\alpha(x^*)) = 1] - \Pr[D(x^*, \alpha\{x^*\}, u) = 1]| \leq \mathsf{negl}(n) \ ,$$

   where $\alpha \leftarrow \{0,1\}^{k(n)}, \alpha\{x^*\} = \mathsf{Punc}(\alpha, x^*)$, and $u \leftarrow \{0,1\}^{m(n)}$.

   We further say that $\mathcal{PRF}$ satisfies δ-indistinguishability if the above negligible indistinguishability gap is smaller than δ.

2-Universal Hashing. We rely on 2-universal families of hash functions, which are known to exist unconditionally [WC81].

Definition 5.6 (2-Universal Hashing)**.** Let k, ℓ, m be polynomially bounded functions. An efficiently computable family of functions

$$2\mathcal{UH} = \left\{ \mathsf{2UH}_\beta : \{0,1\}^{\ell(n)} \to \{0,1\}^{m(n)} \;\middle|\; \beta \in \{0,1\}^{k(n)}, n \in \mathbb{N} \right\} \ ,$$

is 2-universal if for any two distinct $x, x' \in \{0,1\}^{\ell(n)}$

$$\Pr_{\beta \leftarrow \{0,1\}^{k(n)}} \left[\mathsf{2UH}_\beta(x) = \mathsf{2UH}_\beta(x') \right] \leq 2^{-m(n)} \ .$$

5.2 The Construction

We are now ready to state and prove the main result of this section.

Theorem 5.7. *Assuming average-case hardness of* $\mathbf{SD}^{0,1}$ *(or more generally, relaxed perfectly-binding commitments) and the existence of indistinguishability obfuscators and one-way functions that are subexponentially secure, there exists a collision-resistant hash function family.*

Let $\tau(\cdot)$ be an expansion parameter. To get a collision-resistant family with expansion τ, we rely on the following ingredients:

- A relaxed perfectly-hiding commitment scheme $(\mathcal{R}, \mathcal{S})$. We denote by $\ell(n)$ the size of bit commitments.
- An indistinguishability obfuscator $i\mathcal{O}$ with $2^{-\tau(n)\cdot\ell(n)} \cdot \mathsf{negl}(n)$-indistinguishability.
- A puncturable pseudo-random function family $\mathcal{PRF}$ satisfying $2^{-\tau(n)\cdot\ell(n)} \cdot \mathsf{negl}(n)$-indistinguishability.
- A 2-universal hashing family $2\mathcal{UH}$ mapping $\{0,1\}^{\tau(n)\cdot\ell(n)}$ to $\{0,1\}$.

We construct a collision-resistant hashing family

$$\mathcal{CRH} = \left\{ \mathsf{CRH}_\gamma : \{0,1\}^{\tau(n)\cdot\ell(n)} \to \{0,1\}^{\ell(n)} \;\middle|\; \gamma \in \{0,1\}^{k(n)}, n \in \mathbb{N} \right\} \ ,$$

with an associated key generator $\mathsf{Gen}_{\mathcal{CRH}}$.

Construction 5.8 (A Collision-Resistant Hashing Family)**.** *$\mathcal{CRH}$ is given by:*

1. $\mathsf{Gen}_{\mathcal{CRH}}(1^n)$*:*
   - *generate a receiver message* $\sigma \leftarrow \mathcal{R}(1^n)$*,*
   - *sample a key* $\beta \leftarrow \{0,1\}^{k(n)}$ *for a 2-universal hash,*
   - *sample a key* $\alpha \leftarrow \{0,1\}^{k(n)}$ *for a puncturable PRF,*
   - *construct a circuit* $\Pi = \Pi[\sigma, \beta, \alpha]$ *that*
     - *given input* $x \in \{0,1\}^{\tau(n)\cdot\ell(n)}$
     - *computes the hash bit* $\rho_x := \mathsf{2UH}_\beta(x)$*,*

– *outputs a commitment* $\xi_x := \mathcal{S}(\sigma, \rho_x; \mathsf{PRF}_\alpha(x))$.

- *obfuscate* $\gamma \leftarrow i\mathcal{O}(\Pi, 1^n)$ *and output* γ *as the key.*

2. $\mathsf{CRH}_\gamma(x)$*:*

- *parse* γ *as a circuit and output* $\xi_x = \gamma(x)$.

Proposition 5.1. *$\mathcal{CRH}$ is collision-resistant.*

Proof. To prove the proposition, we shall prove the following two claims.

Claim 5.9. *Fix any two keys $\beta_0, \beta_1 \in \{0,1\}^{k(n)}$ for the 2-universal family, and let $\gamma|_{\beta_0}$ (respectively, $\gamma|_{\beta_1}$) be the distributions on CRH keys conditioned on hashing key β_0 (respectively, β_1). Then the two distributions are computationally indistinguishable.*

Claim 5.10. *Assume there exists an efficient A that finds collisions in $\mathcal{CRH}$ with probability δ over the choice of key γ. Then there exists an efficient predictor P that given random β_0, β_1, and $\gamma|_{\beta_b}$ for a random $b \leftarrow \{0,1\}$, predicts b with advantage $\frac{\delta - \mathsf{negl}(n)}{4}$.*

The two claims together imply that a collision finder A for $\mathcal{CRH}$ cannot exist.

Proof of Claim 5.9. First, we note that by the computational indistinguishability of honest receiver messages generated by $\mathcal{R}(1^n)$ and receiver messages generated by $\widetilde{\mathcal{R}}(1^n)$, it is enough to prove the claim for an alternative experiment where γ is sampled as usual except that $\widetilde{\sigma}$ is sampled from $\widetilde{\mathcal{R}}(1^n)$ rather than σ from $\mathcal{R}(1^n)$. In this new experiment, commitments to 0 and 1 are identically distributed. We now prove indistinguishability of $\gamma|_{\beta_0}$ and $\gamma|_{\beta_1}$ based on a standard probabilistic IO argument [CLTV15]. We sketch the argument for the sake of completeness.

For each input $x \in \{0,1\}^{\tau(n)\cdot\ell(n)}$, we consider a hybrid where a circuit $\Pi^x = \Pi^x[\widetilde{\sigma}, \beta_0, \beta_1, \alpha]$ is obfuscated, where

$$\begin{aligned} &\text{for } x' < x: \qquad \Pi^x(x') = \Pi[\widetilde{\sigma}, \beta_0, \alpha](x') = \mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_0}(x'); \mathsf{PRF}_\alpha(x')) \ , \\ &\text{for } x' \geq x: \qquad \Pi^x(x') = \Pi[\widetilde{\sigma}, \beta_1, \alpha](x') = \mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_1}(x'); \mathsf{PRF}_\alpha(x')) \ . \end{aligned}$$

Each two consecutive hybrids differ only at a single point x where one answers according to $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_0}(x); \mathsf{PRF}_\alpha(x))$ and the other according to $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_1}(x); \mathsf{PRF}_\alpha(x))$. We can then puncture α at x and hardwire these outputs, relying on IO, then replace them with truly random samples from $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_0}(x))$ and $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_1}(x))$, relying on pseudo-randomness at punctured points, and finally rely on the fact that the two circuits sample from two identical distributions. Since both $i\mathcal{O}$ and $\mathcal{PRF}$ satisfy $2^{-\tau(n)\ell(n)} \cdot \mathsf{negl}(n)$-indistinguishability, and the two samples from $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_0}(x))$ and $\mathcal{S}(\widetilde{\sigma}, \mathsf{2UH}_{\beta_1}(x))$ are perfectly indistinguishable, we get $O(2^{-\tau(n)\ell(n)} \cdot \mathsf{negl}(n))$-indistinguishability between any two consecutive hybrids. This allows us to deduce a negligible difference between the first and the last hybrid corresponding to $\gamma|_{\beta_0}$ and $\gamma|_{\beta_1}$.[23] □

Proof of Claim 5.10. We start by describing the predictor P.

Given β_0, β_1, and $\gamma = \gamma|_{\beta_b}$, P outputs a prediction b^* (for b) as follows:

[23] Above, since $\widetilde{\sigma}$ may not be efficiently samplable, we formally need to rely on the non-uniform security of $i\mathcal{O}$ and $\mathcal{PRF}$. Alternatively, we can require a less relaxed commitment notion where $\widetilde{\mathcal{R}}$ is also efficient (these still follow from standard perfectly-hiding commitments).

- Apply the collision-finder $A(\gamma)$.
- If A finds $x \neq x'$ such that
  - $\mathsf{CRH}_\gamma(x) = \mathsf{CRH}_\gamma(x')$,
  - $\mathsf{2UH}_{\beta_{b^*}}(x) = \mathsf{2UH}_{\beta_{b^*}}(x')$,
  - $\mathsf{2UH}_{\beta_{1-b^*}}(x) \neq \mathsf{2UH}_{\beta_{1-b^*}}(x')$,

  output b^*. We will below refer to this event as **Col**.
- Otherwise, output a random $b^* \leftarrow \{0,1\}$.

We now show that $b^* = b$ with probability $\frac{1+\delta/2-\mathsf{negl}(n)}{2}$. Indeed, for any $b \in \{0,1\}$, and sampling $\beta_0, \beta_1, \gamma|_{\beta_b}$ as above:

$$\begin{aligned}
&\Pr\left[P(\beta_0,\beta_1,\gamma|_{\beta_b}) = b\right] = \\
&\Pr[\mathbf{Col}] \cdot \Pr\left[P(\beta_0,\beta_1,\gamma|_{\beta_b}) = b \mid \mathbf{Col}\right] + \Pr\left[\overline{\mathbf{Col}}\right] \cdot \Pr\left[P(\beta_0,\beta_1,\gamma|_{\beta_b}) = b \;\middle|\; \overline{\mathbf{Col}}\right] = \\
&\Pr[\mathbf{Col}] \cdot 1 + \Pr\left[\overline{\mathbf{Col}}\right] \cdot \frac{1}{2} = \\
&\frac{1+\Pr[\mathbf{Col}]}{2} \ .
\end{aligned}$$

It is left to note that

$$\Pr[\mathbf{Col}] \geq \delta/2 - \mathsf{negl}(n) \ .$$

Indeed, by our assumption A finds a collision $x \neq x'$ in $\gamma|_{\beta_b}$ with probability δ. Also, by 2-universality, choosing β_{1-b} at random, $\mathsf{2UH}_{\beta_{1-b}}(x) = \mathsf{2UH}_{\beta_{1-b}}(x')$ with probability at most $1/2$. (Note that β_{1-b} is independent of the A's view and thus from x, x'.) To conclude the argument we claim that the probability that x, x' is a collision in $\mathsf{CRH}_{\gamma|_{\beta_b}}$ but not in $\mathsf{2UH}_{\beta_b}$ is negligible by the computational binding of commitment scheme. To see this, note that whenever the latter event occurs, we know that:

$$\mathcal{S}(\sigma, \mathsf{2UH}_{\beta_b}(x), \mathsf{PRF}_\alpha(x)) = \mathcal{S}(\sigma, \mathsf{2UH}_{\beta_b}(x'), \mathsf{PRF}_\alpha(x')) \qquad \text{but} \qquad \mathsf{2UH}_{\beta_b}(x) = \mathsf{2UH}_{\beta_b}(x')$$

□

This completes the proofs of the two claims and the proposition. □

Remark 5.11 (Statistical Hiding instead of Perfect Hiding.). In the above, we have defined and relied on perfectly-hiding commitments. It is natural to ask whether we can relax this requirement to statistical hiding. The bottleneck here is the probabilistic IO argument used in our proof, which requires $2^{-\tau\ell}$-indistinguishability of the commitments where ℓ is the size of a bit commitment and τ is the expansion factor. Accordingly, we can make do with a strong statistical guarantee where the statistical distance is say $2^{-s} \cdot \mathsf{negl}(n)$ where s is the size of a single bit commitment.

In our setting, where the commitment is implemented using the hard statistical difference problem, the above requirement translates to a hard samplable distribution on $\mathbf{SD}^{2^{-s}\cdot\mathsf{negl}(n),1}$ where s is a bound on the output length of samplers in the support of this distribution. This holds for example for known statistically-hiding commitments based on collision-resistant hashing [DPP93, HM96]. However, it cannot be achieved generically by, say by amplifying $\mathbf{SD}^{\frac{1}{3},\frac{2}{3}}$ (via the polarization lemma [SV03] mentioned above), since such amplification increases the size of samples.

Remark 5.12 (Relation with Section 3). We note that the notion of statistical difference hardness required here is stronger than that ruled out in two ways. First, it requires hardness even of $\mathbf{SD}^{0,1}$, whereas in Section 3, we discuss $\mathbf{SD}^{\frac{1}{3},\frac{2}{3}}$. Second, it requires average-case hardness rather than the worst-case hardness considered in Section 3.

We note that the construction of collision-resistant hash functions in this section, in conjunction the result of [AS15] (ruling out fully black-box construction of CRH from OWP and IO for circuits with OWP gates), give an alternative proof to the statement that there is no fully black-box construction of hard on average problems in $\mathbf{SD}^{0,1}$ from OWPs. Indeed, this statement also follows from the more general statement in Section 3, saying that that is no fully black-box construction of worst-case hard problems in $\mathbf{SD}^{\frac{1}{3},\frac{2}{3}}$ from OWPs and IO for circuits with OWP gates. (Roughly speaking, the reason that this alternative proof covers constructions of hard statistical difference problems from OWPs but not from IO is that the result of [AS15] only covers IO for circuits with OWP gates, *but not with IO gates*. Indeed, in our construction the circuits representing the hard statistical difference problem are obfuscated.)

Acknowledgements

We thank Gil Segev, Iftach Haitner and Mohammad Mahmoody for elaborately answering our questions regarding existing separation results in cryptography. We also thank the anonymous FOCS'16 reviewers for their valuable comments.

References

[ABW10] Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 171–180, 2010.

[AGGM06] Adi Akavia, Oded Goldreich, Shafi Goldwasser, and Dana Moshkovitz. On basing one-way functions on np-hardness. In Kleinberg [Kle06], pages 701–710.

[AH91] William Aiello and Johan Hastad. Statistical zero-knowledge languages can be recognized in two rounds. *Journal of Computer and System Sciences*, 42(3):327–345, 1991.

[Ale03] Michael Alekhnovich. More on average case vs approximation complexity. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings* [DBL03], pages 298–307.

[AR04] Dorit Aharonov and Oded Regev. Lattice problems in NP cap conp. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 362–371. IEEE Computer Society, 2004.

[AR16] Benny Applebaum and Pavel Raykov. From private simultaneous messages to zero-information arthur-merlin protocols and back. In Eyal Kushilevitz and Tal Malkin, editors, *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part II*, volume 9563 of *Lecture Notes in Computer Science*, pages 65–82. Springer, 2016.

[AS15] Gilad Asharov and Gil Segev. Limits on the power of indistinguishability obfuscation and functional encryption. In *Symposium on the Foundations of Computer Science*, 2015.

[AS16] Gilad Asharov and Gil Segev. On constructing one-way permutations from indistinguishability obfuscation. In *Theory of Cryptography*, pages 512–541. Springer, 2016.

[Bar13] Boaz Barak. Structure vs. combinatorics in computational complexity. http://windowsontheory.org/2013/10/07/structure-vs-combinatorics-in-computational-complexity/, 2013.

[BB15] Andrej Bogdanov and Christina Brzuska. On basing size-verifiable one-way functions on np-hardness. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part I*, volume 9014 of *Lecture Notes in Computer Science*, pages 1–6. Springer, 2015.

[BBF13] Paul Baecher, Christina Brzuska, and Marc Fischlin. Notions of black-box reductions, revisited. In *Advances in Cryptology - ASIACRYPT 2013 - 19th International Conference on the Theory and Application of Cryptology and Information Security, Bengaluru, India, December 1-5, 2013, Proceedings, Part I*, pages 296–315, 2013.

[BGI+01] Boaz Barak, Oded Goldreich, Russell Impagliazzo, Steven Rudich, Amit Sahai, Salil P. Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs. In Joe Kilian, editor, *Advances in Cryptology - CRYPTO 2001, 21st Annual International Cryptology Conference, Santa Barbara, California, USA, August 19-23, 2001, Proceedings*, volume 2139 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2001.

[BGI14] Elette Boyle, Shafi Goldwasser, and Ioana Ivan. Functional signatures and pseudorandom functions. In Hugo Krawczyk, editor, *PKC*, volume 8383 of *Lecture Notes in Computer Science*, pages 501–519. Springer, 2014.

[BGL+15] Nir Bitansky, Sanjam Garg, Huijia Lin, Rafael Pass, and Sidharth Telang. Succinct randomized encodings and their applications. In *Symposium on Theory of Computing, STOC 2015*, 2015.

[BHZ87] Ravi B Boppana, Johan Hastad, and Stathis Zachos. Does co-np have short interactive proofs? *Information Processing Letters*, 25(2):127–132, 1987.

[BI87] Manuel Blum and Russell Impagliazzo. Generic oracles and oracle classes. In *Proceedings of the 28th Annual Symposium on Foundations of Computer Science*, SFCS '87, pages 118–126, Washington, DC, USA, 1987. IEEE Computer Society.

[BKSY11] Zvika Brakerski, Jonathan Katz, Gil Segev, and Arkady Yerukhimovich. Limits on the power of zero-knowledge proofs in cryptographic constructions. In Ishai [Ish11], pages 559–578.

[BL13] Andrej Bogdanov and Chin Ho Lee. Limits of provable security for homomorphic encryption. In Ran Canetti and Juan A. Garay, editors, *Advances in Cryptology*

- *CRYPTO 2013 - 33rd Annual Cryptology Conference, Santa Barbara, CA, USA, August 18-22, 2013. Proceedings, Part I*, volume 8042 of *Lecture Notes in Computer Science*, pages 111–128. Springer, 2013.

[BM09] Boaz Barak and Mohammad Mahmoody-Ghidary. Merkle puzzles are optimal - an $o(n^2)$-query attack on any key exchange from a random oracle. In Shai Halevi, editor, *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, volume 5677 of *Lecture Notes in Computer Science*, pages 374–390. Springer, 2009.

[BP15] Nir Bitansky and Omer Paneth. Zaps and non-interactive witness indistinguishability from indistinguishability obfuscation. In Yevgeniy Dodis and Jesper Buus Nielsen, editors, *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part II*, volume 9015 of *Lecture Notes in Computer Science*, pages 401–427. Springer, 2015.

[BPR15] Nir Bitansky, Omer Paneth, and Alon Rosen. On the cryptographic hardness of finding a nash equilibrium. In Venkatesan Guruswami, editor, *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 1480–1498. IEEE Computer Society, 2015.

[BPW16] Nir Bitansky, Omer Paneth, and Daniel Wichs. Perfect structure on the edge of chaos - trapdoor permutations from indistinguishability obfuscation. In *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, pages 474–502, 2016.

[Bra79] Gilles Brassard. Relativized cryptography. In *20th Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 29-31 October 1979*, pages 383–391. IEEE Computer Society, 1979.

[BT03] Andrej Bogdanov and Luca Trevisan. On worst-case to average-case reductions for NP problems. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings* [DBL03], pages 308–317.

[BV11] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In Rafail Ostrovsky, editor, *FOCS*, pages 97–106. IEEE, 2011. Invited to SIAM Journal on Computing.

[BW13] Dan Boneh and Brent Waters. Constrained pseudorandom functions and their applications. In Kazue Sako and Palash Sarkar, editors, *ASIACRYPT (2)*, volume 8270 of *Lecture Notes in Computer Science*, pages 280–300. Springer, 2013.

[CDT09] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *J. ACM*, 56(3), 2009.

[CHJV15] Ran Canetti, Justin Holmgren, Abhishek Jain, and Vinod Vaikuntanathan. Succinct garbling and indistinguishability obfuscation for RAM programs. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 429–437, 2015.

[CLTV15] Ran Canetti, Huijia Lin, Stefano Tessaro, and Vinod Vaikuntanathan. Obfuscation of probabilistic circuits and applications. In *TCC*, 2015.

[Cra12] Ronald Cramer, editor. *Theory of Cryptography - 9th Theory of Cryptography Conference, TCC 2012, Taormina, Sicily, Italy, March 19-21, 2012. Proceedings*, volume 7194 of *Lecture Notes in Computer Science*. Springer, 2012.

[DBL00] *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*. IEEE Computer Society, 2000.

[DBL03] *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*. IEEE Computer Society, 2003.

[DGP06] Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The complexity of computing a nash equilibrium. In Kleinberg [Kle06], pages 71–78.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976.

[DHT12] Yevgeniy Dodis, Iftach Haitner, and Aris Tentes. On the instantiability of hash-and-sign RSA signatures. In Cramer [Cra12], pages 112–132.

[DLMM11] Dana Dachman-Soled, Yehuda Lindell, Mohammad Mahmoody, and Tal Malkin. On the black-box complexity of optimally-fair coin tossing. In Ishai [Ish11], pages 450–467.

[DPP93] Ivan Damgård, Torben P. Pedersen, and Birgit Pfitzmann. On the existence of statistically hiding bit commitment schemes and fail-stop signatures. In *Advances in Cryptology - CRYPTO '93, 13th Annual International Cryptology Conference, Santa Barbara, California, USA, August 22-26, 1993, Proceedings*, pages 250–265, 1993.

[ESY84] Shimon Even, Alan L. Selman, and Yacov Yacobi. The complexity of promise problems with applications to public-key cryptography. *Information and Control*, 61(2):159–173, 1984.

[Fis12] Marc Fischlin. Black-box reductions and separations in cryptography. In Aikaterini Mitrokotsa and Serge Vaudenay, editors, *Progress in Cryptology - AFRICACRYPT 2012 - 5th International Conference on Cryptology in Africa, Ifrance, Morocco, July 10-12, 2012. Proceedings*, volume 7374 of *Lecture Notes in Computer Science*, pages 413–422. Springer, 2012.

[For89] Lance Jeremy Fortnow. *Complexity-theoretic aspects of interactive proof systems*. PhD thesis, Massachusetts Institute of Technology, 1989.

[Gen09] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *STOC*, pages 169–178, 2009.

[GG98] Oded Goldreich and Shafi Goldwasser. On the possibility of basing cryptography on the assumption that $p \neq np$. *IACR Cryptology ePrint Archive*, 1998:5, 1998.

[GGH+13a] Sanjam Garg, Craig Gentry, Shai Halevi, Mariana Raykova, Amit Sahai, and Brent Waters. Candidate indistinguishability obfuscation and functional encryption for all circuits. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 40–49. IEEE Computer Society, 2013.

[GGH+13b] Sanjam Garg, Craig Gentry, Shai Halevi, Amit Sahai, Mariana Raikova, and Brent Waters. Candidate indistinguishability obfuscation and functional encryption for all circuits. In *FOCS*, 2013.

[GGKT05] Rosario Gennaro, Yael Gertner, Jonathan Katz, and Luca Trevisan. Bounds on the efficiency of generic cryptographic constructions. *SIAM J. Comput.*, 35(1):217–246, 2005.

[GGM86] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986.

[GK93] Oded Goldreich and Eyal Kushilevitz. A perfect zero-knowledge proof system for a problem equivalent to the discrete logarithm. *Journal of Cryptology*, 6(2):97–116, 1993.

[GKLM12] Vipul Goyal, Virendra Kumar, Satyanarayana V. Lokam, and Mohammad Mahmoody. On black-box reductions between predicate encryption schemes. In Cramer [Cra12], pages 440–457.

[GKM+00] Yael Gertner, Sampath Kannan, Tal Malkin, Omer Reingold, and Mahesh Viswanathan. The relationship between public key encryption and oblivious transfer. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA* [DBL00], pages 325–335.

[GM82] Shafi Goldwasser and Silvio Micali. Probabilistic encryption and how to play mental poker keeping secret all partial information. In Harry R. Lewis, Barbara B. Simons, Walter A. Burkhard, and Lawrence H. Landweber, editors, *Proceedings of the 14th Annual ACM Symposium on Theory of Computing, May 5-7, 1982, San Francisco, California, USA*, pages 365–377. ACM, 1982.

[GMM07] Yael Gertner, Tal Malkin, and Steven Myers. Towards a separation of semantic and CCA security for public key encryption. In Salil P. Vadhan, editor, *Theory of Cryptography, 4th Theory of Cryptography Conference, TCC 2007, Amsterdam, The Netherlands, February 21-24, 2007, Proceedings*, volume 4392 of *Lecture Notes in Computer Science*, pages 434–455. Springer, 2007.

[GMR85] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The knowledge complexity of interactive proof-systems (extended abstract). In Robert Sedgewick, editor, *Proceedings of the 17th Annual ACM Symposium on Theory of Computing, May 6-8, 1985, Providence, Rhode Island, USA*, pages 291–304. ACM, 1985.

[GMR01] Yael Gertner, Tal Malkin, and Omer Reingold. On the impossibility of basing trapdoor functions on trapdoor predicates. In *42nd Annual Symposium on Foundations of Computer Science, FOCS 2001, 14-17 October 2001, Las Vegas, Nevada, USA*, pages 126–135. IEEE Computer Society, 2001.

[GMW91] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity for all languages in NP have zero-knowledge proof systems. *J. ACM*, 38(3):691–729, 1991.

[Gol06] Oded Goldreich. On promise problems: A survey. In *Theoretical Computer Science, Essays in Memory of Shimon Even*, pages 254–290, 2006.

[GT00] Rosario Gennaro and Luca Trevisan. Lower bounds on the efficiency of generic cryptographic constructions. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA* [DBL00], pages 305–313.

[GV99] Oded Goldreich and Salil P. Vadhan. Comparing entropies in statistical zero knowledge with applications to the structure of SZK. In *Proceedings of the 14th Annual IEEE Conference on Computational Complexity, Atlanta, Georgia, USA, May 4-6, 1999*, page 54, 1999.

[Has88] Johan Hastad. Dual vectors and lower bounds for the nearest lattice point problem. *Combinatorica*, 8(1):75–81, 1988.

[HH09] Iftach Haitner and Thomas Holenstein. On the (im)possibility of key dependent encryption. In Omer Reingold, editor, *Theory of Cryptography, 6th Theory of Cryptography Conference, TCC 2009, San Francisco, CA, USA, March 15-17, 2009. Proceedings*, volume 5444 of *Lecture Notes in Computer Science*, pages 202–219. Springer, 2009.

[HHRS15a] Iftach Haitner, Jonathan J Hoch, Omer Reingold, and Gil Segev. Finding collisions in interactive protocols—tight lower bounds on the round and communication complexities of statistically hiding commitments. *SIAM Journal on Computing*, 44(1):193–242, 2015.

[HHRS15b] Iftach Haitner, Jonathan J. Hoch, Omer Reingold, and Gil Segev. Finding collisions in interactive protocols - tight lower bounds on the round and communication complexities of statistically hiding commitments. *SIAM J. Comput.*, 44(1):193–242, 2015.

[HM96] Shai Halevi and Silvio Micali. Practical and provably-secure commitment schemes from collision-free hashing. In *Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings*, pages 201–215, 1996.

[HR04] Chun-Yuan Hsiao and Leonid Reyzin. Finding collisions on a public road, or do secure hash functions need secret coins? In *Advances in Cryptology - CRYPTO 2004, 24th Annual International CryptologyConference, Santa Barbara, California, USA, August 15-19, 2004, Proceedings*, pages 92–105, 2004.

[IKO05] Yuval Ishai, Eyal Kushilevitz, and Rafail Ostrovsky. Sufficient conditions for collision-resistant hashing. In *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, pages 445–456, 2005.

[IR89] Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 44–61. ACM, 1989.

[Ish11] Yuval Ishai, editor. *Theory of Cryptography - 8th Theory of Cryptography Conference, TCC 2011, Providence, RI, USA, March 28-30, 2011. Proceedings*, volume 6597 of *Lecture Notes in Computer Science*. Springer, 2011.

[Kle06] Jon M. Kleinberg, editor. *Proceedings of the 38th Annual ACM Symposium on Theory of Computing, Seattle, WA, USA, May 21-23, 2006*. ACM, 2006.

[KLW15] Venkata Koppula, Allison Bishop Lewko, and Brent Waters. Indistinguishability obfuscation for turing machines with unbounded memory. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 419–428, 2015.

[KMN+14] Ilan Komargodski, Tal Moran, Moni Naor, Rafael Pass, Alon Rosen, and Eylon Yogev. One-way functions and (im)perfect obfuscation. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 374–383. IEEE Computer Society, 2014.

[KPTZ13] Aggelos Kiayias, Stavros Papadopoulos, Nikos Triandopoulos, and Thomas Zacharias. Delegatable pseudorandom functions and applications. In Ahmad-Reza Sadeghi, Virgil D. Gligor, and Moti Yung, editors, *CCS*, pages 669–684. ACM, 2013.

[KSS11] Jeff Kahn, Michael E. Saks, and Clifford D. Smyth. The dual BKR inequality and rudich's conjecture. *Combinatorics, Probability & Computing*, 20(2):257–266, 2011.

[KST99] Jeong Han Kim, Daniel R. Simon, and Prasad Tetali. Limits on the efficiency of one-way permutation-based hash functions. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 535–542. IEEE Computer Society, 1999.

[LLJS90] Jeffrey C Lagarias, Hendrik W Lenstra Jr, and Claus-Peter Schnorr. Korkin-zolotarev bases and successive minima of a lattice and its reciprocal lattice. *Combinatorica*, 10(4):333–348, 1990.

[LV16] Tianren Liu and Vinod Vaikuntanathan. On basing private information retrieval on np-hardness. In Eyal Kushilevitz and Tal Malkin, editors, *Theory of Cryptography - 13th International Conference, TCC 2016-A, Tel Aviv, Israel, January 10-13, 2016, Proceedings, Part I*, volume 9562 of *Lecture Notes in Computer Science*, pages 372–386. Springer, 2016.

[MM11] Takahiro Matsuda and Kanta Matsuura. On black-box separations among injective one-way functions. In *Theory of Cryptography*, pages 597–614. Springer, 2011.

[MP91] Nimrod Megiddo and Christos H. Papadimitriou. On total functions, existence theorems and computational complexity. *Theor. Comput. Sci.*, 81(2):317–324, 1991.

[MV03] Daniele Micciancio and Salil P. Vadhan. Statistical zero-knowledge proofs with efficient provers: Lattice problems and more. In Dan Boneh, editor, *Advances in Cryptology - CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings*, volume 2729 of *Lecture Notes in Computer Science*, pages 282–298. Springer, 2003.

[OV08] Shien Jin Ong and Salil P. Vadhan. An equivalence between zero knowledge and commitments. In *Theory of Cryptography, Fifth Theory of Cryptography Conference, TCC 2008, New York, USA, March 19-21, 2008.*, pages 482–500, 2008.

[OW93a] Rafail Ostrovsky and Avi Wigderson. One-way fuctions are essential for non-trivial zero-knowledge. In *ISTCS*, pages 3–17, 1993.

[OW93b] Rafail Ostrovsky and Avi Wigderson. One-way fuctions are essential for non-trivial zero-knowledge. In *ISTCS*, pages 3–17, 1993.

[Pap94] Christos H. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *J. Comput. Syst. Sci.*, 48(3):498–532, 1994.

[Pas06] Rafael Pass. Parallel repetition of zero-knowledge proofs and the possibility of basing cryptography on np-hardness. In *21st Annual IEEE Conference on Computational Complexity (CCC 2006), 16-20 July 2006, Prague, Czech Republic*, pages 96–110. IEEE Computer Society, 2006.

[Pas13] Rafael Pass. Unprovable security of perfect NIZK and non-interactive non-malleable commitments. In *TCC*, pages 334–354, 2013.

[RAD78] R. Rivest, L. Adleman, and M. Dertouzos. On data banks and privacy homomorphisms. In *Foundations of Secure Computation*, pages 169–177. Academic Press, 1978.

[RSA78] Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, 1978.

[RSS16] Alon Rosen, Gil Segev, and Ido Shahaf. Can PPAD hardness be based on standard cryptographic assumptions? *Electronic Colloquium on Computational Complexity (ECCC)*, 23:59, 2016.

[RTV04] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In *Theory of Cryptography, First Theory of Cryptography Conference, TCC 2004, Cambridge, MA, USA, February 19-21, 2004, Proceedings*, pages 1–20, 2004.

[Rud88] Steven Rudich. *Limits on the Provable Consequences of One-Way Functions.* PhD thesis, University of California, Berkeley, 1988.

[Rud91] Steven Rudich. The use of interaction in public cryptosystems (extended abstract). In Joan Feigenbaum, editor, *Advances in Cryptology - CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1991, Proceedings*, volume 576 of *Lecture Notes in Computer Science*, pages 242–251. Springer, 1991.

[Sim98] Daniel R Simon. Finding collisions on a one-way street: Can secure hash functions be based on general assumptions? In *Advances in CryptologyEUROCRYPT'98*, pages 334–345. Springer, 1998.

[SV03] Amit Sahai and Salil Vadhan. A complete problem for statistical zero knowledge. *Journal of the ACM (JACM)*, 50(2):196–249, 2003.

[SW14] Amit Sahai and Brent Waters. How to use indistinguishability obfuscation: deniable encryption, and more. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 475–484. ACM, 2014.

[Vad99] Salil Pravin Vadhan. *A study of statistical zero-knowledge proofs*. PhD thesis, Massachusetts Institute of Technology, 1999.

[Wat15] Brent Waters. A punctured programming approach to adaptively secure functional encryption. In *Advances in Cryptology - CRYPTO 2015 - 35th Annual Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2015, Proceedings, Part II*, pages 678–697, 2015.

[WC81] Mark N. Wegman and Larry Carter. New hash functions and their use in authentication and set equality. *J. Comput. Syst. Sci.*, 22(3):265–279, 1981.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de