Correlation Intractability vs. One-wayness

Tamer Mour[*]
Weizmann Institute of Science
tamer.mour@weizmann.ac.il

December 2020

Abstract

Correlation intractability is an important cryptographic notion that is used for establishing soundness of Fiat-Shamir over public-coin protocols. In this work, we show that symmetric-key cryptography is neither sufficient nor essential for obtaining correlation intractability.

Specifically, we prove a bidirectional fully black-box separation between one-way functions (OWFs) and correlation-intractable hash (CIH). In the first direction, we show that CIH for relations as simple as degree-3 polynomials cannot be based solely on OWFs. In the other direction, we show that there exists no fully black-box construction of OWF from CIH for all sparse relations. Consequently, we infer that computationally sound Fiat-Shamir over *any* specific constant-round proof system does not necessarily require one-way functions.

[*] Supported by the Binational Science Foundation (Grant No. 2016726), and by the European Union Horizon 2020 Research and Innovation Program via ERC Project REACT (Grant 756482) and via Project PROMETHEUS (Grant 780701), and by the Azrieli Fellows Program.

Contents

1 Introduction

The Fiat-Shamir transform [FS87] is a popular generic technique for eliminating interaction in interactive public-coin protocols. Fiat-Shamir was first applied to 3-round identification protocols to obtain non-interactive signature schemes [FS87]. Since its introduction, this methodology has had a substantial impact on modern cryptography through several lines of research. Fiat-Shamir was found to be very useful, both for achieving new theoretical feasibility results and for designing communication-efficient practical solutions. In particular, among its noticeable applications are non-interactive zero knowledge protocols (NIZKs) [KRR17, CCH+19, PS19, BKM20], succint non-interactive arguments (SNARGs) [Kil92, Mic00, BSCS16, BSBHR19], and complexity-theoretic hardness results [CHK+19, LV20a, JKKZ20].

The Fiat-Shamir transform over an interactive public-coin protocol uses a hash function H. The transform obtains a non-interactive protocol by letting the prover locally simulate the interaction with the verifier while computing the public-coin challenges using H, without any interaction. More specifically, in each round of the simulation, the prover computes the verifier's message as the hash of all messages produced in the simulated interaction so far.

While it is usually straight-forward to show that Fiat-Shamir preserves some properties of the original interactive protocol, e.g. completeness and zero-knowledge, it is not clear that it preserves soundness using any hash function H. This is since, intuitively, the prover has some control over the computed challenges. In fact, in most applications, the soundness of Fiat-Shamir is based on *heuristics*. Namely, the soundness of the non-interactive protocol is assumed per se, and is not based on the hardness of a well-studied mathematical problem. Indeed, in practice, H is instantiated by an "unstructured" function, e.g. SHA-2. The soundness of such an instantiation is theoretically justified through the random oracle model [BR94]: by modeling the hash function as a random oracle, which both parties have access to, one can prove that the Fiat-Shamir transform is secure as long as a cheating prover does not make unreasonably many queries to the oracle. Thus, if the hash function behaves like a random function in the eyes of a bounded adversary, then the non-interactive protocol is sound.

Although the random oracle model provides a clean theoretical framework, it is not clear that a sound Fiat-Shamir under the random oracle is a strong enough evidence that provably sound Fiat-Shamir in the plain model exists. In fact, Goldwasser and Kalai [GK03] show that there exists a computationally sound protocol (i.e. argument) on which the Fiat-Shamir transform is never sound when instantiated with any actual efficient hash function, even though it is sound in the random oracle model. Further, Bitansky et al. [BDSG+13] rule out the possibility of constructing a "universal" Fiat-Shamir hash function for all 3-message public-coin protocols based on standard assumptions, or even basing the soundness of Fiat-Shamir for some specific protocols on any falsifiable assumption.

This gap between the conjectured soundness of Fiat-Shamir using "sufficiently unstructured" functions and its provability under cryptographic assumptions in the plain model led Canetti, Goldreich and Halevi [CGH04] to introduce the notion of *Correlation Intractability*. Essentially, correlation intractability captures the computational hardness needed from a Fiat-Shamir hash function in order to prove the soundness of the transform. We say that H is a correlation-intractable hash for a relation class $\mathcal{R}$ (CIH for $\mathcal{R}$) if, for any relation $R \in \mathcal{R}$, it is computationally hard given a random hash key k to find an input x such that $(x, H(k, x)) \in R$. Roughly speaking, in order to show that a Fiat-Shamir instantiation is sound for a given protocol, we would require that the underlying hash function is correlation-intractable for the relation between partial protocol transcripts and "bad" verifier challenges that allow for soundness error. Based on this outline, it is known [BLV06, CCR16, KRR17] that a CIH for *all sparse relations* (i.e. relations where any x is in relation with at most a negligible fraction of all y's – see Definition 4.1) is sufficient for Fiat-Shamir over *any* constant-round public-coin proof (the special case of 3-message protocols has appeared already in [DNRS03, HT06]).

While Canetti et al. [CGH04] show that obtaining correlation intractability in its most general form is impossible, an extensive line of work has eventually led to CIH constructions that are useful for a wide class of protocols, including zero knowledge [CCR16, KRR17, CCH+19], statistical ZAP arguments [BFJ+20, GJJM20] and other special-purpose protocols [CHK+19, JKKZ20, LV20a]. Overall, the state-of-the-art constructions of CIH are based on well-studied cryptographic primitives which are, in turn, provably secure under standard assumptions such as LWE [PS19, LV20b] (through special fully-homomorphic commitments or shiftable shift-hiding functions [PS18]) and DDH [BKM20] (through trapdoor hash functions [DGI+19]).

In this work, we seek to understand the relation between correlation intractability and a different cryptographic notion of hardness, arguably the most fundamental of all: *One-wayness.* One-way functions (OWF) [DH76] are functions that are easy to compute but hard to invert. OWFs constitute a central building block in modern cryptography, and were shown to be essential and sufficient for obtaining basic symmetric-key cryptographic notions (a.k.a. Impagliazzo's "Minicrypt" [Imp95]), such as pseudorandom generators [HIL99], pseudorandom functions [GGM85], symmetric encryption [GM84], commitment schemes [Nao91], zero knowledge [OW93], and more.

1.1 Our Results

We investigate the two directions of this relation and explore the possibilities, but mainly the impossibilities, in obtaining CIH from OWF and OWF from CIH. More specifically, we consider *fully black-box reductions* between the two primitives and establish a two-directional separation using a well-studied framework that was developed in prior work [IL89, HR04, RTV04]. We then discuss some implications of our findings to the complexity of the Fiat-Shamir transform. We elaborate below.

Correlation Intractability from One-wayess. While it is almost trivial that one-way functions imply restricted notions of correlation intractability, such as CIH for all relations $R_a = \{(x, h(x) + a)\}$ (where h is any arbitrary fixed function and addition is over a finite field)[1], such CIH are too weak to realize any interesting applications, in particular Fiat-Shamir for useful protocols. It is also known [HL18] that exponentially-secure OWF imply *output-intractability*, which is a special case of correlation-intractability for relations R where the membership $(x, y) \in R$ is determined solely by the value of y (but is more general in the sense that it considers tuples of such outputs), and has different applications. In contrast, known useful CIH constructions, for input-output relations, are either based on public-key cryptographic primitives [CCH+19, PS19, BKM20, LV20b], or based on exponentially secure OWF and additionally assume the existence of indistinguishability obfuscation (iO) [HL18], or based on sub-exponentially secure one-way permutations and iO [LV20b]. In the first part of this work, we attempt to understand whether the limitations in basing correlation intractability on OWF are inherent. In general, we ask the following:

Can we construct correlation-intractable hash for useful classes of relations based merely on OWF?

We give a negative answer by showing that there exists no fully-black-box construction of CIH for any sufficiently "expressive" class of relations based on OWF. More specifically, our impossibility result captures any *3-wise universal* class of relations. This is a class where there exists a distribution over the relations such that the membership of input-output pairs in a random relation is 3-wise independent and is equally probable (over all pairs).

Theorem 1.1 (Informal)**.** *There exists no fully-black-box construction of correlation-intractable hash (CIH) for 3-wise universal relations from one-way permutations.*

For instance, the class of all relations searchable by degree-3 polynomials over a finite field [2] is 3-wise universal (and, therefore, so is any class that contains it). For comparison, known Fiat-Shamir applications require CIH for far more expressive classes such as relations searchable by cryptographic algorithms (e.g. decryption in a PKE) which may be general polynomial-time circuits [HL18, CCH+19] or "noisy" constant-degree polynomials [BKM20].

[1] The hash function $H(k, x) = f(x) + h(x) + k$, where f is a OWF, is correlation intractable for $\{R_a\}$. An adversary that breaks the correlation intractability of H for some R_a inverts f at a random image y when given the random key $k = a - y$.

[2] A relation R is searchable by a polynomial p if $(x, y) \in R$ implies $y = p(x)$.

One-wayness from Correlation-Intractability (or Fiat-Shamir). Given the current state of CIH constructions, and the fact that one-way functions are insufficient to obtain correlation intractability (at least in a fully black-box manner), it seems that the notion of correlation intractability is strictly stronger than one-wayness. Indeed, as pointed out in [CLMQ20], there exists a simple construction of OWF based on CIH for a relatively small class of sparse relations. While this demonstrates the necessity of OWFs for general correlation intractability and universal Fiat-Shamir hash [BLV06, CCR16], it does not give any indication regarding the complexity of instantiating Fiat-Shamir over a *specific* protocol and, in particular, it leaves open the following important question:

Is there a protocol Π *such that OWFs are necessary to establish the soundness of Fiat-Shamir over* Π*?*

Not surprisingly, we are able to give a meaningful answer to this question through studying the black-box reducability of one-way functions to correlation intractability. Our main result in this part is a *fully black-box separation* of OWFs from CIH for all sparse relations.

Theorem 1.2 (Informal)**.** *There exists no fully-black-box construction of one-way functions from correlation intractable hash for all sparse relations.*

Our definition of a fully black-box construction in this context captures any construction of a OWF F from a CIH H, where the security reduction breaks the correlation intractability of H for a *pre-determined* sparse relation R using *any* adversary that inverts F. That is, we require that the reduction "chooses" the relation R, with respect to which it breaks H, independently of the given adversary against F. While this notion of fully black-box constructions is rather limited and, in particular, does not capture the aforementioned known construction [CLMQ20][3], such a separation has interesting consequences.

Specifically, building on the connection between correlation intractability and Fiat-Shamir for statistically-sound protocols [BLV06, CCR16] we imply that Fiat-Shamir for *any* such bounded-interaction protocol Π is not sufficient to imply the existence of one-way functions, at least not in a relativizable manner (i.e. relative to any idealized world).

Theorem 1.3 (Informal)**.** *For any statistically-sound constant-round public-coin protocol* Π*, there is an oracle relative to which there exists computationally sound Fiat-Shamir over* Π *but no one-way functions.*

We stress that the separation considers an oracle relative to which the Fiat-Shamir transform is *not* statistically sound (in fact, the Fiat-Shamir simply uses a random hash function). That is, although the Fiat-Shamir transform is based on computational hardness, such notion of hardness does not imply one-way functions relative to any oracle.

1.2 Related Work

We hereby discuss prior related work.

Related Work on the Plausibility of Correlation Intractability from OWF. Already in 1999, Hada and Tanaka [HT99] ask whether the existence of OWFs implies correlation intractability for all sparse relations. They show that if such an implication is true, then all languages with 3-round auxiliary-input zero-knowledge proof systems are easy to approximate. This statement is believed to be difficult to prove unconditionally and, therefore, it is unlikely that one can base strong correlation intractability on one-way functions. While [HT99] give strong evidence against the possibility of obtaining CIH for all sparse relations from OWFs, we definitively rule out any fully black-box construction of CIH, even for degree-3 polynomials, from OWFs.

[3] The construction from [CLMQ20] is only *semi black-box*: their (implicit) security reduction can never commit to a "target" relation R w.r.t. which it breaks the CIH given any OWF inverter. As a matter of fact, a semi-black box separation would imply that OWFs are not necessary even for a "universal" Fiat-Shamir hash (which provides soundness for a wide class of protocols, simultaneously). This is incorrect due to Dodis et al. [DRV12].

Related Work on the Complexity of Fiat-Shamir. The necessity of OWFs (and, in general, cryptographic hardness) for Fiat-Shamir was first studied in the recent work of Chen et al. [CLMQ20], *"Does Fiat-Shamir Require a Cryptographic Hash Function?"*. The authors show that this question may have different answers when considering different types of protocols. On the one hand, they show that there exist useful protocols, possibly in idealizied models (e.g. Schnorr's identification scheme [Sch90] in the generic group model), where a simple "non-cryptographic" function is sufficient for a sound Fiat-Shamir – we note that this does not follow the correlation-intractability framework but rather relies on the soundness of the specific underlying protocols. On the other hand, they show there exist protocols where a sound Fiat-Shamir requires a hash function that posses some computational hardness, such as "dependency" on a generic group oracle or, in the standard model, a notion they call "mix-and-match" resistance. To summarize, [CLMQ20] focus on understanding the complexity of *designing* a Fiat-Shamir hash function, through analyzing a variety of typical use cases, whereas we are interested in the *theoretic-complexity* of Fiat-Shamir (relative to OWFs), and our results apply to Fiat-Shamir over any protocol.

Lastly, Choudhuri et al. [CHK$^+$19] show that if there exists sound Fiat-Shamir for the sumcheck protocol then there exists an efficiently samplable distribution of **PPAD** instances which is hard-on-average given $\#\mathbf{P}$ is hard. Even though an efficiently samplable hard-on-average distribution does not necessarily imply OWFs, we stress that our separation methodology does not apply to their proof technique since their proof is *unrelativizable*, whereas our results presume a world relative to an oracle. Specifically, their reduction builds on the sumcheck protocol which serves as an interactive proof for a $\#\mathbf{P}$-complete language. While this is true in the plain model, the sumcheck protocol is inherently unrelativizable since it uses arithmetization and, therefore, it cannot be used for a $\#\mathbf{P}^{\mathcal{O}}$-complete language relative to any oracle $\mathcal{O}$.

1.3 Technical Overview

We now give an overview of our proofs and the underlying techniques. Our main result is a two-directional fully-black-box separation between correlation intractability and one-wayness.

In the first direction, we show that correlation intractability, even for relations as simple as degree-3 polynomials, cannot be based on one-way permutations in a fully black-box manner. Our proof uses techniques from the work of Bitansky and Degwekar [BD19], where they separate collision-resistant hash from OWFs. In the other direction, we show that there exists no fully-black-box construction of one-way functions based on correlation-intractable hash as secure as a random oracle, namely correlation intractable for all sparse relations.

First, let us briefly recall the fully black-box separation framework which we follow in our proofs.

Fully Black-box Separations. We say that a construction P of a cryptographic primitive $\mathbf{P}$ from a different primitive $\mathbf{Q}$ is *fully black-box* [RTV04] if the construction makes only black-box use of $\mathbf{Q}$ (that is, any *instantiation* of $\mathbf{Q}$, independently of its implementation) and, further, there is a black-box security reduction R which breaks P if it is given a black-box access to *any* adversary $\mathcal{A}$ that breaks the underlying instantiation of $\mathbf{Q}$. A *fully black-box separation* of $\mathbf{P}$ from $\mathbf{Q}$ simply means that fully black-box constructions of $\mathbf{P}$ from $\mathbf{Q}$ are impossible. As observed by Reingold et al. [RTV04], most constructions of cryptographic primitives in the literature are fully black-box and, therefore, it is insightful to understand when such constructions are possible.

A well-developed method to establish a fully black-box separation of $\mathbf{P}$ from $\mathbf{Q}$ is often referred to as the "Two-Oracle Methodology" [Sim98, HR04, AS16, BD19]. In this method, it is shown that there exists an oracle Q, which models an "ideal" implementation of $\mathbf{Q}$, and an oracle A that models an adversary against $\mathbf{P}$, such that (i) A breaks any black-box construction of $\mathbf{P}$ from Q, yet, (ii) Q is still secure (as per the security definition of $\mathbf{Q}$) in the presence of A. Given such oracles Q and A exist, any fully black-box reduction R fails in breaking Q using the adversary A and, hence, no fully black-box construction of $\mathbf{P}$ from $\mathbf{Q}$ exists.

1.3.1 Correlation Intractability from OWP

Inspired by ideas from prior work [BD19], we prove our first result, which is informally stated in Theorem 1.1: a fully-black-box separation of CIH form OWP. The separation proof applies for CIH for any class of relations which is 3-wise universal. Namely, we require that for a random relation R (according to some distribution), the events $\{(z,w) \in R \mid z \in \{0,1\}^m, w \in \{0,1\}^n\}$ are 3-wise independent and all occur with the same probability $p(n)$ (see Definition 3.2 for the formal requirement).

Background: CRH from OWP. We recall the known separation of collision-resistant hash (CRH) from OWP, specifically, the proof from Bitansky-Degwekar [BD19]. Following the two-oracle methodology described above, in order to prove that a fully-black-box construction of CRH from OWP is impossible, [BD19] show that there exists a collision-finding oracle that breaks any CRH, under which a random permutation is still one-way. Consequently, any construction of a CRH using a permutation f (in particular, a random permutation) cannot be proven secure based merely on the one-wayness of f.

At a high-level, the collision-finding oracle takes as input an oracle-aided circuit C^f and outputs a collision (z, z') (s.t. $C^f(z) = C^f(z')$) with high probability, given C is sufficiently compressing. The oracle is designed such that each of z and z' has uniform marginal distribution (over the choice of a random oracle).

It is quite straight forward to see that no CRH exists under the collision-finding oracle and any f: to break a candidate CRH, the adversary simply calls the collision-finding oracle with the circuit describing the computation of the hash. To that end, all existing CRH-from-OWP separation proofs share the same outline. The more challenging part is to show that a random permutation f is one-way under the collision-finder.

The Smoothening Technique. Bitansky and Degwekar [BD19] use a hybrids argument in order to show that any adversary $\mathcal{A}$ with access to f and the collision-finder fails in inverting f at a random image y. At its core, their proof is based on the fact that any query to the collision-finding oracle can be simulated by $\mathcal{A}$ using a *smooth* query, without harming $\mathcal{A}$'s success probability. A smooth query consists of a circuit C such that, for any $x \in \{0,1\}^*$, the probability that C on a uniform input calls f at x is negligible. Consequently, we may reduce our goal to showing one-wayness against *smooth adversaries*, i.e. adversaries that make smooth queries only. This is useful since the marginal distribution of any answer z given by the collision-finder (as part of a collision) is uniform and, therefore, the probability that a smooth query to the collision-finder will output z such that $C^f(z)$ calls $f^{-1}(y)$ is negligible. Intuitively, this means that the collision-finder cannot help a smooth adversary to find the pre-image.

The smoothening of any adversary $\mathcal{A}$ is done as follows: whenever $\mathcal{A}$ is supposed to call the collision finder at C, he evaluates C on sufficiently many random inputs and records all queries made to f through these evaluations in a table. He then produces a circuit C' which emulates C, with the table hardwired in it, and sends it to the collision-finder. Observe that the random evaluations aim to detect any "heavy" f-input x and hardwire it in the circuit C' so it is never called during its evaluation. If sufficiently many random evaluations are made, every heavy input will be detected with high probability and, therefore, C' will be smooth.

Designing a Correlation-Finding Oracle. To separate CIH from OWP, we seek to construct a *correlation-finding oracle* $\mathsf{CorrFinder}$ which takes as input a circuit C and a key k and finds a correlation w.r.t. some relation R, i.e. an input z such that $(z, C(k,z)) \in R$ [4]. Suppose we follow the above outline, then, we would like to claim that if $C^f(k,z)$ makes some f-query x with noticeable probability, where z is a random answer returned by $\mathsf{CorrFinder}(C,k)$, then x is almost surely detected by the smoothening simulation. Recall that this was true with the collision-finding oracle because its answers are marginally uniform, thus, if x is "heavy" among the answers of the collision-finder, then it is "detectable" by the simulation (which evaluates the circuit on many random inputs). Unfortunately, a correlation-finder's answer cannot be a uniform input by the mere fact that it must be a correlation. It appears, then, that we now have a gap between an input

[4] Notice that the correlation-intractability game allows the relation R to depend on the circuit C but not on k.

being “heavy” among oracle answers – in our case, the correlation-finding oracle – and being “detectable” by the smoothening simulation, i.e. heavy w.r.t. a uniform input.

More specifically, we consider a correlation-finder that samples a uniformly random correlation, and we say that an f-input x is *heavy* (among correlations) if $C^f(k,z)$ calls f at x with noticeable probability for a random correlation z (equivalently, for a random z answered by $\mathsf{CorrFinder}$). We say that x is *detectable* if this is true when z is a uniformly random input. Since we want to prevent the correlation-finder from incurring heavy f-inputs in smooth queries, we define it to abort whenever it samples a correlation z such that $C^f(k,z)$ calls an input which is heavy but is not detectable.

Finding Correlations using the Correlation-Finder. In order to show that the oracle $\mathsf{CorrFinder}$ described above indeed breaks any CIH for a relation class $\mathcal{R}$, we must show that for any circuit C (corresponding to a CIH candidate), there exists a relation $R \in \mathcal{R}$ such that the probability that $\mathsf{CorrFinder}(C,k)$ outputs a correlation z (s.t. $(z, C(k,z)) \in R$) is large enough. In fact, we show that this is true for a random $R \leftarrow \mathcal{R}$ with non-zero probability, assuming $\mathcal{R}$ is 3-universal.

Observe that $\mathsf{CorrFinder}$ fails either if there exist no correlations at all (this happens only if C is statistically correlation-intractable, which is impossible for interesting classes of relations), or if the oracle samples a correlation z such that $C^f(k,z)$ makes an f-query which is heavy but not detectable. It suffices, then, to bound the probability that for a random relation R, random key k, and a random correlation z, $C^f(k,z)$ calls f on a heavy but non-detectable input. Equivalently, we may consider an experiment where k and z are sampled according to their original marginal distribution and then, once they are fixed, we sample a relation R conditioned on $(z, C^f(k,z)) \in R$. Thus, we are interested in answering the following question: for any fixed k and z, what is the probability that a non-detectable f-query made by $C^f(k,z)$ is heavy w.r.t. such a random R? Using Chebyshev’s concentration bound, we show that if the distribution of R is such that the events $\{(z', C^f(k,z')) \in R \mid z' \in \{0,1\}^m\}$ are pairwise-independent, then the occurrence of any f-query among the correlations (which is a random subset of the inputs, defined by R) gives a good estimation of its occurrence among all inputs. Put differently, it is unlikely that a non-detectable f-query is heavy among correlations w.r.t. a random R. By requiring that our relation class is 3-universal, where the events $\{(z', w) \in R \mid z' \in \{0,1\}^m, w \in \{0,1\}^n\}$ are 3-wise independent, we derive the pairwise-independence property for the conditioned distribution over relations.

One-wayness of a Random Permutation under $\mathsf{CorrFinder}$. To complete the separation proof, it remains to show that a random permutation f is one-way, also against an adversary with access to the correlation-finding oracle. As already mentioned, we follow the framework laid by [BD19] and reduce to the case where the adversary is assumed to be smooth. We consider a series of hybrids transforming the inversion experiment, where a random input x is sampled, and the smooth adversary $\mathcal{A}^{f,\mathsf{CorrFinder}}$ is challenged to find x given $f(x)$, to an experiment where the adversary has to find x given an independently random image y. It is clear that the adversary cannot win with noticeable probability in the latter case. To establish indistinguishability between the two experiments, we essentially require that an adversary is unable to distinguish between access to the oracle $\mathsf{CorrFinder}$ and access to a *punctured oracle* $\mathsf{CorrFinder}_x$, given $f(x)$. The punctured oracle $\mathsf{CorrFinder}_x$ behaves similarly to $\mathsf{CorrFinder}$ except it aborts also if $C^f(k,z)$ calls f at x (where z is the random correlation sampled by the oracle). We rely on the smoothness of $\mathcal{A}$ to show that

$$|\Pr[\mathcal{A}^{f,\mathsf{CorrFinder}}(f(x))] - \Pr[\mathcal{A}^{f,\mathsf{CorrFinder}_x}(f(x))]| < \mathsf{negl} \tag{1}$$

as follows. For any query (C,k) that $\mathcal{A}$ makes to $\mathsf{CorrFinder}$, if x is heavy then, since x is not detectable by smoothness, $\mathsf{CorrFinder}$ already aborts when $C^f(k,z)$ calls f at x, therefore, $\mathsf{CorrFinder}(C,k) = \mathsf{CorrFinder}_x(C,k)$. Otherwise, if x is not heavy, then the probability that $C^f(k,z)$ calls f at x for a random correlation z is negligible and, therefore, so is the probability that $\mathsf{CorrFinder}(C,k) \neq \mathsf{CorrFinder}_x(C,k)$.

Passive Smoothening. The rest of the hybrid argument is completed following the outline in [BD19]. There remains, however, a subtle issue with smoothening the adversary, specifically, in the correctness of the simulation (which is essential for the reduction). Recall that we simulate any adversary using a smooth adversary, by replacing any $\mathsf{CorrFinder}$-query (C, k) with a query (C', k), where C' emulates C while using a hardwired table to locally answer some of its f-queries. Notice that the table may possibly contain a non-detectable and heavy query. In such a case, $\mathsf{CorrFinder}(C, k)$ and $\mathsf{CorrFinder}(C', k)$ do not have equal distributions and, therefore, the simulation may not be correct. We observe, however, that the indistinguishability argument from (1) only requires that any $\mathsf{CorrFinder}$-query made by $\mathcal{A}$ is smooth at x (the pre-image that $\mathcal{A}$ is challenged to find), and not necessarily on all inputs. This leads us to a change the smoothening strategy as follows: the simulation still evaluates $C(k, \cdot)$ on a bunch of random inputs however it does not change $\mathcal{A}$'s queries to $\mathsf{CorrFinder}$ at all. We also let the simulation to halt immediately and output the pre-image x if it finds it during the random evaluations. It is not hard to see that this "passive" simulation is already smooth at x since if some query is not smooth at x, then x is almost surely detected by the random evaluations before the query is made. It is also straight-forward that the success probability of the simulation in inverting f is at least as good as the success probability of the adversary $\mathcal{A}$. In conclusion, by considering the above relaxed notion of smoothness, which is still useful for the one-wayness proof, we obtain a reduction to a smooth adversary that preserves correctness, and by this complete the separation proof.

1.3.2 One-way Functions from Correlation-Intractable Hash

Having shown a black-box separation of CIH from one-way permutations (and, therefore, OWF), we next present our separation result in the other direction (see Theorem 1.2): there exists no fully-black-box construction of one-way functions from correlation-intractable hash. The separation holds even from a CIH that is as strong as the correlation intractability of a random oracle, namely, a CIH for all sparse relations.

Recall that our notion of fully black-box constructions captures constructions of OWF where the security reduction breaks the correlation intractability of the underlying CIH for a relation R which is pre-determined and independent in the OWF inversion algorithm. That is, we require that there exists a reduction R and a relation R such that the reduction breaks the CIH w.r.t. R given access to *any* adversary $\mathcal{A}$ that breaks the candidate OWF.

Our separation proof has the following high-level outline. We consider a random hash function h as an "idealized" CIH. We then show that for any sparse relation R, there exists an *inversion oracle* that breaks any OWF construction from h and, yet, any security reduction R with access to the inversion oracle is unable to break the CIH of h for R.

The Inversion Oracle. Consider an inversion oracle Inv which, on input a function F (represented as an oracle-aided circuit) and an image y, outputs a random pre-image x s.t. $\mathsf{F}^h(x) = y$. It is clear that any F is invertible on any image under such an oracle. It is not the case, however, that a random hash function h is CI for any relation R under this straight-forward Inv. In particular, we can consider the case where F^h maps any input z, s.t. $(z, h(k, z)) \in R$, to 0. The reduction may use the inversion oracle to invert F at the image 0, and obtain a correlation.

To prevent such attacks, we define the inversion oracle w.r.t. a relation R to sample a random pre-image x such that $\mathsf{F}^h(x)$ does not make a *correlation query* to h, i.e. a query (k, z) s.t. $(z, h(k, z)) \in R$[5]. As a sanity check, notice that the reduction described above will fail under such an oracle since the query it makes will be always answered by $\perp$ because $\mathsf{F}^h(x)$ always makes a correlation query for any x in the pre-image of 0.

We first show that the inversion oracle indeed breaks the one-wayness of any candidate F^h. In fact, this is not true for all F and all h since one may imagine a case where $\mathsf{F}^h(x)$ queries a fixed correlation at the start of the computation on any input x (in such a case, the inverter will never return a pre-image). Instead, we prove that the claim is true for all F and a random hash function h with high probability ($> \frac{2}{3}$). This will be sufficient for the separation. Roughly speaking, if the inverter fails with noticeable probability at inverting a random challenge y (where $y = \mathsf{F}^h(x)$ for a random x) then this means that, with noticeable probability, all pre-images of such a y incur a correlation query under F^h. In this case, we can break the

[5] Note that due to our definition of fully-black-box constructions, we may separate by defining an inverter depending on R.

correlation intractability of h by simply sampling a random x and computing $\mathsf{F}^h(x)$. Since a random h is correlation-intractable for any sparse relation, this is impossible to do using an efficient F. This implies, through an averaging argument, that Inv succeeds in inversion with good probability for a strong majority of functions h, more accurately, for at least $\frac{2}{3}$-fraction. To summarize, we derive the success of Inv from the correlation-intractability of a random h (in the "plain model").

A Random Hash is Correlation-Intractable under Inv. It remains to show that, for any relation R, an adversary $\mathcal{A}$ with access to Inv cannot break the correlation intractability of a random oracle h w.r.t. R. We consider a different experiment, where the adversary $\mathcal{A}$ is given access to a random *statistically correlation-intractable* hash h', i.e. a hash function sampled at random conditioned on $(z, h'(k, z)) \notin R$ for all k and z. We claim that if the adversary $\mathcal{A}$ breaks the correlation intractability of h for R under Inv, then he breaks the correlation intractability of h' for R under Inv' (the inversion oracle defined w.r.t. h'), which is impossible since h' never contains any correlations. We establish the reduction based on the following indistinguishability argument

$$|\Pr[\mathcal{A}^{h,\mathsf{Inv}}(k) = 1] - \Pr[\mathcal{A}^{h',\mathsf{Inv}'}(k) = 1]| < \mathsf{negl}.$$

To see this, imagine sampling h and Inv in the original experiment as follows: First, sample a random stat.-CI hash h' and its corresponding inversion oracle Inv'. Then, for any h-input (k, z) assign $h(k, z) = h'(k, z)$ with probability $\Pr_h[(z, h(k, z)) \notin R]$ and, otherwise, sample a new $h(k, z)$ conditioned on $(z, h(k, z)) \in R$. We set Inv to behave like Inv' while making the necessary changes: if $\mathsf{F}(\mathsf{Inv}'(\mathsf{F}, y)) \neq y$ for some y, sample a fresh $\mathsf{Inv}(\mathsf{F}, y)$ at random from all pre-images x s.t. $\mathsf{F}^h(x)$ does not make a correlation query. By careful inspection, we show that sampling h and Inv as described is equivalent to sampling a random h and a corresponding Inv. Further, notice that the adversary's view when given h' and Inv' rather than h and Inv changes only if it contains an h-query (k, z) for which $h(k, z) \neq h'(k, z)$. From the sparseness of R, this happens with negligible probability for all z and, thus, we obtain the indistinguishability.

By applying another averaging argument, we argue that since any adversary breaks the CI of a random h with negligible probability, then this is true for at least $\frac{2}{3}$ of fixed functions h. Since we have shown that Inv breaks any OWF candidate for at least $\frac{2}{3}$-fraction of functions h, then there must exist a function h (in fact, at least $\frac{1}{3}$-fraction of all functions) that separates OWFs from CIH.

Consequence: Fiat-Shamir does not Necessarily Imply OWFs. Having shown that correlation intractability does not imply one-way functions in a fully black-box manner, we use the well-studied connection between Fiat-Shamir and correlation intractability to separate one-way functions from the soundness of Fiat-Shamir transform over any given constant-round protocol. More specifically, we show that, for any statistically sound constant-round public-coin protocol Π, there exists an oracle relative to which there is Fiat-Shamir over Π which is sound against any efficient adversary but is not statistically sound, yet no one-way functions exist.

First, we establish a fully black-box reduction from the Fiat-Shamir soundness over any statistically sound constant-round protocol to correlation intractability. For any such protocol Π, we identify a sparse relation R_Π where breaking the Fiat-Shamir over Π using a hash function h reduces to breaking the correlation intractability of h for the relation R_Π. Given such a reduction, Theorem 1.3 follows almost immediately from the separation of OWFs from CIH: relative to a random hash function h and the inversion oracle corresponding to the relation R_Π, h is correlation intractable – therefore, due to the reduction, Fiat-Shamir over Π is sound – however, every OWF candidate is broken.

1.4 Paper Organization

The next section contains some preliminaries. In Section 3 we prove the fully black-box separation of correlation intractable hash from OWPs and, in Section 4, we show the separation in the opposite direction and its consequence to the complexity of the Fiat-Shamir transform.

2 Preliminaries

Let us introduce some basic notation and conventions, and recall some preliminary definitions and facts.

Notation. We denote by $\mathcal{F}_{n\to m}$ the set of all functions $f : \{0,1\}^n \to \{0,1\}^m$. For a distribution X, we write $x \in X$ to say that x is in the support of X, and $x \leftarrow X$ to denote that x is sampled from the distribution X. We overload the notation for sets and write $x \leftarrow S$ when x is sampled uniformly at random from the set S. $\mathbf{SD}(X,Y)$ denotes statistical distance between distributions X and Y. For an oracle-aided algorithm $\mathcal{A}$, an oracle Ψ, and a Ψ-input z, we denote by $\mathcal{A}^\Psi(x) \overset{\Psi}{\rightharpoondown} z$ the event where $\mathcal{A}$, on input x, calls the oracle Ψ at z, and use $\mathcal{A}^\Psi(x) \overset{\Psi}{\not\rightharpoondown} z$ to indicate the opposite. We extend this notation for sets: $\mathcal{A}^\Psi(x) \overset{\Psi}{\rightharpoondown} Z$ if $\mathcal{A}^\Psi(x)$ calls some $z \in Z$, and $\mathcal{A}^\Psi(x) \overset{\Psi}{\not\rightharpoondown} Z$ otherwise.

2.1 Coupling and Statistical Distance

Coupling is a useful tool for bounding the statistical distance between two probability measures.

Definition 2.1 (Coupling). *Let X and Y be two random variables (i.e., distributions) over $\mathcal{X}$ and $\mathcal{Y}$ (resp.). We say that a distribution $X'Y'$ over $\mathcal{X} \times \mathcal{Y}$ is a* coupling *of X and Y if the marginal distributions of X' and Y' are identical to the distributions of X and, respectively, Y.*

We denote by $\mathcal{P}_{X,Y}$ the set of all couplings of X and Y.

Proposition 2.2 (Statistical Distance through Coupling). *Given any two distributions X, Y over $\mathcal{X}$,*

$$\mathbf{SD}(X,Y) = \inf_{X',Y' \in \mathcal{P}_{\mathcal{X},\mathcal{Y}}} \Pr_{(x,y)\leftarrow X'Y'}[x \neq y]$$

2.2 Chebyshev's Inequality

We hereby recall Chebyshev's tail bound.

Proposition 2.3 (Chebyshev's Inequality). *Let X be a random variable with excpected value μ and non-zero variance σ^2. Then, for any $k \in \mathbb{R}$,*

$$\Pr[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2}.$$

2.3 The Borel-Cantelli Lemma

Proposition 2.4 (The Borel-Cantelli Lemma). *Let $\{E_n\}_{n\in\mathbb{N}}$ be a sequence of events in some probability space such that the sum $\sum_n \Pr[E_n]$ converges. Then, the event where E_n occurs for infinitely many $n \in \mathbb{N}$ has probability measure 0.*

3 Separating Correlation Intractability from OWP

In this section we prove our first separation result, showing the impossibility of fully-black-box constructions of correlation-intractable hash based on OWP. First, let us define a fully-black-box construction of a CIH from OWP.

Definition 3.1 (Black-box Construction of CIH from OWP). *Let $\kappa := \kappa(n)$ and $m := m(n)$ be implicit length parameters, and let $\mathcal{R}$ be a class of relations. A (q,ϵ)-*fully black-box construction of Correlation Intractable Hash (CIH) for $\mathcal{R}$ from One-way Permutations (OWP)*, for functions $q := q(\lambda)$ and $\epsilon := \epsilon(\lambda)$, consists of a pair of PPT oracle-aided algorithms* $\mathsf{CIH} = (\mathsf{Gen}, \mathsf{Hash})$ *and an oracle-aided reduction* R*, satisfying the following properties:*

- ***Correctness:*** *For any permutation ensemble $f = \{f_\lambda : \{0,1\}^\lambda \to \{0,1\}^\lambda\}_{\lambda\in\mathbb{N}}$, any $n \in \mathbb{N}$ and $k \in \mathsf{Gen}(1^n)$, it holds that*
$$\mathsf{Hash}^f(k,\cdot) : \{0,1\}^m \to \{0,1\}^n$$

- ***Black-box Security Reduction:*** *For any permutation ensemble $f = \{f_\lambda : \{0,1\}^\lambda \to \{0,1\}^\lambda\}_{\lambda\in\mathbb{N}}$ and any probabilistic oracle-aided adversary $\mathcal{A}$, if there exists a relation $R \in \mathcal{R}$ such that*
$$\mathbf{Adv}^{\mathbf{ci}}_{\mathsf{CIH},\mathcal{A},R}(n) := \Pr_{\substack{k\leftarrow\mathsf{Gen}(1^n)\\ z\leftarrow\mathcal{A}^f(1^n,k)}}[(z,\mathsf{Hash}^f(k,z)) \in R] > \frac{1}{2}$$
for infinitely many $n \in \mathbb{N}$, then,
$$\Pr_{\substack{x\leftarrow\{0,1\}^\lambda\\ \mathcal{A},\mathsf{R}}}[\mathsf{R}^{f,\mathcal{A}}(f_\lambda(x)) = x] \geq \epsilon(\lambda)$$
for infinitely many $\lambda \in \mathbb{N}$.

- ***Reduction Efficiency:*** *For any $\lambda \in \mathbb{N}$ and $y \in \{0,1\}^\lambda$, $\mathsf{R}^{f,\mathcal{A}}(y)$ makes at most $q(\lambda)$ queries to the oracles f and $\mathcal{A}$, and for every $\mathcal{A}$-query $(1^n,k)$ made by $\mathsf{R}(y)$, $\mathsf{Hash}^f(k,\cdot)$ makes at most $q(\lambda)$ queries to f on any input.*

Our impossibility result captures CIH for any relation class where the events $(x,y) \in R$ for any (x,y) and a random relation R are 3-wise independent and all occur with the same probability p. Such a class is said to be *3-wise universal*. More generally, we define t-wise universality as follows.

Definition 3.2 (Universal Relations)**.** *Let $t \in \mathbb{N}$ and $p := p(n) > 0$. We say that a relation class $\mathcal{R}$ is* t-wise p-universal *if there exists a distribution over relations in $\mathcal{R}$ (which we ambiguously denote by $\mathcal{R}$) such that for any $n \in \mathbb{N}$, any $t' \leq t$, and any $(x_1,y_1),\dots,(x_{t'},y_{t'}) \in \{0,1\}^* \times \{0,1\}^n$,*
$$\Pr_{R\leftarrow\mathcal{R}}[(x_1,y_1) \in R \wedge \cdots \wedge (x_{t'},y_{t'}) \in R] = p^{t'}(n)$$

We hereby note a useful property satisfied by universal classes of relations.

Proposition 3.3. *Let $\mathcal{R}$ be a t-wise p-universal relation class. For any $(z,w) \in \{0,1\}^* \times \{0,1\}^*$, the relation class $\mathcal{R}_{z,w} = \{R \in \mathcal{R} \mid (z,w) \in R\}$ is a $(t-1)$-wise p-universal relation class.*

We now formally state our separation result: any fully-black-box reduction from CIH to OWP is either highly inefficient or incurs an almost total security loss.

Theorem 3.4 (Black-box Impossibility of CIH from OWP)**.** *Let $\mathcal{R}$ be a 3-wise p-universal class of relations for $p(n) \geq 2^{-n}$. Let $\mathsf{CIH} = (\mathsf{Gen},\mathsf{Hash})$ be a (q,ϵ)-fully black-box construction of CIH for $\mathcal{R}$ from OWP where, for any $k \in \{0,1\}^{\kappa(n)}$, $\mathsf{Hash}^f(k,\cdot)$ makes at most $2^{m-n}/n^2$ queries to f. Then, either*

1. *$q(\lambda) > O(2^{\lambda/9})$, or*
2. *$\epsilon(\lambda) \leq O(2^{-\lambda/9})$.*

Observe that the separation captures any CIH satisfying a reasonable trade-off between compression and query-efficiency. For instance, it applies to any polynomially-efficient CIH with input length $m > n+\omega(\log n)$.

3.1 Detectability of Correlation-Heavy Images

Our design of the separation oracles, specifically the correlation-finding oracle, is based at its heart on the observation that if an f-image y is likely to be *heavy* under correlations, i.e. if $f^{-1}(y)$ is frequently called by the circuit C (any CIH candidate) among the set of correlations between C and a random relation R, then y is *detectable* by random evaluations of C, that is, $f^{-1}(y)$ is frequently called among *all* inputs. Speaking ahead, heavy images are images that a correlation-finding oracle may be useful to invert. Therefore, we would like to design a correlation-finder that avoids outputting such images unless they are anyway detectable – in which case an adversary is able to invert also without the help of the correlation-finder.

First, we define the *set of correlations* between a circuit and a relation, and use Chebyshev's inequality to bound the concentration of its size, for a random 2-wise universal relation.

Definition 3.5 (Set of Correlations)**.** *Let $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ be a Ψ-aided circuit. Fix a key $k \in \{0,1\}^\kappa$ and let R be a relation. The* set of (R, C_k^Ψ)-correlations *is defined as*

$$\mathsf{Corr}_{R,C_k} = \{z \mid (z, C^\Psi(k,z)) \in R\}.$$

We sometimes omit R, C and k from notation when clear by context.

Proposition 3.6. *Let $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ be a Ψ-aided circuit and let $\mathcal{R}$ be a 2-wise p-universal class of relations. Then, for any fixed oracle Ψ and key $k \in \{0,1\}^\kappa$, any set of inputs $S \subseteq \{0,1\}^m$, and any $\alpha > 0$, it holds that*

$$\Pr_R[||\mathsf{Corr} \cap S| - p|S|| > \sqrt{p|S|/\alpha}] < \alpha$$

where $R \leftarrow \mathcal{R}$ is sampled from the 2-wise universal distribution.

Proof. For any $z \in S$, let $\mathsf{Corr}(z)$ be the predicate for $z \in \mathsf{Corr}$. Then, we have

$$\mathbb{E}_R[|\mathsf{Corr} \cap S|] = \sum_{z \in S} \mathbb{E}_R[\mathsf{Corr}(z)] = p|S|$$

and, from pairwise-independence of $\{\mathsf{Corr}(z)\}_{z \in S}$ (due to 2-wise universality of $\mathcal{R}$),

$$\mathbf{Var}_R[|\mathsf{Corr} \cap S|] = \sum_{z \in S} \mathbf{Var}_R[\mathsf{Corr}(z)] = (p - p^2)|S| < p|S|.$$

Thus, the proposition follows from Chebyshev's inequality. □

Next, we define detectable images and heavy images.

Definition 3.7 (Detectable Images)**.** *Let $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ be a Ψ-aided circuit. Fix a permutation Ψ and a key $k \in \{0,1\}^\kappa$. We say that a Ψ-image $y \in \{0,1\}^*$ is* δ-detectable under C_k^Ψ*, for $\delta > 0$, if*

$$\Pr_{z \leftarrow \{0,1\}^m}[C^\Psi(k,z) \xrightarrow{\Psi} \Psi^{-1}(y)] > \delta.$$

Definition 3.8 (Heavy Images)**.** *Let $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ be a Ψ-aided circuit. Fix a permutation Ψ, a key $k \in \{0,1\}^\kappa$, and a relation R. We say that a Ψ-image $y \in \{0,1\}^*$ is* ϵ-heavy among (R, C_k^Ψ)-correlations *(or simply* ϵ-heavy *when unambiguous), for $\epsilon > 0$, if*

$$\Pr_{z \leftarrow \mathsf{Corr}}[C^\Psi(k,z) \xrightarrow{\Psi} \Psi^{-1}(y)] > \epsilon.$$

We now proceed to prove that if an image is not detectable, then it is unlikely to be heavy w.r.t. a random 2-wise universal relation.

Lemma 3.9. *Let $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ be a Ψ-aided circuit and let $\mathcal{R}$ be a 2-wise p-universal class of relations. Let $\epsilon, \delta > 0$ be such that $\epsilon > 2(\delta + \sqrt{\delta})$. For any permutation Ψ, any $k \in \{0,1\}^\kappa$, and any $y \in \{0,1\}^*$, if y is not δ-detectable under C_k^Ψ then*

$$\Pr_{R \leftarrow \mathcal{R}}[y \textit{ is } \epsilon\textit{-heavy among } (R, C_k^\Psi)\textit{-correlations}] < \frac{17}{p2^m}$$

Proof. Fix some permutation Ψ, $k \in \{0,1\}^\kappa$ and $y \in \{0,1\}^*$, and assume that y is not δ-detectable, i.e.

$$\Pr_{z \leftarrow \{0,1\}^m}[C^\Psi(k,z) \overset{\Psi}{\to} \Psi^{-1}(y)] \leq \delta.$$

.

Consider $R \leftarrow \mathcal{R}$ sampled at random and the corresponding random set of (R, C_k^Ψ)-correlations Corr. Denote

$$\mathsf{Hits}_y = \{z \mid C^\Psi(k,z) \overset{\Psi}{\to} \Psi^{-1}(y)\}$$

and observe that $|\mathsf{Hits}_y| < \delta 2^m$ and, further,

$$\Pr_{z \leftarrow \mathsf{Corr}}[C^\Psi(k,z) \overset{\Psi}{\to} \Psi^{-1}(y)] = |\mathsf{Hits}_y \cap \mathsf{Corr}|/|\mathsf{Corr}|.$$

First, we use Proposition 3.6 with $S = \{0,1\}^m$ to obtain the following bound

$$\Pr[|\mathsf{Corr}| < p2^{m-1}] < \frac{16}{p2^m}. \tag{2}$$

From (2) and the definition of an ϵ-heavy image, we have that

$$\Pr_R[y \text{ is } \epsilon\text{-heavy}] < \Pr_R[|\mathsf{Hits}_y \cap \mathsf{Corr}| > \epsilon p2^{m-1}] + \frac{16}{p2^m}. \tag{3}$$

We use Proposition 3.6 again, this time with $S = \mathsf{Hits}_y$, and get that

$$\Pr[|\mathsf{Hits}_y \cap \mathsf{Corr}| > \epsilon p2^{m-1}] < \frac{\delta p2^m}{((\epsilon/2 - \delta)p2^m)^2} = \frac{\delta}{(\epsilon/2-\delta)^2 p2^m} < \frac{1}{p2^m}$$

and conclude $\Pr_R[y \text{ is } \epsilon\text{-heavy}] < \frac{17}{p2^m}$. □

3.2 The Correlation-Finding Oracle

We now define our correlation-finding oracle (more accurately, a distribution over oracles). The oracle is defined with respect to a relation R. On inputs an oracle-aided circuit C^Ψ and a key k, the oracle samples a uniformly random correlation z such that $(z, C^\Psi(k,z)) \in R$ and outputs it, unless $C^\Psi(k,z)$ calls the oracle Ψ (which models a OWP) on an input which is heavy among correlations w.r.t. R but is not detectable.

Definition 3.10 (Correlation-Finding Oracle)**.** *Let $\delta > 0$ be an implicit parameter. Let R be a relation and let $\Psi : \{0,1\}^* \to \{0,1\}^*$ be a permutation oracle. We define the corresponding correlation-finding oracle $\mathsf{CorrFinder}_R^\Psi(\cdot,\cdot)$ as follows.*

- ***Randomness:*** *For any Ψ-aided circuit $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ and any key $k \in \{0,1\}^\kappa$, the oracle assigns a correlation $z_{C,k} \leftarrow \mathsf{Corr}$ chosen uniformly at random from the set of (R, C_k^Ψ)-correlations (if $\mathsf{Corr} = \emptyset$, set $z_{C,k} = \perp$).*
- ***Query:*** *On inputs $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ and $k \in \{0,1\}^\kappa$, the oracle $\mathsf{CorrFinder}_R^\Psi$ outputs $z_{C,k}$ if for every $y \in \{0,1\}^\lambda$ such that $C_k^\Psi(z_{C,k}) \overset{\Psi}{\to} \Psi^{-1}(y)$, y is either δ-detectable or not $3\sqrt{\delta}$-heavy and, otherwise, outputs $\perp$.*

By relying on Lemma 3.9, we show that the oracle $\mathsf{CorrFinder}$, defined w.r.t. a random 3-wise universal relation, finds a correlation with high probability when given a circuit and a random key. Later on, we use this fact to show that any CIH candidate can be broken by an instantiation of $\mathsf{CorrFinder}$ (w.r.t. some relation in a 3-wise universal class).

Lemma 3.11. *Let $\mathcal{R}$ be a class of 3-wise p-universal relations, and let Ψ be a permutation oracle. For any Ψ-aided circuit $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$, that makes at most q queries on any input, and any key $k \in \{0,1\}^\kappa$,*

$$\Pr_{\substack{R \leftarrow \mathcal{R} \\ \mathsf{CorrFinder}^\Psi_R}} [\mathsf{CorrFinder}^\Psi_R(C,k) = \bot] < O(q/p2^m).$$

Proof. Fix a Ψ-aided circuit C^Ψ and let $R \leftarrow \mathcal{R}$ be sampled at random from the 3-wise universal distribution.

Recall that $\mathsf{CorrFinder}$ outputs $\bot$ on inputs C and k only if $\mathsf{Corr}_{R,C_k} = \emptyset$ or $C^\Psi(k, z_{C,k})$ makes a query which is $3\sqrt{\delta}$-heavy but not δ-detectable.

In fact, we consider a slightly different experiment. Instead of sampling a relation $R \leftarrow \mathcal{R}$ and then an oracle $\mathsf{CorrFinder}$, which consists of a random $z_{C,k} \leftarrow \mathsf{Corr}$, we consider an experiment where we first sample $R' \leftarrow \mathcal{R}$, then $z_{C,k} \leftarrow \mathsf{Corr}'$ (where Corr' is the set of (R', C^Ψ_k)-correlations) and then sample $R \leftarrow \mathcal{R}_{z_{C,k},w}$, where $w = C^\Psi(z_{C,k})$ and $\mathcal{R}_{z,w}$ is the conditioned distribution as defined in Proposition 3.3. It is evident that the joint distribution of R and $z_{C,k}$ in the new experiment is identical to that in the original correlation-intractability experiment, and therefore, the outcome of the experiment, which is determined solely by the fixed key k, and R and $z_{C,k}$, is identically distributed.

First, we show that, with high probability, there exists at least one (R', C_k)-correlation. This may be implied using Proposition 3.6 with $S = \{0,1\}^m$ to get

$$\Pr_{k,R'}[|\mathsf{Corr}'| < 1] < \frac{1}{p2^{m-1}}. \tag{4}$$

We now proceed to bound the probability that $\mathsf{CorrFinder}$ outputs $\bot$.

For any $k \in \{0,1\}^\kappa$ and relation R, we denote by $\mathsf{Bad}_{R,k}$ the set of all $y \in \{0,1\}^*$ such that y is not δ-detectable under C^ψ_k and is $3\sqrt{\delta}$-heavy among (R, C^ψ_k)-correlations. Due to the observations above, we may bound the probability that $\mathsf{CorrFinder}$ aborts in the original experiment as follows

$$\begin{aligned}
&\Pr_{R,\mathsf{CorrFinder}}[\mathsf{CorrFinder}(C,k) = \bot] \\
&= \Pr_{R' \leftarrow \mathcal{R}}[\mathsf{Corr}' = \emptyset] + \Pr_{\substack{R' \leftarrow \mathcal{R} \\ z_{C,k} \leftarrow \mathsf{Corr}' \\ R \leftarrow \mathcal{R}_{z_{C,k},w}}} [\exists y : C^\Psi(k, z_{C,k}) \xrightarrow{\Psi} \Psi^{-1}(y) \wedge y \in \mathsf{Bad}_{R,k} \mid \mathsf{Corr}' \neq \emptyset] \\
&< \frac{1}{p2^{m-1}} + \max_{z \in \{0,1\}^m} \Pr_{R \leftarrow \mathcal{R}_{z,w}} [\exists y : C^\Psi(k,z) \xrightarrow{\Psi} \Psi^{-1}(y) \wedge y \in \mathsf{Bad}_{R,k}] \\
&\leq \frac{1}{p2^{m-1}} + q \cdot \max_{y \in \{0,1\}^*} \Pr_{R \leftarrow \mathcal{R}_{z,w}} [y \in \mathsf{Bad}_{R,k}]
\end{aligned}$$

Using Lemma 3.9, and since $\mathcal{R}_{z,w}$ is 2-wise p-universal for any z, w (by Proposition 3.3), we conclude that

$$\Pr_{R,\mathsf{CorrFinder}}[\mathsf{CorrFinder}(C,k) = \bot] < \frac{17q+2}{p2^m}.$$

□

3.3 Finding Correlations using $\mathsf{CorrFinder}$

We now build on the above to prove the first step of the separation: for any CIH candidate, there exists a relation R such that $\mathsf{CorrFinder}_R$ breaks the correlation intractability of the hash for R.

We say that an oracle-aided adversary $\mathcal{A}^{\Psi,\mathsf{CorrFinder}^\Psi}$ is a (q, q', q'')-*adversary* if, on any input $(1^n, k)$, it makes at most $q(n)$ queries to the oracle Ψ and at most $q'(n)$ queries to $\mathsf{CorrFinder}$ and, for every $\mathsf{CorrFinder}$-query (C, k) that $\mathcal{A}$ makes, $C^\Psi(k, \cdot)$ makes at most $q''(n)$ queries to Ψ on any input.

Lemma 3.12. *Let $\mathcal{R}$ be a class of 3-wise p-universal relations for $p(n) \geq 2^{-n}$, and fix a permutation oracle Ψ. Let $\kappa := \kappa(n)$ and $m := m(n)$, and let $C^\Psi = \{C_n^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$ be any Ψ-aided circuit ensemble where, for any $n \in \mathbb{N}$, C_n^Ψ makes at most $q(n) < 2^{m-n}/n^2$ queries on any input. Then, there exists a relation $R \in \mathcal{R}$ and a $(0, 1, q)$-adversary $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}$ such that*

$$\mathbf{Adv}^{\mathbf{ci}}_{C,\mathcal{A},R}(n) := \Pr_{\substack{k \leftarrow \mathsf{Gen}(1^n) \\ \mathsf{CorrFinder}^\Psi_R \\ z \leftarrow \mathcal{A}^{\psi,\mathsf{CorrFinder}}(1^n,k)}} [(z, C^\psi(k, z)) \in R] \geq \frac{1}{2}$$

for infinitely many $n \in \mathbb{N}$.

Proof. Fix any circuit ensemble C^Ψ. We consider a random $R \leftarrow \mathcal{R}$, and show that the claim holds for R with non-zero probability. We construct a straight-forward adversary $\mathcal{A}$ that attacks the correlation intractability of C^Ψ for R by calling $\mathsf{CorrFinder}^\Psi_R(C, k)$ and returning the answer.

We show that the construction indeed breaks the correlation intractability of C with high probability for infinitely many $n \in N$. Notice that, by construction of $\mathsf{CorrFinder}$ (see Definition 3.10), the output of $\mathcal{A}$ is necessarily either a correlation $z \in \mathsf{Corr}$ or $\perp$. Thus, we aim to bound the probability of $\mathsf{CorrFinder}(C, k) = \perp$.

Fix some $n \in \mathbb{N}$, and consider the failure probability of $\mathcal{A}$ as a function of R,

$$B_n(R) = \Pr_{\substack{k \leftarrow \{0,1\}^{\kappa(n)} \\ \mathsf{CorrFinder}}} [\mathsf{CorrFinder}^\Psi_R(C, k) = \perp],$$

and its value in expectation

$$\mathbb{E}_R[B_n(R)] = \Pr_{R,k,\mathsf{CorrFinder}} [\mathsf{CorrFinder}(C, k) = \perp].$$

Observe that, due to Lemma 3.11, it holds

$$\mathbb{E}_R[B_n(R)] < \max_k \Pr_{R,\mathsf{CorrFinder}} [\mathsf{CorrFinder}(C, k) = \perp] < O(q/p2^m)$$

and, therefore, through a standard averaging argument, we get

$$\Pr_R \left[B_n(R) > \frac{1}{2} \right] < O(q/p2^m).$$

Since $\sum_n q/p2^m$ converges, we may apply the Borel-Cantelli lemma to conclude that the event $B_n(R) \leq \frac{1}{2}$ holds for infinitely many $n \in \mathbb{N}$ except for measure 0 of relations $R \leftarrow \mathcal{R}$. Hence, there exists a relation $R \in \mathcal{R}$ such that

$$\mathbf{Adv}^{\mathbf{ci}}_{C,\mathcal{A},R}(n) \geq \frac{1}{2}$$

for infinitely many $n \in \mathbb{N}$. □

3.4 One-wayness against Smooth Adversaries

Having shown that any CIH candidate is broken by an instance of $\mathsf{CorrFinder}$, in this section we prove that there exists a one-way permutation under any instance of $\mathsf{CorrFinder}$, hence derive a separation.

Specifically, letting f be a random permutation, our goal is to show that an adversary $\mathcal{A}^{f,\mathsf{CorrFinder}^f}$ fails to invert f on a given random image y. First, we consider a special case, where the adversary $\mathcal{A}$ is *smooth* and *canonical*. It is easy to see that any successful adversary can be assumed to be canonical w.l.o.g. and, further, we show later in Section 3.5 how to transform any adversary to a smooth adversary, while maintaining the inversion success probability.

Definition 3.13 (Smooth Queries). *Fix an oracle Ψ and let $\delta > 0$.*

A $\mathsf{CorrFinder}$*-query, which consists of a Ψ-aided circuit $C^\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ and a key $k \in \{0,1\}^\kappa$, is (Ψ, δ)-*smooth *at $x \in \{0,1\}^*$ if $\Psi(x)$ is not δ-detectable under C_k^Ψ.*

*We say that (C^Ψ, k) is (Ψ, δ)-*smooth *at a set $X \subset \{0,1\}^*$ if (C^Ψ, k) is smooth at any $x \in X$.*

Definition 3.14 (Smooth Adversaries). *Fix an oracle Ψ and let $\delta := \delta(\lambda)$ and $\sigma := \sigma(\lambda)$. We say that an oracle-aided adversary $\mathcal{A}$ is (Ψ, δ, σ)-*smooth *if for any fixed oracle $\mathsf{CorrFinder}_R^\Psi$, any $\lambda \in \mathbb{N}$ and $y \in \{0,1\}^\lambda$, and any $i \in \mathbb{N}$, letting a_i be the i^{th} query made by $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(1^\lambda, y)$ to* $\mathsf{CorrFinder}$ *and X_i be the set of all Ψ-queries made by $\mathcal{A}$ before a_i,*

$$\Pr_{\mathcal{A}(1^\lambda,y)}[a_i \textit{ is } (\Psi,\delta)\textit{-smooth at } \overline{X_i}] > \sigma$$

Definition 3.15 (Canonical Adversaries). *We say that an oracle-aided adversary $\mathcal{A}$ is* canonical *if for any fixed Ψ and $\mathsf{CorrFinder}_R^\Psi$, any $\lambda \in \mathbb{N}$ and any $y \in \{0,1\}^\lambda$, if $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(1^\lambda, y)$ calls Ψ at $\Psi^{-1}(y)$, he never makes any further queries to either oracles.*

The one-wayness proof is based on a hybrids argument that transforms the original inversion experiment against a smooth adversary to an experiment which is statistically impossible to win. The transformation goes through hybrid experiments, where the adversary is given access to a *punctured* correlation-finding oracle, which is defined as follows.

Definition 3.16 (Punctured Correlation-Finding Oracle). *Let R be a relation and Ψ be a permutation oracle. Let $S \subseteq \{0,1\}^*$ be a set of Ψ-inputs. We define the* punctured correlation-finding oracle $\mathsf{CorrFinder}_{R,S}^\Psi(\cdot,\cdot)$ *similarly to the oracle $\mathsf{CorrFinder}_R^\Psi$ (see Definition 3.10) with the exception that the punctured oracle $\mathsf{CorrFinder}_{R,S}^\Psi$ outputs $\perp$ on inputs C^Ψ and k also if there exists $x \in S$ such that $C_k^\Psi(z_{C,k}) \xrightarrow{\Psi} x$.*

In the following, we show that a smooth adversary cannot possibly distinguish between the case where he is given the oracle $\mathsf{CorrFinder}$ and the case where he is given a punctured oracle $\mathsf{CorrFinder}_{\{x\}}$.

Lemma 3.17. *Fix a permutation oracle Ψ and let R be any relation. Let $\sigma := \sigma(\lambda)$, $\delta := \delta(\lambda)$. For any $\lambda \in \mathbb{N}$ and $x \in \{0,1\}^\lambda$, consider the coupling $\mathcal{P}$ of the correlation-finding oracle $\mathsf{CorrFinder}_R^\Psi$ and its punctured analog $\mathsf{CorrFinder}_{R,\{x\}}^\Psi$ that insantiates both oracles with the same randomness $\{z_{C,k}\}$ (see Definitions 3.10 and 3.16). Then, for any canonical (Ψ, δ, σ)-smooth (q, q', q'')-adversary $\mathcal{A}$,*

$$\Pr_{\mathcal{P},\mathcal{A}}[\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(1^\lambda, \Psi(x)) \neq \mathcal{A}^{\Psi,\mathsf{CorrFinder}_{\{x\}}}(1^\lambda, \Psi(x))] \leq (3\sqrt{\delta} + (1-\sigma))q'$$

Proof. Denote by $(a_1, b_1), \ldots, (a_{q+q'}, b_{q+q'})$ and $(a'_1, b'_1), \ldots, (a'_{q+q'}, b'_{q+q'})$ all oracle queries made by $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(1^\lambda, \Psi(x))$ and, respectively, $\mathcal{A}^{\Psi,\mathsf{CorrFinder}_{\{x\}}}(1^\lambda, \Psi(x))$, and their corresponding answers. Then, it holds that

$$\Pr_{\mathcal{P},\mathcal{A}}[\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(\Psi(x)) \neq \mathcal{A}^{\Psi,\mathsf{CorrFinder}_{\{x\}}}(\Psi(x))] \leq \sum_{i=1}^{q+q'} \Pr_{\mathcal{P},\mathcal{A}} \Pr[(a_i, b_i) \neq (a'_i, b'_i) \mid (a_{<i}, b_{<i}) = (a'_{<i}, b'_{<i})]$$

For any $i \in [q+q']$, if a_i is a Ψ-query, then clearly $b_i = b'_i$ assuming $b_{<i} = b'_{<i}$. Otherwise, a_i is a $\mathsf{CorrFinder}$-query of the form $a_i = (C, k)$ and, since $\mathcal{A}$ is canonical, then $x \notin a_{<i}$. From the (Ψ, δ, σ)-smoothness of $\mathcal{A}$, we have that

$$\begin{aligned}
&\Pr_{\mathcal{P},\mathcal{A}}[(a_i, b_i) \neq (a'_i, b'_i) \mid (a_{<i}, b_{<i}) = (a'_{<i}, b'_{<i})] \\
&\leq \Pr_{\mathcal{P},\mathcal{A}}[(a_i, b_i) \neq (a'_i, b'_i), a_i \text{ is } (\Psi, \delta)\text{-smooth at } x \mid (a_{<i}, b_{<i}) = (a'_{<i}, b'_{<i})] \\
&\ + \Pr_{\mathcal{P},\mathcal{A}}[a_i \text{ is not } (\Psi, \delta)\text{-smooth at } \overline{a_{<i}}] \\
&\leq \Pr_{\mathcal{P},\mathcal{A}}[(a_i, b_i) \neq (a'_i, b'_i) \mid (a_{<i}, b_{<i}) = (a'_{<i}, b'_{<i}), a_i \text{ is } (\Psi, \delta)\text{-smooth at } x] + (1 - \sigma)
\end{aligned}$$

We now bound the probability that $(a_i, b_i) \neq (a'_i, b'_i)$ assuming $\mathcal{A}$ has produced a query $a_i = (C, k)$ which is smooth at x. We consider two cases. If $\Psi(x)$ is $3\sqrt{\delta}$-heavy among (R, C_k^{Ψ})-correlations, then observe that, since $\Psi(x)$ is not δ-detectable (due to smoothness), the oracles $\mathsf{CorrFinder}$ and $\mathsf{CorrFinder}_{\{x\}}$ are equivalent (as they both abort when $C(k, z_{C,k}) \xrightarrow{\Psi} x$) and therefore $(a_i, b_i) = (a'_i, b'_i)$ always. Otherwise, if $\Psi(x)$ is not $3\sqrt{\delta}$-heavy, then the two oracles behave differently only when $C(k, z_{C,k}) \xrightarrow{\Psi} x$ and, since $\Psi(x)$ is not heavy, this occurs with probability at most $3\sqrt{\delta}$ over the choice of $z_{C,k} \leftarrow \mathsf{Corr}$. This completes the proof. □

We are now prepared to prove that a smooth adversary is unable to invert a random permutation f with noticeable probability and imply that f is one-way under the correlation-finding oracle. We first define the oracle f.

Definition 3.18 (The Oracle f). *The oracle $f = \{f_\lambda\}$ takes as input $x \in \{0,1\}^*$ and outputs $f_{|x|}(x)$, where, for any $\lambda \in \mathbb{N}$, $f_\lambda : \{0,1\}^\lambda \to \{0,1\}^\lambda$ is a random permutation.*

Lemma 3.19. *Let $\sigma := \sigma(\lambda)$, $\delta := \delta(\lambda)$. For any relation R, and any (Ψ, δ, σ)-smooth (q, q', q'')-adversary $\mathcal{A}$, and any $\lambda \in \mathbb{N}$,*

$$\Pr_{f,\mathsf{CorrFinder},x}[\mathcal{A}^{f,\mathsf{CorrFinder}^f_R}(1^\lambda, f(x)) = x] < O\left(2^{-\lambda}(q + q'q'') + (\sqrt{\delta} + (1-\sigma))q'\right)$$

Proof. First, without loss of generality, we assume that $\mathcal{A}$ is canonical (the trivial transformation preserves smoothness and may only increase inversion probability). We bound the inversion success probability of such a canonical and smooth $\mathcal{A}$ through a series of hybrid experiments, as follows.

– Hybrid_1: This is the original inversion experiment:

   1. Sample a pre-image $x \leftarrow \{0,1\}^n$ and oracles f and $\mathsf{CorrFinder}^f_R$.
   2. $x' \leftarrow \mathcal{A}^{f,\mathsf{CorrFinder}}(f(x))$.
   3. $\mathcal{A}$ wins if $x' = x$.

– Hybrid_2: Similar to Hybrid_1, except $\mathcal{A}$ is given access to a punctured oracle $\mathsf{CorrFinder}^f_{R,\{x\}}$.

By using a coupling-based argument, since $\mathcal{A}$ is smooth and canonical and relying on Lemma 3.17, we may bound the statistical distance between Hybrid_1 and Hybrid_2, and get that

$$|\Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_1] - \Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_2]| \leq (3\sqrt{\delta} + (1-\sigma))q'. \tag{5}$$

Denote by $f_{\hat{x}\to\hat{y}}$ the oracle which is identical to f except that $\hat{x}$ is mapped to $\hat{y}$. We define the next hybrid as follows

– Hybrid_3: Similar to Hybrid_2, except now we sample an additional pre-image $\hat{x} \leftarrow \{0,1\}^n$, and $\mathcal{A}$ is given access to $f_{\hat{x}\to f(x)}$ and a punctured oracle $\mathsf{CorrFinder}^f_{R,\{x,\hat{x}\}}$.

Notice that the output of $\mathcal{A}$ in Hybrid_3 is identical to its output in Hybrid_2 unless $\hat{x}$ appears in its transcript, either in a direct f-query or in an f-query made by $C^f(k, z)$ for some $\mathsf{CorrFinder}$-query (C, k) with an answer z. Since $\hat{x}$ is sampled uniformly at random, independently of $\mathcal{A}$'s transcript in Hybrid_2, then the probability of this event may be bound by $2^{-\lambda}(q + q'q'')$ and, hence,

$$|\Pr[\mathsf{S} \text{ wins in } \mathsf{Hybrid}_2] - \Pr[\mathsf{S} \text{ wins in } \mathsf{Hybrid}_3]| \le 2^{-\lambda}(q + q'q''). \tag{6}$$

Next, notice that in Hybrid_3, the pre-images x and $\hat{x}$ are completely symmetric and, therefore, we may switch to a hybrid where $\mathcal{A}$'s goal is to find $\hat{x}$, i.e.

– Hybrid_4:

1. Sample pre-images $x, \hat{x} \in \{0,1\}^n$ and oracles f and $\mathsf{CorrFinder}^f_{R,\{x,\hat{x}\}}$.
2. $x' \leftarrow \mathcal{A}^{f_{\hat{x}\to f(x)}, \mathsf{CorrFinder}^f_{R,\{x,\hat{x}\}}}(f(x))$.
3. $\mathcal{A}$ wins if $x' = \hat{x}$.

and obtain

$$\Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_3] = \Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_4]. \tag{7}$$

To complete the proof, we perform the same steps that took us from Hybrid_1 to Hybrid_3. This time we start with Hybrid_4 and go "backwards", while sticking with the same winning condition:

– Hybrid_5: Similar to Hybrid_4, except $\mathcal{A}$ is given access to f and $\mathsf{CorrFinder}^f_{R,\{x\}}$.

– Hybrid_6: Similar to Hybrid_5, except $\mathcal{A}$ is given access to f and $\mathsf{CorrFinder}^f_{R}$.

Based on the same arguments from above, we may bound

$$|\Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_4] - \Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_6]| \le 2^{-\lambda}(q + q'q'') + (3\sqrt{\delta} + (1-\sigma))q'. \tag{8}$$

Lastly, observe that in Hybrid_6, $\hat{x}$ is independent from the output of $\mathcal{A}$ and, therefore,

$$\Pr[\mathcal{A} \text{ wins in } \mathsf{Hybrid}_6] \le 2^{-\lambda}. \tag{9}$$

The proof is complete by combining (5), (6), (7), (8) and (9). □

3.5 Smoothening any Adversary

Lastly, we confirm the generality of smooth adversaries via the following smoothening lemma.

Lemma 3.20 (The Smoothening Lemma). *For any (q, q', q'')-query algorithm $\mathcal{A}$ and any $\beta := \beta(\lambda)$, there exists a $(q + \beta q', q', q'')$-adversary S such that the following two properties hold:*

1. ***Correctness:*** *for any relation R and fixed oracles Ψ and $\mathsf{CorrFinder} := \mathsf{CorrFinder}^{\Psi}_{R}$, $\mathsf{S}^{\Psi,\mathsf{CorrFinder}}$ perfectly simulates $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}$ on any input.*
2. ***Smoothness:*** *S is $(\Psi, \delta, 1 - 2^{-\delta\beta + \log(q''/\delta)})$-smooth.*

Proof. For any adversary $\mathcal{A}$, we construct the simulator S as follows. On inputs 1^λ and y, $\mathsf{S}^{\Psi,\mathsf{CorrFinder}}$ runs $\mathcal{A}^{\Psi,\mathsf{CorrFinder}}(1^\lambda, y)$ and, whenever $\mathcal{A}$ calls $\mathsf{CorrFinder}$ with input (C^Ψ, k), S evaluates $C^\Psi(k, \cdot)$ on β uniformly random inputs and only then calls $\mathsf{CorrFinder}(C, k)$, forwards the answer to $\mathcal{A}$, and proceeds with the simulation.

Correctness and complexity of S are straight-forward. For smoothness, fix an oracle $\mathsf{CorrFinder}$ and $y \in \{0,1\}^\lambda$. We bound the probability that $\mathsf{S}(1^\lambda, y)^{\Psi,\mathsf{CorrFinder}}$ makes a query $a = (C, k)$ which is not (Ψ, δ)-smooth at a Ψ-query which was not made by S before a. Observe that this occurs if there exists x s.t. $\Psi(x)$ is δ-detectable under C^Ψ_k but x is not queried by S beforehand. However, notice if $\Psi(x)$ is detectable, x will be queried by S during the β random evaluations of $C^\Psi(k, \cdot)$ except with probability at most $(1-\delta)^\beta \le 2^{-\delta\beta}$. A simple counting argument shows that, for any such C, there exist at most q''/δ such δ-detectable images. Therefore, by applying a union bound over all detectable images under C^Ψ_k, we get that S queries all of them with probability at least $1 - 2^{-\delta\beta} \cdot q''/\delta$. □

3.6 Proof of Theorem 3.4

Fix a candidate construction of correlation intractable hash from OWP, $\mathsf{CIH} = (\mathsf{Gen}, \mathsf{Hash})$. Let $f = \{f_\lambda\}$ be sampled at random as in Definition 3.18 and let R and $\mathcal{A}$ be the relation and, resp., the $(0, 1, q)$-adversary from Lemma 3.12 that satisfy

$$\mathbf{Adv}^{\mathbf{ci}}_{C,\mathcal{A},R}(n) \geq \frac{1}{2}$$

for infinitely many $n \in \mathbb{N}$, with respect to the circuit ensemble $C^f = \{C^f_n\}$ where, for any $n \in \mathbb{N}$, C^f_n computes $\mathsf{Hash}^f(\cdot, \cdot)$ on inputs of length $\kappa(n)$ and $m(n)$, respectively.

To prove Theorem 3.4, it is sufficient to show that any efficient reduction $\mathsf{R}^{f,\mathcal{A}}(1^\lambda, \cdot)$ fails in inverting f_λ except with low probability. Observe that if R is efficient and $\mathcal{A}$ is a $(0, 1, q)$-adversary, then $\mathsf{R}^{f,\mathcal{A}}$ may be emulated by a (q, q, q)-adversary $\mathcal{B}^{f,\mathsf{CorrFinder}}$. By applying Lemma 3.20 (with $\delta(\lambda) = 2^{-2\lambda/3}$ and $\beta(\lambda) = 2^{2\lambda/3}\lambda$), we may bound $\mathcal{B}$'s success probability in inverting f by the success probability of the best $(\Psi, 2^{-2\lambda/3}, 1 - 2^{-\lambda/3}q)$-smooth $(q + 2^{2\lambda/3}\lambda q, q, q)$-adversary. And, thus, due to Lemma 3.19, we have

$$\Pr_{f,x}[\mathsf{R}^{f,\mathcal{A}}(1^\lambda, f(x)) = x] < O\left(2^{-\lambda/3}q^2\right)$$

Hence, if $\epsilon > O(2^{-\lambda/9})$ then $q > O(2^{\lambda/9})$.

4 Separating One-wayness from Correlation Intractable Hash

In this section, we show separation in the opposite direction and prove that correlation intractability does not imply one-way functions in a fully-black-box manner. Namely, we show that there exists no construction of OWF that only makes a fully-black-box use of CIH. We are able to prove the statement even when we consider correlation intractability for *sparse* relations, which is satisfied by a random hash function.

We first define sparse relations, then formalize the notion of fully-black-box construction of OWF from CIH.

Definition 4.1 (Sparsity). *A binary relation R is said to be* $\mu(\cdot)$-sparse *if for any $n \in \mathbb{N}$ and $z \in \{0,1\}^*$,*

$$\Pr_{w \leftarrow \{0,1\}^n}[(z, w) \in R] \leq \mu(n)$$

Definition 4.2 (Fully-Black-box Construction of OWF from CIH). *Let $\mathcal{R}$ be a class of relations. A (q, ϵ)-*fully black-box construction of a One-way Function (OWF) from Correlation Intractable Hash (CIH) *for $\mathcal{R}$, for functions $q := q(n)$ and $\epsilon := \epsilon(n)$, consists of an ensemble of oracle-aided algorithms $\mathsf{F} = \{\mathsf{F}_\lambda\}_{\lambda\in\mathbb{N}}$ and an oracle-aided reduction R, satisfying the following properties:*

- ***Black-box Security Reduction:*** *For any $\kappa := \kappa(n)$, $m := m(n)$ and keyed-hash ensemble $h = \{h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$, there exists a relation $R \in \mathcal{R}$ such that for any probabilistic oracle-aided adversary $\mathcal{A}$, if*

  $$\mathbf{Adv}^{\mathbf{owf}}_{\mathsf{F},\mathcal{A}}(\lambda) := \Pr_{\substack{x \leftarrow \{0,1\}^\lambda \\ y = \mathsf{F}^h_\lambda(x)}}[\mathsf{F}^h_\lambda(\mathcal{A}^h(1^\lambda, y)) = y] \geq \frac{1}{2}$$

  for infinitely many $\lambda \in \mathbb{N}$, then

  $$\Pr_{\substack{k \leftarrow \{0,1\}^\kappa \\ z \leftarrow \mathsf{R}^{h,\mathcal{A}}(1^n, k)}}[(z, h(k, z)) \in R] \geq \epsilon(n)$$

  for infinitely many $n \in \mathbb{N}$.

- ***Reduction Efficiency:*** *For any $\kappa := \kappa(n)$, $m := m(n)$, any keyed-hash ensemble $h = \{h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$, and any oracle-aided adversary $\mathcal{A}$, the reduction $\mathsf{R}^{h,\mathcal{A}}$, on any input $(1^n, k)$, makes at most $q(n)$ queries to h and, for every $\mathcal{A}$-query $(1^\lambda, y)$ made by $\mathsf{R}(1^n, k)$, F^h_λ makes at most $q(n)$ queries to h on any input.*

The following theorem constitutes the main result of this section. We show that any fully-black-box reduction from OWF to correlation intractability, either has high query complexity – as high as a brute-force attack against the underlying CIH – or has success probability almost as low as a trivial attack against a random CIH.

Theorem 4.3 (Black-box Impossibility of OWF from CIH)**.** *Let F be a (q, ϵ)-fully black-box construction of OWF from CIH for $\mu(\cdot)$-sparse relations. Then, either*

1. $q(n) \geq 1/6\mu(n)$*, or*
2. $\epsilon(n) < 6q^2(n) \cdot \mu(n)$.

4.1 The Inversion Oracle

We start by introducing the inversion oracle w.r.t. a relation R. On inputs an oracle-aided circuit $C^\Psi : \{0,1\}^\lambda \to \{0,1\}^\ell$ and an image $y \in \{0,1\}^\ell$, the inversion oracle outputs a random pre-image $x \leftarrow C^{-1}(y)$ such that $C^\Psi(x)$ does not call Ψ at a correlation (w.r.t. R).

Definition 4.4 (The Inversion Oracle)**.** *Fix an oracle $\Psi : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ and a $\mu(\cdot)$-sparse relation R. For any $q \in \mathbb{N}$, denote by Corr_q the set of all correlations between R and Ψ of length n such that $\mu(n) < 1/6q$, i.e.*

$$\mathsf{Corr}_q^\Psi = \{(k,z) \in \{0,1\}^{\kappa(n)} \times \{0,1\}^{m(n)} \mid \mu(n) < 1/6q, (z, \Psi(k,z)) \in R\}.$$

We define the inversion oracle w.r.t. R, $\mathsf{Inv}_R^\Psi(\cdot,\cdot)$, as follows.

- ***Randomness:*** *For any oracle-aided circuit $C^\Psi : \{0,1\}^\lambda \to \{0,1\}^\ell$ and $y \in \{0,1\}^\ell$, the oracle assigns a uniformly random $x_{C,y}$ sampled from the pre-image of y under C^Ψ, on which C^Ψ does not make a correlation query in Corr_q, where q is a tight upper bound on the query complexity of C on any input, i.e.*
$$x_{C,y} \leftarrow \{x \mid C^\Psi(x) = y \ \wedge \ C^\Psi(x) \overset{\Psi}{\not\rightharpoonup} \mathsf{Corr}_q\}$$
- ***Query:*** *On inputs an oracle-aided circuit C and an image y, the oracle outputs $x_{C,y}$.*

We next define the random oracle h, which models a correlation-intractable hash in our separation proof.

Definition 4.5 (The Oracle h)**.** *Let $\kappa := \kappa(n)$ and $m := m(n)$ be implicit length parameters. The oracle $h = \{h_n\}$ takes as inputs a key $k \in \{0,1\}^\kappa$ and an input $z \in \{0,1\}^m$ and outputs $h_n(k,z)$, where, for any $n \in \mathbb{N}$, $h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n$ is a random hash function.*

In the following lemma, we show that any black-box OWF construction from the oracle h is invertible, with high probability, using the inversion oracle defined above.

Lemma 4.6. *Let h be sampled as in Definition 4.5, let $\mathsf{F}^h = \{\mathsf{F}_\lambda^h\}_{\lambda \in \mathbb{N}}$ be an ensemble of oracle-aided algorithms and let R be a $\mu(\cdot)$-sparse relation. Then, there exists a $(0,1,q)$-adversary $\mathcal{A}^{h,\mathsf{Inv}_R^h}$, such that*

$$\Pr_{h,\mathsf{Inv},y}[\mathsf{F}_\lambda^h(\mathcal{A}^{h,\mathsf{Inv}}(1^\lambda, y)) \neq y] < \frac{1}{6}$$

for all $\lambda \in \mathbb{N}$, where $y = \mathsf{F}_\lambda^\Psi(x)$ for a uniform $x \leftarrow \{0,1\}^\lambda$.

Proof. Consider the straight-forward adversary which, on input $(1^\lambda, y)$, outputs $x \leftarrow \mathsf{Inv}_{\mathsf{F},R}^h(\mathsf{F}_\lambda, y)$ (where F_λ is represented by a circuit). From the definition of the inversion oracle, the only case where the inversion oracle may fail to invert F_λ^h is when the set $\{x \mid \mathsf{F}_\lambda(x) = y \wedge \mathsf{F}_\lambda^h \overset{h}{\not\rightharpoonup} \mathsf{Corr}_q\}$ is empty, where q is the query complexity of F_λ. Since y is sampled as the image of a random input x, we have

$$\Pr_{h,\mathsf{Inv},y}[\mathsf{F}_\lambda(\mathcal{A}^{h,\mathsf{Inv}}(1^\lambda, y)) \neq y] = \Pr_{h,\mathsf{Inv},y}[\{x \mid \mathsf{F}_\lambda(x) = y \wedge \mathsf{F}_\lambda \overset{h}{\not\rightharpoonup} \mathsf{Corr}_q\} = \emptyset] \leq \Pr_{h,x}[\mathsf{F}_\lambda^h(x) \overset{h}{\rightharpoonup} \mathsf{Corr}_q]$$

We bound the above probability as follows. For any fixed x, and any query $(k, z) \in \{0,1\}^{\kappa(n)} \times \{0,1\}^{m(n)}$ that $\mathsf{F}^h_\lambda(x)$ makes for n such that $\mu(n) < 1/6q$, it holds, from sparseness of R, that $(z, h(k,z)) \in R$ with probability at most $1/6q$ over a random h. By applying a union bound over all queries made by $\mathsf{F}_\lambda(x)$, we obtain

$$\Pr_{h,x}[\mathsf{F}_\lambda(x) \overset{\psi}{\rightharpoondown} \mathsf{Corr}_q] < \frac{1}{6}.$$

□

4.2 Correlation Intractability under Inv

We have shown that, under the inversion oracle, any candidate OWF is broken with high probability over the choice of h. Next, we prove that a random oracle h is correlation intractable with high probability. Looking ahead, we will eventually use these two statements to imply the existence of a fixed function h which is correlation intractable under the inversion oracle yet, using which, the OWF candidate is broken under the oracle.

To prove the correlation-intractability of h, we show that the correlation-intractability experiment, where the adversary is challenged to find a correlation in h w.r.t. a relation R, given access to h and the inverter, is in fact indistinguishable from the experiment where the adversary is challenged to find a correlation in a statistically-correlation-intractable random hash, i.e. a random hash function that does not contain any correlations w.r.t. R. Such an experiment is clearly statistically hard to win.

First, we define the statistically-CI oracle $h_{R,n}$, then prove that it is indistinguishable from h under the inversion oracle.

Definition 4.7 (The stat-CI Oracle). *Let $\kappa := \kappa(n)$ and $m := m(n)$ be length parameters and let R be a relation. The* statistically correlation-intractable oracle w.r.t. R*, $h_{R,n}(\cdot,\cdot)$, takes as input $(k,z) \in \{0,1\}^{\kappa(n')} \times \{0,1\}^{m(n')}$ for some $n' \in \mathbb{N}$ and outputs $w_{z,k}$, where $w_{z,k} \leftarrow \{w \in \{0,1\}^n \mid (z,w) \notin R\}$ if $n = n'$ and $w_{z,k} \leftarrow \{0,1\}^{n'}$ otherwise.*

Lemma 4.8. *Fix a $\mu(\cdot)$-sparse relation R. For any (q, q', q'')-adversary $\mathcal{A}$ such that $q''(n) < 1/6\mu(n)$ for all $n \in \mathbb{N}$, and any $n \in \mathbb{N}$,*

$$|\Pr[\mathcal{A}^{h,\mathsf{Inv}}(1^n) = 1] - \Pr[\mathcal{A}(1^n)^{h',\mathsf{Inv}'} = 1]| < (q + q'q'')\mu$$

where h and $h' := h_{R,n}$ are sampled as in Definitions 4.5 and 4.7 (resp.), and $\mathsf{Inv} := \mathsf{Inv}^h_R$ and $\mathsf{Inv} := \mathsf{Inv}^{h'}_R$ as in Definition 4.4.

Proof. We show that the outcomes of the two experiments $\mathcal{A}^{h,\mathsf{Inv}}(1^n)$ and $\mathcal{A}^{h',\mathsf{Inv}'}(1^n)$ are statistically close through a coupling between the distributions (h, Inv) and (h', Inv'). Fix $n \in \mathbb{N}$. The coupling distribution $\mathcal{P}_n$ first samples h' with randomness $\{w_{k,z}\}$ (see Definition 4.7) and an oracle Inv' with randomness $\{x_{C,y}\}$ (see Definition 4.4). The corresponding (h, Inv) are then sampled as follows:

- The oracle h: For any input $(k,z) \in \{0,1\}^{\kappa(n')} \times \{0,1\}^{m(n')}$ for $n' \neq n$, we set $h(k,z) = h'(k,z) = w_{k,z}$. For any input $(k,z) \in \{0,1\}^{\kappa(n)} \times \{0,1\}^{m(n)}$, we set $h(k,z) = h'(k,z)$ with probability $p_z = \Pr_{w \leftarrow \{0,1\}^n}[(z,w) \notin R]$ and with probability $1 - p_z$ we sample $h(k,z) \leftarrow \{w \mid (z,w) \in R\}$ at random.
- The oracle Inv: For any oracle-aided circuit $C : \{0,1\}^\lambda \to \{0,1\}^\ell$ and $y \in \{0,1\}^\ell$, we set $\mathsf{Inv}(C,y) = \mathsf{Inv}'(C,y) = x_{C,y}$ if $C^h(x_{C,y}) \overset{h}{\not\rightharpoondown} \mathsf{Corr}^h_q$ (where Corr^h_q is defined as in Definition 4.4) and, otherwise, we sample a fresh $\mathsf{Inv}(C,y)$ at random from $\{x \mid C^h(x) = y, C^h(x) \overset{h}{\not\rightharpoondown} \mathsf{Corr}^h_q\}$.

We first confirm that the coupling is valid. Letting $((h, \mathsf{Inv}), (h', \mathsf{Inv}')) \leftarrow \mathcal{P}_n$, it is straightforward by construction that (h', Inv') distribute identically to their original distribution in the experiment. Further, notice that the marginal distribution of $h(k,z)$, for any $(k,z) \in \{0,1\}^{\kappa(n')} \times \{0,1\}^{m(n')}$ is uniform over $\{0,1\}^{n'}$: in the case where $n \neq n'$ this is straight-forward and, the case of $n = n'$ follows from the choice of p_z. It remains to analyse the marginal distribution of Inv. For any circuit C (with query complexity q) and y, let $X_{C,y} = \{x \mid C^{h'}(x) = y, C_\lambda^{h'}(x) \overset{h'}{\not\rightharpoonup} \mathsf{Corr}_q^{h'}\}$ – this is the set from which $x_{C,y}$ is sampled at random in Inv'. Further, let $S_{C,y} = \{x \mid C^{h'}(x) = y, C^h(x) \overset{h}{\rightharpoonup} \mathsf{Corr}_q^h\}$ and observe that, by construction of $\mathcal{P}_n$ and since $\mathsf{Corr}_q^{h'} = \mathsf{Corr}_q^h \setminus \{0,1\}^n$, then $X_{C,y} \setminus S_{C,y} = \{x \mid C^h(x) = y, C^h(x) \overset{h}{\not\rightharpoonup} \mathsf{Corr}_q^h\}$. Hence, we may describe the marginal of $\mathsf{Inv}(C,y)$ as being sampled as follows: Sample $x \leftarrow X_{C,y}$ and, if $x \in S_{C,y}$ then sample $x \leftarrow X_{C,y} \setminus S_{C,y}$. This is equivalent to sampling $\mathsf{Inv}(C,y) \leftarrow X_{C,y} \setminus S_{C,y}$ and, therefore, to the distribution of Inv from Definition 4.4.

Having shown that $\mathcal{P}_n$ produces the proper marginal distribution, we may apply Lemma 2.2 to bound

$$|\Pr[\mathcal{A}^{h,\mathsf{Inv}}(1^n) = 1] - \Pr[\mathcal{A}(1^n)^{h',\mathsf{Inv}'} = 1]| < \Pr_{\mathcal{P}_n}[\mathcal{A}^{h,\mathsf{Inv}}(1^n) \neq \mathcal{A}^{h',\mathsf{Inv}'}(1^n)].$$

Now, observe that for any $((h,\mathsf{Inv}),(h',\mathsf{Inv}')) \in \mathcal{P}_n$, it holds that $\mathcal{A}^{h,\mathsf{Inv}}(1^n) \neq \mathcal{A}^{h',\mathsf{Inv}'}(1^n)$ only if one of the following occurs: either $\mathcal{A}^{h',\mathsf{Inv}'}$ makes a query to h' or Inv' which is answered differently by h or, resp. Inv. This happens only if one of the following two events occur:

- Either $\mathcal{A}^{h',\mathsf{Inv}'}$ makes an h'-query (z,k) such that $h(z,k) \neq h'(z,k)$. This happens with probability at most $1 - p_z < \mu$. Or,
- $\mathcal{A}^{h',\mathsf{Inv}'}$ makes an Inv'-query (C,y) and is answered by x where $C^{h'}(x) \overset{h'}{\rightharpoonup} (z,k)$ for some (z,k) such that $h(z,k) \neq h'(z,k)$. Via union bound over all queries made by $C^{h'}(x)$, we bound the probability of this event by $q''(1-p_z) < q''\mu$.

The proof is complete by applying union bound again over all q h'-queries and all q' Inv'-queries that are possibly made by $\mathcal{A}$. □

We now use the indistinguishability argument from Lemma 4.8 to derive the correlation intractability of h under Inv.

Lemma 4.9. *Fix a $\mu(\cdot)$-sparse relation R, and let $q(n) < 1/6\mu(n)$. Then for any (q,q,q)-adversary $\mathcal{A}$ and any $n \in \mathbb{N}$,*

$$\Pr_{\substack{k \leftarrow \{0,1\}^\kappa \\ z \leftarrow \mathcal{A}^{h,\mathsf{Inv}}(1^n,k)}}[(z, h(k,z)) \in R] < 2q^2\mu$$

where h and $\mathsf{Inv} := \mathsf{Inv}_R^h$ are sampled as in Definitions 4.5 and 4.4, resp..

Proof. Assume there exists a (q,q,q)-adversary $\mathcal{A}$ that breaks the correlation intractability of h for some $n \in \mathbb{N}$, specifically, suppose

$$\Pr_{k, z \leftarrow \mathcal{A}(1^n,k)}[(z, h(k,z)) \in R] > 2q^2\mu. \tag{10}$$

We use $\mathcal{A}$ to construct a $(q+1,q,q)$-adversary $\mathcal{B}$ that breaks Lemma 4.8 by distinguishing between a case where he is given access to a random h (and a corresponding Inv) and a case where he is given access to a statistically correlation-intractable $h' := h_{R,n}$ (and a corresponding Inv') as follows. On input 1^n, the algorithm $\mathcal{B}^{H,\mathsf{Inv}}$ samples $k \leftarrow \{0,1\}^\kappa$ then calls $z \leftarrow \mathcal{A}^{H,\mathsf{Inv}}(1^n)$ and outputs 1 if and only if $(z, H(k,z)) \in R$.

From (10), if $H = h$, then $\mathcal{B}$ outputs 1 with probability at least $2q^2\mu$ whereas, if $H = h'$, then since h' is statistically correlation-intractable, i.e. produces no correlations at all, then $\mathcal{B}$ outputs 1 with probability 0. Hence, $\mathcal{B}$'s distinguishing advantage is at least $2q^2\mu > (q+1+q^2)\mu$, in contradiction to Lemma 4.8. □

4.3 Proof of Theorem 4.3

Let F be any (q,ϵ)-fully black-box construction of OWF from correlation-intractable hash for all μ-sparse relations, with a corresponding oracle-aided reduction R. Consider a hash oracle h sampled at random as in Definition 4.5. For any relation $R \in \mathcal{R}$, let $\mathcal{A}$ be the adversary from Lemma 4.6 that inverts F^h with probability all but 1/6, over a random h. Using an averaging argument, such an adversary satisfies

$$\Pr_h\left[\Pr_{\mathcal{A},y}[\mathsf{F}^h_\lambda(\mathcal{A}(1^\lambda, y)) = y] \geq \frac{1}{2}\right] \geq \frac{2}{3}.$$

for all $\lambda \in \mathbb{N}$, where $y = \mathsf{F}^h_\lambda(x)$ for a uniform $x \leftarrow \{0,1\}^\lambda$. Thus, for at least $\frac{2}{3}$-fraction of all oracles h, and any relation R, there exists an adversary $\mathcal{A}$ such that $\mathbf{Adv}^{\mathbf{owf}}_{\mathsf{F},\mathcal{A}}(\lambda) \geq \frac{1}{2}$.

We now show that if $q(n) \leq 1/6\mu(n)$, then for at least $\frac{2}{3}$-fraction of random hash functions h, the reduction $\mathsf{R}^{h,\mathcal{A}}$ fails in finding R-correlations with sufficiently large probability. This completes the proof as it implies the existence of an oracle h relative to which there exists an $\mathcal{A}$ that breaks the one-wayness of F while $\mathsf{R}^{h,\mathcal{A}}$ fails in breaking the correlation-intractability of h.

Observe that if R is efficient and, since $\mathcal{A}$ is a $(0,1,q)$-adversary, then the computation $\mathsf{R}^{h,\mathcal{A}}$ can be expressed as a (q,q,q)-adversary $\mathcal{B}^{h,\mathsf{Inv}}$. Therefore, via Lemma 4.9, we obtain that

$$\Pr_{\substack{h,\mathcal{A} \\ k\leftarrow\{0,1\}^\kappa \\ z\leftarrow\mathsf{R}^{h,\mathcal{A}}(1^n,k)}}[(z, h(k,z)) \in R] < 2q^2\mu$$

and, through a standard averaging argument,

$$\Pr_h\left[\Pr_{\substack{\mathcal{A},k \\ z\leftarrow\mathsf{R}^{h,\mathcal{A}}(1^n,k)}}[(z, h(k,z)) \in R] > 6q^2\mu\right] \geq \frac{2}{3}.$$

We conclude that either $q(n) > 1/6\mu(n)$ or $\epsilon(n) \leq 6q^2\mu$.

4.4 Fiat-Shamir does not Imply OWFs

In this section, we discuss the implication of the fully black-box separation from Theorem 3.4 to the complexity of the Fiat-Shamir transform, compared to one-way functions. We first confirm that the well-known reduction from Fiat-Shamir over any constant-round protocol to correlation intractability is indeed fully black-box. Using such a fully black-box reduction, we then show that the separation proof from above provides us with an oracle, relative to which there exists *computationally* sound Fiat-Shamir for the given protocol, i.e. a Fiat-Shamir transform that is based on computational hardness, rather than statistical impossibility (which may be achieved by an idealized hash function), but still, there exist no one-way functions relative to the oracle.

First, let us recall the Fiat-Shamir transform and set up some notation.

The Fiat-Shamir Transform. Let $\Pi = (\mathsf{P}_\Pi, \mathsf{V}_\Pi)$ be an r-round public-coin protocol (possibly oracle-aided and/or in the CRS model), where we denote by a_i and b_i the prover's message and the verifier's (public-coin) challenge at the i^{th} round, respectively. Let $h := \{h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$ be an oracle. We denote by $\mathsf{FS}^h(1^n, \Pi)$ an instantiation of the Fiat-Shamir transform over Π using the hash function h_n. More specifically, $\mathsf{FS}^h(\Pi) = (\mathsf{P}_{\mathsf{FS}}, \mathsf{V}_{\mathsf{FS}})$ is a non-interactive protocol where, on security parameter $n \in \mathbb{N}$, the CRS contains a random hash key $k \leftarrow \{0,1\}^\kappa$ and possibly a CRS for Π, crs. The prover $\mathsf{P}_{\mathsf{FS}}(1^n, (k, \mathsf{crs}), x)$ locally emulates the interaction between P_Π and V_Π corresponding to crs and the instance x, while computing the public coins in every round as the hash of the current partial transcript including x and crs, under $h_n(k,\cdot)$ (w.l.o.g we assume that n public coins are sent in each round). Upon receiving $\tau = (a_1, b_1, \ldots, a_r)$, the verifier V_{FS} accepts if and only if $b_i = h_n(k, (\mathsf{crs}, x, a_1, b_1, \ldots, a_i))$ for all i and $\mathsf{V}_\Pi(1^n, \mathsf{crs}, x, \tau) = 1$.

Definition 4.10 (Fiat-Shamir Soundness)**.** *Let Π be a public-coin protocol for a language L and $h = \{h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$ be a distribution over oracles. Let $q := q(n)$ and $\epsilon := \epsilon(n)$. We say that the*

Fiat-Shamir transform over Π *using* h, $\mathsf{FS}^h(\Pi) = (\mathsf{P}_{\mathsf{FS}}, \mathsf{V}_{\mathsf{FS}})$, *is* (q, ϵ)-sound *if for any* q*-query adversary* $\mathcal{A}^h$ *and any* $\{x_n\} \notin L$, *it holds that*

$$\Pr_{h,k,\mathsf{crs}}[\mathsf{V}_{\mathsf{FS}}(1^n, (k, \mathsf{crs}), x, \mathcal{A}^h(1^n, (k, \mathsf{crs}))) = 1] < \epsilon$$

for infinitely many $n \in \mathbb{N}$.

We say that $\mathsf{FS}^h(\Pi)$ *is* statistically sound *if it is* (∞, ϵ)*-sound for some negligible* ϵ.

We say that $\mathsf{FS}^h(\Pi)$ *is* computationally sound *if it is* not *statistically sound yet there exists a negligible* ϵ *such that, for any polynomial* q, $\mathsf{FS}^h(\Pi)$ *is* (q, ϵ)*-sound.*

Next, we define a fully black-box reduction from Fiat-Shamir soundness to correlation intractability, which has the same flavor of black-boxness from Definition 4.2, then show that such a reduction exists.

Definition 4.11 (Fully-Black-box Fiat-Shamir Transform from CIH)**.** *Let* $\mathcal{R}$ *be a class of relations and let* Π *be an interactive protocol. A fully black-box (tight) reduction from Fiat-Shamir soundness over* Π *to CIH for* $\mathcal{R}$ *is an oracle-aided algorithm* R *satisfying the following:*

- ***Black-box Soundness Reduction:*** *For any* $\kappa := \kappa(n)$, $m := m(n)$ *and keyed-hash ensemble* $h = \{h_n : \{0,1\}^\kappa \times \{0,1\}^m \to \{0,1\}^n\}$, *there exists a relation* $R \in \mathcal{R}$ *such that for any probabilistic oracle-aided adversary* $\mathcal{A}$, *if* $\mathcal{A}(1^n, \cdot)$ *breaks the soundness of* $\mathsf{FS}^h(1^n, \Pi)$ *with probability* $\epsilon(n)$ *for infinitely many* $n \in \mathbb{N}$, *then* $\mathsf{R}^{h,\mathcal{A}}$ *breaks the correlation intractability of* h *for* R *with the same probability, i.e.*

$$\Pr_{\substack{k \leftarrow \{0,1\}^\kappa \\ z \leftarrow \mathsf{R}^{h,\mathcal{A}}(1^n, k)}}[(z, h(k, z)) \in R] \geq \epsilon(n)$$

for infinitely many $n \in \mathbb{N}$. *Further, for any* $n \in \mathbb{N}$, $\mathsf{R}^{h,\mathcal{A}}(1^n, \cdot)$ *only makes a single query to* $\mathcal{A}(1^n, \cdot)$.

Theorem 4.12. *Let* Π *be any* r*-round public-coin interactive protocol with soundness error* $\sigma(n)$. *Then, there exists a fully-black-box tight reduction from Fiat-Shamir soundness over* Π *to CIH for all* $\sigma^{1/r}$*-sparse relations.*

Proof. Fix an r-round public-coin protocol Π which is statistically sound with soundness error $\sigma(n)$. For any crs, $x \notin L$ and partial transcript $\tau = (a_1, b_1, \ldots, b_i)$, we define the *soundness error w.r.t.* (crs, x, τ) as

$$e_{\mathsf{crs},x,\tau} = \max_{\mathsf{P}^*} \Pr_{b_{i+1},\ldots,b_{r-1}}[\Pi_\tau(\mathsf{crs}, x, \mathsf{P}^*, (b_{i+1}, \ldots, b_{r-1})) = 1]$$

where $\Pi_\tau(\mathsf{crs}, x, \mathsf{P}^*, (b_{i+1}, \ldots, b_{r-1}))$ is an invocation of the protocol Π starting from round $i+1$ given the partial transcript τ, with a CRS crs, an instance x, and a (possibly dishonest) prover strategy P^*, where the public coins are $b_{i+1}, \ldots, b_{r-1}$. Intuitively, one can think of e_τ as the guaranteed soundness error, given τ is the transcript so far.

We construct a reduction R that, for any h, attacks the correlation intractability of h for the relation R defined by:

$$((\mathsf{crs}, x, a_1, b_1, \ldots, b_{i-1}, a_i), b_i) \in R \quad \text{iff} \quad e_{\mathsf{crs},x,(a_1,\ldots,b_i)} > \sigma^{-1/r} \cdot e_{\mathsf{crs},x,(a_1,\ldots,b_{i-1})}.$$

First, we show that this relation is $\sigma^{1/r}$-sparse. Notice that, by definition of the soundness error, for any fixed crs, x, and $\tau = (a_1, \ldots, b_{i-1}, a_i)$, it holds that

$$e_{\mathsf{crs},x,(a_1,\ldots,b_{i-1})} \geq \mathbb{E}_{b_i}[e_{\mathsf{crs},x,(a_1,\ldots,a_i,b_i)}]$$

and, therefore, via an averaging argument,

$$\Pr_{b_i}[((\mathsf{crs}, x, \tau), b_i) \in R] = \Pr_{b_i}[e_{\mathsf{crs},x,(a_1,\ldots,b_i)} > \sigma^{-1/r} e_{\mathsf{crs},x,(a_1,\ldots,b_{i-1})}] < \sigma^{1/r}$$

Next, we describe the reduction R, which builds on an adversary $\mathcal{A}$ that breaks the soundness of $\mathsf{FS}^h(1^n, \Pi)$ for infinitely many $n \in \mathbb{N}$ (see Definition 4.10). On inputs 1^n and k, R samples a CRS crs at random then calls $\mathcal{A}^h(1^n, (k, \mathsf{crs}))$ to obtain a proof $\tau = (a_1, b_1, \dots, a_r)$. From the correctness of $\mathcal{A}$, we know that, for infinitely many $n \in \mathbb{N}$, there exists $x \notin L$ such that $b_i = h_k(a_1, \dots, a_i)$ for all i and $\mathsf{V}_\Pi(1^n, \mathsf{crs}, x, \tau) = 1$, with probability at least σ. We claim that for any such τ, there exists $i \in [r]$ such that $((\mathsf{crs}, x, a_1, b_1, \dots, a_i), b_i) \in R$. This is sufficient since R can detect the instance x and the index i for which the above is satisfied (without any additional queries), then output $(\mathsf{crs}, x, (a_1, b_1, \dots, a_i))$, which is a correlation under R.

The existence of such an i follows from a simple argument: Let $e_i = e_{\mathsf{crs},x,\tau_i}$ denote the soundness error w.r.t. the partial transcript $\tau_i = (a_1, b_1, \dots, b_i)$. Then, by the soundness of Π, we know that $e_0 < \sigma(n)$ and, since τ is an accepting transcript, then $e_{r-1} = 1$. Hence, there must exist $i \in [r]$ such that $e_i > \sigma^{-1/r} \cdot e_{i-1}$. □

We now prove the main result in this section, building on the reduction from Theorem 4.11, and the oracles used in proving the separation of OWFs from CIH.

Theorem 4.13. *Let Π be any statistically-sound constant-round protocol with negligible soundness error, such that the Fiat-Shamir over Π using a random hash function is* not *statistically sound. Then, there exists a (randomized) oracle $\mathcal{O}$ relative to which there exists a computationally sound Fiat-Shamir over Π but no one-way functions.*

Proof. Denote by $\sigma := \sigma(n)$ the negligible soundness error of Π and by r the number of rounds in the protocol. Let R be the fully black-box reduction from Fiat-Shamir soundness over Π to correlation-intractability from Theorem 4.11. Consider the random oracle $\mathcal{O} = (h, \mathsf{Inv}_R)$, where h is a random hash function (see Definition 4.5) and $\mathsf{Inv} := \mathsf{Inv}^h_{R_h}$ is the inversion oracle w.r.t. the relation R_h which is the $\sigma^{1/r}$-sparse relation from Theorem 4.11, corresponding to R and h (it holds that R breaks the correlation intractability of h for R_h given any adversary against $\mathsf{FS}^h(\Pi)$).

By Lemma 4.6, any one-way function candidate is broken with high probability under $\mathcal{O}$. On the other hand, from Lemma 4.9, any q-query adversary fails to break the correlation intractability of h for R_h with probability larger than $O(q^2\sigma^{1/r})$. This implies that there exists no polynomial-query adversary $\mathcal{A}^{\mathcal{O}}$ that breaks $\mathsf{FS}^h(\Pi)$ with non-negligible probability since, otherwise, this breaks the correlation intractability of h for R_h through the reduction R. Further, it is given that $\mathsf{FS}^h(\Pi)$ is not statistically sound under such a random h and, therefore, the Fiat-Shamir transform using h is computationally sound. □

Acknowledgements

We thank Zvika Brakerski for comments and discussion, and Venkata Koppula for his helpful feedback.

References

[AS16] Gilad Asharov and Gil Segev. Limits on the power of indistinguishability obfuscation and functional encryption. *SIAM J. Comput.*, 45(6):2117–2176, 2016.

[BD19] Nir Bitansky and Akshay Degwekar. On the complexity of collision resistant hash functions: New and old black-box separations. In Dennis Hofheinz and Alon Rosen, editors, *Theory of Cryptography*, pages 422–450, Cham, 2019. Springer International Publishing.

[BDSG+13] Nir Bitansky, Dana Dachman-Soled, Sanjam Garg, Abhishek Jain, Yael Tauman Kalai, Adriana López-Alt, and Daniel Wichs. Why "fiat-shamir for proofs" lacks a proof. In Amit Sahai, editor, *Theory of Cryptography*, pages 182–201, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[BFJ+20] Saikrishna Badrinarayanan, Rex Fernando, Aayush Jain, Dakshita Khurana, and Amit Sahai. Statistical zap arguments. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, pages 642–667, Cham, 2020. Springer International Publishing.

[BKM20] Zvika Brakerski, Venkata Koppula, and Tamer Mour. Nizk from lpn and trapdoor hash via correlation intractability for approximable relations. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology – CRYPTO 2020*, pages 738–767, Cham, 2020. Springer International Publishing.

[BLV06] Boaz Barak, Yehuda Lindell, and Salil Vadhan. Lower bounds for non-black-box zero knowledge. *J. Comput. Syst. Sci.*, 72(2):321–391, March 2006.

[BR94] Mihir Bellare and Phillip Rogaway. Entity authentication and key distribution. In Douglas R. Stinson, editor, *Advances in Cryptology — CRYPTO' 93*, pages 232–249, Berlin, Heidelberg, 1994. Springer Berlin Heidelberg.

[BSBHR19] Eli Ben-Sasson, Iddo Bentov, Yinon Horesh, and Michael Riabzev. Scalable zero knowledge with no trusted setup. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 701–732, Cham, 2019. Springer International Publishing.

[BSCS16] Eli Ben-Sasson, Alessandro Chiesa, and Nicholas Spooner. Interactive oracle proofs. In Martin Hirt and Adam Smith, editors, *Theory of Cryptography*, pages 31–60, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

[CCH+19] Ran Canetti, Yilei Chen, Justin Holmgren, Alex Lombardi, Guy N. Rothblum, Ron D. Rothblum, and Daniel Wichs. Fiat-shamir: From practice to theory. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 1082–1090, New York, NY, USA, 2019. Association for Computing Machinery.

[CCR16] Ran Canetti, Yilei Chen, and Leonid Reyzin. On the correlation intractability of obfuscated pseudorandom functions. In *Proceedings, Part I, of the 13th International Conference on Theory of Cryptography - Volume 9562*, TCC 2016-A, page 389–415, Berlin, Heidelberg, 2016. Springer-Verlag.

[CGH04] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557–594, July 2004.

[CHK+19] Arka Rai Choudhuri, Pavel Hubácek, Chethan Kamath, Krzysztof Pietrzak, Alon Rosen, and Guy N. Rothblum. Finding a nash equilibrium is no easier than breaking fiat-shamir. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, page 1103–1114, New York, NY, USA, 2019. Association for Computing Machinery.

[CLMQ20] Yilei Chen, Alex Lombardi, Fermi Ma, and Willy Quach. Does fiat-shamir require a cryptographic hash function? Cryptology ePrint Archive, Report 2020/915, 2020. https://eprint.iacr.org/2020/915.

[DGI+19] Nico Döttling, Sanjam Garg, Yuval Ishai, Giulio Malavolta, Tamer Mour, and Rafail Ostrovsky. Trapdoor hash functions and their applications. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 3–32, Cham, 2019. Springer International Publishing.

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

[DNRS03] Cynthia Dwork, Moni Naor, Omer Reingold, and Larry J. Stockmeyer. Magic functions. *J. ACM*, 50(6):852–921, 2003.

[DRV12] Yevgeniy Dodis, Thomas Ristenpart, and Salil Vadhan. Randomness condensers for efficiently samplable, seed-dependent sources. In Ronald Cramer, editor, *Theory of Cryptography*, pages 618–635, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[FS87] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. In Andrew M. Odlyzko, editor, *Advances in Cryptology — CRYPTO' 86*, pages 186–194, Berlin, Heidelberg, 1987. Springer Berlin Heidelberg.

[GGM85] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. On the cryptographic applications of random functions. In *Proceedings of CRYPTO 84 on Advances in Cryptology*, page 276–288, Berlin, Heidelberg, 1985. Springer-Verlag.

[GJJM20] Vipul Goyal, Abhishek Jain, Zhengzhong Jin, and Giulio Malavolta. Statistical zaps and new oblivious transfer protocols. In Anne Canteaut and Yuval Ishai, editors, *Advances in Cryptology – EUROCRYPT 2020*, pages 668–699, Cham, 2020. Springer International Publishing.

[GK03] S. Goldwasser and Y. T. Kalai. On the (in)security of the fiat-shamir paradigm. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 102–113, 2003.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270 – 299, 1984.

[HIL99] Johan Hastad, Russell Impagliazzo, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM Journal on Computing*, 28, 02 1999.

[HL18] J. Holmgren and A. Lombardi. Cryptographic hashing from strong one-way functions (or: One-way product functions and their applications). In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 850–858, 2018.

[HR04] Chun-Yuan Hsiao and Leonid Reyzin. Finding collisions on a public road, or do secure hash functions need secret coins? In Matt Franklin, editor, *Advances in Cryptology – CRYPTO 2004*, pages 92–105, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[HT99] Satoshi Hada and Toshiaki Tanaka. A relationship between one-wayness and correlation intractability. In *Public Key Cryptography*, pages 82–96, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[HT06] Satoshi Hada and Toshiaki Tanaka. Zero-knowledge and correlation intractability. *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, E89-A(10):2894–2905, October 2006.

[IL89] R. Impagliazzo and M. Luby. One-way functions are essential for complexity based cryptography. In *30th Annual Symposium on Foundations of Computer Science*, pages 230–235, 1989.

[Imp95] Russell Impagliazzo. Personal view of average-case complexity. pages 134–147, 07 1995.

[JKKZ20] Ruta Jawale, Yael Tauman Kalai, Dakshita Khurana, and Rachel Zhang. Snargs for bounded depth computations and ppad hardness from sub-exponential lwe. *IACR Cryptol. ePrint Arch*, 2020:980, 2020.

[Kil92] Joe Kilian. A note on efficient zero-knowledge proofs and arguments (extended abstract). In *Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing*, STOC '92, page 723–732, New York, NY, USA, 1992. Association for Computing Machinery.

[KRR17] Yael Tauman Kalai, Guy N. Rothblum, and Ron D. Rothblum. From obfuscation to the security of fiat-shamir for proofs. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 224–251, Cham, 2017. Springer International Publishing.

[LV20a] Alex Lombardi and Vinod Vaikuntanathan. Fiat-shamir for repeated squaring with applications to ppad-hardness and vdfs. In Daniele Micciancio and Thomas Ristenpart, editors, *Advances in Cryptology - CRYPTO 2020 - 40th Annual International Cryptology Conference, CRYPTO 2020, Santa Barbara, CA, USA, August 17-21, 2020, Proceedings, Part III*, volume 12172 of *Lecture Notes in Computer Science*, pages 632–651. Springer, 2020.

[LV20b] Alex Lombardi and Vinod Vaikuntanathan. Multi-input correlation-intractable hash functions via shift-hiding. Cryptology ePrint Archive, Report 2020/1378, 2020. https://eprint.iacr.org/2020/1378.

[Mic00] Silvio Micali. Computationally sound proofs. *SIAM J. Comput.*, 30(4):1253–1298, October 2000.

[Nao91] Moni Naor. Bit commitment using pseudorandomness. *Journal of cryptology*, 4(2):151–158, 1991.

[OW93] R. Ostrovsky and A. Wigderson. One-way functions are essential for non-trivial zero-knowledge. In *[1993] The 2nd Israel Symposium on Theory and Computing Systems*, pages 3–17, 1993.

[PS18] Chris Peikert and Sina Shiehian. Privately constraining and programming prfs, the lwe way. In Michel Abdalla and Ricardo Dahab, editors, *Public-Key Cryptography – PKC 2018*, pages 675–701, Cham, 2018. Springer International Publishing.

[PS19] Chris Peikert and Sina Shiehian. Noninteractive zero knowledge for np from (plain) learning with errors. In Alexandra Boldyreva and Daniele Micciancio, editors, *Advances in Cryptology – CRYPTO 2019*, pages 89–114, Cham, 2019. Springer International Publishing.

[RTV04] Omer Reingold, Luca Trevisan, and Salil Vadhan. Notions of reducibility between cryptographic primitives. In Moni Naor, editor, *Theory of Cryptography*, pages 1–20, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[Sch90] C. P. Schnorr. Efficient identification and signatures for smart cards. In Gilles Brassard, editor, *Advances in Cryptology — CRYPTO' 89 Proceedings*, pages 239–252, New York, NY, 1990. Springer New York.

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