On The Black-Box Complexity of Correlation Intractability

Nico Döttling[*]
CISPA Helmholtz Center for Information Security
nico.doettling@gmail.com

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

December 2020

Abstract

Correlation intractability is an emerging cryptographic paradigm that enabled several recent breakthroughs in establishing soundness of the Fiat-Shamir transform and, consequently, basing non-interactive zero-knowledge proofs and succinct arguments on standard cryptographic assumptions. In a nutshell, a hash family is said to be *correlation intractable* for a class of relations $\mathcal{R}$ if, for any relation $R \in \mathcal{R}$, it is hard given a random hash function $h \leftarrow H$ to find an input z s.t. $(z, h(z)) \in R$, namely a correlation.

Despite substantial progress in constructing correlation intractable hash functions, all constructions known to date are based on highly-structured hardness assumptions and, further, are of complexity scaling with the circuit complexity of the target relation class.

In this work, we initiate the study of the barriers for building correlation intractability. Our main result is a lower bound on the complexity of any black-box construction of CIH from collision resistant hash (CRH), or one-way permutations (OWP), for any sufficiently expressive relation class. In particular, any such construction for a class of relations with circuit complexity t must make at least $\Omega(t)$ invocations of the underlying building block.

We see this as a first step in developing a methodology towards broader lower bounds.

[*] Funded by the European Union (ERC, LACONIC, 101041207). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

[†] 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. The Fiat-Shamir (FS) transform [FS87, BR94] is a popular technique for eliminating interaction in interactive public-coin protocols. The technique was first conceived to transform 3-round identification protocols into non-interactive signature schemes [FS87]. Since its introduction, this methodology has had a substantial impact on modern cryptography through several lines of research. The concept gave rise to a number of key innovations in modern cryptography, 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, CJJ21], succinct non-interactive arguments (SNARGs) [Kil92, Mic00, BSCS16, BSBHR19, JJ21], and complexity-theoretic hardness results [CHK+19, LV20a, JKKZ20].

The basic blueprint of the FS-transform, as laid out in [BR94], is to transform a (multi-round) public coin protocol by using a hash function H to generate the verifier's public coin messages deterministically based on the protocol transcript so-far.

While it is usually a straight-forward to show that Fiat-Shamir preserves some properties of the original interactive protocol, e.g. completeness and zero-knowledge, it is typically a lot more challenging to show that it preserves soundness using any hash function H. Intuitively, this complication arises as a malicious prover has some control over the computed challenges, e.g. it may just discard a protocol run and retry. In fact, in most constructions the soundness of FS is based on *heuristics*.

More concretely, the soundness of the transformed protocol is often established in an idealized model such as 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 FS transform is sound 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. The heuristic leap occurs when the random oracle is instantiated by an "unstructured" function such as SHA-2.

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 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.

Correlation Intractability. 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 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) is sufficient for Fiat-Shamir over *any* constant-round public-coin proof system(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, most recently, succinct argument [JKKZ20, CJJ21, HJKS22, KLV23, CGJ+23]. Overall, the state-of-the-art constructions of CIH are based on advanced 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, JJ21] (through trapdoor hash functions [DGI+19]).

Towards Understanding The Complexity of Correlation Intractability. For a complete comprehension of the notion of correlation intractability, it is fundamental to investigate not only the possibilities but also the limitations in basing correlation intractability on existing hardness notions. In this work, we focus on the relation between correlation intractability and two of the most prominent hardness notions in cryptography: *One-wayness* and *collision-resistance*.

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.

A collision resistant hash family (CRH) is a family of hash functions where it is hard to find collisions under a given random hash sampled from the family. CRH is one of the most widely used primitives in cryptography, with applications ranging from the most basic cryptographic tasks [Dam87, HM96] to more advanced ones [Kil92, BEG+91, BG02]. Despite its conceptual simplicity, it has been proven that collision resistance cannot be based on one-way functions [Sim98] or even public-key cryptography [AS16, BD19], at least not in a black-box manner.

While, as noted in [CLMQ20], 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 group)[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, JJ21], or based on (sub-)exponentially secure OWF and additionally assume the existence of indistinguishability obfuscation (iO) [HL18, LV20b].

Whereas the theoretical cryptography literature is rich with proven separations between various cryptographic notions, almost no work had been done on the limitations of reducing correlation intractable hash to other primitives, leaving our understanding of the "reduction complexity" of correlation intractability to be very lacking. The only exceptions are [HT99], who rule out building the strongest possible form of CIH (for all sparse relations, implying a universal Fiat-Shamir hash) on one-way functions, and [CLMQ20], who ask whether we can instantiate some specific use-cases of Fiat-Shamir without (or with very weak) cryptography. What we aim for is a more general and accurate picture where we place correlation intractability among the prominent hardness notions in cryptography, specifically one-wayness and collision resistance.

While it is typically beyond our field's current capabilities to rule out general reductions between different hardness notions, a useful framework, that has been developed along the past decades to facilitate reaching meaningful separation results, considers the special case of *fully-black-box* constructions [RTV04], where (i) the construction makes only black-box use of the underlying primitive, i.e. is oblivious in its implementation, and (ii) the reduction is assumed to use the provided adversary against the base primitive in a black-box manner. Such separations are insightful in particular since they already rule out most of the techniques used constructions in the cryptographic literature. The fully-black-box framework has been shown to be extremely fruitful to obtain fundamental separation results, such as separating CRH from OWFs or public-key cryptography [Sim98, AS16, BD19] and separating key-agreement (and hence, public-key cryptography) from OWFs [IR89].

[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 $k = a - y$.

Restricting our focus to fully-black-box reductions, we propose the following question to initiate a thorough study of the complexity of correlation intractability:

What is the black-box complexity of correlation intractable hashing from CRH?

We believe collision resistance is a natural starting point in this general direction as it is a sufficiently simple and basic notion to constitute a first step towards broader research. There are two items to note here: First, as collision resistance implies OWFs (in a fully-black-box manner), any answer for the above question would immediately imply a similar statement for constructions from OWFs. Second, CRH are a special case of *multi-input correlation intractability*, which is a generalization of correlation intractability where it is hard to find a *tuple* of inputs that satisfy some relation between themselves and their images under the hash (in standard CI relations are over a single input-output pair). While general multi-input CI is clearly stronger than regular CI and implies it, what we ask above is whether multi-input CI for a specific natural (multi-input) relation can be useful to build CI for a more general class of (single-input) relations, that is – whether "multiplicity" of the relation class can be exchanged for "expressivness".

1.1 Our Results

In this work, we explore inherent limitations in constructing correlation intractable hash functions and initiate the study of the black-box complexity of correlation intractability. We draw the following connection between the complexity of any fully-black-box construction of CIH from CRH or OWP and the complexity of the relations we get correlation intractability for.

Theorem 1.1 (Black-box Complexity of CIH from CRH or OWP; Informal)**.** *Any fully-black-box construction of correlation intractable hash for any t-wise independent class of relations from collision-resistant hash, or one-way permutations, must make at least $O(t)$ calls to the underlying base primitive(s).*

A t-wise independent class of relations $\mathcal{R}$ is class of relations where for any $t' \leq t$ pairs (z_1, w_1), ..., $(z_{t'}, w_{t'})$, the events $\{(z_i, w_i) \in R\}$ for a random relation $R \leftarrow \mathcal{R}$ are all independent. One example is the class of all relations searchable by degree t polynomials, i.e. any relation consisting of all pairs $(z, p(z))$ for a degree t polyonomial p specified by the relation. Consequently, as polynomials of degree t can be computed by circuits of size t, we get that the class of relations searchable by t-bounded circuits is $\Omega(t)$-wise independent. Hence, the degree of independence provides a meaningful proxy for the complexity of a class of relations. To give some sense, CIH suitable for cryptographic applications, such as NIZKs or succinct non-interactive arguments, as far as we know requires intractability for relations with complexity proportional to the security parameter.

Our result carries a couple of caveats. First, it holds only for fully-black-box constructions. Although this is already insightful and captures many of the existing and imaginable techniques, there might exist non fully-black-box constructions that circumvent this impossibility. Second, this is not an absolute separation in the sense that it does not entirely rule out building one primitive from another, rather it only sets a lower bound on the efficiency of such constructions. While such a result is partial in nature, we believe that the analysis underlying the proof provides many insights regarding the essence of correlation intractability and its complexity, potentially leading to future work advancing our understanding further, through stronger separations and even new constructions. We elaborate below.

1.2 Discussion and Open Questions

We view our result as initiating the research on the complexity of correlation intractable hashing. While our bottom-line yields a lower bound that is far from what is known or even believed to be possible, our hope is that the techniques and observations introduced in our analysis will eventually lead to a better understanding of the notion of correlation intractability.

For instance, it may not be unlikely that, with some additional effort and insights, our proof can be extended to achieve a similar impossibility for CIH from public-key cryptography; In the work of [BD19] by which our initial ideas were inspired, they are able to show separation of CRH not only from OWPs but also from the combination of OWPs with iO, which, in particular, implies a separation from public-key encryption (PKE) [2]. While extending our result in an analogous manner is doomed to fail due to the existence of (fully black-box) CIH based on OWP and iO, demonstrated in [LV20b], we expect that a more careful adaptation of the developed ideas has the potential to yield a separation from PKE.

A more intriguing direction is to investigate the gap between our limited separation result, that does not entirely rule out constructions of CIH from CRH or OWP, and the state-of-the-art CIH constructions which are known from building blocks that are much more complex. We see it is important to understand whether it is merely an artifact of our proof technique that we were not able to extend it to rule out constructions for *any* (non-trivial) class of relations or whether there is an inherent barrier in proving such a separation. In particular, one may ask

Is it indeed impossible to build non-trivial CIH in Minicrypt [Imp95] (or Hashomania [KNY18])? Which relation classes can we get CIH for, based on one-way functions or collision-resistant hash?

1.3 Technical Overview

We will now discuss the ideas behind our main result, Theorem 1.1, which states that any fully-black-box construction of CIH from CRH, or OWP for relations of complexity t (more accurately, that are t-wise independent) must make $\Omega(t)$ invocations of the underlying base primitive(s).

The starting point of our proof is the work of Bitansky and Degwekar [BD19], which provides a separation of collision-resistant hash functions (CRH) from one-way permutations (OWP). We generalize their framework to the correlation intractability setting and further extend it to capture the separation from CRH. Along the way, we introduce a new notion which facilitates establishing hardness under oracles (e.g. of inversion or finding collisions) which we call *differential indistinguishability.* Proving oracle-relative hardness is always at the core of separations of this theme since, typically, the underlying cryptographic primitive, which is accessible only in a black-box manner, is modelled as an oracle that provably satisfies the corresponding intractability property. Interestingly, through differential indistinguishability, we show how to use techniques resembling those from the differential privacy literature in order to obtain traditional cryptographic hardness relative to an oracle.

For the sake of this overview, we outline the lower bound on CIH constructions from one-way permutations and then briefly discuss how the underlying techniques can be further expanded to obtain the lower bound on constructions from CRH.

Let us briefly recall the fully black-box separation framework which we follow in this work.

Fully Black-box Separations and How to Prove Them. 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 $\mathcal{M}$ 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. In many cases, such as ours, conditioned separations are considered, namely, where it is only argued that fully-black-box constructions that satisfy certain constraints (e.g. efficiency) are impossible.

Similarly to prior work on fully-black-box separations, we follow the "Two-Oracle Methodology" [Sim98, HR04, AS16, BD19] where, to show that is it impossible to build correlation intractable hash from another primitive $\mathbf{Q}$, e.g. OWP or CRH – again, possibly assuming certain efficiency constraints – it is shown that there exists an oracle Q, which models an idealized implementation of $\mathbf{Q}$, and an oracle that models an adversary against correlation intractability, namely, a *correlation finder* CF, such that (i) CF breaks any black-box construction of CIH from Q that satisfies the presumed constraints, yet, (ii) Q is still secure, as

[2] This, in fact, was first established in [AS16]

per the security definition of $\mathbf{Q}$, in the presence of CF. Given that such oracles Q and CF exist, any fully black-box reduction $\mathcal{M}$ fails in breaking Q using CF and, hence, no fully black-box construction of CIH from $\mathbf{Q}$ exists.

We first focus on separating CIH constructions from OWP, as this captures many of the key concepts in the extended result, and only later discuss how to further derive a separation from CRH.

The Challenge in Designing a Correlation Finder. We model our "ideal" OWP via a random permutation oracle $f : \{0,1\}^\lambda \to \{0,1\}^\lambda$. While it is straight-forward to show that it is infeasible to invert a random permutation at a random image given bounded black-box access, our goal is to show this is still infeasible even given access to a successful correlation finder CF. The correlation finder CF takes as input a circuit $C \in H$ describing a hash function with oracle access to f. Here we can think of the set H as abstracting away from the keys in a hash construction. In essence the set H limits the adversary's choice.

One natural way to show that any bounded reduction $\mathcal{M}$ still cannot invert f given CF is to show that $\mathcal{M}$ is able to simulate, with little extra cost, any useful information it receives from CF by itself alone, making the correlation finder redundant and using the one-wayness of a random permutation to complete the proof. This approach has been successful to separate, in particular, CRH from OWP [Sim98, BD19]. An elementary reason is that, under any (sufficiently shrinking) CRH, the marginal distribution of any "half" of a uniformly random collision, is almost uniform. Let $C \in H$ be a circuit with oracle-access to f describing a hash function. Thus, when letting the collision finder, on input such a circuit C^f, simply output a random collision (z, z') (s.t. $C^f(z) = C^f(z')$ and $z \neq z'$), the reduction can simulate the marginals of each of z and z' without the help of the collision-finder. Roughly speaking, as has been shown particularly in [BD19], the marginals capture the only "useful" information the reduction can obtain for inverting f.

Things are not that simple, however, when the goal of the oracle is to return a correlation. Here, the correlation finder CF^f_R depends on an a relation R as well as f. We will omit R and f when the context is clear. On input a circuit $C \in H$ the correlation finder $\mathsf{CF}^f_R(C)$ should produce an input z such that $(z, C^f(z)) \in R$. In this setting, a reduction $\mathcal{M}$ may produce a query to CF where all possible correlations under a chosen relation R, i.e. all "correct" answers that CF may possibly return, coincide with a set of inputs that is most useful for inverting f at any given point. For example, we may think of an $\mathcal{M}$ that, given a challenge y, calls $\mathsf{CF}(C)$ where C is the hash circuit that on any input z, outputs a w s.t. $(z, w) \in R$ (it is reasonable to assume that such a w is efficiently computable) if and only if $f(z) = y$[3]. For this C, there is only one such z satisfying $(z, C^f(z)) \in R$, an hence $CF(C)$ must return this z.

Picky Correlation Finder. Given the inherent tension between correctness of the correlation finder and its usefulness for inverting f, we propose the following way out. We design CF to be *imperfect*, that is, to return a correct answer, say a uniformly random correlation, for most inputs while rejecting to do so for the others. The distinction between functions on which CF may "cooporate" and functions on which CF must reject is made possible by the fact that, in order to break correlation intractability, CF must succeed on some relation R only for an *average-case* hash $\tilde{C} \leftarrow H$. Thus, in the CIH game, which one can think of as the "honest" case, the circuit $\tilde{C}$ that computes the hash function is independent of R and should not exhibit any extraordinary behavior w.r.t. correlations under R. On the other hand, if $\mathcal{M}$ attempts to abuse CF to invert f, then it must produce a "malicious" circuit C which is specifically tailored to be useful for inversion and, therefore, as we argue in our proof, must highly "depend" on R.

Thus, we need to construct a correlation finder CF that is able to tell when a circuit C is likely to be malicious, yet does not overshoot as it still needs to answer for an honest C. To that end, we articulate a measure of "extraordinariness" that captures "usefulness" for inversion, which, roughly speaking, happens to be tightly related to the Rényi divergence of infinite order (this can be thought of as an analog of KL-divergence for min-entropy) between what useful information is obtained from CF and what useful information

[3] We are implicitly assuming that the input spaces for f and C are the same. When this is not the case, the reduction can use any arbitrary 1-1 mapping that maps any C-input w to a corresponding f-input x_w and the implication still holds.

can be simulatable without CF. We let our CF reject any circuit C that is extraordinary w.r.t. R to obtain a *picky correlation finder*, namely a correlation finder that is successful only with high probability.

Detecting Malicious CF-Queries. In the following let $\mathsf{Corr}^f_{R,C}$ denote the set of all z which satisfy $(z, C^f(z)) \in R$. To identify what "useful information" is w.r.t. inverting one-way permutations, we take inspiration from [BD19], where they implicitly show that to invert an oracle f at a random image y^*, it is necessary for the reduction $\mathcal{M}$ to be able to distinguish between black-box access to f and black-box access to a different permutation $f' = f_{x^* \leftrightarrow x'}$ that is obtained from f by swapping the solution pre-image $x^* = f^{-1}(y^*)$ with a uniformly random x'. Given this, a malicious query to CF is then a circuit C for which the distribution of a non-rejecting $\mathsf{CF}^f(C)$, namely the correlation finder's answer to C under f can be distinguished from a $\mathsf{CF}^{f'}(C)$, namely its answer under f'. Only using such malicious queries, the reduction $\mathcal{M}$ can use CF to distinguish between f and f' and, thus, invert f. We observe that C can induce such two distinguishable distributions under functions f and f' only if the swap $x^* \leftrightarrow x'$ significantly affects the set of correlations, from which CF samples its answer. This may occur only if, given a random correlation $z \leftarrow \mathsf{Corr}^f_{R,C}$, the hash function $C^f(z)$ calls any of x^*, x' with noticeable probability or, in other words, only if any of x^*, x' are *heavy among correlations*. Hence, our correlation finder should, in particular, look at the weight of any worst-case x w.r.t. the given query C and the chosen relation R (it is crucial to note that CF has no knowledge of x^*, x' as they exist only in the inversion game and its analysis), which we define as

$$\omega^f_{R,C}(x) = \Pr_{z \leftarrow \mathsf{Corr}^f_{R,C}}[C^f(z) \rightharpoondown x],$$

where $C^f(z) \rightharpoondown x$ denotes the event that the computation $C^f(z)$ calls f at x. This alone is not sufficient, however, since it may be the case that there are heavy inputs also under an honest query C, that does not depend on the relation R. For instance, consider a CIH construction C that *always* calls f at some fixed x_0, regardless of its input being a correlation or not. Then, in such case we have that $\max_x \omega(x)$ takes its maximal value 1 and the correlation finder always rejects and is, therefore, never successful. It is clear that such a query to CF cannot possibly be helpful to invert f since, intuitively speaking, any information that $\mathcal{M}$ may extract from $\mathsf{CF}^f(C)$ regarding the image of the heavy input x_0 it could already extract without calling CF by calling C at random inputs (that are not necessarily a correlation). Keeping our initial outline in mind, we are interested in the *relative* "usefulness" of information obtained form CF compared to information simulatable without CF's help. We refine our basic idea to consider the *relative* weight of any worst-case x among correlations compared to its weight in the entire input space. For that, we define the *scale* of any input x as

$$\sigma^f_C(x) = \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x],$$

and look at the *amplification* in the likelihood of observing x in an execution $C^f(z)$ due to restricting z to be a random correlation compared to being a random input (i.e., a random CF answer compared to a random input which is simulatable without CF), that is,

$$\alpha^f_{R,C}(x) = \omega^f_{R,C}(x) / \sigma^f_C(x).$$

We let CF reject only if there exists an x for which $\alpha^f(x) \gg 1$. This gives us a CF that is successful in the honest case since one can easily see that $\alpha^f(x)$ has an average of 1 when the relation R is sampled independently in C (further, a sufficient tail bound for worst-case $\alpha^f(x)$ can be derived already when R is pairwise independent). On the other hand, CF rejects whenever x^* or x' are too heavy among correlations since $\sigma^f_C(x^*)$ and $\sigma^f_C(x')$ can be assumed to be small: x' is a random input that is sampled in the analysis and is independent in the reduction's choice of C, while x^* is the solution pre-image and, had it been heavy among random inputs, the reduction would have been able to observe it by sampling random inputs to C without the help of CF.

As already mentioned, it turns out that $\max_x \alpha^f(x)$ is precisely the Rényi divergence of order infinity between the distributions over the f-input space induced by the PDFs ω and σ. In Figure 1 we visualize the distinction between honest queries, which give low divergence between ω and σ, and malicious queries.

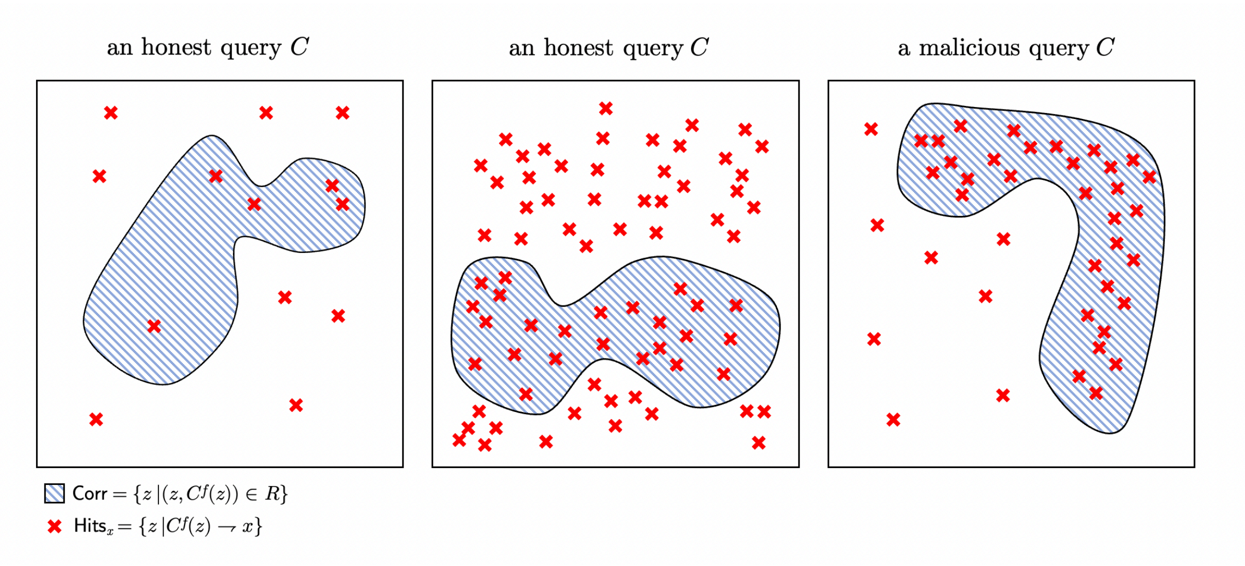
an honest query C
an honest query C
a malicious query C
▨ Corr $= \{z \,|(z, C^f(z)) \in R\}$
✖ Hits$_x = \{z \,|C^f(z) \rightharpoondown x\}$

Figure 1: An illustration of the domain of the circuit C when C is an honest query vs. when it is a malicious one. Notice that C is malicious only if some x is observed with much higher likelihood in the correlation set, compared with the entire space.

While our efforts so far already constitute a major step towards a separation, a new problem pops up: the decision for rejection, namely whether CF rejects or not, might itself give information that is useful for inverting f! This is not merely a hypothetical scenario; One can show, in fact, a concrete "attack" against such CF, namely a reduction $\mathcal{M}$ that, while unable to learn anything from the non-rejection answers, can learn the pre-image entirely based on the rejection decisions of CF (note each such decision conveys at most a single bit of information).

Differential Indistinguishability. To solve the issue raised above, we encapsulate the "usefulness" (or "uselessness") of a correlation finder in inverting f via an novel oracle-relative hardness notion, we call *differential indistinguishability.* At a high level, we say that CF is differentially indistinguishable if its answers and, in particular, its *rejection decision*, do not substantially change when the function f is modified locally (specifically, when two inputs are swapped under the permutation as described above). As already mentioned, the notion of differential indistinguishability is implicitly used in the work of [BD19] and is inspired by their proof. However, the straight-forward collision finder, that returns a random collision, already satisfies differential indistinguishability and, therefore, the main effort in their proof goes to show that differential indistinguishability implies "uselessness" for inversion. One of the contributions in our work is formalizing then extending the implicit framework from [BD19] to capture correlation finders (in fact, any oracle) and additionally show a similar implication regarding "uselessness" for finding collisions, through which we derive the separation from CRH.

A Differentially Indistinguishable Rejection Policy. Given that differential indistinguishability implies that f is hard to invert given CF, it remains to design a differentially indistinguishable CF. To that end, we further refine the picky correlation finder from above and design a "soft" rejection policy that is robust against local changes of the function f. Unsurprisingly, our mechanism for achieving this looks, in retrospect, as if taken from the world of differential privacy (DP) [Dwo06]. In more details, we consider a rejection policy that takes into account not only extraordinary behaviour, namely large $\alpha^f(x)$, w.r.t. the given function f, but also w.r.t. functions f' that are in the "neighborhood" of f, namely functions that can be obtained by performing a limited series of swaps on f. By weighing in the impact of any *extraordinary* *behaviour* in a way that quickly vanishes with increasing distance from f, e.g. an weight function which drops exponentially with the distance from f, we are able to derive differential indistinguishability. Overall, our rejection policy, namely the probability that CF rejects at some circuit query C, is computed by a function similar to the following

$$\rho_R^f(C) = \epsilon \cdot \max_{f',x}[e^{-\Delta(f,f')/d} \cdot \alpha_{R,C}^{f'}(x)],$$

where $\Delta(f, f')$ denotes the *swap distance* [4] between f and f', and $\epsilon \ll 1$ and $d \gg 1$ are carefully chosen normalization parameters; The larger d and ϵ are, the stronger is the differential indistinguishability guarantee, yet the harder it is to show that CF is successful in the honest case. For the sake of this overview, d can be thought of as superpolynomial in the security parameter and ϵ as inverse-superpolynomial.

Having safeguarded intractability of inverting f via differential indistinguishability, we need to make sure we have not broken the subtle balance with the necessary correctness requirement on CF. While it is almost immediate that any circuit C (modelling a random has functionh $\tilde{C} \leftarrow H$ sampled from the CIH candidate) behaves "nicely" under f w.r.t. an independently chosen relation $R \leftarrow \mathcal{R}$, even when relations in $\mathcal{R}$ are only pairwise-independent, it is not clear that this holds under *all* functions f' in the close neighborhoods of f. The straight-forward attempt to apply a union bound over all possible functions in the neighborhood inherently requires that relations in the class have a description of exponential size, which would dramatically weaken our result. Instead, we exploit the fact that the candidate CIH is bounded to make t invocations of f and observe strong dependencies between the $\alpha(x)$ values under different functions across the neighborhood of f. More specifically, since the behavior of C on any input is determined by the images of at most t points in f, we notice that we can represent $\alpha^{f'}(x)$, under any f', as the average of a collection of values $\{\hat{\alpha}^{f'}(x)\}$ where each $\hat{\alpha}^{f'}(x)$ depends solely on $t \ll d$ points in f'. This allows us to apply a much more benign union bound to establish that none of the possible $\hat{\alpha}^{f'}(x)$ is too large. Consequently, we are able to argue that if the relation class is sufficiently "expressive", namely $\Omega(t)$-wise independent, then when the relation R is chosen at random and independently in the hash circuit C, C is indeed behaving well not only under f, but rather under any f' that is sufficiently close to f. This implies that a rejection happens with low probability for an honest query C and, therefore, CF is correct and we have a separation of CIH from OWP.

Extending to Constructions from Collision Resistance. We will now discuss how these results can be extended to capture constructions from CRH as well. At the core of our extended result is the observation that the task of finding collisions in an oracle can be conceptually though of as an "adaptive" inversion task; To find a collision in an oracle f, an adversary must invert an image y for which he had already seen a pre-image under f. The difference from breaking the one-wayness of f, namely inverting f at a random image, is clear: In the collision-finding game, the adversary, in some sense, chooses the images he aims to invert. Hence the "adaptiveness".

Recall that in order to establish a separation of CIH from OWP, we (i) define a notion of differential indistinguishability and prove a random f is hard to invert even given a differentially indistinguishable correlation finder, and (ii) construct a differentially indistinguishable correlation finder (that is successful in the honest case). Based on the above insight, we propose a notion of *adaptive differential indistinguishably*, then prove that it is sufficient to imply collision-resistance of f under the (adaptively differentially indistinguishable) correlation finder. Lastly, we show how to generalize our construction of correlation finder from above to satisfy the required adaptive differential indistinguishably and, by this, finish. We elaborate below.

Re-randomizing Siblings. The reason that differential indistinguishability is sufficient to imply hardness of inversion is that it allows us to re-randomize the target pre-image (that is, swap the original x^* whose image is given as a challenge with a random x') without the adversary noticing that he is given access to a different function $f' = f_{x^* \leftrightarrow x'}$. Thus, the probability that the adversary returns x^* under f is equal to the probability he returns x^* under f', which carries no information about x^*, and therefore he cannot do better than guessing. To adapt this idea to an adversary that is trying to find collisions, we re-randomize,

[4] We define the swap distance between permutations f and f' as the minimal number of times we need to swap outputs between input pairs x_1 and x_2 in order to transform f into f'

as hinted above, the *siblings* of any input x at which the adversary calls f in his executions. The siblings of any x under f are all x' that make a collision with x, i.e. all $x' \neq x$ such that $f(x) = f(x')$. Roughly speaking, since we may assume w.l.o.g. that the adversary calls f at a collision the moment he finds it, then we can assume that a successful adversary must call a sibling of a previously queried input. It would not be sufficient, however, to re-randomize, at every step of the execution, only siblings from previous queries since, hypothetically, the adversary's strategy might be to collect information about "future" siblings, namely siblings of some x before actually making the query to f at x. We must therefore re-randomize *all* siblings induced by the execution at *any* of its steps.

Adaptive Differential Indistinguishability. An inherent difference from the OWP case then arises: In proving intractability of inverting an OWP f, we re-randomize a pre-image x^* which is fixed apriori to the execution of the adversary. In contrast, when re-randomizing siblings, specifically "future" siblings, we are re-randomizing pre-images that are implicitly determined by the adversary's execution. This difference motivates us to define an adaptive analog of differential indistinguishability, where, in a high level, we require that the answers of CF do not change when the function f is swapped even at points chosen adaptively in the answers themselves (essentially, the answers are what constitute the view of the adversary on which he bases his choice of siblings). The new adaptive notion introduces various non-trivial subtleties. For instance, unlike its non-adaptive counterpart, adaptive differential indistinguishability against any general choice of swap sets is impossible to realize while preserving correctness of the correlation finder. To see this, consider an adaptive choice that, given an answer on some query C, namely a correlation $z \leftarrow \mathsf{CF}^f(C)$, chooses to swap the function f at an input x that is called by $C^f(z)$. Swapping x possibly changes the outcome of the computation $C^f(z)$ making z no longer a correlation under the modified function f' and, therefore, no longer a "correct" answer for z. A successful CF will most likely not output z in such a case, practically implying that such a modification of f must cause CF to answer differently with high probability.

Fortunately, we are able to show that, unless our adaptive choice is that "targeted" (that is, chooses to swap inputs that specifically appear in the execution of the query circuit C on CF's answer), then adaptive differential indistinguishability is achievable. On the other hand, we prove that the choice to swap the set of siblings is never such a "targeted" choice under one condition in particular: that the execution of C on the answer $z \leftarrow \mathsf{CF}^f(C)$ does not observe a collision, namely does not make two f-queries that collide, with high probability over CF's randomness. Through these observations, we are able to generalize our correlation finder from above to satisfy the required adaptive notion against any choice of siblings. In particular, our new correlation finder looks at an analog of the amplification values $\alpha = \omega/\sigma$ that we define for *pairs* of inputs and, further, for the "soft" rejection considers the neighborhood of functions that are obtained by swapping between *sets* of inputs (rather than individual points). Overall, we get a correlation finder that is adaptively differentially indistinguishable against siblings, implying a similar separation of correlation intractable hash from collision resistance.

1.4 Paper Organization

We start by introducing some preliminaries in Section 2. In Section 3 we define the notions of fully black-box constructions and separations and in Section 4 we formally state our results. In Section 5 we present a generic framework for proving bounds on constructions of correlation intractability via the notion of (adaptive) differential indistinguishability and, in Section 6 we build a differentially indistinguishable correlation finder that satisfies our requirements. In Section 7 we connect all the peices together and derive our main theorems.

2 Preliminaries

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

Notation. For a distribution $\mathcal{X}$, we write $x \in \mathcal{X}$ to say that x is in the support of $\mathcal{X}$, and $x \leftarrow \mathcal{X}$ to denote that x is sampled from the distribution $\mathcal{X}$. We overload the notation for sets and write $x \leftarrow X$ when x is sampled uniformly at random from a set X. We use $\mathbb{P}(X)$ to denote the power set of X. For an event E, we use $\mathbb{1}(\mathsf{E})$ to denote the binary value which takes 1 if and only if E occurs. $\mathbf{SD}(\mathcal{X}, \mathcal{Y})$ 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 $\mathcal{A}^{\Psi}(x) \xrightarrow{\Psi} Q$ the event where $\mathcal{A}$, on input x, calls the oracle Ψ at Q. We extend this notation for tuples of Ψ-inputs: $\mathcal{A}^{\Psi}(x) \xrightarrow{\Psi} Q_1, \ldots, Q_n$ if $\mathcal{A}^{\Psi}(x)$ calls Q_i for *all* i.

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 $\mathcal{X}$,*

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

Concentration Bounds We hereby state two useful concentration bounds for distributions satisfying "nice" properties. We first recall Chebyshev's inequality.

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 $k \in \mathbb{R}$,*

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

Next is a tail bound from [GGKL21] on the sum of random variables that are "almost" k-wise independent.

Theorem 2.4 (Concentration for almost k-wise independence)**.** *Let $k \in \mathbb{N}$ and $p > 0$. Let $X_1, \ldots, X_n$ be binary random variables such that for any set $I \subseteq [n]$ of size at most k, $\Pr[\prod_{i \in I} X_i = 1] \leq p^{|I|}$. Then, for $c \geq 0$,*

$$\Pr\left[\sum_{i=1}^{n} X_i \geq c \cdot n\right] \leq \min_{0 \leq i \leq k} \left(\frac{p^i \binom{n}{i}}{\binom{cn}{i}}\right).$$

The Borel-Cantelli Lemma. We recall the Borel-Cantelli lemma.

Proposition 2.5 (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.*

A Useful Combinatorial Proposition. In the following simple proposition, we argue that the images of any t inputs under a random order-2 can be thought of as sampled uniformly at random for any practical purpose.

Proposition 2.6. *For any fixed $\ell \in \mathbb{N}$ and distinct $x_1, \ldots, x_t \leftarrow \{0,1\}^{\ell}$, letting $\pi \leftarrow \mathbf{Sym}_2(\{0,1\}^{\ell})$ be a uniformly random order-2 permutation and $y_1, \ldots, y_t$ be uniform over $\{0,1\}^{\ell}$, it holds that*

$$\mathbf{SD}((y_1, \ldots, y_t), (\pi(x_1), \ldots, \pi(x_t))) = O(t^2 \cdot 2^{-\ell}).$$

Proof. Observe that $\pi(x_1), \ldots, \pi(x_t)$ are uniformly random conditioned on $\pi(x_i) \neq \pi(x_j)$ for any $i \neq j$ and $\pi(x_i) = x_j \implies \pi(x_j) = x_i$. These two conditions may be violated by t uniformly random images only in the event where $\pi(x_i) \in \{x_j, \pi(x_j)\}$ for some $i \neq j$. This occurs with probability $O(t^2 \cdot 2^{-\ell})$ and, therefore, the proposition may be derived by Proposition 2.2 via the straight-forward coupling. □

3 Defining Black-box Constructions and Separations

To capture both separations of correlation intractability from OWP and CRH, we define black-box constructions (and separations) of CIH from any intractability assumption following a typical syntax. To that end, we first define a computational task over an idealized cryptographic primitive, namely over an oracle.

Definition 3.1 (Computational Task over Oracle)**.** *We define a* computational task $\mathsf{T} = (\mathcal{D}, \mathbf{A}, V)$ over oracles of the form $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$ *to consist of the following three components:*

- *A* challenge distribution $\mathcal{D}^f = \{\mathcal{D}^f_\lambda\}$ *where, for any* $\lambda \in \mathbb{N}$*,* $\mathcal{D}^f_\lambda$ *is an oracle-aided distribution over* $\{0,1\}^{r(\lambda)}$.
- *An* answer space $\mathbf{A} = \{\mathbf{A}_\lambda\}$ *where, for any* $\lambda \in \mathbb{N}$*,* $\mathbf{A}_\lambda \subseteq \{0,1\}^{s(\lambda)}$.
- *A* winning condition $V^f = \{V^f_\lambda\}$ *where, for any* $\lambda \in \mathbb{N}$*,* V^f_λ *is an oracle-aided relation over* $\{0,1\}^{r(\lambda)} \times \{0,1\}^{s(\lambda)}$.

For any oracle-aided adversary $\mathcal{A}$, we say that $\mathcal{A}$ has advantage $\alpha := \alpha(\lambda)$ *in* T *under a given f if it outputs an answer that satisfies the winning condition with probability at least α, that is, if*

$$\mathbf{Adv}^f_{\mathsf{T}}(\lambda, \mathcal{A}) := \Pr_{c \leftarrow \mathcal{D}_\lambda}[(c, \mathcal{A}^f(c)) \in V^f_\lambda] \geq \alpha(\lambda)$$

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

We now define a fully black-box construction of correlation intractability from any idealized building block over which a given computational task T is assumed to be intractable. We will later choose T to be the task of inverting a given oracle to obtain a separation from OWP and the task of finding collisions under a given oracle to obtain a separation from CRH.

Definition 3.2 (Fully Black-box Construction of CIH from T-Intractable Functions)**.** *Let $m := m(n)$ and $\ell := \ell(\lambda)$ be length parameters, let $\mathcal{R}$ be a class of relations and let T be a computational task over oracles of the form $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$ (as defined in Definition 3.1). A* (t, q, ϵ)-fully black-box construction of Correlation Intractable Hash (CIH) for $\mathcal{R}$ with input length m from T-intractable functions*, for* $t := t(n)$*, $q := q(\lambda)$ and $\epsilon := \epsilon(\lambda)$, is an ensemble of distributions $H = \{H_n\}$ where, for any $n \in \mathbb{N}$, H_n is a distribution over functions mapping m-bit inputs to n-bit outputs, and an oracle-aided reduction $\mathcal{M}$ satisfying the following properties:*

- ***Construction Efficiency:*** *For any $n \in \mathbb{N}$ and any $h \in H_n$, h^f makes at most $t(n)$ queries to f on any input.*
- ***Black-box Security Reduction:*** *For any oracle $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$ and any probabilistic oracle-aided adversary $\mathcal{A}$, if there exists a relation $R \in \mathcal{R}$ such that*

  $$\mathbf{Adv}^H_{\mathbf{ci}}(n, R, \mathcal{A}) := \Pr_{\substack{h \leftarrow H_n \\ z \leftarrow \mathcal{A}^f(1^n, h)}}[(z, h^f(z)) \in R] > \frac{1}{2}$$

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

  $$\mathbf{Adv}^f_{\mathsf{T}}(\lambda, \mathcal{M}^{\mathcal{A}}) \geq \epsilon(\lambda)$$

  for infinitely many $\lambda \in \mathbb{N}$.
- ***Reduction Efficiency:*** *For any $\lambda \in \mathbb{N}$ and $y \in \{0,1\}^\lambda$, $\mathcal{M}^{f,\mathcal{A}}(y)$ makes at most $q(\lambda)$ queries to the oracles f and $\mathcal{A}$, and for every $\mathcal{A}$-query $(1^n, h)$ made by $\mathcal{M}(y)$, it holds that $n < 2^{\lambda/24}$ and $h^f(\cdot)$ makes at most $q(\lambda)$ queries to f on any input.*

Lastly, we define a fully black-box α-separation to embody the impossibility of any fully black-box construction abiding a trade-off (parameterized by α) between the complexity of the underlying reduction and its success probability. A larger value of α gives stronger separation and, in particular, superpolynomial α indicates the impossibility of a reduction that is both polynomial time and has non-negligible advantage, as typically required in the traditional cryptographic setting.

Definition 3.3 (Black-box Separation of CIH from T-Intractability)**.** *Let $m := m(n)$ and $\ell := \ell(\lambda)$ be length parameters and let $\mathcal{R}$ be a class of relations. We say that* t-bounded CIH functions for $\mathcal{R}$ (with input length m) are $\alpha(\lambda)$-fully black-box separated *from T-intractable functions (with input length ℓ), for $t := t(n)$ and $\alpha(\lambda) > 1$, if for any (t, q, ϵ)-fully black-box construction of such CIH from T-intractability, it holds that either*

1. *$q(\lambda) > O(\alpha(\lambda))$, or*
2. *$\epsilon(\lambda) \leq O(1/\alpha(\lambda))$.*

4 Our Results: Statement of Main Theorems

In this section we formally state our separation results of CIH from CRH and OWP. Let us first make few necessary definitions.

Definition 4.1 (Inversion and Collision-Finding Tasks)**.** *We will be particularly interested in two special cases of computational tasks over oracles $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$.*

- ***Inversion:*** *We define the* inversion task $\mathsf{Inv} = (\mathcal{D}, \mathbf{A}, V)$*, where*
  - $\mathcal{D}_\lambda = f(x^*)$ *for a random* x^**,*
  - $\mathbf{A}_\lambda = \{0,1\}^\lambda$*, and*
  - $V^f_\lambda = \{(x, y) \mid f_\lambda(x) = y\}$*.*
- ***Collision Finding:*** *We define the* collision-finding task $\mathsf{CollFind} = (\mathcal{D}, \mathbf{A}, V)$*, where*
  - $\mathcal{D}_\lambda = 1_\lambda$*,*
  - $\mathbf{A}_\lambda = \{0,1\}^\lambda \times \{0,1\}^\lambda$*, and*
  - $V^f_\lambda = \{(1^\lambda, (x_1, x_2)) \mid x_1 \neq x_2,\ f_\lambda(x_1) = f_\lambda(x_2)\}$*.*

Our impossibility result rules out any construction of a CIH for relation classes that, roughly speaking, constitute complexity greater than the black-box complexity of the construction. To articulate the complexity of a given relation class, we refer to the degree of "unpredictability" induced by a random relation. More specifically, we say that a relation is k-wise universal if, in particular, the likelihood of any (z, w) to be in a random relation does not change even given the membership (or non-membership) in the relation of any $k - 1$ pairs. We formalize below.

Definition 4.2 (k-wise Universal Relations)**.** *Let $k : \mathbb{N} \to \mathbb{N}$ and $p : \mathbb{N} \to (0,1)$. We say that a relation class $\mathcal{R} = \{\mathcal{R}_n \subseteq \mathbb{P}(\{0,1\}^m \times \{0,1\}^n)\}$, for $m := m(n)$ is* k-wise p-universal *if, for any $n \in \mathbb{N}$, there exists a distribution over relations in $\mathcal{R}_n$ (which we ambiguously denote by $\mathcal{R}_n$) such that for any $k' \leq k$ and any distinct $(z_1, w_1), \ldots, (z_{k'}, w_{k'}) \in \{0,1\}^m \times \{0,1\}^n$, it holds that*

$$\Pr_{R \leftarrow \mathcal{R}_n}[(z_i, w_i) \in R \quad \forall i \in [k']] = p(n)^{k'(n)}.$$

We now formally state our main separation theorems: In any fully-black-box construction of CIH against a k-wise independent relation class from CRH (or OWP) with non-trivial security, the hash function must invoke the underlying CRH (or OWP) at least $\Omega(k)$ times in its computation. We provide the formal theorems with accurate quantitative details below.

Theorem 4.3 (Black-box Separation of CIH from OWP)**.** *Let $m := m(n)$ and $p : \mathbb{N} \to [0,1]$ be such that $p(n) \geq 4n^2 2^{-m(n)}$ and let $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > 20 \cdot t(n)$ for all $n \in \mathbb{N}$. Then, t-bounded CIH functions, with input length m, for any class of k-wise p-universal relations are $2^{\lambda/10}$-fully black-box separated from OWP.*

Theorem 4.4 (Black-box Separation of CIH from CRH)**.** *Let $m := m(n)$ and $p : \mathbb{N} \to [0,1]$ be such that $p(n) \geq 4n^2 2^{-m(n)}$ and let $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > 25 \cdot t(n)$ for all $n \in \mathbb{N}$. Then, t-bounded CIH functions, with input length m, for any class of k-wise p-universal relations are $2^{\lambda/25}$-fully black-box separated from CRH mapping $\ell(\lambda) = \lambda + O(1)$ bits to λ bits.*

Note that our separation result from collision-resistance considers CRH that shrinks its input only by a constant number of bits. We stress, however, that our proof technique results in equally-merited separations from any CRH with constant multiplicative shrinkage smaller than $\frac{1}{2}$, where we still require $k = \Omega(t)$ and obtain a $2^{\Omega(\lambda)}$-separation. Further, since such a CRH implies collision-resistance with any polynomial shrinkage (via a logarithmic number of sequential invocations), one may derive more general separation results with corresponding parameters.

5 A Generic Framework: Differentially Indistinguishable Correlation Finder

We introduce a generic framework for showing barriers on CIH constructions. Our approach builds on the "Two-Oracle Methodology" [Sim98, HR04, AS16, BD19] where, in order to obtain bounds on cryptographic constructions, one creates an idealized (oracle-relative) world under which such constructions are impossible. In our case, such a world would consist mainly of an ideal oracle representation of a cryptographic primitive (be it a CRH or OWP) and a correlation finder that should be able to break any construction of CIH from the ideal primitive that satisfies certain constraints, e.g. query complexity, yet is useless for breaking the intractability of the underlying oracle (that is, inverting it or finding induced collision).

Our contribution in this section is the formulation of a somewhat unified hardness notion, namely differential indistinguishability, and show that any correlation finder that satisfies it is indeed useless breaking the ideal OWP or CRH. We believe that our approach is sufficiently modular to allow for adaptation in different settings. While we attempt to be as general as possible in our representation, we sometimes adhere to specificity for the sake of brevity and cleanliness.

Setting and Notation. We start by fixing the notation that will be used throughout our proof. Our proof will be centered around two "computational games": In the first, a correlation finder aims to break the correlation intractability of a candidate CIH that maps $m := m(n)$ bits to n bits (n can be thought of as the "security parameter" in this game). In the second game, an adversary is given access to the correlation finder and aims to break the intractability of an idealized primitive that is given as an oracle mapping $\ell := \ell(\lambda)$ bits to λ bits (here, λ is the security parameter). We now list the main playing parts in this settings:

- A *relation class* $\mathcal{R} = \{\mathcal{R}_n\}$ where, for any $n \in \mathbb{N}$, $\mathcal{R}_n$ is a class of relations over $\{0,1\}^{m(n)} \times \{0,1\}^n$. This will denote a relation class which we seek to build (actually, rule out) correlation intractability for.
- A (random) oracle $\mathcal{F} = \{\mathcal{F}_\lambda\}$ where for any $\lambda \in \mathbb{N}$, $\mathcal{F}_\lambda$ is the uniform distribution over regular functions mapping $\ell(\lambda)$-bit inputs to λ-bit outputs (for our purposes, we will always assume that $\ell(\lambda) \geq \lambda$ for any $\lambda \in \mathbb{N}$). This will represent the *idealized cryptographic building block* (OWP or CRH in our case) from which correlation intractable hash is to be constructed separated. We will typically use f to denote a function chosen from $\mathcal{F}$. We say that an algorithm is f-aided if it is given access to an oracle that follows the syntax of $f \in \mathcal{F}$.

- The family of f-aided circuits $\mathcal{C} = \{\mathcal{C}_n\}$ where, for any $n \in \mathbb{N}$, $\mathcal{C}_n$ is the set of all f-aided circuits mapping $m(n)$-bit inputs to n-bit inputs. In particular, a *CIH candidate for $\mathcal{R}$* (from the idealized primitive represented by $\mathcal{F}$) is an ensemble of f-aided circuits $C = \{C_n\}$ where $C_n \in \mathcal{C}_n$ for all $n \in \mathbb{N}$.

- A *correlation finder* $\mathcal{O} = \{\mathcal{O}_R\}$ where, for any $R \in \mathcal{R}$, $\mathcal{O}_R$ is a distribution over f-aided oracles that on input 1^n and an f-aided circuit $C \in \mathcal{C}_n$ (for any $n \in \mathbb{N}$) outputs a C-input of length $m(n)$ bits (which should be a correlation w.r.t. R if successful). We often omit the input 1^n as it is clearly determined by C and sometimes omit the relation R when it is irrelevant in the context.

- An *oracle-aided adversary* $\mathcal{A}$ that is given access to an oracle $f \in \mathcal{F}$ and a correlation finder $\mathsf{O} \in \mathcal{O}$ (which in turn has access to f). We say that $\mathcal{A}$ is a T-adversary if it follows the syntax dedicated by a computational task T. In particular, we will be interested in Inv-adversaries, which take as input an image $y \in \{0,1\}^\lambda$ and return a pre-image $x \in \{0,1\}^{\ell(\lambda)}$ and in $\mathsf{CollFind}$-adversaries which take as input a security parameter 1^λ and return a pair of inputs $(x_1, x_2) \in \{0,1\}^{\ell(\lambda)} \times \{0,1\}^{\ell(\lambda)}$.

5.1 The Two-Oracle Methodology

We now recall the methodology developed in prior separation results [Sim98, HR04, AS16, BD19] and adapt it to separations of correlation intractability from general T-intractability (where T is any computational task T – see Definition 3.1).

Definition 5.1 ((q, q', q'')-Bounded Adversary)**.** *Let $q, q', q'' : \mathbb{N} \to \mathbb{N}$. We say that an oracle-aided adversary $\mathcal{A}$ is* (q, q', q'')-bounded *if, for any fixed correlation finder O and any $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$, $\mathcal{A}^{f,\mathsf{O}^f}$ on any input with security parameter λ makes at most $q(\lambda)$ queries to f_λ and $q'(\lambda)$ queries to O, where each O-query C makes at most $q''(\lambda)$ queries to f_λ on any input.*

Lemma 5.2 (Separation via Correlation-Finding Oracle)**.** *Let $\kappa : \mathbb{N} \to [0,1]$ and $c \in \mathbb{N}$ be a constant. Let $\mathcal{R}$ be a class of relations. Assume there exists a correlation finder $\mathcal{O} = \{\mathcal{O}_R\}_{R \in \mathcal{R}}$ (see* Setting and Notation *above), such that*

– ***$\mathcal{O}$ breaks all CIH (Correctness):*** *For any CIH candidate $H = \{H_n\}$ with query complexity bounded by $t(n)$ (i.e. which satisfies the construction efficiency property in Definition 3.2), it holds that*

$$\mathbb{E}_{\substack{R \leftarrow \mathcal{R} \\ \mathsf{O} \leftarrow \mathcal{O}_R}}[\mathbf{Adv}^H_{\mathbf{ci}}(n, R, \mathsf{O})(n)] = \Pr_{\substack{f \leftarrow \mathcal{F}, h \leftarrow H_n \\ R \leftarrow \mathcal{R}, \mathsf{O} \leftarrow \mathcal{O}_R}}[z \leftarrow \mathsf{O}^f(1^n, h);\ (z, h^f(z)) \in R] > 1 - \frac{1}{2n^2}$$

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

– ***$\mathcal{F}$ is T-Intractable under $\mathcal{O}$ (Security):*** *For any (q, q, q)-bounded $\mathcal{A}$ that on security parameter λ calls O only with queries $(1^n, h)$ s.t. $n < 2^{\lambda/24}$, any $\lambda \in \mathbb{N}$, and any $R \in \mathcal{R}$,*

$$\mathbb{E}_{\substack{f \leftarrow \mathcal{F} \\ \mathsf{O} \leftarrow \mathcal{O}_R}}[\mathbf{Adv}^f_{\mathsf{T}}(\lambda, \mathcal{A}^{\mathsf{O}})] = O(q(\lambda)^c \cdot \kappa(\lambda)).$$

Then, t-bounded CIH functions for $\mathcal{R}$ with input length m are $(\kappa^{-1/(1+c)})$-fully black-box separated from T-intractability.

Proof. Let q, ϵ be such that $q = O(\kappa^{-1/(1+c)})$ and $\epsilon > O(\kappa^{1/(1+c)})$. Fix a candidate (t, q, ϵ)-construction of correlation intractable hash from T-intractable functions, consisting of hash $H = \{H_n\}$ and reduction $\mathcal{M}$ that satisfy the two efficiency properties from Definition 3.3 w.r.t. t and q. To establish separation, we show that there exist $f \in \mathcal{F}$ and $R \in \mathcal{R}$ such that $\mathsf{O}^f \leftarrow \mathcal{O}_R$ breaks correlation intractability with probability at least $1/2$ yet the (q, q, q)-reduction $\mathcal{M}$ cannot invert f given access to such an O with probability bigger than ϵ.

We know, by the first assumption, that

$$\Pr_{\substack{f\leftarrow\mathcal{F},h\leftarrow H_n\\ R\leftarrow\mathcal{R},\mathsf{O}\leftarrow\mathcal{O}_R}}[z\leftarrow \mathsf{O}^f(1^n,h);\ (z,h^f(z))\notin R] < \frac{1}{2n^2}$$

for infinitely many $n\in\mathbb{N}$. Further, $\mathcal{M}^{f,\mathsf{O}^f}$ by the reduction efficiency of $\mathcal{M}$, it is a (q,q,q)-adversary and, therefore, by the second assumption in the lemma, it holds that

$$\Pr_{\substack{f\leftarrow\mathcal{F}\\ R\leftarrow\mathcal{R},\mathsf{O}\leftarrow\mathcal{O}_R\\ x^*\leftarrow\{0,1\}^\ell,\mathcal{M}}}[\mathcal{M}^{f,\mathsf{O}^f}(1^\lambda,f(x^*))=x^*] = O(q(\lambda)^c\kappa(\lambda)/\lambda^2).$$

For any $f\in\mathcal{F}$ and $R\in\mathcal{R}$, we define, for any $n\in\mathbb{N}$,

$$\alpha_{\mathbf{ci}}(n,f,R) = \Pr_{\substack{\mathsf{O}\leftarrow\mathcal{O}_R\\ h\leftarrow H_n}}[z\leftarrow \mathsf{O}^f(1^n,h);(z,h^f(z))\in R]$$

and, for any $\lambda\in\mathbb{N}$,

$$\alpha_{\mathsf{T}}(\lambda,f,R) = \Pr_{\substack{\mathsf{O}\leftarrow\mathcal{O}_R\\ x^*\leftarrow\{0,1\}^\ell,\mathcal{M}}}[\mathcal{M}^{f,\mathsf{O}^f}(1^\lambda,f(x^*))=x^*].$$

Then, by the above, we know that, for infinitely many $n\in\mathbb{N}$ and any $\lambda\in\mathbb{N}$, it holds that

$$\mathbb{E}_{f\leftarrow\mathcal{F},R\leftarrow\mathcal{R}}[1-\alpha_{\mathbf{ci}}(n,f,R)] < 1/n^2 \quad\text{and}\quad \mathbb{E}_{f\leftarrow\mathcal{F},R\leftarrow\mathcal{R}}[\alpha_{\mathsf{T}}(\lambda,f,R)] < O(q(\lambda)^c\kappa(\lambda)/\lambda^2)$$

and, by Markov, we can imply that

$$\Pr_{f,R}[\alpha_{\mathbf{ci}}(n,f,R)<1/2] = \Pr_{f,R}[1-\alpha_{\mathbf{ci}}(n,f,R)>1/2] < 1/n^2 \quad\text{and}\quad \Pr_{f,R}[\alpha_{\mathsf{T}}(\lambda,f,R)>q^c\kappa] < O(1/\lambda^2).$$

We now use the Borel-Cantelli lemma over the above tail bounds, that hold for all $\lambda\in\mathbb{N}$ and infinitely many $n\in\mathbb{N}$ to derive that, with high probability, there are infinitely many $n\in\mathbb{N}$ and, respectively, infinitely many $\lambda\in\mathbb{N}$, for which $\alpha_{\mathbf{ci}}(n,f,R)>\frac{1}{2}$ yet $\alpha_{\mathsf{T}}(\lambda,f,R)<q^c\kappa$.

More formally, let us denote by $B_{\mathbf{ci}}(n,f,R)$ the event that $\alpha_{\mathbf{ci}}(n,f,R)<\frac{1}{2}$ and by $N=\{n_1,n_2,\dots\}\subseteq\mathbb{N}$ the infinite set of all n for which $\Pr[B_{\mathbf{ci}}(n,f,R)]<1/n^2$. We also denote by $B_{\mathsf{T}}(n,f,R)$ the event that $\alpha_{\mathsf{T}}(\lambda,f,R)>q^c\kappa$. Then, since $\sum_{i\in\mathbb{N}}\Pr[B_{\mathbf{ci}}(n_i,f,R)]<\sum_{n\in N}1/n^2$ and $\sum_{\lambda\in\mathbb{N}}\Pr[B_{\mathsf{T}}(\lambda,f,R)]<\sum_{\lambda\in\mathbb{N}}1/\lambda^2$ converge, then by Proposition 2.5, it holds that, for any $\eta>0$, that

$$\Pr[B_{\mathbf{ci}}(n,f,R)\text{ for infinitely many }n\in N]<\eta$$
$$\Pr[B_{\mathbf{inv}}(\lambda,f,R)\text{ for infinitely many }\lambda\in\mathbb{N}]<\eta.$$

Let $B^*_{\mathbf{ci}}(f,R)$ be the event that $B_{\mathbf{ci}}(n,f,R)$ occurs for infinitely many $n\in N$ and $B^*_{\mathbf{inv}}(f,R)$ be the event that $B_{\mathbf{inv}}(\lambda,f,R)$ occurs for infinitely many $\lambda\in\mathbb{N}$. Notice that it is sufficient to show that there exist $f\in\mathcal{F}$ and $R\in\mathcal{R}$ for which neither $B^*_{\mathbf{ci}}(f,R)$ nor $B^*_{\mathbf{inv}}(f,R)$ occur since this means that for infinitely many $n\in N$, $\mathbf{Adv}^H_{\mathbf{ci}}(n,R,\mathsf{O})(n)>\frac{1}{2}$ and, at the same time, for all but finitely many $\lambda\in\mathbb{N}$, $\mathbf{Adv}_{\mathsf{T}}(\lambda,\mathcal{M}^{\mathsf{O}})<q(\lambda)^c\kappa(\lambda)=O(\kappa^{1/(1+c)})<\epsilon$. This contradicts $(H,\mathcal{M})$ being a (t,q,ϵ)-fully black-box construction. Consequently, it would suffice to show that with non-zero probability over the choice of f and R, neither of the two events occur. This follows immediately from the analysis above as follows

$$\Pr_{f,R}[\overline{B^*_{\mathsf{T}}(f,R)}\wedge\overline{B^*_{\mathbf{ci}}(f,R)}]\geq 1-(\Pr_{f,R}[B^*_{\mathsf{T}}(f,R)]+\Pr[B^*_{\mathbf{ci}}(f,R)])>1-2\eta>0.$$

□

5.2 Generic Assumptions on the Candidate, Adversary and the Correlation Finder

To facilitate our proof, we will make few assumption over the structure and behavior of both the correlation finder and the oracle-aided adversary attacking the idealized base primitive. Some of these assumptions evidently hold without loss of generality, while some will require a little effort to manifest generally.

The first assumption we make is from the former type, and will be immediately satisfied by our constructions later on. We will be assuming that the correlation finder answers each of its queries using independent randomness. This is formally captured by the following definition.

Definition 5.3 (Query-Independent Oracle)**.** *We say that a distribution $\mathcal{O} : \mathcal{Q} \to \mathcal{Z}$ over oracles is* query-independent *if the answers of a random oracle $\mathsf{O} \leftarrow \mathcal{O}$ to different queries are independent or, more formally, if*

$$(\mathsf{O}(Q))_{Q\in\mathcal{Q}} \equiv (\mathsf{O}_Q(Q))_{Q\in\mathcal{Q}}$$

where O and $\{\mathsf{O}_Q\}_{Q\in\mathcal{Q}}$ are all sampled independently at random from $\mathcal{O}$.

Next, we will be assuming w.l.o.g. that our candidate H *always* makes exactly t queries on any input. Another simple assumption that we make on the candidate construction of CIH is in the following remark.

Remark 5.4. *We may consider, w.l.o.g., only constructions H that never call f_λ for λ that is too small to give any security. More specifically, we may assume from this point on that, for any $n \in \mathbb{N}$, H_n calls f_λ only if $\lambda > 24 \log n$.*

Proof. The above assumption holds without loss of generality since, were there a correlation finder $\mathcal{O}$ that breaks only such candidate constructions, there exists a correlation finder $\mathcal{O}'$ that breaks any candidate and does not make inversion of f any easier. More specifically, $\mathcal{O}'$ on input a hash function h embeds $f_{\leq} = \{f_\lambda\}_{\lambda\leq 24\log n}$ in the function's circuit (note this is of polynomial size $O(n^{24}\log n)$) to obtain a hash h' where any f_λ-query for $\lambda \leq 24\log n$ is replaced by a lookup in the corresponding hardwired function. Success of $\mathcal{O}'$ follows from the success of $\mathcal{O}$ since h and h' compute the same function under f. Further, if there exists a (q,q,q)-bounded $\mathcal{A}$ such that $\mathcal{A}^{f,\mathsf{O}'}$ is successful in T when $\mathsf{O}' \leftarrow \mathcal{O}'$ then so is $\mathcal{A}^{f,\mathcal{O}}$ for when $\mathsf{O} \leftarrow \mathcal{O}$ by the limitation on $\mathcal{A}$'s queries which guarantees they are answered similarly by $\mathcal{O}$ and $\mathcal{O}'$. □

The rest of the assumptions concern the adversary against the base primitive, which can be any Inv- or $\mathsf{CollFind}$-adversary. We begin with the notion of a *canonical* adversary, which captures a general structure of successful adversaries in the sense that any adversary can be easily made canonical without loss in its advantage or complexity.

Definition 5.5 (Canonical Adversary)**.** *We say that an oracle-aided adversary $\mathcal{A}$ against* Inv *or* $\mathsf{CollFind}$ *is* canonical *if it satisfies the following three properties for any $\mathsf{O} = \{\mathsf{O}_n : \mathcal{C}_{m(n),n} \to \{0,1\}^n\}$ and any f:*

(i) *$\mathcal{A}^{f,\mathsf{O}^f}$ never makes the same oracle query twice.*

(ii) *After any O-query C that $\mathcal{A}^{f,\mathsf{O}^f}$ makes, $\mathcal{A}^{f,\mathsf{O}^f}$ immediately calls f at any x such that $C^f(\mathsf{O}(C)) \rightharpoondown x$.*

(iii) *$\mathcal{A}^{f,\mathsf{O}^f}$ immediately halts and outputs answer if found: In the case of* Inv*, this means if $\mathcal{A}^{f,\mathsf{O}^f}(y^*)$ calls f at some $x \in f^{-1}(y^*)$, he outputs x^* right afterwards, while in the case of* $\mathsf{CollFind}$*, if $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda)$ calls f at some $(x_1,x_2) \in \mathsf{Coll}^f$, (in two separate queries) he outputs the collision immediately after making the second query of the two.*

(iv) *$\mathcal{A}^{f,\mathsf{O}^f}$ always calls f at its final output(s): In the case of* Inv*, this means $\mathcal{A}^{f,\mathsf{O}^f}(y^*) = x$ implies $\mathcal{A}^{f,\mathsf{O}^f}(y^*) \overset{f}{\rightharpoondown} x$, while in the case of* $\mathsf{CollFind}$*, if $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda) = (x_1,x_2)$ then $\mathcal{A}^{f,\mathsf{O}^f}(y^*) \overset{f}{\rightharpoondown} x_i$ for both $i \in \{1,2\}$.*

Crucial to our analysis is one more assumption over the adversary, namely that it is *smooth.* Conceptually, a smooth adversary never calls the correlation finder with queries that already convey sufficient amount of information for succeeding in its task (without the help of the correlation finder). Specifically, these are queries C where the transcript of $C^f(z)$ for a random input z may contain a correct solution with noticeable probability. In the case of inversions, we require that any pre-image x is observed by $C^f(z)$ with negligible probability[5] while, in the case of $\mathsf{CollFind}$, we require that $C^f(z)$ does not observe a collision – neither a "sibling" of a previously observed input nor a colliding pair of inputs.

To formally capture the notion of smoothness, we first define it for sets of f-inputs (these will be later specified to be the "solution" input sets).

Definition 5.6 (Smooth Input Sets)**.** *Fix an oracle f and an f-aided circuit $C \in \mathcal{C}_n$ and let $\gamma : \mathbb{N} \to [0,1]$. We define the family of* γ-smooth input sets w.r.t. C and f *as follows*

$$\mathsf{Smooth}^f_\gamma(C) = \{X \subset \{0,1\}^{\ell(\lambda)} \mid \lambda \in \mathbb{N}, \Pr_{z \leftarrow \{0,1\}^m}[\exists x \in X : \ C^f(z) \rightharpoondown x] \leq \gamma(\lambda)\}.$$

We will further consider the following notion of collision-smoothness, which will be additionally needed when separating from collision-resistance.

Definition 5.7 (Collision-Smooth Circuits)**.** *We say that an f-aided circuit $C : \{0,1\}^m \to \{0,1\}^n$ is* (λ, f, γ)-collision-smooth*, for $f \in \mathcal{F}$, $\lambda \in \mathbb{N}$ and $\gamma \in [0,1]$, if*

$$\Pr_{z \leftarrow \{0,1\}^m}[\exists (x_1, x_2) \in \mathsf{Coll}^f_\lambda : \ C^f(z) \rightharpoondown x_1, x_2] \leq \gamma,$$

where

$$\mathsf{Coll}^f_\lambda = \{(x_1, x_2) \in \{0,1\}^{2 \times \ell(\lambda)} \mid x_1 \neq x_2, \ f(x_1) = f(x_2)\}.$$

Further, for a function $\gamma : \mathbb{N} \to [0,1]$, we denote the family of γ-collision-smooth circuits by

$$\mathcal{C}^*_\gamma = \{\mathcal{C}^*_{\lambda,f,\gamma}\},$$

where, for any $\lambda \in \mathbb{N}$ and $f \in \mathcal{F}$, $\mathcal{C}^*_{\lambda,f,\gamma}$ is the set of all $(\lambda, f, \gamma(\lambda))$-collision-smooth circuits.

To define smoothness for adversaries, we must specify the target input sets for which we will require smoothness. While the target sets for inverting a permutation is straight-forward (it is simply the pre-image of any possible challenge), the target set for finding collisions is not merely the set of all collisions, but also the set of all inputs that collide with f-queries observed by the adversary in its execution.

We begin by defining the function $\mathbf{sib}^f = \{\mathbf{sib}^f_\lambda\}$ that maps any input x to its siblings set under the given function f, i.e.

$$\mathbf{sib}^f_\lambda(x) = \{x' \in \{0,1\}^{\ell(\lambda)} \mid x' \neq x, f_\lambda(x) = f_\lambda(x')\}. \tag{1}$$

Next, for any fixed $f \in \mathcal{F}$, $\mathsf{O} \in \mathcal{O}$, and deterministic oracle-aided adversary $\mathcal{A}$ (we override notation and use $\mathcal{A}$ also to denote the randomness sampled for the adversary, or, equivalently, the random instance of the adversary's deterministic machine when sampling and fixing its randomness), we denote the set of f-queries made by $\mathcal{A}^{f,\mathsf{O}^f}$ by

$$\mathbf{Q}_\lambda(f, \mathsf{O}, \mathcal{A}) = \{x \in \{0,1\}^{\ell(\lambda)} \mid \mathcal{A}^{f,\mathsf{O}^f}(1^\lambda) \overset{f}{\rightharpoondown} x\}. \tag{2}$$

Lastly, we define the function $\mathbf{Sib}_\lambda = \{\mathbf{Sib}_\lambda\}$ that maps any $f \in \mathcal{F}$, $\mathsf{O} \in \mathcal{O}$ and deterministic $\mathcal{A}$ to the set of siblings of all f-queries made by $\mathcal{A}$ in its execution, except for the last query (recall that the last query made by any canonical adversary is an f-query). That is,

$$\mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A}) = \bigcup_{i=1}^{q-1} \mathbf{sib}^f_\lambda(Q_i), \qquad \text{where} \ \ (Q_1, \ldots, Q_q) = \mathbf{Q}_\lambda(f, \mathsf{O}, \mathcal{A}). \tag{3}$$

We are now prepared to define smooth adversaries, both against inversion and against collision resistance.

[5] In fact, we require this only for x that has not been already observed by $\mathcal{A}$ since, by canonicality, if an x has already been observed then this it is not the solution. Notice that, otherwise, the correlation finder has no knowledge about the identity of the targeted pre-image and, therefore, we quantify over all such x's.

Definition 5.8 (Smooth Adversary)**.** *Let $\tau, \gamma : \mathbb{N} \to [0,1]$. We define the following notions of smoothness for oracle-aided adversaries:*

- *We say that an oracle-aided adversary is (τ,γ)-*smooth at unobserved inputs *if, for any f, any fixed correlation finder O, any $\lambda \in \mathbb{N}$, any $\mathcal{A}$-input a, and any $i \in \mathbb{N}$, letting C_i be the (random) i^{th} O-query made by $\mathcal{A}^{f,\mathsf{O}}(a)$ and $X_{<i} \subset \{0,1\}^{\ell(\lambda)}$ be the set of all f_λ-queries made by $\mathcal{A}$ prior to C_i, it holds that*
$$\Pr_{\mathcal{A}}[\forall x \in \{0,1\}^{\ell(\lambda)} \setminus X_{<i},\ \{x\} \in \mathsf{Smooth}^f_\gamma(C_i)] > \tau(\lambda).$$
- *We say that an oracle-aided adversary $\mathcal{A}$ is (τ,γ)-*smooth at siblings *if, for any f, any fixed correlation finder O, any $\lambda \in \mathbb{N}$, any $\mathcal{A}$-input a, and any $i \in \mathbb{N}$, letting C_i be the (random) i^{th} O-query made by $\mathcal{A}^{f,\mathsf{O}}(a)$, it holds that*
$$\Pr_{\mathcal{A}}[\mathbf{Sib}_\lambda(f,\mathsf{O},\mathcal{A}) \in \mathsf{Smooth}^f_\gamma(C_i)] > \tau(\lambda).$$
- *We say that an oracle-aided adversary $\mathcal{A}$ is (τ,γ)-*smooth at collisions *if for any f, any fixed correlation finder O, any $\lambda \in \mathbb{N}$, any $\mathcal{A}$-input a, and any $i \in \mathbb{N}$, letting C_i be the (random) i^{th} query made by $\mathcal{A}^{f,\mathsf{O}}(a)$, it holds that*
$$\Pr_{\mathcal{A}}[C \in \mathcal{C}^*_{\lambda,f,\gamma}] \geq \tau(\lambda).$$

Further, we say that an Inv*-adversary is simply* (τ,γ)-smooth *if it is (τ,γ)-smooth at unobserved inputs and that a* $\mathsf{CollFind}$*-adversary is* (τ,γ)-smooth *if it is (τ,γ)-smooth at siblings* and *(τ,γ)-smooth at collisions.*

We note that the notion of smooth Inv-adversaries at singleton sets is considered already in [BD19] with the difference that we require smoothness only for unobserved inputs – this simplifies the adaption to correlation intractability and is w.l.o.g. assuming canonical adversaries.[6]

Having characterized canonical and smooth adversaries, we proceed to showing that such adversaries are complete, in the sense that it would be sufficient to show intractability against them if we can tolerate a small cost in complexity.

Lemma 5.9 (The Smoothening Lemma)**.** *For any canonical (q,q',q'')-bounded* Inv*- or* $\mathsf{CollFind}$*-adversary $\mathcal{A}$ and any $\beta := \beta(\lambda)$, there exists a canonical $(q+\beta q'q'', q', q'')$-bounded* Inv*- or, resp.,* $\mathsf{CollFind}$*-adversary $\mathcal{B}$ such that the following two properties hold:*

- ***Correctness:*** *for any fixed correlation finder O, any f, and any $\lambda \in \mathbb{N}$,*
$$\mathbf{Adv}^f_{\mathsf{T}}(\lambda, \mathcal{B}^{\mathsf{O}}) \geq \mathbf{Adv}^f_{\mathsf{T}}(\lambda, \mathcal{A}^{\mathsf{O}})$$
where $\mathsf{T} \in \{\mathsf{Inv}, \mathsf{CollFind}\}$ is the corresponding task.
- ***Smoothness:*** *$\mathcal{B}$ is a $(1 - 2^{\log(q''/\gamma)-\gamma\beta}, \gamma)$-smooth* Inv*-adversary or, resp., a $(1-2^{2-\gamma\beta/4},\gamma)$-smooth* $\mathsf{CollFind}$*-adversary, for all $\gamma > 0$ (see Definition 5.8).*

Proof. Given any adversary $\mathcal{A}$, we construct $\mathcal{B}$ to behave as follows, under the constraint to preserve canonicality (specifically properties (i) and (iii) in Definition 5.5). $\mathcal{B}^{f,\mathsf{O}}$ runs $\mathcal{A}^{f,\mathsf{O}}$ on its input and, whenever $\mathcal{A}$ calls O with input C^f, $\mathcal{B}$ evaluates $C^f(\cdot)$ on $\beta(\lambda)$ uniformly random inputs and only then calls $\mathsf{O}(C)$, forwards the answer to $\mathcal{A}$, and proceeds with the simulation. For technicalities, if $\mathcal{B}$ decides to halt (complying to canonicality), we assume it replaces all its further queries by dummy queries (circuits that never call f and are therefore smooth at everything).

Correctness, canonicality and complexity of $\mathcal{B}$ are straight-forward. For smoothness, let us first consider the case of Inv-adversaries. Fix γ, an oracle O, an input to $\mathcal{B}$, an $i \in \mathbb{N}$ and let $C := C_i$ denote $\mathcal{B}$'s i^{th}

[6] More specifically, in [BD19], they make any adversary smooth by modifying his O-queries. Their collision finder is oblivious to these modifications since the functionality of the queries (as f-aided circuits) is preserved. This does not hold for our correlation finder and, therefore, their smoothening method does not preserve advantage in our settings. Hence, we slightly modify the definition, w.l.o.g., to allow for generic smoothening under any oracle, in particular for our correlation finder.

query (which, when exists, is also $\mathcal{A}$'s i^{th} query). We bound the probability that there exists an f_λ-query $x \in \{0,1\}^{\ell(\lambda)}$ which was not made by $\mathcal{B}$ prior to C yet $\{x\} \notin \mathsf{Smooth}^f_\gamma(C)$. Observe that this occurs if there exists $x \in \{0,1\}^\ell$ s.t. $\Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x] > \gamma(\lambda)$ and x is not queried by $\mathcal{B}$ before making the call to O with C. However, notice that in such a case, x will be queried by $\mathcal{B}$ during the β random evaluations of $C^f(\cdot)$ except with probability at most $(1-\gamma(\lambda))^{\beta(\lambda)} \leq 2^{-\gamma\beta}$. A simple counting argument shows that there exist at most q''/γ such non-smooth x's. Therefore, by applying a union bound over all such inputs under C^f, we get that $\mathcal{B}$ queries all of them with probability at least $1 - 2^{-\gamma\beta} \cdot q''/\gamma$.

To show smoothness in the case of $\mathsf{CollFind}$-adversaries, we first derive smoothness at collisions by a similar reasoning to the above; If there exists among $\mathcal{A}$'s queries to O a circuit $C \notin \mathcal{C}^*_{\lambda,f,\gamma}$, then it holds that a collision will be observed by $\mathcal{B}$ during the β random evaluations of $C^f(\cdot)$ except with probability at most $(1-\gamma)^\beta \leq 2^{-\gamma\beta}$. Recall that when $\mathcal{B}$ observes a collision it halts since it is canonical. Therefore, the probability that C is called by $\mathcal{B}$, and not replaced by a dummy call, is at most $2^{-\gamma\beta}$.

It remains to show that smoothness at siblings additionally holds. To that end, we split the set of siblings to three disjoint subsets: (i) the subset $S^*_{\ll i}$ of all siblings of f-queries made prior to making the i^{th} query C and the β evaluations preceding C, (ii) the subset $S^*_{<i}$ of all siblings of f-queries made during the β evaluations preceding C, and (iii) the subset $S^*_{\gg i}$ of all siblings of f-queries x made after making the query C (excluding the last query as per the definition in (3)). Evidently, these three subsets cover $\mathbf{Sib}_\lambda(f,\mathsf{O},\mathcal{A})$ entirely and are disjoint. Hence, it holds that

$$\begin{aligned}&\Pr_{\mathcal{B}}[\mathbf{Sib}_\lambda(f,\mathsf{O},\mathcal{B}) \notin \mathsf{Smooth}^f_\gamma(C)]\\ &\leq \Pr_{\mathcal{B}}[S^*_{\ll i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)] + \Pr_{\mathcal{B}}[S^*_{<i} \notin \mathsf{Smooth}^f_{\gamma/2}(C)] + \Pr_{\mathcal{B}}[S^*_{\gg i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)].\end{aligned}$$

We now analyze each of the above probabilities.

- $\underline{S^*_{\ll i}}$: This case follows by a similar argument since $S^*_{\ll i}$ is independent in the query C and its corresponding "smoothening" queries (namely the β evaluations preceding C). If $S^*_{\ll i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)$ then, it holds that, except with probability $2^{-\gamma\beta/4}$, $\mathcal{B}$ does not observe a sibling in the β evaluations (which are completely independent in the set $S_{\ll i}$). Since when $\mathcal{B}$ does observe a sibling it never reaches the i^{th} query, then with probability only $2^{-\gamma\beta/4}$ we have that $S^*_{\ll i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)$.

- $\underline{S^*_{<i}}$: While the siblings in $S^*_{<i}$ are not at all independent in C and its smoothening, this case still follows quite easily from the same logic. One way to see this is to apply the argument on two separate halves of $S^*_{<i}$ that are independent of each others. Formally, let $S^{(0)}_{<i}$ and $S^{(1)}_{<i}$ denote the siblings of f-queries observed by the first and, resp., second half of the β evaluations preceding C. It holds that

$$\Pr_{\mathcal{B}}[S^*_{<i} \notin \mathsf{Smooth}^f_{\gamma/2}(C)] = \Pr_{\mathcal{B}}[S^{(0)}_{<i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)] + \Pr_{\mathcal{B}}[S^{(1)}_{<i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)].$$

  Now, fix $b \in \{0,1\}$ and $S^{(b)}_{<i}$. By the fact that the evaluations producing $S^{(b)}_{<i}$, are sampled independently of $S^{(1-b)}_{<i}$, it still holds, conditioned on this fixed $S^{(b)}_{<i}$, that $S^{(1-b)}_{<i}$ is the set of siblings of f-queries seen in $\beta/2$ evaluations of C^f on random inputs. Therefore, letting $\mathcal{B}^{(0)}, \mathcal{B}^{(0)}$ denote the random coins used for sampling the two halves of the β evaluations, we have

$$\Pr_{\mathcal{B}}[S^{(b)}_{<i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)] = \Pr_{\mathcal{B}}[S^{(b)}_{<i} \notin \mathsf{Smooth}^f_{\gamma/4}(C) \wedge \mathcal{B} \rightharpoondown C] \leq \max_{S^{(b)}_{<i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)} \Pr_{\mathcal{B}^{(1-b)}}[\mathcal{B} \rightharpoondown C].$$

  Notice that $\mathcal{B}$ calling C implies that $S^{(b)}_{<i}$ has not been observed during the other half of $\beta/2$ evaluations, which again occurs with probability at most $2^{-\gamma\beta/4}$ when $S^{(b)}_{<i}$ is not $\gamma/4$-smooth.

- $\underline{S^*_{\gg i}}$: To analyze this last case, we separate the random coins used by $\mathcal{B}$ to sample the β random evaluations preceding C, which we denote by $\mathcal{B}_i$, from all other randomness used by $\mathcal{B}$ which we denote by $\mathcal{B}_{\text{-}i}$ (this is used for simulating $\mathcal{A}$ and the evaluations corresponding to the other O-queries). Now, observe that the set $S^*_{\gg i}$ is determined by $\mathcal{B}_{\text{-}i}$ up to the event where, during the β evaluations sampled by $\mathcal{B}_i$, a "future sibling" $x \in S^*_{\gg i}$ is observed – in which case x and any input that had been added as a sibling after x, will be "removed" from $S^*_{\gg i}$ (since $\mathcal{B}$ halts once observing the sibling of x and we exclude the last query from **Sib**). More formally, we denote by $S'_{\gg i}$ the set of siblings of all f-queries made by $\mathcal{B}$ after calling C, had it not invoked the β evaluations preceding C (intuitively, this is the $S^*_{\gg i}$ of $\mathcal{B}_{\text{-}i}$). Again, notice that $S'_{\gg i}$ depends only on $\mathcal{B}_{\text{-}i}$ and that $S^*_{\gg i} \subseteq S'_{\gg i}$. Therefore,

$$\Pr_{\mathcal{B}}[S^*_{\gg i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)] \leq \Pr_{\mathcal{B}}[S^*_{\gg i} \notin \mathsf{Smooth}^f_{\gamma/4}(C) \mid S'_{\gg i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)]$$

Now, fix any $S'_{\gg i} \notin \mathsf{Smooth}^f_{\gamma/4}(C)$ by fixing randomness for $\mathcal{B}_{\text{-}i}$. Let us write the elements in $S'_{\gg i}$ by the order they were added to the set as $S'_{\gg i} = \{x_1, \ldots, x_N\}$, and let j be the smallest integer such that $S'_j = \{x_1, \ldots, x_j\} \notin \mathsf{Smooth}^f_{\gamma}(C)$. The identity of S'_j is independent in $\mathcal{B}_i$ since $S'_{\gg i}$ is and, therefore, with probability all but $2^{-\gamma\beta/4}$ it holds that $\mathcal{B}_i$ calls some $x \in S'_j$, in which case $S^*_{\gg i} \subseteq \{x_1, \ldots, x_{j-1}\}$, which is smooth by the choice of j.

□

5.3 One-wayness from Differential Indistinguishability

In this section, we formalize the notion of differential indistinguishability for correlation finders and show that it is hard to invert f under any differentially indistinguishable correlation finder. This allows us to focus our design on obtaining differential indistinguishability to establish separation from OWP (via Lemma 5.2).

We begin by defining this new notion.

Definition 5.10 (Differential Indistinguishability)**.** *Let $\delta, \gamma : \mathbb{N} \to \mathbb{R}^+$ and $q : \mathbb{N} \to \mathbb{N}$. We say that a correlation finder $\mathcal{O} = \{\mathcal{O}_R\}$ is* (non-adaptively) differentially (q, γ, δ)-indistinguishable for $\mathcal{F}$ *if for any $R \in \mathcal{R}$, any $\lambda \in \mathbb{N}$, any $f \in \mathcal{F}$, any f-aided circuit $C \in \mathcal{C}$ which makes at most $q(\lambda)$ queries to f_λ on any input, and any $x^* \in \mathsf{Smooth}^f_{\gamma}(C)$ of length $\ell(\lambda)$, it holds that*

$$\mathbf{SD}(\mathsf{O}^f(C),\ \mathsf{O}^{f'}(C)) \leq \delta(\lambda),$$

where $\mathsf{O} \leftarrow \mathcal{O}_R$ and $f' = f_{x^* \leftrightarrow x'}$ for a uniformly random $x' \leftarrow \{0,1\}^\ell$.

We now state and then prove that it is hard to invert a random permutation f, even when given a differential indistinguishable correlation finder.

Lemma 5.11. *Let $\mathcal{F} = \{\mathcal{F}_\lambda\}$ be the distribution of random permutations over $\{0,1\}^\lambda$ (i.e. $\ell(\lambda) = \lambda$). Let $\delta : \mathbb{N} \to [0,1]$ and let $\mathcal{A}$ be a (q, q', q'')-bounded Inv-adversary that is canonical and (τ, γ)-smooth (see Definitions 5.5 and 5.8). Let $\mathcal{O}$ be a correlation finder that is query-independent (see Definition 5.3) and differentially (q'', γ, δ)-indistinguishable for $\mathcal{F}$. Then, it holds for any $\lambda \in \mathbb{N}$ that*

$$\mathbb{E}_{f,\mathsf{O}}[\mathbf{Adv}^f_{\mathsf{Inv}}(\lambda, \mathcal{A}^{\mathsf{O}})] \leq O((1 - \tau(\lambda)) + q(\lambda) \cdot 2^{-\lambda} + q'(\lambda) \cdot \delta(\lambda)),$$

where $f \leftarrow \mathcal{F}$ and $\mathsf{O} \leftarrow \mathcal{O}$.

Proof. By the definition of Inv (see Definitions 3.1 and 4.1), it holds that

$$\mathbb{E}_{f,\mathsf{O}}[\mathbf{Adv}^f_{\mathsf{Inv}}(\lambda, \mathcal{A}^{\mathsf{O}})] = \Pr_{f,\mathsf{O},x^* \leftarrow \{0,1\}^\lambda, \mathcal{A}}[f(\mathcal{A}^{f,\mathsf{O}^f}(f(x^*))) = x^*].$$

As a first step, since $\mathcal{A}$ is assumed to be smooth, we may condition on the event that $\mathcal{A}$ makes only queries that are smooth at "unobserved inputs" (see Definition 5.8). In particular, by the canonicality of $\mathcal{A}$, notice that x^* is never observed by $\mathcal{A}$ before making any of its O-queries since the moment x^* is observed

$\mathcal{A}$ halts and makes no further queries. Formally, let $\mathsf{E}_{f,\mathsf{O},x^*}$ denote the event that there exists $i \in [q']$ such that $\{x^*\} \notin \mathsf{Smooth}^f_\gamma(C_i)$, where C_i is the i^{th} O-query that $\mathcal{A}^{f,\mathsf{O}}(f(x^*))$ makes. By the (τ,γ)-smoothness of $\mathcal{A}$, we know that, for any fixed f, O and x^*, it holds that

$$\Pr_{\mathcal{A}}[\mathsf{E}_{f,\mathsf{O},x^*}] \le q' \cdot \max_i \Pr_{\mathcal{A}}[x^* \notin \mathsf{Smooth}^f_\gamma(C_i)] \le \Pr_{\mathcal{A}}[\exists x \in \{0,1\}^{\ell(\lambda)} \setminus X_{<i} :\ x \notin \mathsf{Smooth}^f_\gamma(C_i)] \le q' \cdot (1-\tau),$$

where $X_{<i}$ is all observed inputs, as defined in Definition 5.8. Hence,

$$\Pr_{f,\mathsf{O},x^*\leftarrow\{0,1\}^\lambda,\mathcal{A}}[f(\mathcal{A}^{f,\mathsf{O}^f}(f(x^*))) = x^*] \le \Pr_{f,\mathsf{O},x^*\leftarrow\{0,1\}^\lambda,\mathcal{A}}[\mathsf{E}_{f,\mathsf{O},x^*} \ \wedge\ f(\mathcal{A}^{f,\mathsf{O}^f}(f(x^*))) = x^*] + q' \cdot (1-\tau). \quad (4)$$

To show hardness of inversion, we switch the experiment to an experiment where the adversary (and the oracle O) is given access to an oracle f' that statistically hides any information about the pre-image of the given challenge, i.e. x^*, which deems non-trivial success in computing a successful answer virtually impossible. More specifically, we swap f at x^* with a random $x' \leftarrow \{0,1\}^\lambda$, essentially "randomizing" the pre-image of the given challenge under the given function $f' = f_{X^*\leftrightarrow X'}$. By the presumed differential indistinguishability of $\mathcal{O}$ (in particular) we are able to show that such a swap does not affect the view of $\mathcal{A}$ except with a negligible probability. By the symmetry between f and f' given $y^* = f(x^*)$, we may then conclude that f' hides any information about x^* and, hence, inversion is impossible.

More formally, we denote, for every $1 \le i \le q + q'$, by $\mathcal{A}_i$ the algorithm that takes as input the view of $\mathcal{A}$ after making the first $i-1$ queries to its oracles and outputs the i^{th} query (or a final output if halting). Then, recalling the query-independence of $\mathcal{O}$ and the canonicality of $\mathcal{A}$, we may model the probability in (4) with the following experiment

– $\underline{\mathsf{Exp}}$:

1. Sample $x^* \leftarrow \{0,1\}^\lambda$ at random and let $y^* = f(x^*)$.
2. Sample a random permutation $f : \{0,1\}^\lambda \to \{0,1\}^\lambda$.
3. Sample an oracle $\mathsf{O} = (\mathsf{O}_1, \dots, \mathsf{O}_{q'}) \leftarrow \mathcal{O}^{q'}$.
4. For $i = 1 \dots q + q'$, let $Q_i \leftarrow \mathcal{A}_i(y^*, (Q_1, A_1), \dots, (Q_{i-1}, A_{i-1}))$,
   – If Q_i is an f-query: If $Q_i = x^*$ output 1 ($\mathcal{A}$ wins), otherwise let $A_i = f(Q_i)$.
   – If Q_i is an O-query: If $x^* \notin \mathsf{Smooth}^f_\gamma(Q_i)$ output 0, otherwise let $A_i = \mathsf{O}^f_i(Q_i)$.
5. Output 0 ($\mathcal{A}$ fails).

We now consider sampling an independently uniform $x' \leftarrow \{0,1\}^{\ell(\lambda)}$, and swapping f at x^* with x' to obtain the function $f' = f_{x^*\leftrightarrow x'}$ and a new experiment Exp' where the answers we give to the adversary are according to f' (yet, notice that the winning condition remains unchanged).

– $\underline{\mathsf{Exp}'}$:

1. Sample $x^* \leftarrow \{0,1\}^\lambda$ at random and let $y^* = f(x^*)$.
2. Sample a random permutation $f : \{0,1\}^\lambda \to \{0,1\}^\lambda$.
3. Sample a uniform $x' \leftarrow \{0,1\}^\lambda$ and let $f' = f_{x^*\leftrightarrow x'}$.
4. Sample an oracle $\mathsf{O} = (\mathsf{O}_1, \dots, \mathsf{O}_{q'}) \leftarrow \mathcal{O}^{q'}$.
5. For $i = 1 \dots q + q'$, let $Q'_i \leftarrow \mathcal{A}_i(y^*, (Q'_1, A'_1), \dots, (Q'_{i-1}, A'_{i-1}))$,
   – If Q'_i is an f-query: If $Q'_i = x^*$ output 1 ($\mathcal{A}$ wins), otherwise let $A'_i = f'(Q'_i)$.
   – If Q_i is an O-query: If $x^* \notin \mathsf{Smooth}^f_\gamma(Q_i)$ output 0, otherwise let $A'_i = \mathsf{O}^{f'}_i(Q'_i)$.
6. Output 0 ($\mathcal{A}$ fails).

We bound the statistical distance between the two experiments via a coupling argument. We use the straight-forward coupling that samples the same x^*, x', f and randomness for $\mathcal{A}$ for both Exp and Exp', while the randomness for the underlying correlation finding oracles is sampled, for one query at a time, according to the coupling that gives us the lowest probability of discrepancy between the two experiments, which is at most δ by the differential indistinguishability of $\mathcal{O}$ (recall we only care about the case $x^* \in \mathsf{Smooth}^f(Q_i)$ since otherwise the output of the experiments is equal). More formally, we consider the coupling between Exp and Exp' which samples $x^* \leftarrow \{0,1\}^\lambda$, uniform f and x', and randomness for $\mathcal{A}$, then, for $i = 1, \ldots, q + q'$, samples randomness for the correlation finders at query i by $(\mathsf{O}_{i,1}, \mathsf{O}_{i,2}) \leftarrow \mathcal{P}_i$ (to be respectively used by Exp and Exp') where $\mathcal{P}_i$ is the coupling that gives, by Proposition 2.2,

$$\Pr_{(\mathsf{O}_{i,1},\mathsf{O}_{i,2})\leftarrow\mathcal{P}_i}[\mathsf{O}^f_{i,1}(Q_i) \neq \mathsf{O}^{f'}_{i,2}(Q_i)] = \mathbf{SD}(\mathsf{O}^f(Q_i), \mathsf{O}^{f'}(Q_i)) \leq \delta(\lambda)$$

where Q_i is the i^{th} query made by the adversary at $\mathcal{A}^{f,\mathsf{O}^f}(y^*)$ (recall that all relevant information – $f, y^*, \mathcal{A}$ and queries $Q_{<i}$ and their answers $A_{<i}$ – is determined by this point). Notice that by the query-independence of $\mathcal{O}$, sampling randomness for O at the different queries via independent couplings still gives the desired joint marginal distribution for each of $\mathsf{O}_1 = (\mathsf{O}_{i,1})_{i=1}^{q+q'}$ and $\mathsf{O}_2 = (\mathsf{O}_{i,2})_{i=1}^{q+q'}$. Now, since the two experiments are identical except for our answers to the adversary's queries (A_i), it follows that

$$\mathbf{SD}(\mathsf{Exp}, \mathsf{Exp}') \leq \Pr_{\substack{x^*,x',f,\mathcal{A} \\ \forall i,(\mathsf{O}_{i,1},\mathsf{O}_{i,2})\leftarrow\mathcal{P}_i}}[\mathsf{Exp} \neq \mathsf{Exp}'] = \Pr[(A_i)_{i\in[q+q']} \neq (A'_i)_{i\in[q+q']}] = \sum_{i=1}^{q+q'} \Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i]. \tag{5}$$

Keeping in mind that $A_{<i} = A'_{<i}$ implies $Q_i = Q'_i$, we look into the following two cases separately:

- Q_i is an f-query: In such a case, it is evident that the answer to Q_i is different in Exp than in Exp' only when $Q_i = x'$ (notice that when $Q_i = x^*$ the experiment ends before setting A_i, A'_i). Thus, since in Exp (and therefore in its marginal by the coupling) x' is sampled independently of Q_i, it holds

$$\Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i] \leq \Pr[Q_i = Q'_i \wedge A_i \neq A'_i] \leq \Pr_{x^*,x',f,\mathsf{O}_1,\mathcal{A}}[Q_i \in x'] = 2^{-\lambda}.$$

- Q_i is an O-query: In which case, by the definition of $\mathcal{P}_i$, it holds that

$$\Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i] \leq \Pr_{\substack{x^*,x',f,\mathcal{A}, \\ (\mathsf{O}_{i,1},\mathsf{O}_{i,2})\leftarrow\mathcal{P}_i}}[\mathsf{O}^f_{i,1}(Q_i) \neq \mathsf{O}^{f'}_{i,2}(Q_i)] \leq \delta(\lambda).$$

By the above and (5) we conclude

$$\mathbf{SD}(\mathsf{Exp}, \mathsf{Exp}') \leq q(\lambda) \cdot 2^{-\lambda} + q'(\lambda) \cdot \delta(\lambda) \tag{6}$$

and, therefore, it remains to bound the probability that $\mathsf{Exp}' = 1$. On a closer look, Exp' is essentially the inversion game when $\mathcal{A}$ is given access to f' yet wins if inverts w.r.t. f, namely calls some x^*, i.e.

$$\Pr[\mathsf{Exp}' = 1] = \Pr_{x^*,x',f,\mathsf{O},\mathcal{A}}[\mathcal{A}^{f',\mathsf{O}^{f'}}(y^*) \xrightarrow{f'} x^*].$$

Lastly, observe that f and f' are completely symmetric in Exp' given y^*: In steps 1-3 of the experiment, we are sampling (y^*, f, f') at random where for every y^*, f and f' are uniform in the space of all regular function pairs that satisfy $f' = f_{f^{-1}(y^*)\leftrightarrow f'^{-1}(y^*)}$ or, equivalently $f = f'_{f'^{-1}(y^*)\leftrightarrow f^{-1}(y^*)}$ (hence the symmetry). Thus, the marginals of $(y^*, f, x' = f'^{-1}(y^*))$ and $(y^*, f', x^* = f^{-1}(y^*))$ are identical and, in particular, x^* distributes uniformly over $\{0,1\}^\lambda$ and independently of f' and, thus, can be sampled after the output of $\mathcal{A}^{f',\mathsf{O}^{f'}}(y^*)$ is determined. Given this rundown of the experiment we may proceed to conclude

$$\Pr[\mathsf{Exp}' = 1] = \Pr_{y^*,f',x^*}[\mathcal{A}^{f',\mathsf{O}^{f'}}(y^*) \xrightarrow{f'} x^*] \leq q(\lambda) \cdot 2^{-\lambda}.$$

We then finish by combining the above with (4) and (6). □

5.4 Collision Resistance from Adaptive Differential Indistinguishability

We extend the framework from the previous section to allow a separation from collision-resistant hash functions and not only one-way permutations.

Defining Adaptive Differential Indistinguishability. We generalize differential indistinguishability to consider adaptive choices of the input x^* that depend on the random answer of the oracle (recall that in the non-adaptive notion x^* is worst-case but fixed apriori). More specifically, we consider x^* that is computed as a function of $\mathsf{O}^f(C)$, where C is the given O-query and O is the random oracle. Notice that this function defines the dependence of the adversarial choice of x^* on the oracle O and the query C chosen by the environment. Contrary to before, we will not be able to guarantee such an adaptive notion for any worst-case (bounded) C and, therefore, we will be requiring indistinguishability for an average-case query C sampled from a specified distribution C^* (speaking ahead, we will be looking at the distribution of O-queries made by the adversary). Additionally, necessary for our separation from CRH, we generalize the notion to consider sequences of swaps, namely an adaptively-chosen subset X^* that is swapped with a uniformly random X' of the same size. The formal definition is given below.

Definition 5.12 (Adaptive Differential Indistinguishability)**.** *Let $\delta : \mathbb{N} \to \mathbb{R}^+$. Let $\mathsf{C}^* = \{\mathsf{C}^*_{\lambda,f}\}$ be a family of distributions over $\mathcal{C}$ and let $\mathsf{X}^* = \{\mathsf{X}^*_{\lambda,f}\}$ be a probabilistic function $\mathsf{X}^*_{\lambda,f} : \mathcal{C} \times \{0,1\}^* \to \mathbb{P}(\{0,1\}^{\ell(\lambda)})$ (equivalently, distributions over such deterministic functions). We say that a correlation finder $\mathcal{O} = \{\mathcal{O}_R\}$ is* adaptively differentially δ-indistinguishable for $\mathcal{F}$ against $(\mathsf{C}^*, \mathsf{X}^*)$ *if for any $R \in \mathcal{R}$, any $\lambda \in \mathbb{N}$ and any $f \in \mathcal{F}$, it holds that*

$$\mathbf{SD} = \Big((\mathsf{O}, C, \mathsf{O}^f(C)),\ (\mathsf{O}, C, \mathsf{O}^{f'}(C)) \Big) \leq \delta(\lambda),$$

where $\mathsf{O} \leftarrow \mathcal{O}$, $C \leftarrow \mathsf{C}^*_{\lambda,f}$ and $f' = f_{X^* \leftrightarrow X'}$ for $X^* \leftarrow \mathsf{X}^*_{\lambda,f}(C, \mathsf{O}^f(C))$ and a uniformly random subset $X' \subseteq \{0,1\}^{\ell(\lambda)}$ of size $|X^*|$.

Our Adaptive Adversary: The Siblings of Observed Inputs. In order to base collision resistance on adaptive differential indistinguishability, we must first identify a family of adversaries $\mathsf{A} = (\mathsf{C}^*, \mathsf{X}^*)$ against which adaptive differential indistinguishability is required to obtain such a reduction. At a high level, given an adversary $\mathcal{A}$, we consider $(\mathsf{C}^*, \mathsf{X}^*)$ where a random $Q \leftarrow \mathsf{C}^*_{\lambda,f}$ and a random $X^* \leftarrow \mathsf{X}^*_{\lambda,f}(Q, \mathsf{O}^f(Q))$ would imitate the distribution of an O-query made by $\mathcal{A}$ in a random execution and, respectively, the set of siblings corresponding to that execution (recall the definition in (3)). We formalize below.

Definition 5.13 ($(\mathcal{F}, \mathcal{O}, \mathcal{A})$-Siblings)**.** *Let $\mathcal{F}$ be a family of oracles $f = \{f_\lambda : \{0,1\}^{\ell(\lambda)} \to \{0,1\}^\lambda\}$, let $\mathcal{O}$ be an f-aided correlation finder and let $\mathcal{A}$ be an oracle-aided $\mathsf{CollFind}$-adversary. We say that $(\mathsf{C}^*, \mathsf{X}^*)$ is an i^{th}* $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution *for $i \in \mathbb{N}$ if, for any $\lambda \in \mathbb{N}$ and $f \in \mathcal{F}$, it holds that*

$$\mathsf{C}^* \equiv C_i \qquad\qquad \mathsf{X}^*_{\lambda,f}(C, z) \equiv \mathbf{Sib}_\lambda(f, \mathsf{O}_{C,z}, \mathcal{A}_{C,z})$$

where C_i is the i^{th} O-query made by a random $\mathcal{A}^{f,\mathsf{O}}(1^\lambda)$, and $\mathsf{O}_{C,A}$ and $\mathcal{A}_{C,A}$ are deterministic instances of $\mathcal{O}$ and $\mathcal{A}$, that are jointly sampled from their respective distributions conditioned on C is the i^{th} query made by $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda)$ and $\mathsf{O}^f(C) = z$.

Remark 5.14. *Let $\mathcal{O}$ be query-independent. Then, for any i^{th} $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution A, any $f \in \mathcal{F}$ and any $\lambda \in \mathbb{N}$, it holds that*

$$(C^*, \mathsf{O}_i, \mathsf{X}^*_{\lambda,f}(C^*, \mathsf{O}_i^f(C^*))) \equiv (Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})),$$

where $C^* \leftarrow \mathsf{C}^*$, $\mathsf{O} = (\mathsf{O}_j)$ is sampled at random from $\mathcal{O}$ (in a query-independent fashion) and Q_i is the i^{th} query made by $\mathcal{A}^{f,\mathsf{O}^f}$.

Proof. To see the equivalence, notice first that the value of $\mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$ is independent in O_i given the i^{th} answer $A_i = \mathsf{O}_i^f(Q_i)$ and, therefore, we can write $\mathbf{Sib}_\lambda$ as a function of $f, \mathsf{O}_{\neq i}, \mathcal{A}$ and A_i,

$$(Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})) \equiv (Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}_{\neq i}, \mathcal{A}, \mathsf{O}_i^f(Q_i))).$$

Second, since all of Q_i, $\mathsf{O}_{\neq i}$ and $\mathcal{A}$ are independent of O_i, we may think about sampling $\mathsf{O}_{\neq i}$ and $\mathcal{A}$ as first sampling an i^{th} query Q_i according to its marginal distribution in a random execution (that is, according to C^*), then sampling $\mathcal{A}$ and $\mathsf{O}_{\neq i}$ conditioned on Q_i is the i^{th} query made by $\mathcal{A}^{f,\mathsf{O}^f}$ – we denote these (possibly correlated) distributions by $\mathcal{A}_{Q_i}$ and, resp., $\mathsf{O}_{\neq i,Q_i}$) – and only then sampling O_i independently at random.

$$(Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})) \equiv (Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}_{\neq i,Q_i}, \mathcal{A}_{Q_i}, \mathsf{O}_i^f(Q_i))).$$

Lastly, observe that $\mathcal{A}$ and $\mathsf{O}_{\neq i}$ are independent in the answer A_i and, therefore, sampling from the distributions $\mathcal{A}_{Q_i}$ and $\mathsf{O}_{\neq i,Q_i}$ conditioning further on $A_i = \mathsf{O}_i^f(Q_i)$ gives us still the same distributions. Since, additionally, the siblings set is independent of O_i once A_i is fixed, we may replace the random $\mathsf{O}_{\neq i,Q_i}$ and $\mathcal{A}_{Q_i}$ with random $\mathsf{O}_{Q_i,\mathsf{O}_i^f(Q_i)}$ and $\mathcal{A}_{Q_i,\mathsf{O}_i^f(Q_i)}$ (in $\mathsf{O}_{Q_i,\mathsf{O}_i^f(Q_i)}$ we are implicitly sampling "fresh" O_i that anyway does not affect the identity of the siblings set given A_i), to obtain the desired equivalence

$$(Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})) \equiv (Q_i, \mathsf{O}_i, \mathbf{Sib}_\lambda(f, \mathsf{O}_{Q_i,\mathsf{O}_i^f(Q_i)}, \mathcal{A}_{Q_i,\mathsf{O}_i^f(Q_i)}, \mathsf{O}_i^f(Q_i))) \equiv (C^*, \mathsf{O}_i, \mathsf{X}^*_{\lambda,f}(C^*, \mathsf{O}_i^f(C^*))).$$

□

The Reduction Statement. In the lemma below, we argue a reduction from the task of finding collisions under random regular hash functions to the task of breaking adaptive differential indistinguishability against siblings. As a consequence, it would suffice to show that a correlation finder satisfies such adaptive differential indistinguishability in order to establish a separation from collision-resistant hash.

Lemma 5.15. *Let $\mathcal{F} = \{\mathcal{F}_\lambda\}$ be the distribution of random regular functions mapping $\ell(\lambda)$-bit inputs to λ-bit outputs. Let $\delta : \mathbb{N} \to [0,1]$ and let $\mathcal{A}$ be a canonical (q, q', q'')-bounded* CollFind*-adversary. Let $\mathcal{O}$ be a correlation finder that is query-independent (see Definition 5.3), and adaptively differentially δ-indistinguishable for $\mathcal{F}$ against any $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution (see Definition 5.13). Then, it holds for any $\lambda \in \mathbb{N}$ that*

$$\mathbb{E}_{f,\mathsf{O}}[\mathbf{Adv}^f_{\mathsf{CollFind}}(\lambda, \mathcal{A}^{\mathsf{O}})] = O(q^2 \cdot 2^{\ell - 2\lambda} + q' \cdot \delta),$$

where $f \leftarrow \mathcal{F}$ and $\mathsf{O} \leftarrow \mathcal{O}$.

Proof of Lemma 5.15 To show hardness of finding collisions, we show that the adversary would have seen the same view had we switched the function f to a different function f' that hides any information about collisions containing any input x that the adversary calls f at (this is exactly the set $\mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$ as defined in (3)).

More specifically, consider the experiment where we run $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda)$ and record all f-points $\{x_1, \dots, x_t\}$ that $\mathcal{A}$ has observed in its execution via an f-query, i.e. $\mathbf{Q}_\lambda(f, \mathsf{O}, \mathcal{A})$ (see (2)). We take the set of all siblings $S^* = \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$ (recall (3)) and notice that if $\mathcal{A}$ finds a collision, then he must have found an input in S^* and therefore, in this sense, S^* is the "target set" for the adversary and can alone define his winning condition in the game. To "hide" S^*, we swap f at S^* with randomly chosen inputs S' to obtain f'. We claim that, by adaptive differential indistinguishability against canonical siblings, if we were to re-run the execution of $\mathcal{A}$ using the same randomness under the same O, yet replacing the oracle f with f', we would have obtained the same outcome. We then show that S^* distributes as if uniform and independent given the function f' and, thus, finding any input in S^* given access only to f', as happens in the re-execution, is hard for any query-bounded adversary.

Formally, we denote, for every $1 \leq i \leq q + q' + 1$, by $\mathcal{A}_i$ the algorithm that takes as input the view of the adversary after making the first $i - 1$ queries to its oracles and outputs the i^{th} query (or a final output if $i = q + q' + 1$). By additionally exploiting the query-independence of $\mathcal{O}$, we may model the collision-finding experiment as follows

```
– $\underline{\mathsf{Exp}}$:
    1. Sample a random function $f \leftarrow \mathcal{F}$, an oracle $\mathsf{O} = (\mathsf{O}_1, \ldots, \mathsf{O}_{q'}) \leftarrow \mathcal{O}^{q'}$ and randomness for $\mathcal{A}$.
    2. Let $S = \emptyset$.
    3. For $i = 1 \ldots q + q'$, let $Q_i \leftarrow \mathcal{A}_i(1^\lambda, (Q_1, A_1), \ldots, (Q_{i-1}, A_{i-1}))$,
        – If $Q_i$ is an $f$-query:
            – If $Q_i \in S$ output 1 ($\mathcal{A}$ succeeds).
            – Otherwise, let $A_i = f(Q_i)$ and, if $Q_i \in \{0,1\}^{\ell(\lambda)}$, $S \leftarrow S \cup \mathbf{sib}^f(Q_i)$.
        – If $Q_i$ is an $\mathsf{O}$-query, let $A_i = \mathsf{O}_i^f(Q_i)$.
    4. Output 0 ($\mathcal{A}$ fails).
```

As a first step, we notice that by the canonicality of $\mathcal{A}$, we can replace the intermediate value of S in all iterations in Exp by the final value that it eventually takes, which is essentially $\mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$ (see the definition in (3)). This allows us to focus on the following mutation of the experiment above.

```
– $\underline{\mathsf{Exp}^*}$:
    1. Sample a random function $f \leftarrow \mathcal{F}$, an oracle $\mathsf{O} = (\mathsf{O}_1, \ldots, \mathsf{O}_{q'}) \leftarrow \mathcal{O}^{q'}$ and randomness for $\mathcal{A}$.
    2. Let $S^* = \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$.
    3. For $i = 1 \ldots q + q'$, let $Q_i \leftarrow \mathcal{A}_i(1^\lambda, (Q_1, A_1), \ldots, (Q_{i-1}, A_{i-1}))$,
        – If $Q_i$ is an $f$-query:
            – If $Q_i \in S^*$ output 1 ($\mathcal{A}$ succeeds).
            – Otherwise, let $A_i = f(Q_i)$.
        – If $Q_i$ is an $\mathsf{O}$-query, let $A_i = \mathsf{O}_i^f(Q_i)$.
    4. Output 0 ($\mathcal{A}$ fails).
```

Claim 1. $\mathbf{SD}(\mathsf{Exp}^*, \mathsf{Exp}) = 0$.

Proof. We show that the outcomes of Exp and Exp^* are equal for any fixed f, O and $\mathcal{A}$. To see this, observe that the only scenario where the experiments may behave differently is when $Q_i \in S^*$ (and Exp^* output 1 at iteration i) yet $Q_i \notin S_i$, where S_i is the value that S takes during the i^{th} iteration of Exp prior to its update (notice that $S_i \setminus S^* = \emptyset$). However, since $Q_i \in S^* \setminus S_i$, then there must exist $j > i$ for which Q_j is an f-query and $Q_i \in \mathbf{sib}(Q_j)$ and, therefore, $Q_i \in S_j$. In such a case, it holds by the symmetry of $\mathbf{sib}$ that $Q_j \in \mathbf{sib}(Q_i) \subseteq S_j$, hence Exp halts at iteration j and outputs 1 as well. □

We now consider an experiment Exp' similar to Exp^* where the adversary $\mathcal{A}$ and the oracle O are given access a different function f', yet the winning condition remains unchanged. Specifically, we obtain the function f' from f by swapping the inputs in S^* by the order they were added to the set (in fact, any fixed arbitrary order works for us) with an ordered set $S' \subseteq \{0,1\}^\ell$ of the same size as S^*, which we sample independently and uniformly at random.

```
– $\underline{\mathsf{Exp}'}$:
    1. Sample a random function $f \leftarrow \mathcal{F}$, an oracle $\mathsf{O} = (\mathsf{O}_1, \ldots, \mathsf{O}_{q'}) \leftarrow \mathcal{O}^{q'}$ and randomness for $\mathcal{A}$.
    2. Let $S^* = \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})$.
    3. Sample a uniformly random ordered set $S' \subseteq \{0,1\}^{\ell(\lambda)}$ of size $q(2^{\ell-\lambda} - 1)$ and let $f' = f_{S^* \leftrightarrow S'}$.
    4. For $i = 1 \ldots q + q'$, let $Q'_i \leftarrow \mathcal{A}_i(1^\lambda, (Q'_1, A'_1), \ldots, (Q'_{i-1}, A'_{i-1}))$,
        – If $Q'_i$ is an $f$-query:
            – If $Q'_i \in S^*$ output 1 ($\mathcal{A}$ succeeds).
```

    - Otherwise, let $A_i = f(Q'_i)$.
  - If Q'_i is an O-query, let $A_i = \mathsf{O}_i^f(Q'_i)$.
5. Output 0 ($\mathcal{A}$ fails).

Let us define the following deterministic mapping

$$\mathsf{View} : (f, \mathsf{O}, \mathcal{A}) \mapsto ((Q_1, A_1), \ldots, (Q_q, A_q))$$

that maps any f, oracle O and adversary randomness $\mathcal{A}$ to the view of the adversary in Exp when given this randomness, that is, when running $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda)$.

We hereby prove that, with high probability, the views seen by $\mathcal{A}$ under f is identical to that under f'. We use this, in particular, to imply statistical proximity between the outcomes of Exp^* and Exp'.

Claim 2. *For $f, \mathsf{O}, \mathcal{A}, f'$ sampled as in Exp^* and Exp',*

$$\Pr_{f,\mathsf{O},\mathcal{A},S'}[\mathsf{View}(f', \mathsf{O}, \mathcal{A}) \neq \mathsf{View}(f, \mathsf{O}, \mathcal{A})] \leq q^2(\lambda) \cdot 2^{-\lambda} + q'(\lambda) \cdot \delta(\lambda).$$

Proof. It holds by inspection that

$$\Pr_{f,\mathsf{O},\mathcal{A},S'}[\mathsf{View}(f', \mathsf{O}, \mathcal{A}) \neq \mathsf{View}(f, \mathsf{O}, \mathcal{A})] = \Pr[(A_i)_{i\in[q]} \neq (A'_i)_{i\in[q]}] \quad = \sum_{i=1}^{q+q'} \Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i]. \quad (7)$$

Keeping in mind that $A_{<i} = A'_{<i}$ implies $Q_i = Q'_i$, we look into the following two cases separately:

(i) Q_i is an f-query: In such a case, it is evident that the answer to Q_i is different in Exp' than in Exp^* only when $Q_i \in S'$ (notice that when $Q_i \in S^*$ the experiment ends before setting A_i, A'_i in both cases). Thus, since S' is sampled independently of Q_i,

$$\begin{aligned}\Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i] &\leq \Pr[Q_i = Q'_i \wedge A_i \neq A'_i] \leq \Pr_{f,\mathsf{O},\mathcal{A},S'}[Q_i \in S'] \\ &\leq 2^{-\ell(\lambda)} \cdot |S^*|= 2^{-\ell} \cdot q(2^{\ell-\lambda} - 1) \leq q2^{-\lambda}. \qquad (8)\end{aligned}$$

(ii) Q_i is an O-query: In which case, it holds that

$$\Pr[A_{<i} = A'_{<i} \wedge A_i \neq A'_i] \leq \Pr_{f,\mathsf{O},\mathcal{A},S^*,S'}[\mathsf{O}_i^{f'}(Q_i) \neq \mathsf{O}_i^f(Q_i)].$$

We bound the above probability by relying on the assumed adaptive differential indistinguishability of $\mathcal{O}$ against siblings. Specifically, we look at the i^{th} $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution $(\mathsf{C}^*, \mathsf{X}^*)$ (see Definition 5.13). By the equivalence demonstrated in Remark 5.14, we may write

$$\Pr_{f,\mathsf{O},\mathcal{A},S^*,S'}[\mathsf{O}_i^{f'}(Q_i) \neq \mathsf{O}_i^f(Q_i)] = \Pr_{f,Q_i,\mathsf{O}_i,S^*,S'}[\mathsf{O}_i^{f'}(Q_i) \neq \mathsf{O}_i^f(Q_i)] = \Pr_{f,\mathsf{O}_i,C,X^*,X'}[\mathsf{O}_i^{f'}(C) \neq \mathsf{O}_i^f(C)]$$

where f and O_i distribute as in Exp^* (or Exp'), $C \leftarrow \mathsf{C}^*$, $X^* \leftarrow \mathsf{X}^*_{\lambda,f}(C, \mathsf{O}_i^f(C))$, X' is uniformly random of size $|X^*|$ and $f' = f_{X^* \leftrightarrow X'}$. The above probability is bound by δ based on the adaptive differential indistinguishability of $\mathcal{O}$ against A and Proposition 2.2 (the best coupling must sample the same O_i and Q_i).

The proof of the claim is then complete by plugging the above and (8) into (7). □

By Claim 2 and Proposition 2.2 it immediately follows that

$$\mathbf{SD}(\mathsf{Exp}', \mathsf{Exp}^*) \leq q^2 \cdot 2^{-\lambda} + q' \cdot \delta(\lambda).$$

It remains to bound the probability that $\mathsf{Exp}' = 1$. On a closer look, Exp' is essentially the collision finding task when $\mathcal{A}$ is given access to f' yet wins if succeeds w.r.t. f, i.e. finds S^* – siblings of observed points under f. Our goal then is to show that S^* distributes uniformly at random given f', making it virtually impossible to find. To that end, we draw a symmetry between f and f' and prove that (f, f') and (f', f) are indistinguishable. We further notice that the deterministic mapping that maps (f, f') to (f, S') also maps (f', f) to (f', S^*) and therefore, given the symmetry, since S' is independent in f, so is S^* in f'.

Claim 3. *For f, f' sampled as in Exp',*

$$\mathbf{SD}((f, f', \mathsf{O}, \mathcal{A}), (f', f, \mathsf{O}, \mathcal{A})) = O(q^2(\lambda) \cdot (2^{-\lambda} + 2^{\ell - 2\lambda}) + q'(\lambda) \cdot \delta(\lambda)).$$

Proof. For any $\mathsf{O} \in \mathcal{O}$ and fixed deterministic $\mathcal{A}$, let us denote by $\mathcal{V}_{\mathsf{O},\mathcal{A}}$ the marginal distribution of $\mathsf{View}(f, \mathsf{O}, \mathcal{A})$, when f is sampled uniformly at random. For any $V \in \mathcal{V}_{\mathsf{O},\mathcal{A}}$, we denote

$$\mathcal{F}_{V,\mathsf{O},\mathcal{A}} = \{f \mid \mathsf{View}(f, \mathsf{O}, \mathcal{A}) = V\}.$$

We can then think about (f, f') that are produced by Exp' as being generated via sampling random O, $\mathcal{A}$, $V \leftarrow \mathcal{V}_{\mathsf{O},\mathcal{A}}$ and S', then sampling a uniformly random $f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$ and obtaining the corresponding function $f' = f_{S^* \leftrightarrow S'}$ (recall S^* is deterministically fixed by $f, \mathsf{O}, \mathcal{A}$). Given such a representation, we observe that by Claim 2, such an f' is in $\mathcal{F}_{V,\mathsf{O},\mathcal{A}}$ with high probability and, hence, we may replace the uniform S' with a slightly different distribution, that samples S' uniformly conditioned on $f, f' \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$ (by rejection sampling). More formally, we consider a recasting of Exp' which we denote by Exp'_0 (and which evidently produces the same distribution of (f, f') as sampled by Exp') and a modified experiment Exp'_1:

$\underline{\mathsf{Exp}'_0}$:

1. Sample random O,$\mathcal{A}$ and $V \leftarrow \mathcal{V}_{\mathsf{O},\mathcal{A}}$.
2. Sample uniformly random S'.
3. Sample $f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$.
4. Let $S^* = S^*(f, \mathsf{O}, \mathcal{A})$.
5. Output $(f, f' = f_{S^* \leftrightarrow S'}, \mathsf{O}, \mathcal{A})$.

$\underline{\mathsf{Exp}'_1}$:

1. Sample random O,$\mathcal{A}$ and $V \leftarrow \mathcal{V}_{\mathsf{O},\mathcal{A}}$.
2. Sample uniformly random S'.
3. Sample $f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$.
4. Let $S^* = S^*(f, \mathsf{O}, \mathcal{A})$.
5. If $f' = f_{S^* \leftrightarrow S'} \notin \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$, repeat from step 1. Otherwise, output $(f, f', \mathsf{O}, \mathcal{A})$.

Then, by Claim 2, it holds that

$$\begin{aligned}\mathbf{SD}(\mathsf{Exp}'_0, \mathsf{Exp}'_1) \leq \Pr_{\substack{\mathsf{O},\mathcal{A},V,S' \\ f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}}} [f' \notin \mathcal{F}_{V,\mathsf{O},\mathcal{A}}] &= \Pr_{f \leftarrow \mathcal{F},\mathsf{O},\mathcal{A},S'} [f' \notin \mathcal{F}_{\mathsf{View}(f,\mathsf{O},\mathcal{A}),\mathsf{O},\mathcal{A}}] \\ &= \Pr_{f \leftarrow \mathcal{F},\mathsf{O},\mathcal{A},S'} [\mathsf{View}(f', \mathsf{O}, \mathcal{A}) \neq \mathsf{View}(f, \mathsf{O}, \mathcal{A})] \leq q^2 \cdot 2^{-\lambda} + q' \cdot \delta.\end{aligned}$$

We next consider an experiment where, instead of sampling uniformly random S', we derive S' from S^* using a random permutation over the input space of f. In fact, in order to facilitate the symmetry argument, we will be sampling a uniformly random *order-2* permutation, that is, a permutation π where $\pi = \pi^{-1}$ (equivalently, $\pi^2 = I$). By Proposition 2.6, this still gives us S' that is statistically close to uniform.

$\underline{\mathsf{Exp}'_2 :}$

1. Sample random O, $\mathcal{A}$ and $V \leftarrow \mathcal{V}_{\mathsf{O},\mathcal{A}}$.
2. Sample a uniformly random order-2 permutation $\pi \leftarrow \mathbf{Sym}_2(\{0,1\}^\ell)$.
3. Sample $f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$ and let $S^* = S^*(f, \mathsf{O}, \mathcal{A})$ and $S' = \pi(S^*)$.
4. If $f' = f_{S^* \leftrightarrow S'} \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$, output $(f, f', \mathsf{O}, \mathcal{A})$ and, otherwise, repeat from step 1.

Then, by Proposition 2.6, we may bound the distance between Exp'_2 and Exp'_1 as follows

$$\mathbf{SD}(\mathsf{Exp}'_2, \mathsf{Exp}'_1) \leq \mathbf{SD}(\pi(S^*), S') = O(2^{-\ell} \cdot |S^*|^2) = O(q^2 2^{\ell-2\lambda}).$$

Now, letting

$$\mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi) = \{f \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}} \mid f_{S^* \leftrightarrow S'} \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}, \quad \text{where } S^* = S^*(f, \mathsf{O}, \mathcal{A}) \text{ and } S' = \pi(S^*)\},$$

we may look at f is if being sampled uniformly at random from $\mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$. Moreover, we denote by $\mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}$ the deterministic transformation that transforms any $f \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$ to the corresponding f' as defined in Exp'_2. More formally, we define

$\underline{\mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f) :}$ 1. Let $S^* = S^*(f, \mathsf{O}, \mathcal{A})$. 2. Let $S' = \pi(S^*)$. 3. Output $f' = f_{S^* \leftrightarrow S'}$.

and observe that $\mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}$ is in fact 1-1 over $\mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$! To see this, notice that, if $f \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$ then $f' = \mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f)$ satisfies $f' \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}$ and, therefore, $\mathsf{View}(f', \mathsf{O}, \mathcal{A}) = V = \mathsf{View}(f, \mathsf{O}, \mathcal{A})$ and

$$S^*(f', \mathsf{O}, \mathcal{A}) = \bigcup_{x : \mathcal{A}^{f'} \to x} \mathbf{sib}^{f'}(x) = \bigcup_{x : \mathcal{A}^{f} \to x} \mathbf{sib}^{f'}(x) = S'.$$

Further, since π is of order 2, then it holds that $\pi(S') = \pi^{-1}(S') = S^*$. Hence, $f = \mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f')$ and $f' \in \mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$. We may use $\mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}$ then, to look at $\mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$ as the disjoint union of pairs $\mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi) = \dot{\bigcup}_{(f,f') \in \mathcal{P}} \{f, f'\}$ where, for each such pair $(f, f') \in \mathcal{P}$, $f = \mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f')$ or, equivalently, $f' = \mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f)$. By these insights, we may rewrite the last experiment above as follows

$\underline{\mathsf{Exp}'_3 :}$

1. Sample $\mathsf{O} \leftarrow \mathcal{O}$, random $\mathcal{A}$ and $V \leftarrow \mathcal{V}_{\mathsf{O},\mathcal{A}}$.
2. Sample $\pi \leftarrow \mathbf{Sym}_2(\{0,1\}^\ell)$.
3. Sample uniformly random $f \leftarrow \mathcal{F}_{V,\mathsf{O},\mathcal{A}}(\pi)$ and let $f' = \mathsf{Swap}_{\mathsf{O},\mathcal{A},\pi}(f)$ and output $(f, f', \mathsf{O}, \mathcal{A})$.

where it is easy to see that f and f' are symmetric – they are simply a uniformly random pair $(f, f') \leftarrow \mathcal{P}$ from the disjoint union $\mathcal{P}$ defined above (which is deterministically defined by $\mathsf{O}, \mathcal{A}, V, \pi$). This completes the proof of the claim. □

By claim 3 and as hinted above, it immediately follows that

$$\mathbf{SD}((f', S^*), (f, S')) = \mathbf{SD}((f', S^*(f, \mathsf{O}, \mathcal{A})), (f, S^*(f', \mathsf{O}, \mathcal{A}))) = O(q^2 \cdot 2^{\ell-2\lambda} + q' \cdot \delta)$$

and, hence,

$$\begin{aligned}\Pr[\mathsf{Exp}' = 1] &= \Pr_{f', S^*}[\mathcal{A}^{f', \mathsf{O}^{f'}}(1^\lambda) \xrightarrow{f} S^*] \\ &= \Pr_{f, S'}[\mathcal{A}^{f, \mathsf{O}^{f}}(1^\lambda) \xrightarrow{f} S'] + O(q^2 \cdot 2^{\ell-2\lambda} + q' \cdot \delta) \\ &= O(q^2 \cdot 2^{\ell-2\lambda} + q' \cdot \delta)\end{aligned}$$

and the proof of Lemma 5.15 is finished.

5.5 Towards Adaptive Differential Indistinguishability against Siblings

We have reduced our task in Section 5.4 to design a correlation finder with adaptive differential indistinguishability against siblings. As a first step towards this end, we identify in this section two key properties of siblings distributions that make it eventually possible to build such a correlation finder. In fact, it is hard to imagine a correlation finder that is both correct and adaptively differentially indistinguishable against an adversarial $(\mathsf{C}^*, \mathsf{X}^*)$ that does not satisfy these properties.

The first of these properties is *elusiveness*, using which we rule out a scenario where the adaptive adversary swaps the function specifically at inputs that appear in the transcript produced by the oracle's answer, namely in the execution $C^f(\mathsf{O}^f(C))$. We formalize below.

Definition 5.16 (Elusive $(\mathsf{C}^*, \mathsf{X}^*)$)**.** *Let $\mathcal{O} = \{\mathcal{O}_R\}$ be a correlation finder and let $(\mathsf{C}^*, \mathsf{X}^*)$ be an adversarial distribution against adaptive differential indistinguishability, i.e. a distribution C^* over $\mathcal{C}$ and a probabilistic function $\mathsf{X}^* : \mathcal{C} \times \{0,1\}^* \to \mathbb{P}(\{0,1\}^{\ell(\lambda)})$ (see Definition 5.12). We say that $(\mathsf{C}^*, \mathsf{X}^*)$ is $(\mathcal{O}, \eta)$-*elusive *for $\eta : \mathbb{N} \to [0,1]$ if for any $R \in \mathcal{R}$, any $\lambda \in \mathbb{N}$ and any f, it holds that*

$$\Pr_{\substack{\mathsf{O} \leftarrow \mathcal{O}_R, C \leftarrow \mathsf{C}^*_{\lambda,f} \\ X^* \leftarrow \mathsf{X}^*_{\lambda,f}(C, \mathsf{O}^f(C))}} [\exists x^* \in X^* : C^f(\mathsf{O}^f(C)) \rightharpoondown x^*] < \eta(\lambda).$$

The second property is *smoothness*, namely that X^* is smooth under f with high probability.

Definition 5.17 (Smooth $(\mathsf{C}^*, \mathsf{X}^*)$)**.** *Let $\mathcal{O} = \{\mathcal{O}_R\}$ be a correlation finder. We say that an adversarial distribution against adaptive differential indistinguishability $(\mathsf{C}^*, \mathsf{X}^*)$ (see Definition 5.12) is $(\mathcal{O}, \tau, \gamma)$-*smooth *for $\tau, \gamma : \mathbb{N} \to [0,1]$ if for any $R \in \mathcal{R}$, any $\lambda \in \mathbb{N}$ and any f, it holds that*

$$\Pr_{\substack{\mathsf{O} \leftarrow \mathcal{O}, C \leftarrow \mathsf{C}^*_{\lambda,f} \\ X^* \leftarrow \mathsf{X}^*_{\lambda,f}(\mathsf{O}^f(C))}} [X^* \in \mathsf{Smooth}^f_\gamma(C)] \geq \tau(\lambda).$$

Siblings are Smooth. In the following straight-forward lemma, we observe that the siblings distribution corresponding to any canonical and smooth adversary $\mathcal{A}$ is smooth as per the definition above.

Lemma 5.18. *Let $\mathcal{O}$ be a query-independent correlation finder and let $\mathcal{A}$ be a (q, q', q'')-bounded* CollFind*-adversary that is canonical and (τ, γ)-smooth. Then, any $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution* A *(see Definition 5.13) is $(\mathcal{O}, \tau, \gamma)$-smooth.*

Proof. Follows by Remark 5.14 and the definition of a smooth adversary (Definition 5.8). □

Siblings are Elusive under Sound Correlation Finders. While the smoothness of siblings follows by definition, it is not clear whether siblings are always elusive even assuming a canonical and smooth adversary. We hereby define two notions of soundness for correlation finders and show that they are sufficient to imply the elusiveness of siblings. At a high level, soundness bounds the amplification in the likelihood to observe any x when computing C^f on the oracle's answer, compared to when computing C^f on a random input. In another way to look at it, this is the exponent of the Rényi divergence of order infinity (which is to min-entropy what KL-divergence is to Shannon entropy) between the queries made by C^f on $\mathsf{O}^f(C)$, namely *information leaked by* O, and the queries made by C^f on a random input, namely *information simulatable without* O. We also define pairwise-soundness, which simply considers the likelihood to observe pairs rather than individual inputs.

Definition 5.19 (Soundness)**.** *Let $\mathcal{O} = \{\mathcal{O}_R\}$ be a correlation finder for $\mathcal{R}$. We define two notions of soundness for $\mathcal{O}$:*

– *We say that $\mathcal{O}$ has* soundness *$\epsilon : \mathbb{N} \to [0,1]$ if for any $R \in \mathcal{R}$, any f, any $C \in \mathcal{C}$, any $\lambda \in \mathbb{N}$, $\gamma : \mathbb{N} \to [0,1]$ and any $x \in \{0,1\}^{\ell(\lambda)}$, it holds that*

$$\Pr_{\mathsf{O} \leftarrow \mathcal{O}_R}[C^f(\mathsf{O}^f(C)) \rightharpoondown x] \leq \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x]/\epsilon(\lambda).$$

- *We say that $\mathcal{O}$ has* pairwise-soundness $\epsilon : \mathbb{N} \to [0,1]$*, if for any $R \in \mathcal{R}$, any f, any $\lambda \in \mathbb{N}$ and any $x_1, x_2 \in \{0,1\}^{\ell(\lambda)}$, it holds that*

$$\Pr_{\mathsf{O} \leftarrow \mathcal{O}_R}[C^f(\mathsf{O}^f(C)) \rightharpoondown x_1, x_2] \leq \Pr_{z \in \{0,1\}^m}[C^f(z) \rightharpoondown x_1, x_2]/\epsilon(\lambda).$$

Lemma 5.20. *Fix f and let $\mathcal{O}$ be a correlation finder (for any fixed relation) with soundness ϵ. Let $C \in \mathcal{C}$ be an f-aided circuit that makes at most $q(\lambda)$ f_λ-queries for any $\lambda \in \mathbb{N}$ and on any input. Then, for any $\gamma : \mathbb{N} \to [0,1]$, any $\lambda \in \mathbb{N}$ and any subset $S \subseteq \{0,1\}^{\ell(\lambda)}$ such that $S \in \mathsf{Smooth}^f_\gamma(C)$, it holds that*

$$\Pr_{\mathsf{O} \leftarrow \mathcal{O}}[\exists x \in S : \; C^f(\mathsf{O}^f(C)) \rightharpoondown x] \leq q(\lambda) \cdot \gamma(\lambda)/\epsilon(\lambda).$$

Additionally, if $\mathcal{O}$ has pairwise-soundness ϵ and $C \in \mathcal{C}^*_{\lambda,f,\gamma}$, then

$$\Pr_{\mathsf{O} \leftarrow \mathcal{O}}[\exists (x_1, x_2) \in \mathsf{Coll}^f_\lambda : \; C^f(\mathsf{O}^f(C)) \rightharpoondown x_1, x_2] \leq q^2(\lambda) \cdot \gamma(\lambda)/\epsilon(\lambda).$$

Proof. We start with the first statement in the lemma. It holds by the presumed soundness of $\mathcal{O}$ that

$$\Pr_{\mathsf{O} \leftarrow \mathcal{O}}[\exists x \in S : \; C^f(\mathsf{O}^f(C)) \rightharpoondown x] \leq \sum_{x \in S} \Pr_{\mathsf{O} \leftarrow \mathcal{O}}[C^f(\mathsf{O}^f(C)) \rightharpoondown x] \leq \sum_{x \in S} \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x]/\epsilon(\lambda).$$

We can write

$$\sum_{x \in S} \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x] = 2^{-m} \cdot \sum_{\substack{z \leftarrow \{0,1\}^m \\ x \in S}} \mathbb{1}(C^f(z) \rightharpoondown x)$$

where we recall $\mathbb{1}(\cdot) \in \{0,1\}$ denotes the predicate function for its argument. Now, by the assumption on C, there are at most $q := q(\lambda)$ distinct inputs $x \in S$ such that $C^f(z) \rightharpoondown x$. Therefore, assuming w.l.o.g. that C always makes exactly $q(\lambda)$ queries to f_λ, we may bound the above sum by

$$2^{-m} \cdot \sum_{\substack{z \in \{0,1\}^m \\ x \in S}} \mathbb{1}(C^f(z) \rightharpoondown x) \leq q 2^{-m} \cdot \sum_{\substack{z, (x_1, \ldots, x_q): \\ (x_1, \ldots, x_q) \cap S \neq \emptyset}} \mathbb{1}(C^f(z) \rightharpoondown x_1, \ldots, x_q),$$

where $(x_1, \ldots, x_q)$ iterates over all q-tuples of inputs in $\{0,1\}^{\ell(\lambda)}$. We rewrite

$$q 2^{-m} \cdot \sum_{\substack{z, (x_1, \ldots, x_q): \\ (x_1, \ldots, x_q) \cap \mathsf{Coll}^f_\lambda \neq \emptyset}} \mathbb{1}(C^f(z) \rightharpoondown x_1, \ldots, x_q) = q \cdot \mathbb{E}_{z \leftarrow \{0,1\}^m}[\sum_{\substack{(x_1, \ldots, x_q): \\ (x_1, \ldots, x_q) \cap S \neq \emptyset}} \mathbb{1}(C^f(z) \rightharpoondown x_1, \ldots, x_q)].$$

and observe that the sum in the expectation above is always binary; by the q-boundedness of C, at most a single summon takes the value of 1. Moreover, this occurs only when $C^f(z)$ calls some $x \in S$. Thus, we conclude that for any subset S,

$$\sum_{x \in S} \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x] \leq q \cdot \Pr_{z \leftarrow \{0,1\}^m}[\exists x \in S : \; C^f(z) \rightharpoondown x] \tag{9}$$

and finish the proof of the first part of the lemma by the fact that $S \in \mathsf{Smooth}^f_\gamma(C)$. The second part of the lemma, concerning collision smooth circuits, follows by an identical proof, except we rely on the pairwise-soundness of $\mathcal{O}$ and we get an overhead of q^2 since we are counting pairs. □

Lemma 5.21. *Let $\mathcal{O} = \{\mathcal{O}_R\}$ be a query-independent correlation finder with soundness and pairwise-soundness ϵ and let $\mathcal{A}$ be a (q, q', q'')-bounded* $\mathsf{CollFind}$*-adversary that is canonical and (τ, γ)-smooth. Then, any $(\mathcal{F}, \mathcal{O}_R, \mathcal{A})$-siblings distribution $(\mathsf{C}^*, \mathsf{X}^*)$ (see Definition 5.13) is $(\mathcal{O}_R, \eta)$-elusive for $\eta = O((1-\tau) + (q'')^2 \cdot \gamma/\epsilon)$.*

Proof. By the definition of elusiveness (Definition 5.16), we must bound the probability that, for a random $C \in \mathsf{C}^*_{\lambda,f}$, a random oracle $\mathsf{O} \leftarrow \mathcal{O}$ returns an answer $z \in \{0,1\}^{m(n)}$ such that $C^f(z)$ calls an input in the induced set of siblings $S^* = \mathsf{X}^*_{\lambda,f}(C,z)$. Recall that, by Remark 5.14, when $\mathcal{O}$ is query-independent, such an (C, S^*) imitates, for some $i \in \mathbb{N}$, the distribution of the i^{th} query made by $\mathcal{A}$ and the siblings of inputs seen by $\mathcal{A}^{f,\mathsf{O}^f}$. Therefore, roughly speaking, the aforementioned event occurs when $C^f(z)$ calls a sibling of either

- an f-query which $\mathcal{A}$ made prior to making the query C, in which case O_i is independent in the sibling's identity and we may use the soundness property to claim that the sibling is called by $C^f(A)$ with low probability, or
- an f-query which $\mathcal{A}$ makes after making the query C because it appears in the execution $C^f(A)$ (recall $\mathcal{A}$ is canonical), which means that $C^f(A)$ calls a collision and occurs with low probability by the collision-soundness of $\mathcal{O}$, or
- an f-query which $\mathcal{A}$ makes after making the query C and that does not appear in $C^f(A)$, i.e. after finishing calling all x s.t. $C^f(A) \rightharpoondown x$ (see Definition 5.5 of canonical adversaries), a scenario we claim impossible by the canonicality of $\mathcal{A}$.

More formally, by Remark 5.14, it holds for any $(\mathcal{F}, \mathcal{O}, \mathcal{A})$-siblings distribution $(\mathsf{X}^*, \mathsf{C}^*)$ that, for some $i \in \mathbb{N}$ and any $f \in \mathcal{F}$,

$$\Pr_{\substack{\mathsf{O}_i \leftarrow \mathcal{O}, C \leftarrow \mathsf{C}^*_{\lambda,f} \\ S^* \leftarrow \mathsf{X}^*_{\lambda,f,C}(\mathsf{O}_i^f(C))}} [\exists x^* \in S^* : C^f(\mathsf{O}^f(C)) \rightharpoondown x^*] = \Pr_{\substack{\mathcal{A}, \mathsf{O} \leftarrow \mathcal{O} \\ S^* = \mathbf{Sib}_\lambda(f, \mathsf{O}, \mathcal{A})}} [\exists x^* \in S^* : C^f(\mathsf{O}_i^f(C)) \rightharpoondown x^*]$$

Now, denote by $S^*_{<i}$ the subset of S^* corresponding to siblings of f-queries made prior to making the i^{th} query, by $S^*_{=i}$ the subset of S^* corresponding to siblings of any x such that $C^f(\mathsf{O}^f(C)) \rightharpoondown x$, and by $S^*_{>i}$ the subset of S^* corresponding to siblings of f-queries x made after making the i^{th} query such that $C^f(\mathsf{O}^f(C))$ *does not* call f at x. Evidently, these three subsets cover S^* entirely. We now analyze the probability of $C^f(\mathsf{O}^f(C))$ hitting each of these subsets, then finish using a union bound.

- $\underline{S^*_{<i}}$: Notice, first, that O_i is random over $\mathcal{O}$ and independent in $S^*_{<i}$ and the i^{th} query C. Second, by the smoothness of $\mathcal{A}$ at siblings (Definition 5.8), we know that

  $$\Pr_{\mathcal{A}}[S^*_{<i} \in \mathsf{Smooth}^f_\gamma(C)] \geq \tau$$

  and, therefore, by the presumed soundness of $\mathcal{O}$ and Lemma 5.20

  $$\begin{aligned} \Pr_{\mathcal{A},\mathsf{O}}[\exists x^* \in S^*_{<i} : C^f(\mathsf{O}_i^f(C)) \rightharpoondown x^*] &\leq (1-\tau) + \max_{C, S \in \mathsf{Smooth}^f_\gamma(C)} \Pr_{\mathsf{O}_i}[\exists x^* \in S : C^f(\mathsf{O}_i(C)) \rightharpoondown x^*] \\ &\leq (1-\tau) + q'' \cdot \gamma/\epsilon. \end{aligned}$$

- $\underline{S^*_{=i}}$: The case that $C^f(\mathsf{O}^f(C))$ hits $S^*_{=i}$ occurs when $C^f(\mathsf{O}^f(C))$ calls an x and one of its siblings. We similarly bound the probability of such a scenario based on the smoothness at collisions of $\mathcal{A}$, the pairwise-soundness of $\mathcal{O}$ and Lemma 5.20 as follows (recall that O_i is independent in C)

  $$\begin{aligned} &\Pr_{\mathcal{A},\mathsf{O}}[\exists x^* \in S^*_{=i} : C^f(\mathsf{O}_i^f(C)) \rightharpoondown x^*] \\ &\leq \Pr_{\mathcal{A},\mathsf{O}}[\exists (x_1, x_2) \in \mathsf{Coll}^f_\lambda : C^f(\mathsf{O}_i^f(C)) \rightharpoondown x_1, x_2] \\ &\leq (1-\tau) + \max_{C \in \mathcal{C}^*_{\lambda,f,\gamma'}} \Pr_{\mathsf{O}_i}[\exists (x_1, x_2) \in \mathsf{Coll}^f_\lambda : C^f(\mathsf{O}_i^f(C)) \rightharpoondown x_1, x_2] \\ &\leq (1-\tau) + (q'')^2 \cdot \gamma/\epsilon. \end{aligned}$$

– $\underline{S^*_{>i}}$: If $C^f(\mathsf{O}^f(C))$ ever hits an $x^* \in S^*_{>i}$, i.e. an x^* such that there is some x that $\mathcal{A}^{f,\mathsf{O}^f}(1^\lambda)$ calls *after making* C and querying all the induced f-queries, then by the point that $\mathcal{A}$ is about to query f at x, $x^* \in \mathbf{sib}(x)$ has already been called by $\mathcal{A}$. By the symmetry of **sib**, it holds that $x \in \mathbf{sib}(x^*)$ and, hence, x is the last query made by $\mathcal{A}$, after which it immediately halts and outputs (x, x^*) (by canonicality, see Definition 5.5). Recall that we define **Sib** to exclude the siblings of the last query (see (3)). Hence, such a scenario is impossible and has probability zero.

□

The Smoothness and Elusiveness of Non-Adaptive x^*. In the next section we will show a construction of correlation finder that will satisfy adaptive differential indistinguishability against smooth and elusive distributions $(\mathsf{X}^*, \mathsf{C}^*)$. We have demonstrated above that this captures siblings distributions and is therefore sufficient to derive separation from CRH (via Lemma 5.15). For the construction to be sufficient for a separation from OWP as well, we must prove that our non-adaptive notion of differential indistinguishability (see Definition 5.10, Lemma 5.11) is a special case of adaptive differential indistinguishability against smooth and elusive distributions.

Remark 5.22. *If $\mathcal{O}$ is ϵ-sound and adaptively differentially δ-indistinguishable for $\mathcal{F}$ against any $(\mathcal{O}, 1, \gamma)$-smooth and $(\mathcal{O}, \gamma/\epsilon)$-elusive* $(\mathsf{C}^*, \mathsf{X}^*)$ *then it is differentially (q, γ, δ)-indistinguishable for any q.*

Proof. Let $(\mathsf{C}^*, \mathsf{X}^*)$ be any pair where C^* outputs a constant $C \in \mathcal{C}$ such that, for all $\lambda \in \mathbb{N}$, $C \leftarrow \mathsf{C}^*_{\lambda,f}$ makes at most $q(\lambda)$ queries to f_λ on any input, and $\mathsf{X}^* = \{\mathsf{X}^*_{\lambda,f} : \mathcal{C} \times \{0,1\}^* \to \{0,1\}^{\ell(\lambda)}\}$ is independent of its second argument and satisfies $\Pr_{x^* \leftarrow \mathsf{X}^*(C)}[x^* \in \mathsf{Smooth}^f_\gamma(C)] = 1$. Notice that adaptive differential δ-indistinguishability for all such distributions implies (non-adaptive) differential (q, γ, δ)-indistinguishability. These distributions are all $(\mathcal{O}, 1, \gamma)$-smooth by inspection and, further, the fact they are all $(\mathcal{O}, \gamma)$-elusive follows easily by smoothness (notice the answer of O is independent in x^*) and the ϵ-soundness of $\mathcal{O}$. □

6 The Correlation Finder

In this section, we build a correlation finder that is correct and adaptively differentially indistinguishable against elusive and smooth adversarial choices (as defined in Section 5.5). Given the work done so far, in particular in Lemmas 5.15, 5.18 and 5.21, this is sufficient to derive the desired separation.

Our correlation finder follows a natural structure of a *picky correlation finder*, namely a correlation finder that given any input, rejects with some probability (outputs $\perp$[7]) and, otherwise, simply outputs a uniformly random correlation under the target relation. Given this framework, it remains only to specify the *rejection policy* of our correlation finder, that is, the probability with which he rejects for any given input. As a first step, we specify a list of conditions on the rejection policy (which is simply a function from the query space $\mathcal{C}$ to real values in $[0, 1]$) and show that any policy that satisfies these conditions gives a correlation finder that is both correct and (adaptively) differentially indistinguishable. Having these conditions in hand, we then proceed to define our rejection policy and show that it satisfies these conditions.

6.1 Strategy: Picky Correlation Finder

First, we define the *set of correlations* between a circuit and a relation.

Definition 6.1 (Set of Correlations)**.** *Let $f : \{0,1\}^* \to \{0,1\}^*$ be any oracle function and $C^f : \{0,1\}^m \to \{0,1\}^n$ be an f-aided circuit. Let R be a relation. The* set of (R, C^f)-correlations *is defined as*

$$\mathsf{Corr}^f_{R,C} = \{z \mid (z, C^f(z)) \in R\}.$$

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

[7] Although our definition for a correlation finder allows only for outputs in $\{0,1\}^m$, we can always dedicate some $z \in \{0,1\}^m$ (e.g. the all-zeros input) to correspond to $\perp$.

We are now ready to present our generic picky correlation finder, which we will late instantiate with a proper rejection policy.

Construction 6.1 (Picky Correlation Finder)**.** *Let $\mathcal{R}$ be a relation class and $\mathcal{F}$ be a class of oracles. Let $\rho := \{\rho^f_R : \mathcal{C} \to [0,1]\}$ be a an ensemble of functions (namely, a* rejection policy*) that take as input a description of an f-aided circuit and outputs a real value in $[0,1]$ w.r.t. fixed $R \in \mathcal{R}$ and $f \in \mathcal{F}$. We define our* picky correlation finder *with rejection policy ρ, which we denote by $\mathcal{CF}[\rho] = \{\mathcal{CF}_R[\rho_R]\}_{R\in\mathcal{R}}$, such that, for every $R \in \mathcal{R}$, $\mathcal{CF}_R[\rho_R]$ is the distribution over deterministic oracles where, for any C, letting $\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R]$, $\mathsf{CF}^f_R(C)$ is an independent random variable that is equal to the random output of the following algorithm* [8]

$\underline{\mathsf{CF}^f_R(C):}$

Reject *with probability $\rho^f_R(C)$ and, otherwise, output a uniformly random* ***correlation*** *$z_C \leftarrow \mathsf{Corr}^f_{R,C}$ (if $\mathsf{Corr} = \emptyset$, set $z_C = \perp$).*

We point out that $\mathcal{CF}[\rho]$ is by construction query-independent (as by Definition 5.3).

6.2 Sufficient Conditions on the Rejection Policy

In this section, we state a list of conditions on the rejection policy ρ from Construction 6.1 that are sufficient for $\mathcal{CF}[\rho]$ to be correct and adaptively differentially indistinguishable for siblings, paving the way towards a separation via a picky correlation finder.

Lemma 6.2. *Let $t, q, N : \mathbb{N} \to \mathbb{N}$ and $\epsilon, \tau, \gamma, \eta : \mathbb{N} \to [0,1]$. Let $\mathcal{R}$ be a relation class and let $\rho := \{\rho^f_R : \mathcal{C}^f \to [0,1]\}_{R\in\mathcal{R}}$ that satisfies the following properties:*

- ***Correctness:*** *For any $f \in \mathcal{F}$ and any circuit $C = \{C_n \in \mathcal{C}_n\}$ with query complexity bounded by $t(n)$, it holds that*
$$\mathbb{E}_{R\leftarrow\mathcal{R}}[\rho^f_R(C_n)] < \frac{1}{2n^2}$$
for infinitely many $n \in \mathbb{N}$.
- ***Soundness:*** *For any $R \in \mathcal{R}$, any $f \in \mathcal{F}$ and any circuit $C \in \mathcal{C}$, if $\rho^f_R(C) < 1$ then, for any $\lambda \in \mathbb{N}$ and any $x \in \{0,1\}^{\ell(\lambda)}$, it holds that*
$$\Pr_{z\leftarrow\mathsf{Corr}^f}[C^f(z) \rightharpoondown x] < \Pr_{z\leftarrow\{0,1\}^m}[C^f(z) \rightharpoondown x]/\epsilon(\lambda).$$
- ***Worst-case Differential Indistinguishability:*** *For any $R \in \mathcal{R}$, any $f \in \mathcal{F}$, any circuit $C \in \mathcal{C}$, any $\lambda \in \mathbb{N}$ and any $X^*, X' \subseteq \{0,1\}^{\ell(\lambda)}$ such that $|X^*|= |X'|$ and $X^*, X' \in \mathsf{Smooth}^f_\gamma(C)$, it holds that*
$$|\rho^{f'}_R(C) - \rho^f_R(C)|< \delta(\lambda),$$
where $f' = f_{X^*\leftrightarrow X'}$.

Then, $\mathcal{CF}[\rho] = \{\mathcal{CF}_R[\rho_R]\}$ from Construction 6.1 satisfies

➤ ***Correctness*** *as required by Lemma 5.2, for CIH candidates with query complexity bounded by $t(n)$,*

➤ ***Soundness*** *ϵ (as defined in Definition 5.19), and*

[8] Although the described algorithm has access to f, we think of its random coins as being sampled obliviously of f (w.l.o.g.), and therefore $\mathcal{CF}_R[\rho_R]$ is well-defined prior to setting f. One way to sample such an oracle is proposed in the proof of Lemma 6.2.

- ***Adaptive Differential δ'-Indistinguishability** , where*

  $$O((1-\tau)+\delta+(qN/\gamma)\cdot 2^{-\ell}+2q\cdot\gamma/\epsilon+\eta)$$

  ***against any** $(\mathcal{CF}[\rho],\tau,\gamma)$-**smooth and** $(\mathcal{CF}[\rho],\eta)$-**elusive** $(\mathsf{C}^*,\mathsf{X}^*)$ such that any $C\in\mathsf{C}^*_{\lambda,f}$ makes at most $q(\lambda)$ f_λ-queries for any $\lambda\in\mathbb{N}$ on any input and any $X^*\in\mathsf{X}^*_{\lambda,f,C}$ has size at most $N(\lambda)$.*

Additionally, if ρ satisfies

- ***Pairwise-Soundness:** For any $R\in\mathcal{R}$, $f\in\mathcal{F}$ and any circuit $C\in\mathcal{C}$, if $\rho^f_R(C)<1$ then, for any $\lambda\in\mathbb{N}$ and any $x_1,x_2\in\{0,1\}^{\ell(\lambda)}$, it holds that*

  $$\Pr_{z\leftarrow\mathsf{Corr}^f}[C^f(z)\rightharpoondown x_1,x_2]<\Pr_{z\leftarrow\{0,1\}^m}[C^f(z)\rightharpoondown x_1,x_2]/\epsilon(\lambda),$$

then, for any $R\in\mathcal{R}$, $\mathcal{CF}_R[\rho_R]$ satisfies

- ***Pairwise-Soundness** ϵ (as defined in Definition 5.19).*

The following two corollaries result from plugging in Remark 5.22 into Lemma 6.2 and, respectively, Lemmas 5.18 and 5.21 into Lemma 6.2.

Corollary 6.3. *Let ρ be a rejection policy satisfying the assumptions of Lemma 6.2 w.r.t. a relation class $\mathcal{R}$. Let $\mathcal{A}$ be a (q,q',q'')-bounded Inv-adversary that is canonical and (τ,γ)-smooth. Then, $\mathcal{CF}[\rho]$ is differentially (q'',γ,δ')-indistinguishable for permutations, where*

$$\delta'=O(\delta+(q''/\gamma)\cdot 2^{-\lambda}+q''\cdot\gamma/\epsilon).$$

Corollary 6.4. *Let ρ be a rejection policy satisfying the assumptions of Lemma 6.2 w.r.t. a relation class $\mathcal{R}$. Let $\mathcal{A}$ be a (q,q',q'')-bounded $\mathsf{CollFind}$-adversary that is canonical and (τ,γ)-smooth. Then, $\mathcal{CF}[\rho]$ is adaptively differentially δ'-indistinguishable against any $(\mathcal{F},\mathcal{CF}_R[\rho_R],\mathcal{A})$-siblings, where*

$$\delta'=O((1-\tau)+\delta+(qq''/\gamma)\cdot 2^{-\lambda}+(q'')^2\cdot\gamma/\epsilon).$$

In what follows till the end of this section is a proof of Lemma 6.2.

Proof of Lemma 6.2. Soundness is immediate and correctness follows quite straight-forwardly by our construction; Let $H=\{H_n\}$ be a hash family. Then, it holds that

$$\begin{aligned}\Pr_{\substack{f,h\leftarrow H_n\\ R\leftarrow\mathcal{R},\mathsf{CF}\leftarrow\mathcal{CF}_R}}[z\leftarrow\mathsf{CF}^f(h);(z,h^f(z))\notin R]&\le\max_{f,C}\Pr_{\substack{R\leftarrow\mathcal{R}\\ \mathsf{CF}\leftarrow\mathcal{CF}_R}}[z\leftarrow\mathsf{CF}^f(C_n);(z,C^f_n(z))\notin R]\\&=\max_{f,C}\mathbb{E}_{R\leftarrow\mathcal{R}}\Pr_{\mathsf{CF}\leftarrow\mathcal{CF}_R}[z\leftarrow\mathsf{CF}^f(C_n);(z,C^f_n(z))\notin R]=\max_{f,C}\mathbb{E}_{R\leftarrow\mathcal{R}}[\rho^f_R(C_n)]<1/2n^2.\end{aligned}$$

It remains, then, to show adaptive differential indistinguishability holds given the conditions assumed in the lemma. Let us fix $R\in\mathcal{R}$ and omit it from notation. Recall that, by the definition of adaptive differential indistinguishability (see Definition 5.12) and Proposition 2.2, it suffices to show that for any $\lambda\in\mathbb{N}$ and any f, it holds that

$$\Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[\mathsf{CF}^{f'}(C)\neq\mathsf{CF}^f(C)]$$

where $C\leftarrow\mathsf{C}^*_{f,\lambda}$ and $f'=f_{X^*\leftrightarrow X'}$ for $X^*\leftarrow\mathsf{X}^*_{\lambda,f,C}(\mathsf{CF}^f(C))$ and a uniformly random $X'\subseteq\{0,1\}^{\ell(\lambda)}$ of size $|X^*|$.

We first note that by the presumed smoothness of $(\mathsf{C}^*,\mathsf{X}^*)$ we have $\Pr[X^*\notin\mathsf{Smooth}^f_\gamma(C)]\le(1-\tau)$. Second, we use the following claim to deduce that the random X' is also smooth with high probability.

Claim 4. *Fix a circuit $C \in \mathcal{C}$ that makes at most $q''(\lambda)$ queries to f_λ on any input, fix $\lambda \in \mathbb{N}$ and let $X' \subseteq \{0,1\}^{\ell(\lambda)}$ be a uniformly random subset of size N. Then, it holds for any $\gamma : \mathbb{N} \to [0,1]$ that*

$$\Pr_{X'}[X' \notin \mathsf{Smooth}^f_\gamma(C)] \leq q'' N 2^{-\ell(\lambda)}/\gamma.$$

Proof. By definition and Markov,

$$\begin{aligned}\Pr_{X'}[X' \notin \mathsf{Smooth}^f_\gamma(C)] &= \Pr_{X'}[\Pr_{z \leftarrow \{0,1\}^m}[\exists x' \in X' :\ C^f(z) \rightharpoondown x'] > \gamma] \\ &< \mathbb{E}_{X'}[\Pr_{z \leftarrow \{0,1\}^m}[\exists x' \in X' :\ C^f(z) \rightharpoondown x']]/\gamma,\end{aligned}$$

where

$$\mathbb{E}_{X'}[\Pr_{z \leftarrow \{0,1\}^m}[\exists x' \in X' :\ C^f(z) \rightharpoondown x']] = \mathbb{E}_{X',z}[\mathbb{1}(\exists x' \in X' :\ C^f(z) \rightharpoondown x')] \leq \max_z \Pr_{X'}[\exists x' \in X' :\ C^f(z) \rightharpoondown x'].$$

Fixing any $z \in \{0,1\}^m$, notice that the probability, over a random choice of X', that $C^f(z)$ calls any $x' \in X'$ can be bound by $q'' \cdot N 2^{-\ell(\lambda)}$ based on the assumed limitation over C. Therefore,

$$\mathbb{E}_{X'}[\Pr_{z \leftarrow \{0,1\}^m}[\exists x' \in X' :\ C^f(z) \rightharpoondown x']] \leq q'' N 2^{-\ell(\lambda)}$$

and we finish. □

By the above, we may conclude that both X^* and X' are smooth with high probability. Formally, letting E denote the event that $X^*, X' \in \mathsf{Smooth}^f_\gamma(C)$, we know that $\Pr[\mathsf{E}] \geq \tau - q'' N 2^{-\ell}/\gamma$ and, therefore,

$$\Pr_{\substack{\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R] \\ C, X^*, X'}}[\mathsf{CF}^{f'}(C) \neq \mathsf{CF}^f(C)] \leq \Pr_{\substack{\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R] \\ C, X^*, X'}}[\mathsf{E} \wedge \mathsf{CF}^{f'}(C) \neq \mathsf{CF}^f(C)] + (1 - \tau) + q'' N 2^{-\ell}/\gamma \qquad (10)$$

We are now prepared to proceed with the analysis conditioned on E. We frame the sampling of $\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R]$ as sampling, for any possible query C, a uniformly random point $r_C \leftarrow [0,1]$ (i.e., $\Pr[r \in [a,b]] = b - a$ for any $0 \leq a \leq b \leq 1$) and a uniformly random permutation $\pi_C : \{0,1\}^m \to \{0,1\}^m$. Then, when given access to f, the oracle computes the following function on C:
$\underline{\mathsf{CF}^f(C; r, \pi)}$:

1. If $r \leq \rho^f(C)$, reject.
2. Otherwise, output $z = \pi(i)$ where i is the lexicographically smallest element in $\{0,1\}^m$ such that $\pi(i) \in \mathsf{Corr}^f_{R,C}$.

The fact that the above representation of $\mathsf{CF}^f(C; r, \pi)$ matches a random $\mathsf{CF}^f(C)$ as by Construction 6.1 is evident.

Now, observe that $\mathsf{CF}^{f'}(C) \neq \mathsf{CF}^f(C)$ occurs either when rejection occurs in one of the executions but not in the other, or when rejection occurs in neither, yet the answers given by the two are different. The former case occurs only when r is in the range between $\rho^{f'}(C)$ and $\rho^f(C)$, since in any other scenario the rejection decision is similar under f and f'. This happens with probability equal to the expected length of this range and, since we are conditioning on E, we can use the assumed differential indistinguishability of the rejection policy (see statement of Lemma) to derive a bound. Namely,

$$\begin{aligned}
&\Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[\mathsf{E}\wedge\mathsf{CF}^{f'}(C)\neq\mathsf{CF}^{f}(C)]\\
&\le \Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[\mathsf{E}\wedge\mathsf{CF}^{f'}(C)\neq\mathsf{CF}^{f}(C)\mid \mathsf{CF}^{f'}(C),\mathsf{CF}^{f}(C)\neq\bot]+\Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[r\in[\rho^{f'}(C),\rho^{f}(C)]\mid\mathsf{E}]\\
&= \Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[\mathsf{E}\wedge\mathsf{CF}^{f'}(C)\neq\mathsf{CF}^{f}(C)\mid \mathsf{CF}^{f'}(C),\mathsf{CF}^{f}(C)\neq\bot]+\mathbb{E}_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[|\rho^{f'}(C)-\rho^{f}(C)||\;\mathsf{E}]\\
&\le \Pr_{\substack{\mathsf{CF}\leftarrow\mathcal{CF}_R[\rho_R]\\ C,X^*,X'}}[\mathsf{E}\wedge\mathsf{CF}^{f'}(C)\neq\mathsf{CF}^{f}(C)\mid \mathsf{CF}^{f'}(C),\mathsf{CF}^{f}(C)\neq\bot]+\delta. \qquad (11)
\end{aligned}$$

It would be sufficient at this point, then, to bound the probability expression on the left-hand side of (11). We note that, when the oracle does not reject nor under f neither under f', then we have that $\mathsf{CF}^f(C)=\pi(i)$ and $\mathsf{CF}^{f'}(C)=\pi(i')$ for the lexicographically smallest i and i' that satisfy $\pi(i)\in\mathsf{Corr}^f_{R,C}$ and, resp., $\pi(i')\in\mathsf{Corr}^{f'}_{R,C}$. Therefore, the inequality occurs only when $i\neq i'$, which subsequently means that there is some $x\in X^*\cup X'$ that is called by either $C^f(\pi(i))$ or $C^{f'}(\pi(i'))$ (otherwise, these executions are identical under f and f' and, therefore, they must be both corresponding to a correlation, implying $i=i'$). Based on this observation, we bound the probability for $i\neq i'$ via the following two claims and complete the proof.

Claim 5. $\Pr_{\mathsf{CF},C,X^*,X'}[i<i']\le N\cdot q2^{-\ell}+\eta$.

Proof. The event of $i<i'$ occurs only when $C^f(\pi(i))\rightharpoonup x$ for some $x\in X^*\cup X'$ since, otherwise, $\pi(i)$ would have been a correlation also w.r.t. f', implying $i'\le i$. First, since X' is sampled at random independently in any of f,CF and C. Thus, it holds that

$$\Pr_{\mathsf{CF},C,X^*,X'}[\exists x\in X':\ C^f(\mathsf{CF}^f(C))\rightharpoonup x]\le N\cdot q2^{-\ell}.$$

Second, by the elusiveness of $(\mathsf{C}^*,\mathsf{X}^*)$, it holds that

$$\Pr_{\mathsf{CF},C,X^*,X'}[\exists x\in X^*:\ C^f(\mathsf{CF}^f(C))\rightharpoonup x]\le \eta.$$

The proof of the claim concludes by union bound. □

Claim 6. $\Pr_{\mathsf{CF},X^*,X'}[i'<i\ \wedge\ \mathsf{E}]\le 2q\cdot\gamma/\epsilon$.

Proof. As already mentioned, the event $i'>i$ necessarily implies that $C^{f'}(\pi(i'))\rightharpoonup x$ for some $x\in X^*\cup X$ since otherwise the executions $C^{f'}(\pi(i'))$ and $C^f(\pi(i'))$ are identical. Further, notice that if $X^*\cup X'\in\mathsf{Smooth}^f(C)$ then $X^*\cup X'\in\mathsf{Smooth}^{f'}(C)$ since if an execution of C calls an input in $X^*\cup X'$ under f then it must do so under f' as well (these are the only inputs that are altered). Therefore, when conditioning on E, then $X^*\cup X'$ are smooth w.r.t. f' as well and, therefore, we can rely on the soundness of $\mathcal{CF}[\rho]$ to argue that the aforementioned event is improbable. There is, however, a subtle issue in this reasoning that has to be addressed: it is not clear that the random answer of $\mathsf{CF}^{f'}$ on C, namely $\pi(i')$, distributes independently of X^* and X', since the latter are a function of $\mathsf{CF}^f(C)$ which may be correlated with $\mathsf{CF}^{f'}(C)$ through the choice of π.

To overcome this dependency, we observe that, when fixing any C, f and f' such that E holds (notice that f and f' uniquely determine X^* and X' and therefore E), then $\pi(i')$ distributes as if it were a uniformly random $z'\leftarrow\mathsf{Corr}^{f'}\setminus\mathsf{Corr}^f$; It is necessarily in this set since $i'<i$ and it is equal to any of its elements with the same probability by the uniformity of $\pi(i')$. Further, when additionally fixing randomness for X, we see that the event $i'<i$ occurs only when the smallest lexicographically $\hat{i}$ s.t. $\pi(\hat{i})\in\mathsf{Corr}^{f'}\cup\mathsf{Corr}^f$ satisfies

$\pi(\hat{i}) \in \mathsf{Corr}^{f'} \setminus \mathsf{Corr}^{f}$. Since, in such a case, π is uniform conditioned on $\mathsf{X}(C, \pi(i)) = (X^*, X')$, then $\hat{z} = \pi(\hat{i})$ is a uniformly random element in $\mathsf{Corr}^{f'} \cup \mathsf{Corr}^{f}_{X^*,X'}$, where $\mathsf{Corr}^{f}_{X^*,X'}$ is the set of all correlations $z \in \mathsf{Corr}^{f}$ s.t. $\mathsf{X}(C, z) = (X^*, X')$. More formally, by the above and for any fixed C, X, f, f' s.t. E holds, letting π' be the marginal distribution of π under such a fixing, we obtain

$$\begin{aligned}
\Pr_{\pi'}[i' < i] &\leq \Pr_{\pi'}[\exists x \in X^* \cup X' : C^{f'}(\pi'(i')) \rightharpoondown x \wedge i' < i] \\
&= \Pr_{\pi'}[\exists x \in X^* \cup X' : C^{f'}(\pi'(i')) \rightharpoondown x \mid i' < i] \cdot \Pr_{\pi'}[i' < i] \\
&= \Pr_{z' \leftarrow \mathsf{Corr}^{f'} \setminus \mathsf{Corr}^{f}}[\exists x \in X^* \cup X' : C^{f'}(z') \rightharpoondown x] \cdot \Pr_{\hat{z} \leftarrow \mathsf{Corr}^{f'} \cup \mathsf{Corr}^{f}_{X^*,X'}}[\hat{z} \in \mathsf{Corr}^{f'} \setminus \mathsf{Corr}^{f}] \\
&= \Pr_{z' \leftarrow \mathsf{Corr}^{f'} \cup \mathsf{Corr}^{f}_{X^*,X'}}[\exists x \in X^* \cup X' : C^{f'}(z') \rightharpoondown x \wedge z' \in \mathsf{Corr}^{f'} \setminus \mathsf{Corr}^{f}] \\
&\leq \Pr_{z' \leftarrow \mathsf{Corr}^{f'} \cup \mathsf{Corr}^{f}_{X^*,X'}}[\exists x \in X^* \cup X' : C^{f'}(z') \rightharpoondown x \wedge z' \in \mathsf{Corr}^{f'}] \\
&\leq \Pr_{z' \leftarrow \mathsf{Corr}^{f'}}[\exists x \in X^* \cup X' : C^{f'}(z') \rightharpoondown x] \leq q\gamma/\epsilon.
\end{aligned}$$

where the last inequality follows by the soundness ϵ of $\mathcal{CF}[\rho]$ (which we already derived by the soundness assumption in the lemma) and Lemma 5.20 (recall $X^* \cup X'$ is smooth under f' given E).

□

By Claims 5 and 6, we conclude that

$$\Pr_{\substack{\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R] \\ C, X^*, X'}}[\mathsf{E} \wedge \mathsf{CF}^{f'}(C) \neq \mathsf{CF}^{f}(C) \mid \mathsf{CF}^{f'}(C), \mathsf{CF}^{f}(C) \neq \bot] \leq \Pr_{\substack{\mathsf{CF} \leftarrow \mathcal{CF}_R[\rho_R] \\ C, X^*, X'}}[\mathsf{E} \wedge i \neq i'] \leq 2q \cdot \gamma/\epsilon + N \cdot q 2^{-\ell} + \eta.$$

The proof of the adaptive differential indistinguishability part in Lemma 6.2 is then complete by plugging the above into (11) then (10).

6.3 The Rejection Policy

In this section, we present our construction for the rejection policy ρ and show it satisfies the conditions listed in Lemma 6.2, making it utilizable in our proof.

Definitions and Notation. We first lay the groundwork necessary to define our rejection policy. Recall that a rejection probability is calculated for any circuit C with respect to an oracle f and a relation R. In what follows, we often omit C, f and R from notation as they are clear from context.

– **Weights.** For any $\lambda \in \mathbb{N}$, we let

$$\mathbf{X} = \{\mathbf{X}_\lambda\} \qquad\qquad \mathbf{X}_\lambda = \{0,1\}^{\ell(\lambda)} \cup \{(x_1, x_2) \in \{0,1\}^{2 \times \ell(\lambda)} \mid x_1 \neq x_2\}$$

be the set of all points and pairs of f-inputs. Recall that these are the tuples on which soundness and pairwise-soundness hinge. We define the *weight* of any such tuple $\mathbf{x} \in \mathbf{X}$ as

$$\omega^{f}_{R,C}(\mathbf{x}) = \Pr_{z \leftarrow \mathsf{Corr}^{f}_{R,C}}[C^f(z) \rightharpoondown \mathbf{x}] = |\mathsf{Hits}^{f}_{\mathbf{x}} \cap \mathsf{Corr}^{f}_{R,C}| / |\mathsf{Corr}^{f}_{R,C}|, \tag{12}$$

and its *scale* as

$$\sigma^{f}_{C}(\mathbf{x}) = \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown \mathbf{x}] = |\mathsf{Hits}^{f}_{\mathbf{x}}| / 2^m. \tag{13}$$

We will be interested in the ratio between the weight of any given $\mathbf{x}$ and its scale which captures, as the reader may notice, the *amplification* in the likelihood of observing $\mathbf{x}$ due the correlation finder (equivalently, the Rényi divergence of order infinity between the distributions induced by ω and σ as PDFs). We denote

$$\alpha^f_{R,C}(\mathbf{x}) = \omega^f_{R,C}(\mathbf{x})/\sigma^f_{R,C}(\mathbf{x}). \tag{14}$$

– **Neighborhoods.** To guarantee differential indistinguishability, our rejection policy must consider, besides f itself, functions in the d-Neighborhood of f. These are functions that are obtained by applying at most d swaps over f. In fact, we will consider only alterations of the function that consist of a sequence of *smooth swaps*, where at every step, we swap two input sets that are *jointly* smooth relative to the function obtained hitherto. More formally, if we define the set of γ-smooth swaps to be

$$\widetilde{\mathbb{X}}^f_\gamma(\lambda, C) = \{(X^*, X') \mid X^*, X' \subset \{0,1\}^{\ell(\lambda)},\ |X^*| = |X'|,\ X^* \cup X' \in \mathsf{Smooth}^f_\gamma(C)\}$$

and, for any series of $d' \leq d$ *swaps* denoted by

$$\mathbb{X} = ((X_1^*, X_1'), \dots, (X_{d'}^*, X_{d'}')), \qquad \text{where} \quad X_i^*, X_i' \subset \{0,1\}^{\ell(\lambda)},\ |X_i^*| = |X_i'| \quad \text{for all } i,$$

we define the function

$$f_{\mathbb{X}} = f_{\{X_i^* \leftrightarrow X_i'\}_{i \in [d']}},$$

then, the *d-Neighborhood* of f (for smoothness parameter γ) is defined as

$$\mathcal{N}^f_{d,\gamma}(\lambda, C) = \bigcup_{0 \leq d' \leq d} \left\{ f_{\mathbb{X}} \mid \mathbb{X} = ((X_1^*, X_1'), \dots, (X_{d'}^*, X_{d'}')) :\ \forall\, i,\ (X_i^*, X_i') \in \widetilde{\mathbb{X}}^{f_{\mathbb{X}_{i-1}}}_\gamma(\lambda, C) \right\},$$

where $\mathbb{X}_\ell$ denotes the first ℓ swaps in $\mathbb{X}$.

Further, we define the *distance* between functions f and f' as the smallest number of swaps required between smooth inputs in $\{0,1\}^\lambda$ to obtain f' from f (and is set to ∞ if such a transformation is not possible). That is,

$$\Delta_{\gamma,\lambda,C}(f, f') = \min_d f' \in \mathcal{N}^f_{d,\gamma}(\lambda, C).$$

Remark 6.5. *Notice that if $(X^*, X') \in \widetilde{\mathbb{X}}^f_\gamma(\lambda, C)$, then $(X^*, X') \in \widetilde{\mathbb{X}}^{f'}_\gamma(\lambda, C)$ where $f' = f_{X^* \leftrightarrow X'}$. This is because any execution $C^f(z)$ that calls an f-query either in X^* or X' will do so also under f' (and vice-versa). Consequently, for any $f' \in \mathcal{N}^f_{d,\gamma}(\lambda, C)$ it holds that $f \in \mathcal{N}^{f'}_{d,\gamma}(\lambda, C)$ and, hence, the distance function $\Delta_{\gamma,\lambda,C}$ is symmetric.*

The Rejection Policy ρ. We are now prepared to define our rejection policy. As hinted by the above definitions, our strategy is to have the rejection probability be proportional to the worst-case amplification $\alpha(\mathbf{x})$ to obtain soundness and, further, to blur out the difference between adjacent functions to obtain differential indistinguishability, by "spreading out" large amplification factors corresponding to some "bad" f over its neighborhood in the function space. We formally define our rejection policy in Figure 2 below.

In the following three lemmas, we show that our rejection policy ρ satisfies the conditions required by Lemma 6.2: correctness, soundness, pairwise-soundness and adaptive differential indistinguishability. Looking ahead, through their proofs (and via Corollaries 6.3 and 6.4) we will obtain the following consequence.

Corollary 6.6. *Let $\epsilon, \gamma : \mathbb{N} \to [0,1]$ be such that $\epsilon(\lambda) \leq \min(2^{-\lambda/4}, 2^{9\lambda/10}\gamma)$. Let $p : \mathbb{N} \to [0,1]$ be such that $p(n) \geq 4n^2 \cdot 2^{-m(n)}$ and $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > (-\lambda/\log \epsilon(\lambda))(6t(n) + 2)$ for all $n, \lambda \in \mathbb{N}$. Let $\mathcal{R}$ be a k-wise p-universal class of relations (see Definition 4.2). Then, the oracle $\mathcal{CF}[\rho]$ from Construction 6.1, instantiated with the rejection policy from Figure 2 with parameters ϵ, γ, satisfies*

1. ***Correctness** as required by Lemma 5.2, for CIH candidates with query complexity bounded by $t(n)$,*

2. ***Differential** (q'', γ, δ)-**Indistinguishability** where*

$$\delta = O((q''/\gamma) \cdot 2^{-\lambda} + q'' \cdot \gamma/\epsilon).$$

$$\rho = \{\rho_R : \mathcal{C} \to [0,1]\}$$

Parameters: $\epsilon, \gamma : \mathbb{N} \to [0,1]$ (such that $\epsilon \gg \gamma$).
For any $C \in \mathcal{C}$, letting $d = (\epsilon/\gamma t)^2$, where t is the largest number of queries that C makes on any input, we define

$$\rho^f_R(C) = \max_{\substack{\lambda\in\lambda(C),\mathbf{x}\in\mathbf{X}_\lambda \\ f'\in\mathcal{N}^f_{d,\gamma}(\lambda,C)}} [e^{-\Delta_{\gamma,\lambda,C}(f,f')/\sqrt{d}} \cdot \int(\epsilon(\lambda) \cdot \alpha^{f'}_{R,C}(\mathbf{x}))],$$

where $\int(\chi) = \min(1, \chi)$ and

$$\lambda(C) = \{\lambda \in \mathbb{N} \mid \exists z \in \{0,1\}^m, x \in \{0,1\}^{\ell(\lambda)} :\ C^f(z) \rightharpoondown x\}.$$

Figure 2: The Rejection Policy.

*3. **Adaptive Differential δ-Indistinguishability against any $(\mathcal{F}, \mathcal{CF}[\rho], \mathcal{A})$-siblings** where*

$$\delta = O((1-\tau) + (q'')^2 \cdot \gamma/\epsilon + (qq''/\gamma) \cdot 2^{-\lambda}),$$

for any (q, q', q'')-bounded CollFind*-adversary $\mathcal{A}$ that is canonical and (τ, γ)-smooth.*

Soundness. The following straight-forward lemma is sufficient to imply both soundness and pairwise-soundness ϵ of ρ.

Lemma 6.7. *For any $R \in \mathcal{R}$, any f and $C : \{0,1\}^m \to \{0,1\}^n$, if $\rho^f_R(C) < 1$ then, for any $\lambda \in \mathbb{N}$ and any $\mathbf{x} \in \mathbf{X}_\lambda$, it holds that*

$$\Pr_{z\leftarrow\mathsf{Corr}_{R,C}}[C^f(z) \overset{f}{\rightharpoondown} \mathbf{x}] < \Pr_{z\leftarrow\{0,1\}^m}[C^f(z) \overset{f}{\rightharpoondown} \mathbf{x}]/\epsilon(\lambda).$$

Proof. The proof follows by definition; $\rho^f(C) < 1$ implies, in particular, that $\epsilon \cdot \alpha^f(\mathbf{x}) < 1$ and, therefore,

$$\Pr_{z\leftarrow\mathsf{Corr}}[C^f(z) \overset{f}{\rightharpoondown} \mathbf{x}] = \omega^f(\mathbf{x}) = \alpha^f(\mathbf{x}) \cdot \sigma^f(\mathbf{x}) < \sigma^f(\mathbf{x})/\epsilon(\lambda) = \Pr_{z\leftarrow\{0,1\}^m}[C^f(z) \overset{f}{\rightharpoondown} \mathbf{x}]/\epsilon(\lambda).$$

□

Differential Indistinguishability. The proof of differential indistinguishability is also simple and follows almost immediately by construction.

Lemma 6.8. *For any $R \in \mathcal{R}$, any f, any C that makes at most q queries on any input, any $\lambda \in \mathbb{N}$ and any $X^*, X' \subset \{0,1\}^{\ell(\lambda)}$ such that $|X^*| = |X'|$ and $X^*, X' \in \mathsf{Smooth}^f_{\gamma/2}(C)$, it holds that*

$$|\rho^{f'}_R(C) - \rho^f_R(C)| < O(q\gamma/\epsilon),$$

where $f' = f_{X^*\leftrightarrow X'}$.

Proof. It holds, for any $X^*, X' \in \mathsf{Smooth}^f_{\gamma/2}(C)$ that $(X^*, X') \in \widetilde{\mathbb{X}}^f_\gamma(C)$ and, therefore, by our definition of a d-neighborhood, we have that $\Delta_{\gamma,\lambda,C}(f, f') = 1$. Therefore, $\Delta(f'', f') \le \Delta(f'', f) + 1$ for any f'' and

$\mathcal{N}^{f'}_{d,\gamma}(\lambda, C) \subseteq \mathcal{N}^{f}_{d,\gamma}(\lambda, C) \cup \{f'' \mid \Delta(f'', f') = d\}$. Therefore,

$$\begin{aligned}
\rho^f_R(C) &= \max_{\substack{\lambda\in\lambda(C),\mathbf{x}\in\mathbf{X}\\ f''\in\mathcal{N}^{f}_{d,\gamma}(\lambda,C)}} [e^{-\Delta(f,f'')/\sqrt{d}} \cdot \int(\epsilon\cdot\alpha^{f''}(\mathbf{x}))] \\
&\leq \max_{\substack{\lambda\in\lambda(C),\mathbf{x}\in\mathbf{X}\\ f''\in\mathcal{N}^{f}_{d,\gamma}(\lambda,C)}} [e^{-(\Delta(f',f'')-1)/\sqrt{d}} \cdot \int(\epsilon\cdot\alpha^{f''}(\mathbf{x}))] \\
&= e^{1/\sqrt{d}} \cdot \max_{\substack{\lambda\in\lambda(C),\mathbf{x}\in\mathbf{X}\\ f''\in\mathcal{N}^{f}_{d,\gamma}(\lambda,C)}} [e^{-\Delta(f',f'')/\sqrt{d}} \cdot \int(\epsilon\cdot\alpha^{f''}(\mathbf{x}))] \\
&\leq e^{1/\sqrt{d}} \cdot \max_{\substack{\lambda\in\lambda(C),\mathbf{x}\in\mathbf{X}\\ f''\in\mathcal{N}^{f'}_{d,\gamma}(\lambda,C)}} [e^{-\Delta(f',f'')/\sqrt{d}} \cdot \int(\epsilon\cdot\alpha^{f''}(\mathbf{x}))] + e^{-(d-1)/\sqrt{d}} = \rho^{f'}_R(C) + O(1/\sqrt{d}).
\end{aligned}$$

The lemma is derived by symmetry. □

Correctness. The rest of this section is dedicated to prove that our ρ satisfies correctness which, unlike the security properties that follow easily by the definition of our policy, will demand much more effort.

Let us denote, for any CIH candidate $C = \{C_n\} \in \mathcal{C}$, the largest security parameter at which C calls f by $\lambda^*_n(C) = \sup \lambda(C_n)$.

Lemma 6.9. *Let ρ be the rejection policy defined in Figure 2 with $\epsilon(\lambda) \leq \min(2^{-\lambda/4}, 2^{9\lambda/10}\gamma)$. Let $p : \mathbb{N} \to [0,1]$ be such that $p(n) \geq 4n^2 2^{-m(n)}$ and let $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > (-\lambda/\log\epsilon(\lambda))(6t(n)+2)$ for all $n, \lambda \in \mathbb{N}$. Let $\mathcal{R}$ be a $k(n)$-wise $p(n)$-universal relation class. Then, it holds, for any oracle f and any $t(n)$-bounded circuit $C = \{C_n\}$ and any $n \in \mathbb{N}$, that*

$$\mathbb{E}_{R\leftarrow\mathcal{R}}[\rho^f_R(C_n)] < \frac{1}{2n^2}.$$

Proof of Lemma 6.9. Correctness of ρ shall be implied by a bound, for any fixed f, on the probability that under a random relation $R \leftarrow \mathcal{R}$ there exists $f_{\mathbb{X}} \in \mathcal{N}^f_{d,\gamma}$ and $\mathbf{x} \in \mathbf{X}$ for which $\alpha^{f_{\mathbb{X}}}(\mathbf{x})$ is significantly large.

A straight-forward attempt would be to apply a union bound over all possible functions in the neighborhood $\mathcal{N}^f_{d,\gamma}$. Such a bound, however, is doomed to be too wasteful and requires an inverse doubly-exponential bound on the probability of the "bad" event for any fixed $f_{\mathbb{X}}$ to eliminate the doubly-exponential blow-up incurred by the neighborhood's size. This can be satisfied only by a relation class where relations have exponential size description, which is unreasonable to assume in any real-world utilization. Instead, we observe that the quantities $\alpha^{f_{\mathbb{X}}}(\mathbf{x})$ in the neighborhood exhibit strong dependencies by the fact that C is of bounded locality. Roughly speaking, we show that it is possibly to write any $\alpha^{f_{\mathbb{X}}}(\mathbf{x})$, which potentially depends on the $d \gg t$ swaps in $\mathbb{X}$, as an average of corresponding "local" quantities $\{\alpha^f_Q\}$, each depending on the projection of $\mathbb{X}$ on merely t points in f! [9] Consequently, the existence of a large $\alpha^{f_{\mathbb{X}}}$ in the neighborhood must imply the existence of such a small projection that gives a large amplification and, hence, a union bound over all such projections, whose number is exponentially smaller than that of different functions $f_{\mathbb{X}} \in \mathcal{N}^f_{d,\gamma}$, is sufficient.

We achieve such a structure via "pixelating" the space of C-inputs that incur a call to $f_{\mathbb{X}}$ at x, i.e. all $z \in \mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x}}$, by the set of additional $t - |\mathbf{x}|$ queries made by the execution $C^{f_{\mathbb{X}}}(z)$. That is, each "pixel" corresponds to some $Q \subset \{0,1\}^\ell$ of size $t - |\mathbf{x}|$ and contains $\mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x},Q} = \mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x}} \cap \mathsf{Hits}^{f_{\mathbb{X}}}_{Q}$. Based on the fact that such a splitting of the space induces a disjoint union of $\mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x}}$ (recall we assume w.l.o.g. that C always makes exactly t queries), we are able to show that $\alpha^{f_{\mathbb{X}}}(\mathbf{x})$ is essentially the average of all its restrictions over

[9] In contrary to our high-level notation, we will actually have many (yet bounded) local quantities corresponding to any such subset of t points.

these "pixels" (under some fixed distribution). Further, fixing Q, let us define $\mathbb{X}[Q]$ to be the projection of $\mathbb{X}$ on Q, i.e. the set of alterations (not necessarily swaps) induced by $\mathbb{X}$ on the $t - |\mathbf{x}|$ inputs in Q. We notice that for any $z \in \mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x},Q}$, the execution $C^{f_{\mathbb{X}}}(z)$ is identical to the execution $C^{f_{\mathbb{X}[Q]}}(z)$, as it is oblivious in the value of f on any $x' \notin Q \cup \{\mathbf{x}\}$. Therefore, any measure under $f_{\mathbb{X}}$ over the "pixel" corresponding to Q may be reduced to depend only on $\mathbb{X}[Q]$, which is chosen from the space of all t pairs $(x, y) \in \{0,1\}^{\ell(\lambda)} \times \{0,1\}^{\lambda}$.

To formalize the above outline, we extend our notation as follows. For any $\overrightarrow{x} \subseteq \{0,1\}^{\ell(\lambda)}$, define

$$\tilde{\omega}^f_{R,C}(\overrightarrow{x}) = \Pr_{z \leftarrow \mathsf{Corr}^f_{R,C}}[\bigwedge_{x \in \overrightarrow{x}} C^f(z) \rightharpoondown x] = |\mathsf{Hits}_{\overrightarrow{x}} \cap \mathsf{Corr}|/p2^m,$$

$$\sigma^f_{R,C}(\overrightarrow{x}) = \Pr_{z \leftarrow \{0,1\}^m}[\bigwedge_{x \in \overrightarrow{x}} C^f(z) \rightharpoondown x] = |\mathsf{Hits}_{\overrightarrow{x}}|/2^m,$$

and

$$\tilde{\alpha}^f(\overrightarrow{x}) = \tilde{\omega}^f(\overrightarrow{x})/\sigma^f(\overrightarrow{x}).$$

We note that, to facilitate our analysis, we slightly diverge from the definition of ω given in (12), and define $\tilde{\omega}$ as if assuming in its calculation that the size of the correlation set takes its expected average, which is $p2^m$. To justify this, we need first to bound the probability that the size of the correlation set largely deviates, which will assure us that $\tilde{\alpha}$ is expected to be close enough to α.

Claim 7. *Let $m, n \in \mathbb{N}$ and $\mathcal{R} \subseteq \mathbb{P}(\{0,1\}^m \times \{0,1\}^n)$ be a class of pairwise p-universal relations. Then, for any (possibly oracle-aided) circuit $C : \{0,1\}^m \to \{0,1\}^n$ and any $\alpha > 0$, it holds that*

$$\Pr_{R \leftarrow \mathcal{R}}[|\mathsf{Corr}_{R,C}| < p2^{m-1}] < 4/p2^m.$$

Proof. For any $z \in \{0,1\}^m$, we have

$$\mathbb{E}_R[|\mathsf{Corr}|] = \sum_{z \in \{0,1\}^m} \mathbb{E}_R[\mathbb{1}(z \in \mathsf{Corr})] = p2^m$$

and, from pairwise-independence of $\{\mathbb{1}(z \in \mathsf{Corr})\}_z$,

$$\mathbf{Var}_R(|\mathsf{Corr}|) = \sum_{z \in \{0,1\}^m} \mathbf{Var}_R(\mathbb{1}(z \in \mathsf{Corr})) = (p - p^2)2^m < p2^m.$$

Thus, the claim follows immediately from Chebyshev's inequality (Proposition 2.3). □

While the above bound is sufficient to imply that $\tilde{\alpha}^f < 2\alpha^f$, it does not say anything about the value of $\tilde{\alpha}$ compared to α under functions in the neighborhood $f' \in \mathcal{N}^f_{d,\gamma}$. We extrapolate a similar relation over the neighborhood by relying on the bound under f using the conditional bound which we will later show is useful when applied inductively.

Claim 8. *Let $m, n \in \mathbb{N}$ and $T > 0$. For any $\lambda \in \mathbb{N}$, any relation $R \subseteq \{0,1\}^m \times \{0,1\}^n$, any f-aided circuit $C : \{0,1\}^m \to \{0,1\}^n$ that makes at most t queries on any input, and any oracle functions f and f' such that $\Delta_{\lambda,C}(f, f') = 1$, if $\max_{x \in \{0,1\}^{\ell(\lambda)}} \alpha^f_{R,C}(x) \leq T$, then it holds that*

$$|\mathsf{Corr}^{f'}_{R,C}| \geq (1 - 2tT\gamma(\lambda)) \cdot |\mathsf{Corr}^f_{R,C}|.$$

Proof. Let $\mathbb{X} = (X^*, X') \in \widetilde{\mathbb{X}}^f_\gamma(C)$ be the swap that takes f to f', i.e. $f' = f_{X^* \leftrightarrow X'}$ (its existence is implied by the distance 1). Letting $\mathsf{Hits}_{\mathbb{X}} = \bigcup_{x \in X^* \cup X'} \mathsf{Hits}_x$, we have that

$$\begin{aligned} |\mathsf{Corr}^f| &= |\mathsf{Corr}^f \setminus \mathsf{Hits}^f_{\mathbb{X}}| + |\mathsf{Corr}^f \cap \mathsf{Hits}^f_{\mathbb{X}}| = |\mathsf{Corr}^{f'} \setminus \mathsf{Hits}^f_{\mathbb{X}}| + |\mathsf{Corr}^f \cap \mathsf{Hits}^f_{\mathbb{X}}| \\ &\leq |\mathsf{Corr}^{f'}| + \sum_{x \in X^* \cup X'} |\mathsf{Corr}^f \cap \mathsf{Hits}^f_x| = |\mathsf{Corr}^{f'}| + |\mathsf{Corr}^f| \cdot \sum_{x \in X^* \cup X'} \omega^f(x) \\ &\leq |\mathsf{Corr}^{f'}| + |\mathsf{Corr}^f| \cdot T \cdot \sum_{x \in X^* \cup X'} \sigma^f(x) = |\mathsf{Corr}^{f'}| + |\mathsf{Corr}^f| \cdot T \cdot \sum_{x \in X^* \cup X'} \Pr_{z \leftarrow \{0,1\}^m}[C^f(z) \rightharpoondown x] \end{aligned}$$

(note that $\mathsf{Corr}^f \setminus \mathsf{Hits}^f_{\mathbb{X}} = \mathsf{Corr}^{f'} \setminus \mathsf{Hits}^f_{\mathbb{X}}$ since f, f' are identical at all points except for $\mathbb{X}$). We borrow (9) from the proof of Lemma 5.20 to conclude

$$|\mathsf{Corr}^f| \leq |\mathsf{Corr}^{f'}| + |\mathsf{Corr}^f| \cdot T \cdot t \cdot \Pr_{z \leftarrow \{0,1\}^m}[\exists x \in X^* \cup X' : C^f(z) \rightharpoondown x] \leq |\mathsf{Corr}^{f'}| + |\mathsf{Corr}^f| \cdot 2tT\gamma.$$

□

Given the statements proven so far, we are prepared now to show a connection between the rejection probability, which is determined by the maximal α-value in the neighborhood, and the probability of having a large $\tilde{\alpha}$-value in the neighborhood, thus transforming our focus into bounding the $\tilde{\alpha}$'s.

Claim 9. *Let $m, n \in \mathbb{N}$ and $p > 0$ and let $\mathcal{R} \subseteq \mathbb{P}(\{0,1\}^m \times \{0,1\}^n)$ be a class of pairwise p-universal relations. Then, for any oracle function f and any f-aided circuit $C : \{0,1\}^m \to \{0,1\}^n$ that makes at most t queries, it holds that*

$$\mathbb{E}_R[\rho^f(C)] \leq 1/p2^m + 1/4n^2 + t\sqrt{d}\ln(4n^2) \cdot \max_{d', \lambda \in \lambda(C)} \Pr_R[\max_{\substack{\mathbf{x} \in \mathbf{X}_\lambda \\ f' : \Delta(f,f') = d'}} \tilde{\alpha}^{f'}(\mathbf{x}) > 1/\epsilon(\lambda)4n^2].$$

Proof. Let B denote the event where $|\mathsf{Corr}^f| < p2^{m-1}$ and recall that $\Pr_R[B] < 4/p2^m$ by Claim 7. Further, for any $d', \lambda \in \mathbb{N}$ let us denote by $B_{\lambda,d'}$ the event that there exist $\mathbf{x} \in \mathbf{X}_\lambda$ and f' with $\Delta(f, f') = d'$ such that $e^{-d'/\sqrt{d}} \cdot \int(\epsilon \cdot \alpha^{f'}(\mathbf{x})) > 1/4n^2$. Observe that we need to concern ourselves only with $d' \leq \ln(4n^2)\sqrt{d}$ since otherwise it holds that $e^{-d'/\sqrt{d}} < 1/4n^2$. We can then simplify by law of total expectation and union bound as follows

$$\begin{aligned}
\mathbb{E}_R[\rho^f(C)] &= \mathbb{E}_R[\max_{\substack{\lambda \in \lambda(C), \mathbf{x} \in \mathbf{X}_\lambda \\ f' \in \mathcal{N}^f_{d,\gamma}(\lambda, C)}} [e^{-\Delta(f,f')/\sqrt{d}} \cdot \int(\epsilon \cdot \alpha^{f'}(\mathbf{x}))]] \\
&\leq \Pr_R[B] + 1/4n^2 + \Pr_R[\overline{B} \ \wedge \ \exists \lambda \in \lambda(C), d' \leq \ln(4n^2)\sqrt{d} : \ B_{\lambda,d'}] \\
&\leq 1/p2^m + 1/4n^2 + \sum_{\substack{\lambda \in \lambda(C) \\ 0 \leq d' \leq \ln(4n^2)\sqrt{d}}} \Pr_R[\overline{B} \wedge B_{\lambda,d'} \bigwedge_{d'' < d'} \overline{B_{\lambda,d''}}] \qquad (15)
\end{aligned}$$

Looking more closely at the conjunction of events above, we notice that it is possibly satisfied only if: (i) $|\mathsf{Corr}^f| > p2^{m-1}$, and (ii) for any $d'' < d'$ and any function f' of distance d'' from f, it holds that $\max_{\mathbf{x}} \alpha^{f'}(\mathbf{x}) \leq 1/\epsilon 4n^2$. Hence, by inductively applying Claim 8, starting with f and transitioning, swap by swap, to f' of distance d', we may infer that for any such f' with $\Delta(f, f') = d'$, it holds that

$$|\mathsf{Corr}^{f'}| \geq (1 - 2t\gamma/\epsilon 4n^2)^{d'} |\mathsf{Corr}^f| \geq (1 - 2t\gamma/\epsilon 4n^2)^{d'} p2^{m-1} = e^{-2t\gamma d'/\epsilon 4n^2} p2^{m-1}$$

and, hence, in such case

$$\alpha^{f'}(\mathbf{x}) < 2e^{2\gamma d'/\epsilon 4n^2} \cdot \tilde{\alpha}^{f'}(\mathbf{x}) < 2e^{t\gamma d'/\epsilon} \cdot \tilde{\alpha}^{f'}(\mathbf{x})$$

for all $\mathbf{x}$.

Consequently, for any λ and d',

$$\begin{aligned}
\Pr_R[\overline{B} \wedge B_{\lambda,d'} \bigwedge_{d'' < d'} \overline{B_{\lambda,d''}}] &\leq \Pr_R[\max_{\mathbf{x}, f' : \Delta(f,f') = d'} e^{-d'/\sqrt{d}} \cdot \int(\epsilon \cdot 2e^{t\gamma d'/\epsilon} \tilde{\alpha}^{f'}(\mathbf{x})) > 1/4n^2] \\
&\leq \Pr[\max_{\mathbf{x}, f' : \Delta(f,f') = d'} \tilde{\alpha}^{f'}(\mathbf{x}) > e^{d'(1/\sqrt{d} - t\gamma/\epsilon)}/\epsilon 4n^2] \\
&\leq \Pr[\max_{\mathbf{x}, f' : \Delta(f,f') = d'} \tilde{\alpha}^{f'}(\mathbf{x}) > 1/\epsilon 4n^2],
\end{aligned}$$

and, therefore, the desired inequality follows by plugging the above in (15). □

We have reduced our task to bounding large deviations in the values of $\tilde{\alpha}$ in the neighborhood. We now proceed by describing $\tilde{\alpha}^{f'}(\mathbf{x})$ as the average of similar "local" measures, in the sense that each depends on the choice of at most $t - |\mathbf{x}|$ swaps out of those applied to obtain f'.

Claim 10. *Let $m, n \in \mathbb{N}$. For any relation $R \subseteq \{0,1\}^m \times \{0,1\}^n$, any f-aided circuit $C : \{0,1\}^m \to \{0,1\}^n$ that makes at most t queries on any input, any oracle function f, any $\mathbb{X} = ((X_1^*, X_1'), \ldots, (X_{d'}^*, X_{d'}'))$ and $\mathbf{x} \in \mathbf{X}_\lambda$, there exist $\{\nu_Q\}_{\substack{Q \subset \{0,1\}^{\ell(\lambda)} \setminus \{\mathbf{x}\},\\ |Q| = t - |\mathbf{x}|}}$ such that $\nu_Q > 0$ for all Q and $\sum_Q \nu_Q = 1$ and, for any $R \in \mathcal{R}$, it holds that*

$$\tilde{\alpha}^{f_{\mathbb{X}}}(\mathbf{x}) = \sum_{\substack{Q \subset \{0,1\}^{\lambda} \setminus \{x\},\\ |Q| = t - |\mathbf{x}|}} \nu_Q \cdot \tilde{\alpha}^{f_{\mathbb{X}[Q]}}(x, Q),$$

where $\mathbb{X}[Q]$ is projection of $\mathbb{X}$ onto Q, i.e. the series of alterations (not necessarily swaps) that partially applies $\mathbb{X}$, altering only inputs in Q.

Proof. Since $\mathsf{Hits}_{\mathbf{x}} = \dot{\bigcup}_Q \mathsf{Hits}_{\mathbf{x},Q}$, it holds for $\varphi \in \{\sigma, \tilde{\omega}\}$ that

$$\varphi^{f_{\mathbb{X}}}(\mathbf{x}) = \sum_{\substack{Q \subset \{0,1\}^{\ell(\lambda)} \setminus \{\mathbf{x}\}\\ |Q| = t - |\mathbf{x}|}} \varphi^{f_{\mathbb{X}}}(\mathbf{x}, Q).$$

Next, since C is t-bounded, then for any $\mathbf{x}$, Q and $z \in \mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x},Q}$, $C^{f_{\mathbb{X}}}(z)$ makes no queries to $f_{\mathbb{X}}$ at points other than $\mathbf{x}$ and those in Q. Thus, we have that $C^{f_{\mathbb{X}}}(z) = C^{f_{\mathbb{X}[Q]}}(z)$ and, therefore, $\mathsf{Hits}^{f_{\mathbb{X}}}_{\mathbf{x},Q} = \mathsf{Hits}^{f_{\mathbb{X}[Q]}}_{\mathbf{x},Q}$ and $\varphi^{f_{\mathbb{X}}}(\mathbf{x}, Q) = \varphi^{f_{\mathbb{X}[Q]}}(\mathbf{x}, Q)$. It follows, then, that

$$\varphi^{f_{\mathbb{X}}}(\mathbf{x}) = \sum_{\substack{Q \subset \{0,1\}^{\ell(\lambda)} \setminus \{\mathbf{x}\}\\ |Q| = t - |\mathbf{x}|}} \varphi^{f_{\mathbb{X}[Q]}}(\mathbf{x}, Q). \tag{16}$$

The claim follows immediately from (16) by setting

$$\nu_Q = \sigma^{f_{\mathbb{X}}}(\mathbf{x}, Q) / \sigma^{f_{\mathbb{X}}}(\mathbf{x}).$$

□

The above representation of $\tilde{\alpha}$ implies that, if $\tilde{\alpha}^{f'}(\mathbf{x})$ is too large for some $f' = f_{\mathbb{X}}$, then it must be the result of a large local value of the form $\tilde{\alpha}^{f_{\mathbb{X}[Q]}}(\mathbf{x}, Q)$. Such a variable value depends on the choice of $Q \subset \{0,1\}^{\ell(\lambda)}$ of size $t - |\mathbf{x}|$ and the alterations $\mathbb{X}[Q]$ (possibly involving only a subset of Q – those that appear in $\mathbb{X}$). Thus, we are maximizing over all such values corresponding to any choice of an input $\mathbf{x}$, a $(t-|\mathbf{x}|)$-tuple $\overrightarrow{x} = (x_1, \ldots, x_{t-|\mathbf{x}|}) \in \{0,1\}^{(t-|\mathbf{x}|) \times \lambda}$, a subset $S \subseteq [t - |\mathbf{x}|]$ and $|S|$ outputs $\overrightarrow{y_S} = \{y_i\}_{i \in S} \in \{0,1\}^{|S| \times \lambda}$, that correspond to the respective images that $\{x_i\}_{i \in S}$ are mapped to in $f_{\mathbb{X}[Q]}$. This allow to apply the following union bound

$$\begin{aligned}
\Pr_R[\max_{\substack{\mathbf{x} \in \mathbf{X}_\lambda\\ f' : \Delta(f, f') = d'}} \tilde{\alpha}^{f'}(\mathbf{x}) > 1/\epsilon 4n^2] &\leq \Pr_R[\max_{\substack{\mathbf{x} \in \mathbf{X}_\lambda\\ f_{\mathbb{X}} : \Delta(f, f_{\mathbb{X}}) = d'\\ Q \subset \{0,1\}^{\ell(\lambda)} : |Q| = t - |\mathbf{x}|}} \tilde{\alpha}^{f_{\mathbb{X}[Q]}}(\mathbf{x}, Q) > 1/\epsilon 4n^2]\\
&\leq \Pr_R[\max_{\substack{\mathbf{x} \in \mathbf{X}_\lambda\\ S, \overrightarrow{x}, \overrightarrow{y_S}}} \tilde{\alpha}^{f_{\overrightarrow{x_S} \to \overrightarrow{y_S}}}(\mathbf{x}, x_1, \ldots, x_{t-|\mathbf{x}|}) > 1/\epsilon 4n^2]\\
&\leq 2^{t(2\lambda+1)} \cdot \max_{\substack{\mathbf{x}, S\\ \overrightarrow{x}, \overrightarrow{y_S}}} \Pr_R[\tilde{\alpha}^{f_{\overrightarrow{x_S} \to \overrightarrow{y_S}}}(\mathbf{x}, x_1, \ldots, x_{t-|\mathbf{x}|}) > 1/\epsilon 4n^2]\\
&\leq 2^{t(2\lambda+1)} \cdot \max_{f, \mathbf{x}, x_1, \ldots, x_{t-|\mathbf{x}|}} \Pr_R[\tilde{\alpha}^{f}(\mathbf{x}, x_1, \ldots, x_{t-|\mathbf{x}|}) > 1/\epsilon 4n^2]. \qquad (17)
\end{aligned}$$

To finish, we use the k-wise universality of our relation class to bound the probability that any individual "local" $\tilde{\alpha}$-value is too large.

Claim 11. *Let $m, n \in \mathbb{N}$ and let $\mathcal{R} \subseteq \mathbb{P}(\{0,1\}^m \times \{0,1\}^n)$ be an almost k-wise $(\cdot, p)$-universal class of relations. For any circuit $C : \{0,1\}^m \to \{0,1\}^n$, any oracle function f, any $\lambda, t \in \mathbb{N}$ and $\overrightarrow{x} = (x_1, \dots, x_t) \in \{0,1\}^{t \times \lambda}$, it holds for any $T > 0$ that*

$$\Pr_{R \leftarrow \mathcal{R}_n}[\tilde{\alpha}^f(\overrightarrow{x}) > T] < (T/e)^{-k}.$$

Proof. We start by rewriting

$$\Pr_R[\tilde{\alpha}^f(\overrightarrow{x}) > T] = \Pr_R[\tilde{\omega}^f(\overrightarrow{x}) > T \cdot \sigma^f(\overrightarrow{x})] = \Pr_R[|\mathsf{Hits}^f_{\overrightarrow{x}} \cap \mathsf{Corr}^f| > T \cdot p|\mathsf{Hits}^f_{\overrightarrow{x}}||].$$

Now, for any $z \in \mathsf{Hits}^f_{\overrightarrow{x}}$, we define the binary random variable $\mathbb{1}_{\mathsf{Corr}}(z)$ that takes 1 if and only if $z \in \mathsf{Corr}^f$, i.e. $(z, C^f(z)) \in R$. Letting $N = |\mathsf{Hits}^f_{\overrightarrow{x}}|$, it holds by the definition of k-wise universality (Definition 4.2) and Theorem 2.4 that [10]

$$\Pr_R[\tilde{\alpha}^f(\overrightarrow{x}) > T] = \Pr_R[\sum_{z \in \mathsf{Hits}_{\overrightarrow{x}}} \mathbb{1}_{\mathsf{Corr}}(z) > pT \cdot N] \leq \frac{p^k \binom{N}{k}}{\binom{pTN}{k}} \leq \left(\frac{peN/k}{pTN/k}\right)^k = (e/T)^k.$$

□

The proof is then complete by applying Claim 11 to bound the expression in (17) then plugging it in Claim 9; Based on the assumptions stated in the lemma and in particular assuming, w.l.o.g., that $\log n < \lambda/24 < -\log \epsilon/6$ for any $\lambda \in \lambda(C_n)$ (see Remark 5.4), we obtain the following for any $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{E}_R[\rho^f(C)] &\leq 1/p(n)2^{m(n)} + 1/4n^2 + t(n)\ln(4n^2) \cdot \max_{\lambda \in \lambda(C_n)} \sqrt{d(\lambda)} 2^{t(n)(2\lambda+1)} \cdot (\epsilon(\lambda)4en^2)^{k(n)} \\
&\leq 1/2n^2 + \max_{\lambda \in \lambda(C_n)} 2^{\log\log n + \log(\epsilon(\lambda)/\gamma(\lambda)) + 3\lambda \cdot t(n) + k(n) \cdot (\log \epsilon(\lambda) + 2\log n + O(1)} \\
&\leq 1/2n^2 + \max_{\lambda \in \lambda(C_n)} 2^{\log(\epsilon(\lambda)/\gamma(\lambda)) + 3\lambda \cdot t(n) + (k(n)/2) \cdot \log \epsilon(\lambda)} \\
&\leq 1/2n^2 + \max_{\lambda \in \lambda(C_n)} 2^{-\lambda/10} \leq 1/n^2.
\end{aligned}$$

7 Putting Everything Together: Proof of Main Theorems

Finally, we recall our separation results stated in Theorems 4.3 and 4.4, then show how to utilize our correlation finder construction from Section 6 with carefully chosen parameters to obtain them via the differential indistinguishability framework from Section 5.

We begin with the separation of CIH from CRH (Theorem 4.4) since it requires little more attention.

Theorem 4.4 (Black-box Separation of CIH from CRH)**.** *Let $m := m(n)$ and $p : \mathbb{N} \to [0,1]$ be such that $p(n) \geq 4n^2 2^{-m(n)}$ and let $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > 25 \cdot t(n)$ for all $n \in \mathbb{N}$. Then, t-bounded CIH functions, with input length m, for any class of k-wise p-universal relations are $2^{\lambda/25}$-fully black-box separated from CRH mapping $\ell(\lambda) = \lambda + O(1)$ bits to λ bits.*

Proof. By Lemma 5.2, the desired separation can be derived by a correlation finder that satisfies both the correctness and security conditions in the lemma w.r.t. our ideal implementation of CRH as a uniformly random regular ℓ-bit-to-λ-bit hash function $f \leftarrow \mathcal{F}$. We choose our correlation finder to be the picky correlation finder from Construction 6.1 with the rejection policy from Figure 2, instantiated with parameters $\epsilon = 2^{-\lambda/4}$ and $\gamma = 2^{-5\lambda/12}$. Corollary 6.6 immediately implies correctness. For security, we need to bound the advantage in finding collisions of any (q, q, q)-bounded adversary $\mathcal{A}$ that makes no O-queries with $n < 2^{\lambda/24}$.

[10] As noted previously, assuming almost universality, where the probability of a k'-conjunction is merely bounded by p^k, is also sufficient.

By the smoothening lemma (Lemma 5.9, with $\gamma = 2^{-5\lambda/12}$ and $\beta = \lambda/\gamma$), it would be sufficient to bound the advantage of any $(q + (\lambda/\gamma)q^2, q, q)$-bounded $(\tau = 1 - 2^{2-\lambda/4}, \gamma = 2^{-5\lambda/12})$-smooth adversary $\mathcal{B}$. For that, we use Lemma 5.15 and the adaptive differential indistinguishability of our construction for siblings, that is implied by Corollary 6.6, to derive

$$\begin{aligned}
\mathbb{E}_{f,\mathsf{O}}[\mathbf{Adv}^f_{\mathbf{CRH}}(\lambda, f, \mathcal{A}^{\mathsf{O}})] &\leq \mathbb{E}_{f,\mathsf{O}}[\mathbf{Adv}^f_{\mathbf{CRH}}(\lambda, f, \mathcal{B}^{\mathsf{O}})] \\
&= O(((\lambda/\gamma)q^2)^2 2^{-\lambda} + q \cdot ((1-\tau) + q^2 \cdot (\gamma/\epsilon) + (\lambda/\gamma^2)q^3 \cdot 2^{-\lambda}))) \\
&= O(q \cdot ((1-\tau) + q^2 \cdot (\gamma/\epsilon) + (\lambda/\gamma^2)q^3 \cdot 2^{-\lambda}))) \\
&= O(\lambda q^3 \cdot 2^{-\lambda/6}).
\end{aligned}$$

This allows to plug our correlation finder into Lemma 5.2, with $c = 3$ and $\kappa = 2^{-\lambda/6}\lambda^3$, to get $2^{\lambda/24}/\lambda^3 > 2^{\lambda/25}$-separation. □

We similarly derive the separation from OWP, yet with a different choice of parameters, and obtain a slightly stronger separation.

Theorem 4.3 (Black-box Separation of CIH from OWP)**.** *Let $m := m(n)$ and $p : \mathbb{N} \to [0, 1]$ be such that $p(n) \geq 4n^2 2^{-m(n)}$ and let $k, t : \mathbb{N} \to \mathbb{N}$ be such that $k(n) > 20 \cdot t(n)$ for all $n \in \mathbb{N}$. Then, t-bounded CIH functions, with input length m, for any class of k-wise p-universal relations are $2^{\lambda/10}$-fully black-box separated from OWP.*

Proof. The theorem is derived along similar lines to the proof of Theorem 4.4, using the same construction and its (non-adaptive) differential indistinguishability implied by Corollary 6.6. In this case, however, we choose $\gamma = 2^{-2\lambda/3}$ and $\epsilon = 2^{-\lambda/3}$ for the construction and obtain $c = 2$ and $\kappa = \lambda 2^{-\lambda/3}$ in Lemma 5.2. □

Acknowledgements

We thank Zvika Brakerski for comments and discussion, and Venkata Koppula for his helpful feedback on a prior version of this work.

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.

[BEG+91] Manuel Blum, William S. Evans, Peter Gemmell, Sampath Kannan, and Moni Naor. Checking the correctness of memories. In *32nd Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 1-4 October 1991*, pages 90–99. IEEE Computer Society, 1991.

[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.

[BG02] Boaz Barak and Oded Goldreich. Universal arguments and their applications. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, pages 194–203. IEEE Computer Society, 2002.

[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):321391, 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 10821090, 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 389415, Berlin, Heidelberg, 2016. Springer-Verlag.

[CGH04] Ran Canetti, Oded Goldreich, and Shai Halevi. The random oracle methodology, revisited. *J. ACM*, 51(4):557594, July 2004.

[CGJ$^+$23] Arka Rai Choudhuri, Sanjam Garg, Abhishek Jain, Zhengzhong Jin, and Jiaheng Zhang. Correlation intractability and snargs from sub-exponential DDH. In Helena Handschuh and Anna Lysyanskaya, editors, *Advances in Cryptology - CRYPTO 2023 - 43rd Annual International Cryptology Conference, CRYPTO 2023, Santa Barbara, CA, USA, August 20-24, 2023, Proceedings, Part IV*, volume 14084 of *Lecture Notes in Computer Science*, pages 635–668. Springer, 2023.

[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 11031114, New York, NY, USA, 2019. Association for Computing Machinery.

[CJJ21] Arka Rai Choudhuri, Abhishek Jain, and Zhengzhong Jin. Snargs for $\mathcal{P}$ from LWE. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 68–79. IEEE, 2021.

[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.

[Dam87] Ivan Damgård. Collision free hash functions and public key signature schemes. In David Chaum and Wyn L. Price, editors, *Advances in Cryptology - EUROCRYPT '87, Workshop on the Theory and Application of of Cryptographic Techniques, Amsterdam, The Netherlands, April 13-15, 1987, Proceedings*, volume 304 of *Lecture Notes in Computer Science*, pages 203–216. Springer, 1987.

[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.

[Dwo06] Cynthia Dwork. Differential privacy. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[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.

[GGKL21] Nick Gravin, Siyao Guo, Tsz Chiu Kwok, and Pinyan Lu. Concentration bounds for almost k-wise independence with applications to non-uniform security. In Dániel Marx, editor, *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 2404–2423. SIAM, 2021.

[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 276288, 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.

[HJKS22] James Hulett, Ruta Jawale, Dakshita Khurana, and Akshayaram Srinivasan. Snargs for P from sub-exponential DDH and QR. In Orr Dunkelman and Stefan Dziembowski, editors, *Advances in Cryptology - EUROCRYPT 2022 - 41st Annual International Conference on the Theory and Applications of Cryptographic Techniques, Trondheim, Norway, May 30 - June 3, 2022, Proceedings, Part II*, volume 13276 of *Lecture Notes in Computer Science*, pages 520–549. Springer, 2022.

[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.

[HM96] Shai Halevi and Silvio Micali. Practical and provably-secure commitment schemes from collision-free hashing. In Neal Koblitz, editor, *Advances in Cryptology - CRYPTO '96, 16th Annual International Cryptology Conference, Santa Barbara, California, USA, August 18-22, 1996, Proceedings*, volume 1109 of *Lecture Notes in Computer Science*, pages 201–215. Springer, 1996.

[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):28942905, October 2006.

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

[IR89] R. Impagliazzo and S. Rudich. Limits on the provable consequences of one-way permutations. In *Proceedings of the Twenty-First Annual ACM Symposium on Theory of Computing*, STOC '89, page 4461, New York, NY, USA, 1989. Association for Computing Machinery.

[JJ21] Abhishek Jain and Zhengzhong Jin. Non-interactive zero knowledge from sub-exponential DDH. In Anne Canteaut and Franccois-Xavier Standaert, editors, *Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17-21, 2021, Proceedings, Part I*, volume 12696 of *Lecture Notes in Computer Science*, pages 3–32. Springer, 2021.

[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 723732, New York, NY, USA, 1992. Association for Computing Machinery.

[KLV23] Yael Tauman Kalai, Alex Lombardi, and Vinod Vaikuntanathan. Snargs and PPAD hardness from the decisional diffie-hellman assumption. In Carmit Hazay and Martijn Stam, editors, *Advances in Cryptology - EUROCRYPT 2023 - 42nd Annual International Conference on the Theory and Applications of Cryptographic Techniques, Lyon, France, April 23-27, 2023, Proceedings, Part II*, volume 14005 of *Lecture Notes in Computer Science*, pages 470–498. Springer, 2023.

[KNY18] Ilan Komargodski, Moni Naor, and Eylon Yogev. Collision resistant hashing for paranoids: Dealing with multiple collisions. In Jesper Buus Nielsen and Vincent Rijmen, editors, *Advances in Cryptology - EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Tel Aviv, Israel, April 29 - May 3, 2018 Proceedings, Part II*, volume 10821 of *Lecture Notes in Computer Science*, pages 162–194. Springer, 2018.

[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):12531298, 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.

[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.