AICR · OPEN PROBLEM

Black-Box Helpfulness of One-Way Functions for Collision Resistance

The existence of an auxiliary primitive that only works with a one-way function

Status. Statement: AI-written from Geoffroy Couteau, Pooya Farshim and Mohammad Mahmoody, *Black-Box Uselessness: Composing Separations in Cryptography*, Cryptology ePrint Archive 2021, not yet checked by a human. No case is settled: the paper proves only that two relaxations — distributional black-box helpfulness (Theorem 6.8) and class helpfulness (Theorem 6.12) — follow from an unproved amplification conjecture about Simon's oracle, and the complementary statement that one-way functions are black-box *useless* for collision resistance is likewise open.

Category. *Black-Box Separations & Reductions.*

Conjecture (informal). *Simon showed that collision-resistant hash functions cannot be built from one-way functions in a black-box way, by exhibiting a random permutation together with an oracle that returns collisions for any circuit built from it: relative to that pair, one-way permutations exist but collision-resistant hashing does not. This paper asks the sharper question of whether one-way functions are entirely* useless *for collision resistance, meaning that no auxiliary primitive is ever made sufficient for collision resistance by adding a one-way function to it. The authors conjecture the opposite. That is, they conjecture that there exists some auxiliary primitive from which collision-resistant hashing cannot be built in a black-box way, but from which, together with an arbitrary one-way function, it can. This would exhibit a natural inhabitant of the gray zone between separation and uselessness, and would explain why one-way functions might yet play a role in future constructions of collision resistance from apparently weaker ingredients. The paper describes why proving it looks hard: Simon's result is an oracle separation and it is unclear how to relativize it, and the concrete candidate auxiliary primitive the authors propose reduces the conjecture to an unproved statement about the simultaneous hardness of inverting a one-way function and a random permutation.*

1 The setting

Simon [Sim98] separated collision-resistant hash functions from one-way functions (indeed from one-way permutations) by exhibiting a pair of oracles $(\pi, \mathsf{Coll}^{\pi})$, where π is a random permutation and Coll^{π} takes as input a circuit with π-gates and returns a random collision for it, relative to which one-way permutations exist but collision-resistant hashing does not. This paper studies whether such separations *compose*: a primitive Q is called black-box useless for $\mathcal{P}$ if adding Q to *any* auxiliary primitive never enables a black-box construction of $\mathcal{P}$ that the auxiliary primitive did not already enable, and black-box helpful otherwise. The compiling-out separations lift to uselessness in full (Theorems 5.6 and 5.7); the transcript-sampling separations lift only partially, giving uselessness of one-way functions for perfectly correct unbalanced, constant-round imperfect, and one-round key agreement, with the general case left as an open

conjecture. Simon's oracle separation, the authors observe, does not obviously lift at all, since it is not clear how to relativize an oracle separation of this kind. They conjecture that this failure is inherent, and that one-way functions are in fact helpful for collision resistance.

The gray zone between separation and uselessness is known to be non-empty, but only for artificial reasons: taking the joint primitive consisting of a collision-resistant hash function and a trapdoor permutation, neither component alone black-box implies the pair (using [IR89] in one direction and Fischlin [Fis02] in the other), yet each is trivially helpful for it. A proof of this conjecture would give "a natural (and important) example of a primitive which does not black-box imply another primitive, yet is black-box helpful for it", and, as the paper stresses, would be an indication of why one-way functions could feature in future black-box constructions of collision resistance from seemingly weaker assumptions.

The paper offers two routes. The first takes the candidate auxiliary primitive to be Simon's pair $(\pi, \mathsf{Coll}^{\pi})$ itself, and uses the result of Holmgren and Lombardi [HL18] that exponentially secure one-way *product* functions black-box imply collision-resistant hashing; making this work requires showing that for every one-way function F the pair $(\mathsf{F}, (\pi, \mathsf{Coll}^{\pi}))$ is product one-way, which the authors could not prove. The second route uses the backdoored random oracle model of Bauer, Farshim and Mazaheri [BFM18], whose leakage oracle is strong enough to implement Simon's collision finder and where two independent instances combine into a collision-resistant hash function under a communication-complexity conjecture.

Because proving the conjecture seems quite challenging, the paper formulates two relaxations to gauge its plausibility — one using distributional black-box reductions, which need only work for a measure-one set of implementations, and one using class reductions in the sense of Shaltiel [Sha20], which need only work for efficient implementations — and shows that both follow from the amplification conjecture about Simon's oracle. The distributional route falls short of the full conjecture because the measure-one set of auxiliary implementations depends on the one-way function, so no universal auxiliary primitive can be extracted; restricting to efficient one-way functions is what recovers a single implementation via Borel–Cantelli. Reduction terminology throughout follows [RTV04].

2 Notation and parameters

- $\lambda \in \mathbb{N}$ is the security parameter; PPT means probabilistic polynomial time; an *oracle-aided PPT* machine runs in polynomial time relative to every oracle, counting each oracle call as one step.
- $\mathcal{F}$ denotes the primitive of one-way functions, $\mathcal{CRH}$ the primitive of collision-resistant hash functions, and $\mathcal{Z}$ an auxiliary primitive; all three are cryptographic primitives in the sense of Definition 1.
- $(\mathcal{F}, \mathcal{Z})$ denotes the joint primitive whose implementations are pairs (F, Z) of implementations of $\mathcal{F}$ and $\mathcal{Z}$, and which is broken when either component is broken.
- Calligraphic letters $\mathcal{P}, \mathcal{Q}, \mathcal{Z}, \mathcal{F}, \mathcal{CRH}$ denote primitives; sans-serif letters $\mathsf{P}, \mathsf{Q}, \mathsf{Z}, \mathsf{F}, \mathsf{H}$ denote particular implementations of them.

The parameters are:

- λ, the security parameter.
- $\mathcal{Z}$, the auxiliary primitive whose existence is asserted; the whole content of the conjecture is that some such primitive exists.
- $\mathcal{F}$, the primitive of one-way functions (Definition 5).
- $\mathcal{CRH}$, the primitive of collision-resistant hash functions (Definition 6).

3 The conjecture

Definition 1 (Cryptographic primitive). A *cryptographic primitive* is a pair $\mathcal{P} = (F_{\mathcal{P}}, R_{\mathcal{P}})$, where $F_{\mathcal{P}}$ is the set of functions *implementing* $\mathcal{P}$ and $R_{\mathcal{P}}$ is a relation; for $\mathsf{P} \in F_{\mathcal{P}}$, $(\mathsf{P}, A) \in R_{\mathcal{P}}$ means the adversary A *breaks* P. At least one $\mathsf{P} \in F_{\mathcal{P}}$ must be computable by a PPT algorithm.

Definition 2 (Joint primitive). Given $\mathcal{P} = (F_{\mathcal{P}}, R_{\mathcal{P}})$ and $\mathcal{Q} = (F_{\mathcal{Q}}, R_{\mathcal{Q}})$, the joint primitive $(\mathcal{P}, \mathcal{Q})$ has implementations $F_{\mathcal{P}} \times F_{\mathcal{Q}}$, and $((\mathsf{P}, \mathsf{Q}), (A_{\mathcal{P}}, A_{\mathcal{Q}}))$ lies in its breaking relation iff $(\mathsf{P}, A_{\mathcal{P}}) \in R_{\mathcal{P}}$ or $(\mathsf{Q}, A_{\mathcal{Q}}) \in R_{\mathcal{Q}}$.

Definition 3 (Fully black-box reduction). A *fully black-box reduction* of a primitive $\mathcal{P}$ to a primitive $\mathcal{Q}$ is a pair (P, S) of oracle-aided PPT machines such that for every implementation $\mathsf{Q} \in F_{\mathcal{Q}}$:

- (implementation reduction) $\mathsf{P}^{\mathsf{Q}} \in F_{\mathcal{P}}$; and
- (security reduction) for every function A with $(\mathsf{P}^{\mathsf{Q}}, A) \in R_{\mathcal{P}}$, it holds that $(\mathsf{Q}, \mathsf{S}^{\mathsf{Q},A}) \in R_{\mathcal{Q}}$.

We say $\mathcal{Q}$ *fully black-box implies* $\mathcal{P}$ if such a pair exists.

Definition 4 (Fully black-box helpfulness). A primitive $\mathcal{Q}$ is *fully black-box helpful* for a primitive $\mathcal{P}$ if there exists an auxiliary primitive $\mathcal{Z}$ such that the joint primitive $(\mathcal{Q}, \mathcal{Z})$ fully black-box implies $\mathcal{P}$, yet $\mathcal{Z}$ alone does not fully black-box imply $\mathcal{P}$. (This is the negation of $\mathcal{Q}$ being fully black-box useless for $\mathcal{P}$.)

Definition 5 (One-way functions). The primitive $\mathcal{F}$ of one-way functions has as implementations all functions F, and an adversary A breaks F if there is a polynomial p such that for infinitely many $\lambda \in \mathbb{N}$,

$$\Pr_{x \leftarrow \{0,1\}^{\lambda}} \left[\mathsf{F}\big(A(1^{\lambda}, \mathsf{F}(x))\big) = \mathsf{F}(x) \right] \geq \frac{1}{p(\lambda)}.$$

Remark 1 (what the paper actually states). The paper's Definition 2.6 (p. 10) defines a one-way function as an oracle-aided PPT machine relative to an oracle, secure against PPT adversaries with negligible advantage. The breaking relation displayed in Definition 5 is the correct negation of that negligibility requirement, and the paper's Definition 2.2 does quantify over arbitrary functions A; but

the reading of $F_{\mathcal{F}}$ as "all functions F" is a reconstruction from the paper's usage $\mathsf{F} \in F_{\mathcal{F}}$ on p. 31 rather than a definition the paper gives.

Definition 6 (Collision-resistant hash functions). The primitive $\mathcal{CRH}$ of collision-resistant hash functions has as implementations keyed compressing functions H, which map a key $k \in \mathcal{K}_\lambda$ and an input x to a shorter digest $\mathsf{H}(k,x)$; an adversary A breaks H if there is a polynomial p such that for infinitely many $\lambda \in \mathbb{N}$,

$$\Pr_{k \leftarrow \mathcal{K}_\lambda}\left[(x,x') \leftarrow A(1^\lambda,k) \;:\; x \neq x' \;\wedge\; \mathsf{H}(k,x) = \mathsf{H}(k,x')\right] \;\geq\; \frac{1}{p(\lambda)}.$$

Remark 2 (the syntax of $\mathcal{CRH}$ is not fixed by the paper). The paper gives no definition of collision-resistant hash functions: its Section 2.2 (pp. 10–11) defines one-way functions (Definition 2.6), key agreement (Definition 2.7) and (approximate) indistinguishability obfuscation (Definition 2.8) only. Definition 6 is therefore supplied by the drafter of this record, and is deliberately generic, since the paper leaves both the syntax and the compression parameters unspecified and any standard instantiation is intended. For concreteness one may read $\mathcal{K}_\lambda = \{0,1\}^\lambda$ with inputs $x \in \{0,1\}^{2\lambda}$ and digests $\mathsf{H}(k,x) \in \{0,1\}^\lambda$, but those parameters, and the compression ratio they fix, are the drafter's instantiation and not the paper's.

Conjecture 1 (one-way functions are black-box helpful for collision resistance). *There exists a cryptographic primitive $\mathcal{Z}$ such that:*

1. *there exists a fully black-box reduction of $\mathcal{CRH}$ to the joint primitive $(\mathcal{F},\mathcal{Z})$; and*
2. *there exists no fully black-box reduction of $\mathcal{CRH}$ to $\mathcal{Z}$ alone.*

Equivalently: $\mathcal{F}$ is fully black-box helpful *for $\mathcal{CRH}$.*

Remark 3 (the paper's own phrasing). The conjecture is stated

in the paper as "One-way functions are fully black-box helpful for collision-resistant hash functions" (p. 29), and announced already in the introduction as "Indeed, we conjecture the opposite: OWFs are black-box helpful for building CRHFs" (p. 7). The notion is defined there as: "we call primitive Q black-box helpful (BBH) for P, if there exists an auxiliary primitive Z such that there exists a black-box construction of P from (Q, Z), but there exists no black-box construction of P from Z alone" (p. 1). The unfolding in Conjecture 1 and Definition 4 is a reconstruction at the level of primitives: the paper says of its Section 6 that it maintains "a relatively informal discussion", so while the unfolding matches the paper's informal definition and its formal distributional analogue (Definition 6.4), it is not stated by the paper in this form.

Remark 4 (openness). The conjecture is entirely open, and so is its complement. "At first sight, it is not clear how one would extend Simon's strategy to prove that one-way functions are BBU for CRHFs. In fact, we believe that this is inherent: we conjecture that the answer to the above question is no" (p. 29), and "While we believe that Conjecture 6.1 holds, proving it seems quite challenging" (p. 29). Refuting the conjecture means proving that one-way functions are fully black-box *useless* for collision resistance, which is itself a strong and unproved statement; a would-be refuter should be clear which direction is being attacked.

Remark 5 (what is proved). "Towards getting a better understanding of its plausibility, we introduce two natural relaxations of the conjecture, and show that each of these natural relaxations would follow from a plausible conjecture regarding Simon's oracle, which might be of independent interest" (p. 29). Concretely, Theorem 6.8 shows that the distributional relaxation (Conjecture 6.5) follows from the amplification conjecture about Simon's oracle (Conjecture 6.7), and Theorem 6.12 shows that the class-reduction relaxation (Conjecture 6.11) follows from the same conjecture, using a Borel–Cantelli argument over the countable class of efficient one-way functions to extract a single auxiliary implementation that works for all of them. The distributional statement does not upgrade to the full conjecture because, in the paper's words (p. 32), "the measure 1 of implementations Z in F_Z needs not be the same for every F, which is the reason why we cannot extract from this proof a

universal primitive". Section 6.4 records a second, independent route via backdoored random oracles [BFM18], conditional on a communication-complexity conjecture from that work.

Remark 6 (caveats for a reviewer). Four things a reviewer should check. First, the formalization of $\mathcal{CRH}$ is the drafter's (Remark 2); nothing in the conjecture should turn on keyed versus unkeyed hashing, or on the compression ratio, and that should be confirmed. Second, the primitive-level unfolding of "fully black-box helpful" is a reconstruction from a section the paper itself calls informal (Remark 3). Third, the direction of attack matters: a refutation is a uselessness theorem, not merely the absence of a helpful $\mathcal{Z}$ (Remark 4). Fourth, no resolution since January 2021 is known to the author of this record, but follow-up work on one-way product functions and on Simon's oracle has not been exhaustively surveyed.

4 Bibliography

[BFM18] Balthazar Bauer, Pooya Farshim, and Sogol Mazaheri. *Combiners for backdoored random oracles.* Advances in Cryptology — CRYPTO 2018, Part II, volume 10992 of Lecture Notes in Computer Science, pages 272–302, Springer, 2018.

[Fis02] Marc Fischlin. *On the impossibility of constructing non-interactive statistically-secret protocols from any trapdoor one-way function.* Topics in Cryptology — CT-RSA 2002, volume 2271 of Lecture Notes in Computer Science, pages 79–95, Springer, 2002.

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

[IR89] Russell Impagliazzo and Steven Rudich. *Limits on the provable consequences of one-way permutations.* 21st Annual ACM Symposium on Theory of Computing, pages 44–61, ACM Press, 1989.

[RTV04] Omer Reingold, Luca Trevisan, and Salil P. Vadhan. *Notions of reducibility between cryptographic primitives.* TCC 2004: 1st Theory of Cryptography Conference, volume 2951 of Lecture Notes in Computer Science, pages 1–20, Springer, 2004.

[Sha20] Ronen Shaltiel. *Is it possible to improve Yao's XOR lemma using reductions that exploit the efficiency of their oracle?* APPROX/RANDOM 2020, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2020.

[Sim98] Daniel R. Simon. *Finding collisions on a one-way street: Can secure hash functions be based on general assumptions?* Advances in Cryptology — EUROCRYPT'98, volume 1403 of Lecture Notes in Computer Science, pages 334–345, Springer, 1998.