CONJURA · OPEN PROBLEM

Black-Box Uselessness of One-Way Functions for Key Agreement

Arbitrary protocols, no restriction on rounds or query counts

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. Three restricted classes are settled — perfectly correct key agreement in which one party makes a constant number of one-way-function queries (Theorem 4.1), constant-round imperfect key agreement in which both parties make a constant number of one-way-function queries (Theorem 4.9), and Merkle-type protocols, in which all one-way-function queries precede all messages, with no query bound (Theorem 4.12) — while the general case, with arbitrary rounds and arbitrary polynomial query counts, is open and is named by the paper as its central open problem.

Category. *Black-Box Separations & Reductions.*

Conjecture (informal). *Impagliazzo and Rudich showed that one-way functions alone cannot be used to build key agreement in a black-box way. That leaves open a weaker-sounding but much more useful possibility: perhaps one-way functions, while useless on their own, still help when combined with some other primitive that is itself insufficient for key agreement. The paper calls a primitive* black-box useless *for a target if it never helps in this sense: for every auxiliary primitive you might add, if the pair yields a black-box construction of the target then the auxiliary primitive already yielded one by itself. The conjecture is that one-way functions are black-box useless for key agreement in exactly this sense, for arbitrary two-party protocols. The paper proves this only for restricted protocol classes: when one party queries the one-way function only a constant number of times and correctness is perfect; when both parties query it a constant number of times and the protocol has a constant number of rounds; and when all queries to the one-way function are made before any message is sent. The obstruction to the general case is named: the known attacks repeatedly sample protocol views consistent with the transcript, and making those samplers efficient in the presence of the auxiliary primitive requires recursively invoking an inverter, whose query complexity can blow up so fast that only constantly many sampling rounds can be afforded.*

1 The setting

Impagliazzo and Rudich [IR89] showed that relative to a random oracle (together with a PSPACE oracle) one-way functions exist but key agreement does not, ruling out fully black-box constructions of key agreement from one-way functions. Such a separation says nothing, however, about whether one-way functions can *help*: it leaves open that some auxiliary primitive $\mathcal{Z}$, itself insufficient for key agreement, becomes sufficient once one-way functions are added. The framework of *black-box uselessness* introduced in this paper is designed exactly to rule this out, and it composes: if Q_1 and Q_2 are each black-box useless for $\mathcal{P}$, so is the joint primitive (Q_1, Q_2). The notions of reduction used are those of Reingold, Trevisan and Vadhan [RTV04].

The paper establishes three restricted instances of the conjecture. (i) Infinitely-often one-way functions are [semi $\to$ $\forall\exists$-semi]

black-box useless for infinitely-often *perfectly correct* key agreement whenever one of the two parties makes only a constant number of queries to the one-way function (arbitrarily many queries to $\mathcal{Z}$ are allowed). (ii) The same holds for imperfect key agreement when *both* parties make a constant number of one-way-function queries *and* the protocol has a constant number of rounds. (iii) One-way functions are black-box useless for *Merkle-type* protocols [Mer78], in which both parties ask all of their one-way-function queries before exchanging any message, with no bound on the number of queries.

All three proofs run the same case distinction. Fix the auxiliary implementation Z. Either an infinitely-often one-way function exists relative to Z, in which case it can be plugged into the candidate construction and the key agreement is obtained from Z alone; or no such function exists relative to Z, in which case efficient inverters exist for every efficient Z-aided function (by hardness amplification [Yao82], and in the distributional form of [IL89] as stated in [BHT14]), and these inverters are used to make the inefficient transcript-consistent-view-sampling attacks of [BKSY11] (for perfect correctness) and [IR89, BM09, BM17] (for imperfect correctness) efficient in the presence of Z.

The obstruction to the general case is explicitly named. The sampling attacks are adaptive: each round of sampling is applied to a function that itself invokes the inverter from the previous round, so the query complexity to Z can blow up (a sampler needing $O(n^2)$ queries to invert an n-query function leads to $O(n^4)$ after one step), and only a constant number of such steps can be afforded. The adaptivity-reduction technique of [MMV11] does not remove the constant-round restriction, because the low-adaptivity attacker it produces must simulate the original attacker in its head, which is not efficient for protocols that are not constant round.

2 Notation and parameters

- $\lambda \in \mathbb{N}$ is the security parameter; $\mathrm{negl}(\lambda)$ denotes a negligible function; PPT means probabilistic polynomial time.
- An *oracle-aided PPT* machine A^{O} is a PPT machine with oracle access to O, where each oracle call counts as one step, and which runs in polynomial time for *every* oracle O.

- Calligraphic letters $\mathcal{P}, \mathcal{Q}, \mathcal{Z}, \mathcal{F}, \mathcal{KA}$ denote cryptographic primitives; sans-serif letters $\mathsf{P}, \mathsf{Q}, \mathsf{Z}, \mathsf{F}$ denote particular implementations of them.
- io-$\mathcal{F}$ is the primitive of infinitely-often one-way functions and io-$\mathcal{KA}$ the primitive of infinitely-often key agreement (defined below).
- For an oracle O, $\langle \mathsf{Alice}^{\mathsf{O}}(1^\lambda), \mathsf{Bob}^{\mathsf{O}}(1^\lambda) \rangle$ denotes the distribution of triples (T, K_A, K_B) over the randomness of the two parties, where T is the transcript and K_A, K_B are the parties' outputs.

The parameters are:

- λ, the security parameter.
- $\mathcal{Z}$, the auxiliary primitive, universally quantified over all cryptographic primitives.
- ε, the correctness function of the key agreement; unrestricted (Definition 6, the paper's Definition 2.7, places no lower bound on it).
- q_A, q_B, the number of oracle queries made by Alice and Bob; unrestricted (any polynomial) in the conjecture, restricted to constants in the paper's theorems.

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 a set of functions (the *implementations* of $\mathcal{P}$) and $R_{\mathcal{P}}$ is a relation. For $\mathsf{P} \in F_{\mathcal{P}}$, the statement $(\mathsf{P}, A) \in R_{\mathcal{P}}$ means that the adversary A *breaks* the implementation P. It is required that at least one $\mathsf{P} \in F_{\mathcal{P}}$ is computable by a PPT algorithm.

Definition 2 (Joint primitive). Given primitives $\mathcal{P} = (F_{\mathcal{P}}, R_{\mathcal{P}})$ and $\mathcal{Q} = (F_{\mathcal{Q}}, R_{\mathcal{Q}})$, the *joint primitive* $(\mathcal{P}, \mathcal{Q}) = (F_{(\mathcal{P},\mathcal{Q})}, R_{(\mathcal{P},\mathcal{Q})})$ is defined by $F_{(\mathcal{P},\mathcal{Q})} := F_{\mathcal{P}} \times F_{\mathcal{Q}}$, and for

$\mathsf{G} = (\mathsf{P}, \mathsf{Q})$ and $A = (A_{\mathcal{P}}, A_{\mathcal{Q}})$ one sets $(\mathsf{G}, A) \in R_{(\mathcal{P},\mathcal{Q})}$ iff $(\mathsf{P}, A_{\mathcal{P}}) \in R_{\mathcal{P}}$ or $(\mathsf{Q}, A_{\mathcal{Q}}) \in R_{\mathcal{Q}}$.

Definition 3 (Semi- and ∀∃-semi-black-box reductions). There is a *semi-black-box reduction* of $\mathcal{P}$ to $\mathcal{Q}$ if there is an oracle-aided PPT machine P such that for every implementation $\mathsf{Q} \in F_{\mathcal{Q}}$:

- (implementation) $\mathsf{P}^{\mathsf{Q}} \in F_{\mathcal{P}}$; and
- (security) if there exists an oracle-aided PPT A with $(\mathsf{P}^{\mathsf{Q}}, A^{\mathsf{Q}}) \in R_{\mathcal{P}}$, then there exists an oracle-aided PPT S with $(\mathsf{Q}, \mathsf{S}^{\mathsf{Q}}) \in R_{\mathcal{Q}}$.

If the order of quantifiers for the implementation is switched, namely if for every $\mathsf{Q} \in F_{\mathcal{Q}}$ there exists an oracle-aided PPT P for which the two conditions above hold, then the reduction is called a ∀∃-*semi-black-box reduction.*

Definition 4 ([semi → ∀∃-semi] black-box uselessness). A primitive $\mathcal{Q}$ is [semi → ∀∃-*semi*] *black-box useless* for a primitive $\mathcal{P}$ if for every auxiliary primitive $\mathcal{Z}$, whenever there exists a semi-black-box reduction of $\mathcal{P}$ to the joint primitive $(\mathcal{Q}, \mathcal{Z})$, there also exists a ∀∃-semi-black-box reduction of $\mathcal{P}$ to $\mathcal{Z}$ alone.

Definition 5 (Infinitely-often one-way functions). An oracle-aided PPT machine F is an *infinitely-often one-way function* relative to an oracle O if for every oracle-aided PPT adversary A there is a negligible function negl and an infinite set $\Lambda \subseteq \mathbb{N}$ such that for all $\lambda \in \Lambda$,

$$\Pr_{x \leftarrow \{0,1\}^{\lambda}} \left[\mathsf{F}^{\mathsf{O}}\left(A^{\mathsf{O}}(1^{\lambda}, \mathsf{F}^{\mathsf{O}}(x))\right) = \mathsf{F}^{\mathsf{O}}(x) \right] = \mathrm{negl}(\lambda).$$

The primitive io-$\mathcal{F}$ has as implementations all functions F, and A breaks F iff the displayed probability is at least $1/p(\lambda)$ for some polynomial p and all but finitely many λ.

Definition 6 (Infinitely-often key agreement). Let $\varepsilon : \mathbb{N} \to [0,1]$. An *$\varepsilon$-key agreement* relative to an oracle O is an interactive protocol between oracle-aided PPT machines Alice and Bob such that

$$\Pr\left[(T, K_A, K_B) \leftarrow \langle \mathsf{Alice}^{\mathsf{O}}(1^\lambda), \mathsf{Bob}^{\mathsf{O}}(1^\lambda)\rangle \;:\; K_A = K_B\right] \geq \varepsilon(\lambda) \quad \text{for all } \lambda.$$

It is *secure* if for every PPT adversary Eve, every polynomial p and all sufficiently large λ,

$$\Pr\left[(T, K_A, K_B) \leftarrow \langle \mathsf{Alice}^{\mathsf{O}}(1^\lambda), \mathsf{Bob}^{\mathsf{O}}(1^\lambda)\rangle, K_E \leftarrow \mathsf{Eve}^{\mathsf{O}}(1^\lambda, T) \;:\; K_E = K_B\right] \leq \frac{\varepsilon(\lambda)}{p(\lambda)}.$$

If the security condition is only required to hold for infinitely many λ (in the sense that for every polynomial p and every adversary there is an infinite set of λ for which the winning probability is below $\varepsilon(\lambda)/p(\lambda)$), the protocol is an *infinitely-often key agreement*. The primitive io-$\mathcal{KA}$ has as implementations all such pairs (Alice, Bob), with Eve breaking an implementation iff the security condition fails.

The unqualified primitives $\mathcal{F}$ and $\mathcal{KA}$ named in Conjecture 1 are the plain counterparts of Definitions 5 and 6, in which the security requirement is imposed for all sufficiently large λ rather than only on an infinite set of them.

Conjecture 1 (one-way functions are black-box useless for key agreement). *The primitive $\mathcal{F}$ of one-way functions is* [semi $\to$ $\forall\exists$*-semi*] *black-box useless for the primitive $\mathcal{KA}$ of key agreement: for every cryptographic primitive $\mathcal{Z}$, if there is a semi-black-box reduction of $\mathcal{KA}$ to $(\mathcal{F}, \mathcal{Z})$, then there is a $\forall\exists$-semi-black-box reduction of $\mathcal{KA}$ to $\mathcal{Z}$ alone.*
No restriction whatsoever is placed on the key-agreement constructions considered: the correctness function ε may be any function, the number of rounds of interaction may be any polynomial in λ, and each of

Alice *and* Bob *may make any polynomial number of queries both to the implementation of* $\mathcal{F}$ *and to the implementation of* $\mathcal{Z}$.

Remark 1 (the infinitely-often variant). The paper's theorems are proved for the infinitely-often primitives io-$\mathcal{F}$ and io-$\mathcal{KA}$ of Definitions 5 and 6, and the corresponding statement — that io-$\mathcal{F}$ is [semi $\rightarrow$ $\forall\exists$-semi] black-box useless for io-$\mathcal{KA}$ — is strictly weaker than Conjecture 1: Remark 4.2 (p. 11) records that "$\mathcal{A}$ is black-box useless for $\mathcal{C}$" implies "io-$\mathcal{A}$ is black-box useless for io-$\mathcal{C}$". The paper names that weaker form only parenthetically, when it speaks on p. 18 of "(infinitely-often) OWFs ... for more general forms of key agreement". The central open problem, and hence Conjecture 1, is the unqualified one.

Remark 2 (what is settled). Theorems 4.1, 4.9 and 4.12 give the three restricted results described in Section 1. The paper also records why the natural route to the general case stalls: the recursive use of inverters inside the sampler blows up query complexity to the auxiliary oracle, limiting the argument to a constant number of sampling steps — "this in particular implies that the recursive argument above can be applied for only a constant number of steps" (p. 4) — and the adaptivity-reduction technique of [MMV11] cannot be substituted, because "this technique requires the attacker with lower adaptivity to simulate in 'its head' the original attacker which cannot be done efficiently if the protocol was not constant-round" (p. 18).

Remark 3 (openness). The paper calls this its central open problem: "the central open problem left by our work is that of black-box uselessness of OWFs for arbitrary key agreement protocols" (p. 6), adding that "given our BBU results for special classes of KA protocols, this conjecture may well be within reach", while "a straightforward generalization seems to require a refined sampler technique with low adaptivity". Section 4 repeats this: the "dream result" would be "to show that one-way functions are black-box useless for any key agreement", and the paper leaves it "as an intriguing open question" (p. 11).

Remark 4 (caveats for a reviewer). Three things a reviewer should check. First, the flavour: the abstract states the conjecture with no

flavour attached, while the paper's theorems and its Section 4.2 formulation are in the [semi → ∀∃-semi] flavour; Conjecture 1 follows the theorems in flavour but the abstract in taking plain rather than infinitely-often primitives, and Remark 1 records the relation between the two readings. Second, the resolution status: no resolution of the general case since January 2021 is known to the author of this record, but follow-up work has not been exhaustively surveyed. Third, Theorem 4.12 already removes the query restriction in the Merkle-type case, so a candidate general proof must subsume that case, not merely reprove it.

4 Bibliography

[BHT14] Itay Berman, Iftach Haitner, and Aris Tentes. *Coin flipping of any constant bias implies one-way functions.* 46th Annual ACM Symposium on Theory of Computing, pages 398–407, ACM Press, 2014.

[BKSY11] Zvika Brakerski, Jonathan Katz, Gil Segev, and Arkady Yerukhimovich. *Limits on the power of zero-knowledge proofs in cryptographic constructions.* TCC 2011: 8th Theory of Cryptography Conference, volume 6597 of Lecture Notes in Computer Science, pages 559–578, Springer, 2011.

[BM09] Boaz Barak and Mohammad Mahmoody-Ghidary. *Merkle puzzles are optimal — an $O(n^2)$-query attack on any key exchange from a random oracle.* Advances in Cryptology — CRYPTO 2009, volume 5677 of Lecture Notes in Computer Science, pages 374–390, Springer, 2009.

[BM17] Boaz Barak and Mohammad Mahmoody. *Merkle's key agreement protocol is optimal: An $O(n^2)$ attack on any key agreement from random oracles.* Journal of Cryptology, 30(3):699–734, 2017.

[IL89] Russell Impagliazzo and Michael Luby. *One-way functions are essential for complexity based cryptography (extended abstract).* 30th Annual Symposium on Foundations of Computer Science, pages 230–235, IEEE Computer Society Press, 1989.

[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.

[Mer78] Ralph C. Merkle. *Secure communications over insecure channels.* Communications of the ACM, 21(4):294–299, 1978.

[MMV11] Mohammad Mahmoody, Tal Moran, and Salil P. Vadhan. *Time-lock puzzles in the random oracle model.* Advances in Cryptology —

CRYPTO 2011, volume 6841 of Lecture Notes in Computer Science, pages 39–50, Springer, 2011.

[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.

[Yao82] Andrew Chi-Chih Yao. *Theory and applications of trapdoor functions (extended abstract).* 23rd Annual Symposium on Foundations of Computer Science, pages 80–91, IEEE Computer Society Press, 1982.