Breaking Fully Quantum Time-Lock Puzzles in the Quantum Random Oracle Model

Both the puzzle generator and the honest solver are quantum, and so is the attacker

Status. Statement: AI-written from Abtin Afshar, Kai-Min Chung, Yao-Ching Hsieh, Yao-Ting Lin and Mohammad Mahmoody, *On the (Im)possibility of Time-Lock Puzzles in the Quantum Random Oracle Model*, Cryptology ePrint Archive 2023 (no venue or paper number is printed on the PDF itself), not yet checked by a human. Settled in that paper: a classical generator with a quantum solver (for arbitrary completeness error, and round-optimally under perfect completeness), and a quantum generator with a classical solver (under perfect completeness); open, and conjectured below, is the fully quantum case, where the paper presents no attack and shows that any attack asking only classical queries would prove the Aaronson–Ambainis simulation conjecture.

Category. *Idealized Models & Non-Uniformity.*

Conjecture (informal). *A time-lock puzzle locks a secret into a puzzle so that unwrapping it takes a long time even for someone with many computers running in parallel, while making the puzzle is cheap. If the only source of hardness is a random oracle, one can measure "time" by the number of rounds of oracle queries an algorithm needs, and "parallelism" by how many queries it may ask inside one round. Classically such puzzles cannot exist: whatever the generator does with its queries, an attacker recovers the solution using a number of query rounds proportional to the number of queries the generator made, no matter how many rounds the honest solver needs. The source paper extends this impossibility to a quantum honest solver querying in superposition, and separately to a quantum generator with a classical solver. The remaining case is the one where generator and solver are* both *quantum; there the paper gives no attack, and it shows why its technique stalls, by exhibiting a scheme — based on a long-standing conjecture about approximating quantum query algorithms classically being false — that no attacker restricted to classical queries can break even with polynomially many queries and unlimited parallelism. The conjecture is that the impossibility nevertheless holds once the attacker is allowed quantum queries: for every such puzzle there is a quantum attacker that recovers the solution about as well as the honest solver does, using polynomially many quantum queries arranged into a number of rounds proportional to the generator's query count.*

1 The setting

Time-lock puzzles [RSW96] let a generator cheaply produce a puzzle P hiding a solution s so that recovering s takes a long *sequential* time, even for an adversary with unbounded parallelism. Whether such puzzles can be built from symmetric-key primitives alone is naturally studied in the random oracle model, where "time" has a clean information-theoretic proxy: the number of *rounds* of oracle queries an algorithm needs, with arbitrarily many queries allowed inside a single round.

In that model, Mahmoody, Moran and Vadhan [MMV11] ruled out time-lock puzzles outright: if the generator asks n oracle queries and the honest solver asks m, there is always an attacker that recovers s using only $O(n)$ rounds of queries and $\mathrm{poly}(n, m)$ queries in total,

however sequential the honest solver is. They also showed n rounds is the best one can hope for, via a "pseudo-chain" puzzle that genuinely needs n rounds.

Since random oracles can also be given in superposition [BDF11], the barrier splits into three settings, according to which of the generator and the solver are quantum. The present paper closes two of them. When the generator is classical and the solver quantum, it exhibits an attacker asking $O(n)$ rounds of *classical* queries and $\mathrm{poly}(n, m)$ queries total, so the solver's quantum power cannot be what makes the puzzle hard; a variant of the attack even runs in quantum polynomial time, which rules out ROM constructions that lean on additional computational assumptions. When the generator is quantum and the solver classical, it gives an attack in exactly n rounds and $m \cdot n$ queries for perfectly complete schemes, by combining certificate-complexity ideas with the polynomial representation of a quantum party's state used in the key-agreement impossibility of [ACC22]. It also shows the pseudo-chain still needs n rounds against a solver making parallel *quantum* queries, so $O(n)$ rounds remains optimal in the quantum world.

The remaining case — generator and solver both quantum — is left open, and the paper explains precisely what blocks the technique it uses elsewhere. Every one of its attacks asks only classical queries, and classical-query attacks cannot suffice unconditionally: assuming the paper's asymptotic Quantum Polynomial-Query Simulation Conjecture (Conjecture 6.1, p. 29) — a weakening of the folklore simulation conjecture formally stated as Conjecture 4 of Aaronson and Ambainis [AA14] — is *false*, the paper builds a fully quantum time-lock puzzle that no classical-query adversary breaks, even with polynomially many queries and unlimited parallelism. It does so by amplifying the conditional one-way-communication quantum key agreement of [ACC22] to negligible completeness and soundness error and reading the resulting transcript as a puzzle. The simulation conjecture — that the acceptance probability of an n-quantum-query algorithm can be approximated on most oracles using $\mathrm{poly}(n)$ classical queries — is known only in the case where that probability is always 0 or 1 [OSSS05] and has resisted proof for over a decade, so the barrier is unlikely to be removed soon. What is left, and what this conjecture asks for, is an attack that itself uses quantum queries.

The stakes reach beyond time-lock puzzles. Under the reading of

a key-agreement transcript as a puzzle, observed in [MMV11] and used here, round-efficient attacks on puzzles with quantum parties translate into round-efficient attacks on key agreement in the quantum ROM [ACC22], and the paper notes that its techniques plausibly bear on the impossibility of verifiable delay functions from random oracles [MSW20]. Conversely, a resolution in the negative would be a genuine positive result: proofs of sequential work *are* achievable against quantum adversaries from random oracles [CFHL21], so the boundary between what sequential primitives survive quantum access and what do not is exactly what is at issue.

Progress to date. The paper settles the two mixed settings. Theorem 3.1 gives, for a classical generator and quantum solver, an attacker asking $4n^2m^2/(\varepsilon\delta^2)$ classical queries in n/ε rounds with success at least $1-\rho-\varepsilon-\delta$, and a second variant with $2n$ rounds; Theorem 3.12 makes the attack run in quantum polynomial time. Theorem 4.10 gives a round-optimal attack in at most n rounds and at most mn queries with success probability 1 for perfectly complete schemes with a classical generator and quantum solver, and Theorem 4.6 does the same for a quantum generator and classical solver. Theorem 5.3 shows the pseudo-chain puzzle of Mahmoody, Moran and Vadhan is still secure against solvers making q rounds of k-parallel quantum queries, up to $O(\sqrt{kq^3/2^\kappa})$, so $O(n)$ rounds cannot be improved. For the fully quantum case there is no attack, only the conditional barrier of Theorem 6.2 against classical-query attackers.

2 Notation and parameters

- $\kappa \in \mathbb{N}$ is the security parameter. All asymptotics are in κ.
- $\mathcal{X}$ and $\mathcal{Y}$ are finite nonempty sets with $\mathcal{Y}$ carrying a group operation $\oplus$ (for concreteness one may take $\mathcal{X} = \mathcal{Y} = \{0,1\}^\kappa$ with bitwise XOR). The random oracle is $H\colon \mathcal{X} \to \mathcal{Y}$, drawn uniformly from the set of all such functions.
- Quantum access to H is via the unitary $O_H\colon |x\rangle|y\rangle \mapsto |x\rangle|y \oplus H(x)\rangle$ on $\mathbb{C}^{\mathcal{X}} \otimes \mathbb{C}^{\mathcal{Y}}$. For $k \geq 1$ the k-parallel query operator is $O_H^{\otimes k}$, acting on k independent query/answer register pairs.
- n is the number of oracle queries of the puzzle generator, m

the number of oracle queries of the honest puzzle solver, and ρ the completeness error of the puzzle (the honest solver fails with probability at most ρ).

- $\varepsilon, \delta \in (0,1]$ are slack parameters of the attack: ε controls its round budget and δ its additional loss in success probability.
- r_G denotes the randomness of the puzzle generator, P the puzzle, s the correct solution, s'' the attacker's output.
- $\not{p}$ denotes a polynomial in four variables.

3 The conjecture

Definition 1 (rounds of quantum queries). Let $H : \mathcal{X} \to \mathcal{Y}$. A quantum oracle algorithm $\mathsf{A}^{(\cdot)}$ *asks at most q queries in at most d rounds* if, on every input, its execution relative to H can be written as

$$U_d \, O_H^{\otimes k_d} \, U_{d-1} \cdots U_1 \, O_H^{\otimes k_1} \, U_0$$

applied to a fixed initial state, followed by a measurement of a designated output register, where the unitaries $U_0, \ldots, U_d$ do not depend on H and the widths satisfy $k_1, \ldots, k_d \geq 0$ and $\sum_{i=1}^{d} k_i \leq q$. Randomized algorithms are included by letting U_0 prepare a uniform superposition over a randomness register that is measured at the start. A *classical-query* algorithm is the special case in which every query register is measured in the computational basis immediately before each application of $O_H^{\otimes k_i}$.

Definition 2 (fully quantum time-lock puzzle in the quantum random oracle model). Let $n = n(\kappa)$, $m = m(\kappa)$ and $\rho = \rho(\kappa)$. A *fully quantum time-lock puzzle with generator query complexity n, solver query complexity m and completeness error ρ* is a pair $\Pi = (\mathsf{Gen}, \mathsf{Sol})$ of quantum oracle algorithms such that:

- $\mathsf{Gen}^H(r_G)$, on randomness r_G, makes at most n quantum queries to H and outputs a pair of classical strings (P, s), the puzzle and its solution;

- $\mathsf{Sol}^{H}(\mathsf{P})$ makes at most m quantum queries to H and outputs a classical string s';
- (completeness)

$$\Pr\left[\mathsf{s} = \mathsf{s}' \;:\; (\mathsf{P},\mathsf{s}) \leftarrow \mathsf{Gen}^{H}(r_G),\ \mathsf{s}' \leftarrow \mathsf{Sol}^{H}(\mathsf{P})\right] \geq 1-\rho,$$

over the uniform choice of H, the choice of r_G, and the measurements of Gen and Sol.

No restriction is placed on the number of rounds in which Gen and Sol ask their queries; in particular the honest solver may need all m of its queries to be sequential, which is what a time-lock puzzle intends to exploit.
The requirement that the puzzle P and the solution s be *classical* is a restriction, and the conjecture below covers only puzzles of that form. It is the form taken by the source paper's own formal definition (Definition 2.6, p. 11) and by its fully quantum construction (Construction 6.13, p. 33, where $\mathsf{P} = c_3$ and $\mathsf{s} = \mathsf{key}_{A_3}$), and hence the form in which its Theorem 6.2 barrier is proved. The paper explicitly contemplates, but never formalizes, the case in which the puzzle itself is a quantum object; that case is a strictly more general question which this statement does not address.

Conjecture 1 (breaking fully quantum time-lock puzzles with quantum queries). *There is a fixed polynomial p, independent of everything below, such that the following holds.*
Let $\Pi = (\mathsf{Gen},\mathsf{Sol})$ be any fully quantum time-lock puzzle in the quantum random oracle model in the sense of Definition 2, with generator query complexity $n = n(\kappa)$, solver query complexity $m = m(\kappa)$, and completeness error $\rho = \rho(\kappa)$. Then for every $\varepsilon,\delta \in (0,1]$ there exists a quantum oracle algorithm E (allowed to depend on Π, on ε and on δ, and otherwise unrestricted in its internal computation, in particular not required to be efficient) such that:

1. E *makes at most $p(n,m,1/\varepsilon,1/\delta)$ quantum queries to H;*

2. *those queries are arranged in at most* $\lceil n/\varepsilon \rceil$ *rounds in the sense of Definition 1; and*

3. $$\Pr\big[\mathsf{s}'' = \mathsf{s} \;:\; (\mathsf{P},\mathsf{s}) \leftarrow \mathsf{Gen}^{H}(r_G),\ \mathsf{s}'' \leftarrow \mathsf{E}^{H}(\mathsf{P})\big] \;\geq\; 1-\rho-\varepsilon-\delta,$$

where the probability is over the uniform choice of $H : \mathcal{X} \to \mathcal{Y}$*, the generator randomness* r_G*, and the measurements of* E*.*

Here n*,* m *and* ρ *may be arbitrary functions of the security parameter* κ*, and the statement is asserted for all sufficiently large* κ*.*

Corollary 1 (consequence, if the conjecture holds). *Taking* $\varepsilon = \delta = 1/4$ *and* $\rho \leq 1/4$ *would give, for every such* Π*, a quantum attacker asking* $\mathrm{poly}(n,m)$ *quantum queries in* $O(n)$ *rounds that recovers* s *with probability at least* $1/4$*, so no fully quantum time-lock puzzle in the quantum random oracle model would have negligible soundness error against* $O(n)$*-round quantum attackers.*

Remark 1 (the quantitative shape is the sharp form). The source paper's own words are only that "Theorem 1.1 could potentially be extended to the fully quantum setting using a quantum adversary", without pinning down the round and query bounds. The shape asserted above — $\lceil n/\varepsilon \rceil$ rounds, $p(n,m,1/\varepsilon,1/\delta)$ queries, success $1-\rho-\varepsilon-\delta$ — is therefore Theorem 3.1's shape transplanted verbatim to the new setting rather than a bound the paper writes down for it. An attacker achieving, say, $\mathrm{poly}(n)$ rounds, or merely non-negligible rather than near-honest success, would already answer what the authors are asking and would refute nothing here; Conjecture 1 is the sharp form.

Remark 2 (scope). As recorded at the end of Definition 2, the puzzle and its solution are classical here. The paper observes that in the fully quantum setting "one can imagine the puzzle itself to be a quantum object", and the restriction is deliberate: it matches the paper's Definition 2.6 and is the setting of its Theorem 6.2 barrier. It does mean that the quantum-puzzle variant is a strictly more general question, on which Conjecture 1 is silent.

4 Bibliography

[AA14] S. Aaronson, A. Ambainis. *The need for structure in quantum speedups.* 2014. (No venue is given in the source reference list.)

[ACC22] P. Austrin, H. Chung, K.-M. Chung, S. Fu, Y.-T. Lin, M. Mahmoody. *On the impossibility of key agreements from quantum random oracles.* Advances in Cryptology — CRYPTO 2022: 42nd Annual International Cryptology Conference, Proceedings, Part II, pages 165–194, Springer, 2022.

[BDF11] D. Boneh, O. Dagdelen, M. Fischlin, A. Lehmann, C. Schaffner, M. Zhandry. *Random oracles in a quantum world.* Advances in Cryptology — ASIACRYPT 2011, volume 7073 of Lecture Notes in Computer Science, pages 41–69, Springer, Heidelberg, 2011.

[CFHL21] K.-M. Chung, S. Fehr, Y.-H. Huang, T.-N. Liao. *On the compressed-oracle technique, and post-quantum security of proofs of sequential work.* Advances in Cryptology — EUROCRYPT 2021, Part II, volume 12697 of Lecture Notes in Computer Science, pages 598–629, Springer, Heidelberg, 2021.

[MMV11] M. Mahmoody, T. Moran, S. 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, Heidelberg, 2011.

[MSW20] M. Mahmoody, C. Smith, D. J. Wu. *Can verifiable delay functions be based on random oracles?* ICALP 2020: 47th International Colloquium on Automata, Languages and Programming, volume 168 of LIPIcs, pages 83:1–83:17, Schloss Dagstuhl, 2020.

[OSSS05] R. O'Donnell, M. Saks, O. Schramm, R. A. Servedio. *Every decision tree has an influential variable.* 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05), pages 31–39, IEEE, 2005.

[RSW96] R. L. Rivest, A. Shamir, D. A. Wagner. *Time-lock puzzles and timed-release crypto.* Massachusetts Institute of Technology, Laboratory for Computer Science, 1996.