On the Impossibility of Key Agreements from Quantum Random Oracles

Per Austrin[1⋆], Hao Chung[2⋆⋆], Kai-Min Chung[3⋆⋆⋆], Shiuan Fu[3†], Yao-Ting Lin[3‡], and Mohammad Mahmoody[4 §]

[1] KTH Royal Institute of Technology, Stockholm
[2] Carnegie Mellon University, USA
[3] Academia Sinica, Taiwan
[4] University of Virginia, USA

Abstract. We study the following question, first publicly posed by Hosoyamada and Yamakawa in 2018. Can parties A, B with quantum computing power and classical communication rely only on a random oracle (that can be queried in quantum superposition) to agree on a key that is private from eavesdroppers?

We make the first progress on the question above and prove the following.

- When only *one* of the parties A is classical and the other party B is quantum powered, as long as they ask a total of d oracle queries and agree on a key with probability 1, then there is always a way to break the key agreement by asking $O(d^2)$ number of *classical* oracle queries.
- When both parties can make quantum queries to the random oracle, we introduce a natural conjecture, which if true would imply attacks with $\mathrm{poly}(d)$ *classical* queries to the random oracle. Our conjecture, roughly speaking, states that the multiplication of any two degree-d real-valued polynomials over the Boolean hypercube of influence at most $\delta = 1/\mathrm{poly}(d)$ is nonzero. We then prove our conjecture for exponentially small influences, which leads to an (unconditional) classical $2^{O(md)}$-query attack on any such key agreement protocol, where m is the oracle's output length.
- Since our attacks are classical, we then ask whether it is always possible to find classical attacks on key agreements with imperfect completeness in the quantum random oracle model. We prove a barrier for this approach, by showing that if the folklore "Simulation Conjecture" (first formally stated by Aaronson and Ambainis in 2009) about the possibility of simulating efficient-query quantum algorithms using efficient-query classical algorithms is false, then there is in fact such a secure key agreement in the quantum random oracle model that cannot be broken classically.

⋆ austrin@kth.se.
⋆⋆ haochung@andrew.cmu.edu. Supported by Packard Fellowship and NSF award 2044679. Part of the work was done when working at Academia Sinica.
⋆⋆⋆ kmchung@iis.sinica.edu.tw. Partially supported by the 2021 Academia Sinica Investigator Award (AS-IA-110-M02) and Executive Yuan Data Safety and Talent Cultivation Project (AS-KPQ-110-DSTCP).
† rubik.sf@gmail.com.
‡ 1213tonylin@gmail.com.
§ mohammad@virginia.edu. Supported by NSF grants CCF-1910681 and CNS1936799.

1 Introduction

In a course project, now known as "Merkle Puzzles", Merkle [Mer74] proposed the first ever nontrivial key agreement protocol between two parties using an ideal hash function. This protocol can be formally analyzed in the random oracle model (ROM) to prove that Alice and Bob can ask d queries to a random oracle h and agree on a key, while an eavesdropper Eve, who can see the exchanged messages t, needs $\Omega(d^2)$ queries to h to find the key. Shortly after, seminal works [DH76,RSA78] showed how to achieve a super-polynomially secure key agreement protocol by relying on number theoretic assumptions. In comparison, Merkle's protocol suffers from only offering polynomial security. However, after all the years of research and newly developed candidate constructions for public-key encryption and key agreements (see the survey [Bar17] for such works), Merkle's protocol enjoys a qualitative advantage: it only relies on an idealized *symmetric primitive*, namely a random function without any structure. Indeed, basing public-key encryption or key agreement on symmetric key primitives is still one of the most fundamental open questions in cryptography.

Merkle's protocol led to the following natural question (also attributed to Merkle by [IR89]). Is there any d-query key agreement protocol in the ROM with larger security $\omega(d^2)$, or is the $O(d^2)$ bound optimal?[5] Impagliazzo and Rudich were the first to prove an upper bound on the security of key agreement protocols in the ROM. They showed that all such protocols can be broken by an attacker who asks $\tilde{O}(dr)^3$ queries, where r is the round complexity of the protocol. This result, in particular, showed that there is no "black-box" way of obtaining key agreements from one-way functions, because roughly speaking a random oracle is one-way with high probability. Finally, Barak and Mahmoody [BM17] showed that every key agreement in the ROM can be broken by $O(d^2)$ queries, showing that Merkle's protocol was indeed optimal.

Key agreement in a quantum world. Merkle's protocol and attacks of [IR89, BM17] are all classical. With the growing interest in understanding the power and limitations of quantum computation, this brings up the following natural question. What if parties can perform quantum computation? Bennett and Brassard [BB84] showed that when parties can communicate quantum bits, then there is an information-theoretically secure key agreement protocol. This still leaves out the case of protocols with classical communication, which is the focus of our work. Classical-communication protocols are particularly attractive as they can be used over the current infrastructure (e.g., the Internet). In this model, all the quantum computation is done locally by the parties who exchange classical messages and aim to establish a private key. We refer to this model as the quantum-computation classical-communication (QCCC) model.

Quantum random oracle. A QCCC protocol in the quantum random oracle model (QROM) allows a quantum-powered party to ask superposition queries to the oracle. This party could be the honest parties or the attacker. Brassard and Salvail [BS08] and Biham, Goren and Ishai [BGI08] revisited the security of Merkle's protocol against quantum

[5] Note that a sufficiently large *polynomial gap* could still be a meaningful fine-grained security, particularly because this cap can only mean *more* security when the CPU clocks get shorter. In particular, with faster computers, Alice and Bob can pick a *larger* d, while running in the same time as before, while Eve now needs d times more running time than Alice and Bob.

adversaries and showed that Merkle's protocol can be broken by a quantum eavesdropper (essentially, Grover's search [Gro96]) that asks $O(d)$ number of quantum queries to the random oracle. This showed that Merkle's protocol gives no super-linear security over d against quantum attackers. Brassard and Salvail [BS08] then showed how to regain a super-linear gap by having Alice and Bob also leverage quantum queries to the oracle. Their protocol had the extra property that only *one* of the parties Alice and Bob needs to run a quantum algorithm.[6] Brassard et al. [BHK+15] further improved this result and showed that a quantum Alice and Bob can agree on a key by d queries, while even a quantum attacker would require $\approx d^2$ number of queries to break it.

All of these works seek lower bounds on the gap between the query complexity of quantum algorithms Alice/Bob and the adversary Eve. However, no previous work has shown an *upper* bound on the achievable security. In fact, our current knowledge about the limitations of security in the QROM is consistent with the possibility that QCCC protocols can establish a key agreement over a *classical* channel, while it would take exponentially many queries to the oracle (even by a quantum attacker) to find the key. This brings up the main question of this work, which was also posed by Hosoyamada and Yamakawa [HY20].[7]

> *Is there a key agreement protocol using* classical *communication, in which Alice and Bob ask* d quantum *queries to a random oracle, while the eavesdropper needs a* super-*polynomial* $d^{\omega(1)}$ *number of queries to find the key?*

1.1 Our Results

In this work, we present the first barriers against obtaining super-polynomially secure QCCC key agreement protocols in the QROM model.

Classical Alice Quantum Bob (CAQB). Our first result shows that when one of the parties Alice is classical, the quadratic gap achieved by Merkle is optimal, *even against classical* adversaries. This is an interesting setting on its own, as it can model unbalanced parties. For example, suppose Google wants to agree on a key with a typical user, who does not have any quantum computing power, over the Internet. Then, our result shows that there is a limit to how much security such protocols can achieve.

Theorem 1.1 (Attacking CAQB protocols – informal). *Suppose Π is a QCCC d-query key agreement protocol with perfect completeness in the QROM. If Alice is classical and only Bob uses quantum queries to the random oracle, then there is a* classical *adversary who can find the key by asking $O(d^2)$ classical queries to the oracle.*

Note that the above result assumes that the two parties agree on a key with probability one, and this is the case for all of our attacks in this work; extending them to allow imperfect completeness is an intriguing question for future work.

Quantum Alice and Quantum Bob (QAQB). We then turn to study protocols in which Alice and Bob both have quantum access to the oracle. For this more general setting, we

[6] In comparison, Theorem 1.1 shows that such protocols (with a classical party and a quantum party) cannot offer more than quadratic security when the protocol has perfect completeness.

[7] To the best of our knowledge, the question was first asked in 2018 [HY18].

show a *conditional result* based on a conjecture about multilinear polynomials, which will also prove for some extreme cases.

Some basic notions. We first recall some basic notions about polynomials. Suppose

$$f = \sum_{\mathcal{S} \subseteq [N]} \alpha_{\mathcal{S}} \prod_{i \in \mathcal{S}} x_i$$

is a multilinear polynomial over binary variables $x_i \in \{\pm 1\}, i \in [N]$ and real coefficients $\alpha_{\mathcal{S}} \in \mathbb{R}, \mathcal{S} \subseteq [N]$. The degree of f is $\max_{\alpha_{\mathcal{S}} \neq 0} |\mathcal{S}|$ and the ℓ_2 norm of f is $\|f\|_2 = \mathbb{E}_{\mathbf{x} \leftarrow \{\pm 1\}^N}[f(\mathbf{x})^2]$. The influence of x_i on f is defined as $\mathrm{Inf}_i(f) = \sum_{i \in \mathcal{S}} \alpha_{\mathcal{S}}^2$, and more generally for a distribution F over such multilinear polynomials, we let $\mathrm{Inf}_i(F) = \mathbb{E}_{f \leftarrow F}[\mathrm{Inf}_i(f)]$ denote the *expected* influence.

Conjecture 1.2 (Polynomial Compatibility) *There is a function $\delta(d) = 1/\operatorname{poly}(d)$, such that the following holds for all $d \in \mathbb{N}$. Suppose F, G are distributions over multilinear polynomials of degree d with variables $x_1, \ldots, x_N \in \{\pm 1\}$ and ℓ_2-norm 1 and bounded influences $\mathrm{Inf}_i(F), \mathrm{Inf}_i(G) \leq \delta(d)$ for all $i \in [N]$. Then, there exist $f \in \operatorname{supp}(F)$, $g \in \operatorname{supp}(G)$ and $\mathbf{x} \in \{\pm 1\}^N$ such that $f(\mathbf{x}) \cdot g(\mathbf{x}) \neq 0$.*

All assumptions are needed. In Appendix B of the full version [ACC+22] we show, through constructive examples, that for Conjecture 1.2 to be true one needs *both* F, G to have *both* of the low-degree and low-influence conditions. Furthermore, we give an example showing that relation between δ and the degree d must satisfy $\delta < \frac{1}{2d}$, otherwise the conjecture is false.

We then prove the following conditional result. We state the group structure $\mathbb{Z}_2^m$ to clarify how the answers are read by the quantum algorithm. In particular, the oracle answers are added (in $\mathbb{Z}_2^m$) to the answer registers.

Theorem 1.3 (Attacking QAQB protocols – informal). *If Alice and Bob ask a total of d quantum queries to a random oracle $h \colon [N] \to \mathbb{Z}_2^m$ and agree on a key k with probability 1, and if Conjecture 1.2 holds, then there is an attacker who asks $\operatorname{poly}(d, m)$* classical *queries to h and finds the key k with probability 0.9.*

More generally, we show that if the Polynomial Compatibility Conjecture holds with respect to an influence δ, then for any d-query key agreement protocol using the random oracle $h : [N] \to \{0,1\}^m$, there is an attacker who asks $\operatorname{poly}(dm/\delta)$ number of queries and finds the key with high probability. Thus while we are unable to prove Conjecture 1.2 as stated, this motivates trying to prove it for some smaller influence δ which is independent of the size of the input space $N = 2^\kappa$ for security parameter κ.

Random oracles using other groups for answers. Random oracles can be defined with an arbitrary Abelian group G (other than $\mathbb{Z}_2^m$). We further extend Theorem 1.3 in two directions. We first generalize the Polynomial Compatibility Conjecture (see Conjecture 5.5) that is parameterized by an Abelian group G_1, such that when $G_1 = \mathbb{Z}_2$, then this becomes Conjecture 1.2. We then show (see Theorem 4.8) that if this conjecture holds for any constant-size Abelian group G_1, then for all Abelian groups G_2 we can get $\operatorname{poly}(d, \log |G_2|)$-query (classical) attacks on perfectly complete key agreement

protocols that use a random oracle $h\colon [N] \to G_2$. Note that this reduction allows the size of the group elements in G_2 to grow polynomially with the security parameter κ, while we still get a $\mathrm{poly}(\kappa)$-query (classical) attack.

Proving the conjecture for exponentially small influence. We then prove (a variant of) Conjecture 1.2 where δ is exponentially small $\delta(d) < O(2^{-d}/d)$ as a function of d instead of polynomially small. As a result, we obtain an $O(2^{dm} \cdot d^2)$-query (classical) attack on any key agreement in the QROM. Note that this is a nontrivial upper bound on the security, only when the input length n is sufficiently larger than m (e.g., when $n = m^2$, or that the input space is $\{0,1\}^*$, while the outputs have fixed length m).

Learning heavy queries for quantum protocols. One of the major contributions of our work in proving Theorem 1.3 is to generalize the "heavy-queries learner" of Barak and Mahmoody [BMG09a] to the quantum setting. In fact, doing so is crucial for us even to come up with *any candidate attack* in the QAQB model, regardless of proving it to be successful. Our quantum-heavy query learner could pave the way for proving more separations in the quantum random oracle model.

Implications to quantum black-box separations. The $\mathrm{poly}(d)$-query attacks of [IR89, BM17] were used to obtain black-box separations for key agreement from one-way functions. The same argument extends to the case of QCCC key agreements with perfect completeness. Our Theorem 1.1 also leads to a $\mathrm{poly}(d, m) \leq \mathrm{poly}(\kappa)$-query attack, and hence can be used to obtain similar separations with respect to "quantum black-box" constructions, for the case of perfect completeness and classical Alice. In a quantum black-box construction [HY20] the reductions (to implement the primitive and prove its security) can have quantum superposition access to the oracles they use. Our Theorem 1.3 implies a similar separation for QCCC key agreement protocols from one-way functions, but based on the assumption that Polynomial Compatibility Conjecture holds. See Theorem 6.3 of the full version [ACC+22] for a formalization.

Attacking other primitives. Once we obtain polynomial-query attacks on QCCC key agreement in the QROM model, we also immediately obtain further corollaries about the impossibility of using quantum random oracles for realizing other primitives such as public-key encryption and oblivious transfer, or more generally, any primitive $\mathcal{P}$ that implies key agreement in a black-box way, when the communication and the inputs are classical. For example, since oblivious transfer implies key agreement [GKM+00], our Theorems 1.1 and 1.3 also extend to rules out the possibility of OT protocols with perfect completeness in the QCCC model using random oracles. Similarly, our separations extend to similar separations from other primitives, such as Oblivious Transfer, that imply key agreements in a black-box way.

Connection to the Simulation Conjecture. Since our attacks on perfectly complete key agreement protocols in the QROM model are classical, it is natural to ask if such classical attacks can be extended to all such protocols, even against protocols with imperfect completeness. We show that obtaining such attacks would resolve a basic and long-standing open question about the power of quantum vs. classical algorithms. That means obtaining such classical attacks *unconditionally* might be quite challenging. More specifically, a folklore conjecture, which we refer to as the "Simulation Conjecture", states that for any $\mathrm{poly}(\kappa)$-query quantum algorithm Q^h using a random oracle h, and for any $\varepsilon = 1/\mathrm{poly}(\kappa)$, there is another $\mathrm{poly}(\kappa)$-query *classical* algorithm S^h

that can approximate the acceptance probability $\Pr[Q^h = 1]$ with $\pm\varepsilon$ additive error, for $1 - \varepsilon$ fraction of oracles h. Aaronson and Ambainis (see Conjecture 4 in [AA09]) formalized this conjecture and showed that it is implied by a Fourier-analytic conjecture, now known as the Aaronson-Ambainis conjecture, that has some resemblance to our Polynomial Compatibility Conjecture but also with key differences (see Section 1.3).

In this work, we observe that the Simulation Conjecture is in fact necessary for extending classical attacks on key agreement protocols in the QCCC model using quantum random oracles and with negligible completeness error. Doing so shows that proving an unconditional classical attack of $\mathrm{poly}(\kappa)$ query complexity on QCCC key agreements in the QROM are not possible, unless one resolves the Simulation Conjecture positively.

Theorem 1.4 (QCCC key agreement against classical adversaries – informal). *If the Simulation Conjecture is false, then there is a key agreement in the QCCC model in which quantum powered parties Alice and Bob use a random oracle to agree on a bit b with probability $1 - \mathrm{negl}(\kappa)$, while for an infinite set of security parameters κ, the protocol is secure against all classical $\mathrm{poly}(\kappa)$-query eavesdropping algorithms.*

See Theorem 7.6 of the full version [ACC+22] for a formalization of the theorem above, and see the next section below for a highlight of the ideas behind its proof.

1.2 Technical Overview

In this section, we highlight the ideas behind Theorems 1.1, 1.3, and 1.4.

Our starting point is the work of Brakerski et al. [BKSY11] that showed a simpler attack and analysis than that of [IR89, BM17], to break any key agreement *with perfect completeness* in the ROM using $O(d^2)$ queries. To obtain our results, we start by modifying the attack of [BKSY11] to a version that is more robust so that it can be adapted to the quantum setting. We start by describing this attack for the setting that *both* Alice and Bob are *classical*. We then discuss, step by step, the new ideas that are introduced to extend the attack to the case of quantum parties.

Case of Classical Alice and Classical Bob. Let $h : [N] \to \{0,1\}^m$ be the random oracle. Suppose t is the (classical) transcript of the protocol, and P_{A} (resp. P_{B}) is the partial function that defines the set of queries asked by Alice (resp. Bob) and their answers. Let $Q_{\mathsf{A}} = \mathrm{dom}(P_{\mathsf{A}})$ (resp. $Q_{\mathsf{B}} = \mathrm{dom}(P_{\mathsf{B}})$) be the set of queries asked by Alice (resp. Bob). Also, let k be the key that Alice and Bob agree upon.

Attacking CACB protocols. The adversary Eve E is given the transcript t and wants to find out the key k. Our simple attack follows the "heavy query learning" approach of [IR89, BM17]. Eve maintains a partial function L that defines the answers to the queries Q_L that are asked by Eve has asked so far. (At the beginning $L = \emptyset$.) During the attack, Eve asks any query $x \notin Q_L$ that is "ε-heavy for being in Q_{A}" conditioned on what Eve knows so far: (L, t). More formally, x is called ε-heavy if $\Pr[x \in Q_{\mathsf{A}} | L, t] \geq \varepsilon$. Whenever Eve reaches a point that there is no heavy query left to ask, Eve simply samples a full (fake) view V'_{A} for Alice in her head and outputs the key k'_{A} that is implied by V'_{A}. We claim that the attack is both efficient and successful. Namely, Eve asks an expected number of at most d/ε queries, and that it finds the key $k'_{\mathsf{A}} = k$ with probability at least $1 - \varepsilon d$. Then, by taking $\varepsilon \approx 1/d$ we obtain the desired result.

Efficiency of the attack. It is easy to prove, using the linearity of expectation, that $\mathbb{E}[|L|] \leq d/\varepsilon$. This is roughly because every query asked by Eve has at least ε-chance of being in Q_{A}, and that there are a limited $|Q_{\mathsf{A}}| \leq d$ possible queries in Q_{A}.

Success of the attack. Perhaps the more interesting aspect is the success of the attack, which is argued based on two facts.

- Independence: For every fixed oracle h and transcript t, the random variables V_{A} and V_{B} that describe the views of Alice and Bob conditioned on h and t are independent random variables (i.e., they have a product distribution).
- Consistency: If (1) the views V_{A} and V_{B} are each consistent with the transcript t, and (2) their partial functions $P_{\mathsf{A}}, P_{\mathsf{B}}$ are also consistent partial functions, then one can conclude that there is an oracle h that is consistent with each of the views $V_{\mathsf{A}}, V_{\mathsf{B}}$. The second condition is equivalent to saying that there is a partial function L such (1) L is consistent with both $P_{\mathsf{A}}, P_{\mathsf{B}}$, and (2) $(Q_{\mathsf{A}} \setminus Q_L) \cap (Q_{\mathsf{B}} \setminus Q_L) = \emptyset$.[8]

The above two facts can be used to argue the success of the attack as follows. Let us fix Bob's (real) view V_{B}. Let $x \in Q_{\mathsf{B}}$ be any particular query asked by Bob that is *not* in Q_L, and hence not learned by Eve. Any such query shall be ε-light (otherwise it was learned by Eve and hence in Q_L). Therefore, the probability that x is in Q'_{A}, where Q'_{A} is the set of queries in the fake view V'_{A} sampled by Eve, is at most ε. By a union bound, with probability at least $1 - d\varepsilon$, it holds that P'_{A} and P_{B} are consistent (where P'_{A} is the partial function of the view of the fake Alice V'_{A} sampled by Eve). This means that there is a full oracle h that is consistent with both of $V'_{\mathsf{A}}, V_{\mathsf{B}}$. Then, by perfect completeness, this means the key $k = k_{\mathsf{B}}$ for Bob should match the key $k_E = k'_{\mathsf{A}}$ output by Eve.

Case of Classical Alice and Quantum Bob. Here we describe what steps would be different when attacking protocols with a quantum Bob (but still classical Alice). Interestingly, the attack description remains exactly the same as before. First note that, because Alice is classical it is well-defined to talk about whether $x \in Q_{\mathsf{A}}$ or not at the end of the protocol as once a query is asked by Alice it would belong to Q_{A} forever.[9] The efficiency analysis of the attack also remains the same as the CACB case above. Below, we describe the key differences in the analysis of the success of the attack.

Success of the attack. At a high level, we prove quantum variants for both of the Independence and Consistency properties.

- Independence: We show that, even if Alice and Bob are both quantum, then their "views" (i.e., the measurement of their registers) would be independent conditioned on the fixed classical transcript t and oracle h. More generally, we show that the *joint quantum state* of Alice and Bob, conditioned on h, t is a *product state*.
- Consistency: Again, we first prove a result that applies to the more general case of *two* quantum parties. We start by using two ideas that were popularized following the breakthrough work of Zhandry [Zha19]. First, we use a purified quantum random oracle h that is in the uniform superposition over all possible classical oracles (which is equivalent to using a classical random oracle). Second, we represent the oracle's answers in the Fourier domain, and denote the oracle $\hat{h}$.

[8] In [IR89, BM17], this condition is referred to as having no "intersection queries" outside L.

[9] One cannot say the same thing for quantum algorithm Bob, as it might choose to "forget" things about oracle as it proceeds.

We show that if parties ask a total of d queries to the oracle, then the joint quantum state $|\phi\rangle$ that describes both Alice's and Bob's registers W and the oracle $\hat{h}$ (using registers H) is "d-sparse" over its oracle part H, in the sense that $\hat{h}$ can be represented with a degree d multi-linear polynomial f with variables $x_i, i \in [2^n]$. Finally, we show that in the case when Alice is classical, then if Alice's fake queries Q'_{A} do not intersect with the "queries" in $\mathcal{S}$, where $\mathcal{S}$ is a (maximal) monomial $\prod_{i\in\mathcal{S}} x_i$ in f of $\deg(f)$, then there exists an oracle h such that (1) h is consistent with the real views $|\phi\rangle$, and (2) h is also consistent with Alice's fake view V'_{A}.

The above generalization of the Consistency condition allows us to now basically apply the same argument used in the CACB case by treating the variables in the maximal monomial $\mathcal{S}$ as Bob's queries. In particular, once $Q'_{\mathsf{A}} \cap \mathcal{S} = \emptyset$, then we conclude that there is an oracle h that is consistent with each of V'_{A} and the the real (quantum) Alice and Bob. Then, by the Independence property, h is consistent with V'_{A} and real Bob *at the same time*, and hence by perfect completeness the key implied by Alice's fake view V'_{A} sampled by Eve shall match that of the real Bob.

Case of Quantum Alice and Quantum Bob. When it comes to the case of quantum Alice and Bob, we can no longer use the classical attack of the CACB setting, as both Alice and Bob can now ask *superposition* queries to the oracle (e.g., all of their queries might have non-zero amplitude for *all* possible oracle queries). Hence, we need to change the attack and its analysis. In this case, without loss of generality, we focus on the simpler case that the key k is a bit.

Description of the attack. In the previous case of CAQB, we described how we choose to represent the (now quantum) random oracle $\hat{h}$ in the Fourier domain. Roughly speaking, in the Fourier domain, an oracle answer $\hat{0}$ to a query x, means that it has uniform distribution (when measured in the computational basis), and any other answer $\hat{y} \neq \hat{0}$ refers to non-uniform answers. Therefore, a "non-uniform" $\hat{y} \neq \hat{0}$ answer here means that either Alice or Bob have (at least partially) "read" the answer to x at some point. More precisely, conditioned on all Eve knows, let p_x be the probability that after measuring the answer to the query x in the Fourier basis, we obtain an answer other than $\hat{0}$. Then, informally speaking, we interpret p_x as the "probability that either Alice or Bob has read x from the oracle". In that case, if $p_x \geq \varepsilon$, then Eve will call x *quantum ε-heavy*. In the new attack, Eve goes ahead and asks any (classical query) x that is quantum ε-heavy (under the new definition) and updates L as before. When no "quantum ε-heavy query" is left, Eve outputs the more likely key $k \in \{0, 1\}$.

Efficiency. We generalize the efficiency argument for the classical case to the quantum regime. Namely, if Alice and Bob ask a total of d queries to the oracle, then the quantum ε-heavy learner Eve will stop after asking $|L|$ queries, where we have $\mathbb{E}[|L|] \leq d/\varepsilon$.

Success of the attack. Our goal is to show that once no quantum ε-heavy query is left, then conditioned on Eve's knowledge (t, L), at least one of the possible keys $k \in \{0, 1\}$ is much more likely to be the key chosen by Alice and bob. In that case, Eve will indeed succeed in finding the true key with high probability. For sake of contradiction, suppose after learning L and conditioned on (t, L) both values of $k \in \{0, 1\}$ have probabilities $\approx 1/2$. We would like to show that this situation violates perfect completeness. As explained in the previous case of CAQB, once we view the oracle $\hat{h}$ in the Fourier domain, after Alice and Bob ask d oracle queries, the joint state of the oracle and the registers

of Alice and Bob corresponds to a *distribution* F over degree-d multi-linear polynomials like f. The distribution is obtained by measuring the work registers of Alice and Bob.[10] Below, we further analyze this distribution over low-degree polynomials, while for simplicity we assume that we deal with *one* fixed polynomial f.

Because at the end of the attack Eve has learned all the quantum ε-heavy queries of the oracle, it can be shown that any unlearned query x, which corresponds to a variable in the polynomial f, has influence (as defined prior to Conjecture 1.2) at most ε. Putting things together, the polynomial f has the following properties: (1) f has ℓ_2 norm 1, because of representing a quantum state, (2) f has degree d, and (3) the influence of every variable in f is bounded by ε. Furthermore, if we let f_b be the polynomials that represent the "conditional states" of the oracle and Alice-Bob registers *conditioned* on the key being $k = b$, then by the fact that the key k is still unbiased (in Eve's view) we can conclude that f_0, f_1 both essentially inherit all the properties of f (the only difference being that the influences increase to $\approx 2\varepsilon$ instead of ε).

Our Conjecture 1.2 states that when ε is sufficiently small, any two polynomials f_0, f_1 with properties stated above would have a nonzero product. This implies that there exists an oracle h that is consistent with two very different executions with two outcomes for the final key. By the Independence property, we can now choose Alice's view from the execution leading to the key 0 and choose Bob's view from the execution leading to key 1, but this violates the perfect completeness.

Obtaining exponentially small influences. To prove the weaker variant of Conjecture 1.2 where the influences are less than $2^{-d}/d$ rather than the desired $1/\operatorname{poly}(d)$, the high level idea is as follows. Take any maximum-degree term appearing in f, and consider what happens when we fix all variables except the $\leq d$ ones in the term. Clearly, the resulting restriction of f is not a constant function so there is always some assignment to the remaining d variables that makes f non-zero, regardless of how the first variables were fixed. We show that, if g has all influences less than $2^{-d}/d$ then there is some assignment to the variables outside the term such that g is non-zero for all assignments to the remaining d variables, yielding an $\mathbf{x}$ such that $f(\mathbf{x}) \cdot g(\mathbf{x}) \neq 0$. To prove this property of g, we show that in expectation over a random assignment of the variables outside the term, the resulting restriction of g has a constant term that dominates all the non-constant terms. The exponential loss of 2^d essentially comes from the fact that there are 2^d non-constant terms in this restriction of g.

Ideas behind Theorem 1.4. We now sketch some of the ideas behind the proof of Theorem 1.4. We start by assuming that Q is a quantum algorithm accessing a random oracle h that asks $\operatorname{poly}(\kappa)$ queries, while there is $\varepsilon = 1/\operatorname{poly}(\kappa)$ such that any $\operatorname{poly}(\kappa)$-query classical algorithm will fail to approximate $\Pr[Q^h = 1]$ within $\pm\varepsilon$ additive error for *at least* ε fraction of the sampled random oracles h. Note that even though a classical algorithm cannot do so, a quantum algorithm (e.g., Alice or Bob) can indeed approximate $\Pr[Q^h]$ within an arbitrarily small additive error $\delta = 1/\operatorname{poly}(\kappa)$. As a result, quantum Alice and Bob can access the "same" number (approximately) that is, at least sometimes, not as accessible by the classical Eve. Therefore, roughly speaking, the quantum

[10] As expected, the formulation of our Polynomial Compatibility Conjecture is such that, to use the conjecture for obtaining attacks, it does not matter in which basis the work registers of Alice and Bob are measured.

parties can leverage on this "source of shared unpredictable" numbers and bootstrap it to a full fledged key agreement that is secure against classical Eve in the QROM.

In more detail, we first show that the above argument leads to a "weak" key agreement such that the key cannot be guessed with probability $1-\delta$ for some $\delta = 1/\operatorname{poly}(\kappa)$. We then use a careful number of repetitions to agree on a longer key that is much harder for the adversary to guess. The proof of this steps relies on the fact that *concurrent* composition of interactive *proofs* (rather than arguments) decrease the soundness error optimally. Then, one approach is to use the Goldreich-Levin technique to extract a uniform key from the "unpredictable key", and then bootstrap the completeness to $1 - \operatorname{negl}(\kappa)$ using the amplification technique of Holenstein [Hol05]. More conveniently, we use a tool from the recent work of Haitner et al. [HMST21] that combines the last two steps.

Complexity of our attacks. When one aims to use *only* a random oracle for security, then it means that the security is defined based on the number of adversary queries, regardless of how computationally hard it is to run such attacks. Indeed, if one adds computational intractability assumptions, one can ignore the random oracle all together and run a computationally secure protocol. In this work, we also primarily focus on studying the feasibility of key agreements from quantum random oracles in the QCCC model, while the implications to *fully* black-box separations are also discussed in Section 6 of the full version [ACC+22]. For sake of completeness, here we also comment on the computational complexity of our attack. In the classical setting, an $\mathbf{NP}$ oracle can be used to "uniformly invert" efficient processes that do not use an $\mathbf{NP}$ oracle themselves [BGP00]. This allows the adversary Eve to find the heavy queries, as needed in the attack of [BMG09b], through repeated sampling of the views conditioned on the transcript.[11]). In the quantum setting, we can use a "post-selection" gate [Aar05] to do the same thing. More formally, first we observe that Zhandry's compressed oracle lets us efficiently simulate the quantum random oracle while we maintain the "sampled oracle answers" in the Fourier basis using a list of polynomial size. Then, using post-selection one can sample oracle queries that are queried conditioned on the given transcript. Finally, by repeated sampling, we can again efficiently find the heavy queries.

1.3 Related Work

Black-box separations. Impagliazzo and Rudich [IR89] initiated the field of "black-box separations" by proving the existence of an oracle relative which one-way functions exist but secure key agreement protocols do not. The notions of black-box reductions, in various forms, were later formalized by Reingold, Trevisan, and Vadhan [RTV04].

Quantum black-box separations. The work of Hosoyamada and Yamakawa [HY20] initiated the study of "quantum black-box" separations by formalizing the notion of quantum black-box constructions (for primitives with non-interactive adversaries) and showing that even quantum black-box constructions cannot base collision resistant hash functions on one-way functions. Their work extended the previous result of Haitner et al. [HHRS07] about classical constructions to the quantum setting. Cao and Xue [CX21] proved quantum black-box separation of one-way permutations from one-way functions. Their work extended the previous result of Rudich [Rud88] and Kahn et

[11] See Remark 3.2 in https://www.boazbarak.org/Papers/merkle.pdf.

al. [KSS00] about classical constructions and classical security proofs, to the setting of allowing quantum reductions of security.

The QCCC model. The model of classical communication and quantum-powered parties is also used in other lines of work. One such recent body of work aims to classically verify a quantum computation [Mah18, CCY20, ACGH20, BKVV20, Zha21, Bar21]. More generally, an active line of work aims for designing on post-quantum security (e.g., see the recent works [BS20, BKS21, ABG+21, ACP21]) in which we deal with quantum powered adversaries, while the honest parties are fully classical. However, in our setting, honest parties are also quantum powered.

Limitations of random oracles. Haitner et al. [HOZ16], and Mahmoody et al. [MMP14] studied the limitations of using random oracles for *secure multiparty computation*. It was shown in [HOZ16] that inputless functionalities cannot rely on ROM to get security (unless they are trivially possible). [MMP14] showed that non-trivial and non-complete two-party functionalities cannot be based on random oracles. The work of Haitner et al. [HMO+21] studies the *communication* complexity of key agreement from random oracles. It is interesting to see whether similar lower bounds on the communication complexity of key agreement hold in the *Q*ROM model.

Comparison with the Aaronson-Ambainis Conjecture. As mentioned above, Aaronson Ambainis [AA09] proved that if a Fourier-theoretic conjecture, with resemblance to our Polynomial Compatibility Conjecture holds, then the Simulation Conjecture holds as well. The AA Conjecture states that any *bounded* degree d polynomial $f : \{-1,1\}^n \to [0,1]$ with variance ε has a variable with influence at least $\mathrm{poly}(\varepsilon/d)$. In a language closer to our Polynomial Compatibility Conjecture, the contrapositive of the AA Conjecture says that for any degree d polynomial f with constant variance and polynomially small influences $\mathrm{poly}(\mathrm{Var}[f]/d)$, there must exist an $\mathbf{x} \in \{0,1\}^n$ such that $|f(\mathbf{x})| > 1$. One interesting similarity is that both conjectures hold, when we assume *exponentially* small influences [DFKO06]. Despite that, our conjecture and the AA conjecture do not seem to be directly comparable, and it would be interesting to prove implications in either direction between them. For the application to key agreements, the implications of the two conjectures also seem incomparable. Our conjecture is tailored for perfect completeness and can be applied when there *is* communication. On the contrary, the AA conjecture can be applied to give an attack in the setting of imperfect completeness, but (as far as we can see) it is limited to the case of no communication. Furthermore, the "intersection" of these, i.e., the case of no communication and perfect completeness, can be proved without a conjecture [OSSS05].

2 Preliminaries and Notation

2.1 Quantum Computation

Let Σ be a finite and nonempty set of classical states. The finite dimensional Hilbert space associated with a *register* X is defined to be $\mathbb{C}^{|\Sigma|}$ for Σ being the state set of X. A *quantum state* of a register X is a unit vector in $\mathbb{C}^{|\Sigma|}$. We use standard bra-ket

notation for vectors and their adjoint. That is, we can write $|\psi\rangle_X \in \mathbb{C}^{|\Sigma|}$ as a vector

$$|\psi\rangle_X = \sum_{i\in\Sigma} \alpha_i |i\rangle_X,$$

where $\sum_{i\in\Sigma} |\alpha_i|^2 = 1$, and $\{|i\rangle\}_{i\in\Sigma}$ is an orthonormal basis of $\mathbb{C}^{|\Sigma|}$. We define $\langle\psi|_X$ as the row vector that is conjugate to $|\psi\rangle_X$. The inner product between $|\phi\rangle_X$ and $|\psi\rangle_X$ is denoted by $\langle\phi|\psi\rangle_X$. We sometimes neglect the subscripts when the corresponding registers are clear form the context.

For combined registers $Y = (X_1, \ldots, X_n)$, where Σ_i is the state set for each X_i, the state set of Y is defined as $\Sigma = \Sigma_1 \times \cdots \times \Sigma_n$. The finite dimensional Hilbert space associated with Y is defined to be $\mathbb{C}^{|\Sigma_1|} \otimes \cdots \otimes \mathbb{C}^{|\Sigma_n|}$. Since every register is labeled by a distinct name, we sometimes permute the order of tensor product for ease of expression. A quantum state $|\psi\rangle_{AB}$ over registers A, B is called a *product state* if and only if it can be written as $|\psi\rangle_{AB} = |\phi_1\rangle_A \otimes |\phi_2\rangle_B$.

The evolution of a quantum state $|\psi\rangle \in \mathbb{C}^{|\Sigma|}$ is governed by a unitary operator $U : \mathbb{C}^{|\Sigma|} \to \mathbb{C}^{|\Sigma|}$. The state becomes $|\psi'\rangle = U|\psi\rangle$. The measurement operator corresponding to a finite nonempty set of outcomes Γ is a set of operators $\{M_i\}_{i\in\Gamma}$ which satisfies $\sum_{i\in\Gamma} {M_i}^\dagger M_i = I$, where $(\cdot)^\dagger$ denotes Hermitian conjugation and I is the identity operator. The probability of obtaining i by measuring $|\psi\rangle$ is $\|M_i|\psi\rangle\|_2^2$, and the post-measurement state then collapses to $\frac{M_i|\psi\rangle}{\|M_i|\psi\rangle\|_2}$, where $\|\cdot\|_2$ denotes the Euclidean norm. An operator $\Pi_X : \mathbb{C}^{|\Sigma|} \to \mathbb{C}^{|\Sigma|}$ is called a projection operator (or projector) if it satisfies $\Pi_X^2 = \Pi_X$. For projection operators acting on multiple registers of the form $\Pi_{X_1X_2} = \Pi_{X_1} \otimes I_{X_2}$, we write only the non-trivial part Π_{X_1} for convenience. We say an operator A commutes with another operator B if $AB = BA$.

A quantum circuit consists of registers, unitary gates and measurements. By the deferred measurement principle, all intermediate measurements can be delayed at the end of the circuit by introducing ancillary registers. Without loss of generality, we assume that at the end all the registers are measured in the computational basis. Indeed (efficient) classical algorithms can be simulated using quantum circuits (efficiently).

Some of the components of our analysis rely on ideas inspired by the Compressed Oracle technique of Zhandry [Zha19]. The following preliminary follows closely to the formalization in Section 3 of [CFHL21].

The computational and the Fourier bases. Let $\mathcal{Y}$ be a finite Abelian group of cardinality $|\mathcal{Y}|$. Let $\{|y\rangle\}_{y\in\mathcal{Y}}$ be an orthonormal basis of $\mathbb{C}^{|\mathcal{Y}|}$, where the basis vectors are labeled by the elements of $\mathcal{Y}$. We refer to this basis as the *computational basis*. Let $\hat{\mathcal{Y}}$ be the dual group of $\mathcal{Y}$, which consists of all group homomorphisms $\mathcal{Y} \to \{\omega \in \mathbb{C} \mid |\omega| = 1\}$ and is known to be isomorphic to $\mathcal{Y}$, and thus to have cardinality $|\mathcal{Y}|$ as well.[12] We consider $\hat{\mathcal{Y}}$ to be an additive group; the neutral element is denoted $\hat{0}$. The *Fourier basis* $\{|\hat{y}\rangle\}_{\hat{y}\in\hat{\mathcal{Y}}}$ of $\mathbb{C}^{|\mathcal{Y}|}$ is defined by the transformations below, where $(\cdot)^*$ is complex conjugation.

$$|\hat{y}\rangle = \frac{1}{\sqrt{|\mathcal{Y}|}} \sum_{y\in\mathcal{Y}} \hat{y}(y)^* |y\rangle \qquad\qquad |y\rangle = \frac{1}{\sqrt{|\mathcal{Y}|}} \sum_{\hat{y}\in\hat{\mathcal{Y}}} \hat{y}(y) |\hat{y}\rangle.$$

[12] We do not rely on $\hat{\mathcal{Y}}$ and $\mathcal{Y}$ being isomorphic and think of them simply as disjoint sets.

An elementary property of the Fourier basis is the following.

Fact 2.1 *The operator defined by* $|y\rangle|y'\rangle \mapsto |y+y'\rangle|y'\rangle$ *for all* $y, y' \in \mathcal{Y}$ *is the same as the operator defined by* $|\hat{y}\rangle|\hat{y}'\rangle \mapsto |\hat{y}\rangle|\hat{y}' - \hat{y}\rangle$ *for all* $\hat{y}, \hat{y}' \in \hat{\mathcal{Y}}$.

Functions and their (quantum) representations. Let $\mathcal{H}$ be the set of all functions $h : \mathcal{X} \to \mathcal{Y}$ and $\hat{\mathcal{H}}$ be the set of all functions $\hat{h} : \mathcal{X} \to \hat{\mathcal{Y}}$. For any $h \in \mathcal{H}$, we define its *quantum representation* to be $|h\rangle_H := \bigotimes_{x\in\mathcal{X}} |h(x)\rangle_{H_x}$ in the computational basis, where the register H_x is associated with $\mathbb{C}^{|\mathcal{Y}|}$ for all $x \in \mathcal{X}$, and the register H is compounded of all H_x. One can view $|h\rangle_H$ as the vector representing the truth table of h. Similarly, for any $\hat{h} \in \hat{\mathcal{H}}$ we define $|\hat{h}\rangle_H := \bigotimes_{x\in\mathcal{X}} |\hat{h}(x)\rangle_{H_x}$ in the Fourier basis. Both $\{|h\rangle_H\}_{h\in\mathcal{H}}$ and $\{|\hat{h}\rangle_H\}_{\hat{h}\in\hat{\mathcal{H}}}$ are orthonormal bases of $\mathbb{C}^{|\mathcal{Y}|^{|\mathcal{X}|}}$.

Superposition oracle. In the quantum random oracle model, an oracle-aided quantum algorithm A consists of the *query* register X, the *answer* register Y and ancillary register Z. For convenience, we let $W := (X, Y, Z)$ denote the *internal registers* of A. Initially, a function $h : \mathcal{X} \to \mathcal{Y}$ is sampled from $\mathcal{H}$ uniformly at random, and A begins with the state $|0\rangle_W$. The algorithm A is able to ask adaptive quantum queries. Between the queries, A can apply unitaries and perform measurements on its registers. The query operation O is defined as the following unitary mapping in the computational basis.

$$|x\rangle_X|y\rangle_Y|h\rangle_H \mapsto |x\rangle_X|y + h(x)\rangle_Y|h\rangle_H$$

Since quantum operators are reversible, we assume the algorithm has access to $O^\dagger$ as well. By default, O acts as identity on registers other than X, Y and H.

We define the quantum state $|\Phi_0\rangle_H$ to be a uniform superposition over all $h \in \mathcal{H}$

$$|\Phi_0\rangle_H := \sum_{h\in\mathcal{H}} \frac{1}{\sqrt{|\mathcal{H}|}} |h\rangle_H = \bigotimes_{x\in\mathcal{X}} |\hat{0}\rangle_{H_x}. \tag{1}$$

The sampling of h is equivalent to measuring $|\Phi_0\rangle_H$ in the computational basis. Since the unitary operators and measurements performed by A commute with the measurement on $|\Phi_0\rangle_H$, and the fact that registers in H are used only as control-bits for O, we can delay the measurement on $|\Phi_0\rangle_H$ to the end of the computation.

Now, we analyze the behavior of the superposition oracle in the Fourier basis. By Fact 2.1, O becomes

$$|x\rangle_X|\hat{y}\rangle_Y|\hat{h}\rangle_H \mapsto |x\rangle_X|\hat{y}\rangle_Y \bigotimes_{x'\in\mathcal{X}} |\hat{h}(x') - \hat{y}\cdot\delta_{x,x'}\rangle_{H_{x'}} \tag{2}$$

in the Fourier basis, where $\delta_{x,x'} = 1$ when $x = x'$ and $\delta_{x,x'} = 0$ otherwise.

2.2 Key Agreement Using Quantum Computation and Classical Communication

A key agreement protocol in the Quantum-Computation Classical-Computation (QCCC) model is a protocol in which two quantum algorithms, Alice and Bob, can query the oracle, apply quantum operation on their internal registers, and send classical strings over the public channel to the other party. We also refer to this model as the *Quantum-Alice*

Quantum-Bob model. The sequence of the strings sent during the protocol is called the *transcript* of the protocol. Let W_A and W_B be Alice's and Bob's internal registers, respectively. Before the protocol starts, an oracle function h is chosen from $\mathcal{H}$ uniformly at random, and query operation O_h given the oracle h is defined as

$$O_h : |x\rangle|y\rangle \mapsto |x\rangle|y + h(x)\rangle.$$

When we consider the case that Alice and Bob are both quantum algorithms, they start with a product state $|0\rangle_{W_A} \otimes |0\rangle_{W_B}$. When Alice is a classical algorithm and Bob is a quantum algorithm, Alice is given a random tape at the beginning. That is, Alice and Bob start with a product state $|r_A\rangle_{W_A} \otimes |0\rangle_{W_B}$, where $r_A \in \{0,1\}^*$ is uniform.

Apart from the *real* execution, we can take not only W_A, W_B but also the oracle register H initialized as $|\Phi_0\rangle_H$ into account. As we mentioned, the sampling of h can be postponed at the end. Additionally, by the deferred measurement principle, all the intermediate measurements can be delayed as well. Now, the joint state of W_A, W_B and H remains as a *pure* state during the protocol. Importantly, such a switching of viewpoints could display several non-trivial properties providing better leverage while still being perfectly indistinguishable from the previous one. Therefore, the analysis will be done in the so-called *purified view* in the following sections. In other word, whenever any classical information appears, the joint state collapses to the corresponding post-measurement state and stays pure. For any key agreement protocol, we define its *purified version* as follows:

- Start with $|0\rangle_{W_A}|0\rangle_{W_B}|\Phi_0\rangle_H$.
- Alice and Bob runs the protocol in superposition, that is, all the measurements (including those used for generating the transcript[13]) are delayed and the query operator O_h is replaced by O.
- Let $|\Psi\rangle_{W_A W_B H}$ denote the state at the end of the protocol and $|\Psi_t\rangle_{W_A W_B H}$ be its post-measurement that is consistent with the transcript t.

Definition 2.2 (Nonzero queries in Fourier basis). *For any $\hat{h} \in \hat{\mathcal{H}}$, we define the set*

$$Q_{\hat{h}} \coloneqq \{x : x \in \mathcal{X}, \hat{h}(x) \neq \hat{0}\}$$

and the size of $\hat{h}$ by

$$|\hat{h}| \coloneqq |\{x : x \in \mathcal{X}, \hat{h}(x) \neq \hat{0}\}| = |Q_{\hat{h}}|.$$

Definition 2.3 (Oracle support). *For any vector $|\phi\rangle_{WH} = \sum_{w,\hat{h}\in\hat{\mathcal{H}}} \alpha_{w,\hat{h}}|w\rangle_W|\hat{h}\rangle_{\hat{H}}$, we define the* oracle support in the Fourier basis *of $|\phi\rangle$ as*

$$\widehat{\mathrm{supp}}^H(|\phi\rangle) \coloneqq \{\hat{h} : \exists w \text{ s.t. } \alpha_{w,\hat{h}} \neq 0\}.$$

[13] By delaying the measurement for the transcript, one can view it as applying an CNOT gate, where the controlled bit is the register that supposed to sent and the target bit is an ancilla. Then, one sends the ancilla bit, and in the rest of the computation, the ancilla bits are served only as control bits for Alice's and Bob's computation. The ancilla bits (transcript) remain unchanged throughout the computation. Thus, it is equivalent to sending classical information, and it is consistent with QCCC model.

We denote the largest $\hat{h}$ in $\widehat{\mathrm{supp}}^H(|\phi\rangle)$ as

$$\hat{h}^H_{\max}(|\phi\rangle) \coloneqq \operatorname*{arg\,max}_{\hat{h}\in\widehat{\mathrm{supp}}^H(|\phi\rangle)} |\hat{h}|.$$

(If the choice is not unique, then choose the alphabetically first one.) When the oracle registers H are clear, we simply denote this by $\hat{h}_{\max}(|\phi\rangle)$. Similarly, if we write the oracle part in the computational basis $|\phi\rangle_{WH} = \sum_{w,h\in\mathcal{H}} \beta_{w,h}|w\rangle_W|h\rangle_H$, then we define the oracle support in the computational basis *of $|\phi\rangle$ as*

$$\mathrm{supp}^H(|\phi\rangle) \coloneqq \{h : \exists w \text{ s.t. } \beta_{w,h} \neq 0\}.$$

Lemma 2.4 (Sparse representation). *If* A *asks at most d queries to the superposition oracle, then for all possible outcomes of* A*'s intermediate measurements, the joint state $|\psi\rangle_{WH}$ conditioned on the outcome satisfies $|\hat{h}_{\max}(|\psi\rangle)| \leq d$.*

Proof. We prove the lemma by induction on the number of queries asked by A, denoted by q. For the base case $q = 0$, the joint state $|\psi_0\rangle_{WH} = |0\rangle_W|\Phi_0\rangle_H$ satisfies the statement. Assume that the joint state $|\psi_k\rangle_{WH}$ satisfies $|\hat{h}_{\max}(|\psi_k\rangle)| \leq k$ for some k.

For the induction step, since the unitaries and measurements act only on W, the size of the state never increases. Moreover, for every $x \in \mathcal{X}, \hat{y} \in \hat{\mathcal{Y}}$ and $\hat{h} \in \hat{\mathcal{H}}$, by the observation in Equation (2), the size of $\hat{h}$ increases at most by one after the query operation. Therefore, the size of the state increases at most by one. By induction hypothesis the resulting state $|\psi_{k+1}\rangle_{WH}$ satisfies $|\hat{h}_{\max}(|\psi_{k+1}\rangle)| \leq k+1$. □

Definition 2.5. *A partial oracle L is a partial function from $\mathcal{X}$ to $\mathcal{Y}$. The domain of L is denoted by $Q_L = \mathrm{dom}(L)$. Equivalently, we view L as a finite set of pairs $(x, y_x) \in \mathcal{X}\times\mathcal{Y}$ such that for all $(x, y_x), (x', y'_x) \in L$, $x \neq x'$.*

Note that our partial oracles are always in the computational basis. We say a partial oracle L is consistent with $h : \mathcal{X} \to \mathcal{Y}$ if and only if $h(x) = y_x$ holds for all $x \in Q_L$.

Definition 2.6. *For any partial oracle L, we define the associated projector Π_L by*

$$\Pi_L \coloneqq \bigotimes_{x\in Q_L} |y_x\rangle\langle y_x|_{H_x} \bigotimes_{x\notin Q_L} I_{H_x},$$

where I_{H_x} is the identity operator acting on H_x. It holds that $\Pi_L|h\rangle_H = |h\rangle_H$ if h is consistent with L, and $\Pi_L|h\rangle_H = 0$ otherwise.

Lemma 2.7. *Given a state $|\psi\rangle_{WH}$ and a partial oracle L, the state $\Pi_L|\psi\rangle_{WH}$ can be written as*

$$\Pi_L|\psi\rangle_{WH} = \sum_{w\in\mathcal{W},\hat{h}\in\hat{\mathcal{H}}'} \alpha'_{w,\hat{h}}|w\rangle_W \bigotimes_{x\notin Q_L} |\hat{h}(x)\rangle_{H_x} \bigotimes_{x\in Q_L} |y_x\rangle_{H_x},$$

where $\hat{\mathcal{H}}'$ is the set of functions from $\mathcal{X}\setminus Q_L$ to $\hat{\mathcal{Y}}$. Furthermore, if $|\hat{h}^H_{\max}(|\psi\rangle)| \leq d$, then $|\hat{h}^{H'}_{\max}(\Pi_L|\psi\rangle)| \leq d$, where H' is the set of registers corresponding to $\mathcal{X}\setminus Q_L$.

3 Attacking Classical-Alice Quantum-Bob Protocols

In this section, we consider the case where A is a classical algorithm and B is a quantum algorithm and prove the following theorem.

Theorem 3.1 (Breaking CAQB protocols). *Let (A, B) be a two-party protocol in which algorithm classical A communicates with a quantum algorithm B and they both have access to a random oracle $h\colon \mathcal{X} \to \mathcal{Y}$, and at the end they agree on a key k with probability 1. Suppose Alice asks at most d_{A} classical oracle queries, while Bob asks at most d_{B} quantum oracle queries. Then, there is an eavesdropper E who, after receiving the transcript t, asks at most $d_{\mathsf{A}} \cdot d_{\mathsf{B}}/\lambda$ queries to h after receiving the classical transcript t and finds the key k with probability $1 - \lambda$.*

Note that in the above theorem, the adversary's query complexity is $d_{\mathsf{A}} \cdot d_{\mathsf{B}}/\lambda$ rather than the simpler (still correct) bound of d^2/λ where $d = d_{\mathsf{A}} + d_{\mathsf{B}}$. Even though, when $d_{\mathsf{A}} = \Theta(d_{\mathsf{B}})$, it also holds that $d_{\mathsf{A}} \cdot d_{\mathsf{B}} = \Theta(d^2)$, when the query complexity of the parties are unbalanced, e.g., when $d_{\mathsf{A}} = \sqrt{\kappa}, d = \kappa$ for security parameter κ, our attacker's query complexity will be $O(\kappa^{1.5})$ rather than $O(\kappa^2)$. This is particularly a natural scenario when the quantum-powered party is more powerful and can ask many more queries. Later on, we will give a concrete construction of the adversary (Theorem 3.5) in the proof. Notice that the adversary is actually a classical algorithm, where it only makes classical queries.

The rest of this section will be dedicated to proving the theorem. Before constructing the attacker and analyzing it, we introduce some useful lemmas.

3.1 Useful Lemmas

Lemma 3.2 (Independence of quantum views in the QCCC model). *Suppose two quantum algorithms A and B interact classically in the quantum random oracle model. Let W_{A} and W_{B} denote their registers respectively. Then, at any time during the protocol, conditioned on the transcript t and the fixed oracle $h \in \mathcal{H}$, the joint state of the registers W_{A} and W_{B} conditioned on t and h is a product state.*

Proof. We prove the lemma by induction on the round index r. For the base case $r = 0$, A and B's joint state $|0\rangle_{W_{\mathsf{A}}} \otimes |0\rangle_{W_{\mathsf{B}}}$. Suppose for some k, A and B's joint state after k rounds is a product state conditioned on the transcript t and oracle h. For the induction step, in the $(k+1)$-th round, one of them will apply "deterministic" local unitaries and query operators O_h conditioned on t and h. Therefore, further conditioned on the message generated in this round, the resulting joint state is still a product state. □

Lemma 3.3 (Consistency). *Given a state $|\psi\rangle_H$, if L is a partial oracle such that $Q_{\hat{h}_{\max}(|\psi\rangle)} \cap Q_L = \emptyset$, then $\|\Pi_L|\psi\rangle\|_2^2 > 0$. Equivalently, there exists at least one oracle $h \in \mathcal{H}$ such that (i) h is consistent with L and (ii) $h \in \mathrm{supp}^H(|\psi\rangle)$.*

Proof. For convenience, we write $\hat{h}_{\max}$ to denote $\hat{h}_{\max}(|\psi\rangle)$, and we represent $|\psi\rangle_H = \sum_{\hat{h}} \gamma_{\hat{h}} |\hat{h}\rangle$ in the Fourier basis. The proof directly comes from the following two claims:

Claim. $\gamma_{\hat{h}_{\max}} \Pi_L |\hat{h}_{\max}\rangle$ is not a zero vector.

Proof of Section 3.1. Since $Q_{\hat{h}_{\max}} \cap Q_L = \emptyset$ and $\gamma_{\hat{h}_{\max}} \neq 0$ by definition, we have

$$\gamma_{\hat{h}_{\max}} \Pi_L |\hat{h}_{\max}\rangle = \frac{\gamma_{\hat{h}_{\max}}}{\sqrt{|\mathcal{Y}|^{|Q_L|}}} \bigotimes_{x \in Q_L} |y_x\rangle_{H_x} \bigotimes_{x \notin Q_L} |\hat{h}_{\max}(x)\rangle_{H_x},$$

which is not a zero vector. □

Claim. For all $\hat{h} \in \widehat{\text{supp}}^H(|\psi\rangle) \setminus \{\hat{h}_{\max}\}$, it holds that $\Pi_L|\hat{h}_{\max}\rangle$ is orthogonal to $\Pi_L|\hat{h}\rangle$. As a corollary, we have that $\gamma_{\hat{h}_{\max}} \Pi_L|\hat{h}_{\max}\rangle$ is orthogonal to $\sum_{\hat{h} \neq \hat{h}_{\max}} \gamma_{\hat{h}} \Pi_L |\hat{h}\rangle$ since the latter is a linear combination of vectors which are orthogonal to the former.

Proof of Section 3.1. Since $\hat{h}_{\max}$ is maximal and $Q_{\hat{h}_{\max}} \cap Q_L = \emptyset$, for all $\hat{h} \in \widehat{\text{supp}}^H(|\psi\rangle) \setminus \{\hat{h}_{\max}\}$, it holds that

$$|\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}_{\max}(x) \neq \hat{0}\}| \geq |\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}(x) \neq \hat{0}\}|.$$

For the case of $|\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}_{\max}(x) \neq \hat{0}\}| > |\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}(x) \neq \hat{0}\}|$, there exist an $x' \in \mathcal{X} \setminus Q_L$ such that $\hat{h}(x') = \hat{0}$ and $\hat{h}_{\max}(x') \neq \hat{0}$. Therefore, we have

$$\langle \hat{h} | \Pi_L | \hat{h}_{\max} \rangle = \bigotimes_{x \in Q_L} \langle \hat{h}(x) | y_x \rangle \langle y_x | \hat{h}(x) \rangle \bigotimes_{x \notin Q_L} \langle \hat{h}(x) | \hat{h}_{\max}(x) \rangle = 0,$$

since $\langle \hat{h}(x') | \hat{h}_{\max}(x') \rangle = 0$.

For the case of $|\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}_{\max}(x) \neq \hat{0}\}| = |\{x : x \in \mathcal{X} \setminus Q_L, \hat{h}(x) \neq \hat{0}\}|$, suppose there exists an $\hat{h}$ such that $\hat{h}(x) = \hat{h}_{\max}(x)$ holds for all $x \in \mathcal{X} \setminus Q_L$. There are two possible cases. First, For all $x \in Q_L$, it holds that $\hat{h}(x) = \hat{0}$. Because $Q_{\hat{h}_{\max}} \cap Q_L = \emptyset$, we have $\hat{h}_{\max}(x) = 0$ for all $x \in Q_L$. Consequently, we have $\hat{h} = \hat{h}_{\max}$ which contradicts to $\hat{h} \neq \hat{h}_{\max}$. Second, there exists $x \in Q_L$ such that $\hat{h}(x) \neq \hat{0}$. It implies $|\hat{h}| > |\hat{h}_{\max}|$ which contradicts to the maximal size of $\hat{h}_{\max}$. Therefore, for all $\hat{h}$ of the second case, there exists an $x' \in \mathcal{X} \setminus Q_L$ such that $\hat{h}(x') \neq \hat{h}_{\max}(x')$. It implies $\langle \hat{h} | \Pi_L | \hat{h}_{\max} \rangle = 0$. □

Finally, by Section 3.1 and Section 3.1 we can conclude that

$$\|\Pi_L|\psi\rangle\|_2^2 = \|\gamma_{\hat{h}_{\max}} \Pi_L |\hat{h}_{\max}\rangle\|_2^2 + \Big\| \sum_{\hat{h} \neq \hat{h}_{\max}} \gamma_{\hat{h}} \Pi_L |\hat{h}\rangle \Big\|_2^2 \geq \|\gamma_{\hat{h}_{\max}} \Pi_L |\hat{h}_{\max}\rangle\|_2^2 > 0.$$

□

The proof of the following lemma could be found in the full version [ACC+22].

Lemma 3.4 (Bounding the classical heavy queries). *Let Q be a random variable over subsets of universe $\mathcal{U}$. Suppose $z_1, x_1, z_2, x_2, \ldots$ is a finite sequence of random variables that are correlated with Q, and we have $x_i \in \mathcal{U} \cup \{\perp\}$ for all i. Suppose $x_i = x_j$ for $i \neq j$, then $x_i = x_j = \perp$. (Namely, no nontrivial x_i gets repeated). For a full sample $z_1, x_1, z_2, x_2, \ldots$, call x_i ε-*heavy *(conditioned on $z_1, x_1, \ldots, z_i$) if $\Pr[x_i \in Q \mid z_1, x_1, \ldots, z_i] \geq \varepsilon$, and for the same sequence, define $\mathcal{S} = \{x_i \mid x_i \text{ is } \varepsilon\text{-heavy}\}$. (Note that $\mathcal{S}$ is also a random variable correlated with Q.) Then, $\mathbb{E}[|\mathcal{S}|] \leq \mathbb{E}[|Q|]/\varepsilon$.*

3.2 The Attack and Its Analysis

Notation and basic notions. For a classical algorithm A (perhaps in a multi-party protocol) in an oracle model, we use $V_{\mathsf{A}} = (r_{\mathsf{A}}, t, P)$ to denote Alice's view in an execution, which consists of Alice's randomness r_{A}, the transcript t, and the partial oracle P of query-answer pairs that Alice encounters during her execution. By f_{A} we denote the function which takes V_{A} as input and outputs A's key k_{A}. We use $Q_{\mathsf{A}} = Q_P$ to refer to the set of queries asked by A. Given transcript t and some partial knowledge about the oracle h encoded by a partial oracle L, we call x an ε-heavy query for Alice (conditioned on (t, L)) if $\Pr[x \in Q_{\mathsf{A}} \mid t, L] \geq \varepsilon$, where the probability is over Alice's randomness and the oracle answers outside L.

Construction 3.5 (Attacking Classical-Alice Quantum-Bob protocols) *Let (A, B) be a key agreement protocol in which A (Alice) is classical and B (Bob) is quantum and they both have access to a random oracle h. Given the transcript t, the attacking algorithm E (Eve) is parameterized by ε and works as follows.*

- *Let $L = \emptyset$.*
- *While there is any ε-heavy query for Alice conditioned on (t, L), do the following.*
  - *Ask the lexicographically first ε-heavy query for Alice from the oracle h.*
  - *Update L by adding $(x, h(x))$ to L.*
- *Sample Alice's view V'_{A} conditioned on (t, L), and output the key $k'_{\mathsf{A}} = f_{\mathsf{A}}(V'_{\mathsf{A}})$.*

Lemma 3.6 (Efficiency). *The expected number of queries asked by Eve in Construction 3.5 is at most $d_{\mathsf{A}}/\varepsilon$, where d_{A} is the maximum number of queries asked by Alice.*

Proof. The proof is identical to the efficiency argument of the attack from [BM17]. More formally, we can use the abstract Lemma 3.4 to derive the claim by letting Q model Alice's set of queries, x_i be the ith query asked by E, and letting z_i be the information E receives about Q after asking x_{i-1}. In particular z_1 is the transcript, and z_i is the oracle answer to the query x_{i-1}, in case it is asked, and $x_j = \perp$ if no heavy query is left after asking x_i for $i < j$. In this case, all the queries Q_L asked by Eve E are ε-likely to be in Q_{A} conditioned on the transcript and the previously revealed information encoded in L, and so at the end we have $\mathbb{E}[|L|] \leq |d_{\mathsf{A}}|/\varepsilon$. □

Lemma 3.7 (Success). *If Alice and Bob, respectively, ask a total of $d_{\mathsf{A}}, d_{\mathsf{B}}$ oracle queries (where Bob's queries can be quantum queries) and agree on a key with probability 1, then Eve of Construction 3.5 outputs a key k_{E} such that $\Pr[k_{\mathsf{E}} = k] \geq 1 - \varepsilon d_{\mathsf{B}}$, where k is the key agreed by Alice and Bob.*

Proof. For the proof, we need to define a "quantum extension" of Alice's algorithm, which is denoted by QA. QA basically runs A by making "pure" quantum queries to the oracle h, and measuring Alice's quantum registers W_{A} would reveal the answers to the oracles queries of the original Alice who is emulated by QA.

Let QAB be the combined party of QA and B. Let W be all the registers of QA and B. Let $\mathcal{W}$ be the set of all possible outcomes of measuring registers W in the computational basis. Below, let $d = d_{\mathsf{A}} + d_{\mathsf{B}}$ be the total number of oracle queries.

For simplicity of presentation, we first give a proof with a looser probability $1 - \varepsilon d$ of finding the key. See the full version for the full proof for the tighter bound.

Loose analysis. Consider the purified version of the protocol execution, let $|\Psi_t\rangle_{WH}$ be the state conditioned on the transcript t. Since there is at most d queries in total, it holds that $|\hat{h}^H_{\max}(|\Psi_t\rangle)| \leq d$ by Lemma 2.4. Suppose the attacker E asks her queries from the oracle, starting from the transcript t, and obtains the partial oracle L where for every x asked by E we have $(x, y_x) \in L$. After she learns the first (x, y_x), the state becomes the post-measurement state corresponding to measuring $|\Psi_t\rangle_{WH}$ on register H_x with the outcome y_x. In this sense, for any t and L we can define the state conditioned on them, denoted by $|\Psi_{t,L}\rangle_{WH}$. Similarly, by Lemma 2.7 it holds that $|\hat{h}^{H'}_{\max}(|\Psi_{t,L}\rangle)| \leq d$. Since the oracle registers corresponding to Q_L are now measured, we can consider the "truncated" version of $|\Psi_{t,L}\rangle_{WH}$ by discarding those registers. Let H' be the set of remaining registers, that is, $H' = \{H_x\}_{x \in \mathcal{X} \setminus Q_L}$. By $|\Psi_{t,L}\rangle_{WH'}$ we denote the truncated $|\Psi_{t,L}\rangle_{WH}$. In the following analysis, we further assume that QAB measure the internal registers $W = (W_{\mathsf{A}}, W_{\mathsf{B}})$ at the end of the protocol and then obtain the outcome w in the computational basis. The resulting state is denoted by $|\Psi_{t,L,w}\rangle_{WH'}$. By Lemma 2.7, for any w it holds that $|\hat{h}^{H'}_{\max}(|\Psi_{t,L,w}\rangle)| \leq d$. In the following proof, we will show that for every (t, L, w), E will find the correct key in (t, L, w) with probability at least $1 - \varepsilon d$. From now on, we fix an arbitrary (t, L, w) and define $Q_{\max} := Q_{\hat{h}^{H'}_{\max}(|\Psi_{t,L,w}\rangle)}$.

Recall that Alice A was a classical algorithm and all the ε-heavy queries of A were already learned by the attacker E, and hence for any $x \notin Q_L$ we have $\Pr[x \in Q_{\mathsf{A}} \mid t, L] \leq \varepsilon$. In particular, this holds for every $x \in Q_{\max}$. Therefore, by a union bound, with probability at least $1 - \varepsilon|Q_{\max}| \geq 1 - \varepsilon d$, it holds that $Q'_{\mathsf{A}} \cap Q_{\max} = \emptyset$, where Q'_{A} is the set of queries in the fake view V'_{A} of Alice sampled by Eve. All we have to show is that for any Q'_{A} such that $Q'_{\mathsf{A}} \cap Q_{\max} = \emptyset$, it holds that Eve finds Bob's key: $f_{\mathsf{A}}(V'_{\mathsf{A}}) = k_{\mathsf{B}}$. (By perfect completeness, it also holds that $k_{\mathsf{B}} = k_{\mathsf{A}}$.)

Let P'_{A} be the set of query-answer pairs in the view V'_{A}. We now apply Lemma 3.3 with L and H in Lemma 3.3 set to be P'_{A} and H', respectively. Then, Lemma 3.3 shows that there exists an oracle $|h\rangle$ in the computational basis that is simultaneously consistent with L, t, P'_{A} (and hence Alice's fake view V'_{A}) and the measurements w of real Alice and Bob. Hence, we have the following:

- The probability of obtaining h as the oracle and V'_{A} as Alice's view is nonzero.
- The probability of obtaining h as the oracle and $w = w_{\mathsf{A}}, w_{\mathsf{B}}$ as the views of Alice and Bob is nonzero. In particular, the probably of obtaining (h, w_{B}) is nonzero.

By Lemma 3.2, we conclude that the probability of obtaining $(V'_{\mathsf{A}}, h, w_{\mathsf{B}})$ is nonzero. Then, by the perfect completeness, the key output by V'_{A} and w_{B} should be equal, and this finishes the proof of the weak bound, showing that Eve finds the key with probability $1 - \varepsilon d = 1 - \varepsilon(d_{\mathsf{A}} + d_{\mathsf{B}})$. □

4 Attacking Quantum-Alice Quantum-Bob Protocols

In this section, we consider the case where both A and B are quantum algorithms in the QCCC model. In this general setting, we show a *conditional result* based on a conjecture, that any QCCC key agreement protocol with perfect completeness can be broken with an expected polynomial number of queries. While we have so far been unable to prove the conjecture, we can prove a weaker version of the conjecture with exponentially worse parameters, which still leads to a non-trivial attack on QCCC key agreement

protocols. We present the conjecture and the variant that we can prove in Section 4.1. In Section 4.2, we state the main result, which gives an efficient attack when combined with the conjecture and a non-trivial attack when combined with the weak variant we can prove. In Section 4.3, we prove the necessary lemma for our main result.

4.1 Main Conjecture and Related Notions

Let $\mathcal{Y}$ be an Abelian group of order $|\mathcal{Y}|$ and $\hat{\mathcal{Y}}$ be the dual group. Let $\mathcal{H}$ be the set of all functions $h : \mathcal{X} \to \mathcal{Y}$ and $\hat{\mathcal{H}}$ be the set of all functions $\hat{h} : \mathcal{X} \to \hat{\mathcal{Y}}$.

Definition 4.1 (**$(\mathcal{Y}, \delta, d, N)$-state).** *Let H be a register over the Hilbert space $\mathcal{Y}^N$. A quantum state $|\psi\rangle$ over registers W and H is a $(\mathcal{Y}, \delta, d, N)$-state if it satisfies the following two conditions:*

- *d-**sparsity:** $|\hat{h}^H_{\max}(|\psi\rangle)| \leq d$.*
- *δ-**lightness:** For every $x \in \mathcal{X}$, if we measure the H_x register of $|\psi\rangle$ in the Fourier basis, the probability of getting $\hat{0}$ is at least $1 - \delta$.*

The first item above is equivalent to saying that for any measurement of registers H in the Fourier basis, and W in any basis, the oracle support in the Fourier basis (as defined in Definition 2.3) is at most d. Also, looking ahead, the second property above is equivalent to saying that $|\psi\rangle$ has no δ-heavy queries as defined in Definition 4.9.

Definition 4.2 (Compatibility). *Two quantum states $|\psi\rangle$ and $|\phi\rangle$ over registers W and H are* compatible *if $\mathrm{supp}^H(|\psi\rangle) \cap \mathrm{supp}^H(|\phi\rangle) \neq \emptyset$, i.e., if their oracle supports in the computational basis (as defined in Definition 2.3) have non-empty intersection.*

In general, we pose the following question. *How small should δ be, as a function of $|\mathcal{Y}|$ and d, in order to guarantee that any two $(\mathcal{Y}, \delta, d, N)$-states are compatible?* Our main conjecture is as follows.

Conjecture 4.3 *There exists a finite Abelian group $\mathcal{Y}$ and $\delta = 1/\operatorname{poly}(d)$ such that for any $d, N \in \mathbb{N}$, it holds that any two $(\mathcal{Y}, \delta(d), d, N)$-states $|\psi\rangle$ and $|\phi\rangle$ are compatible.*

Readers may notice that we introduce Conjecture 1.2 in terms of polynomials, while Conjecture 4.3 is formulated in terms of quantum states. In Section 5.1, we will show that two formulations are equivalent. We found that the one in quantum states is more natural to use, while the one in polynomials has a clearer mathematical statement.

While we do not have a proof of Conjecture 4.3, we can prove the following theorem when the influences are exponentially small. The proof is deferred to Section 5.2.

Theorem 4.4. *For all groups $\mathcal{Y}$, $d, N \in \mathbb{N}$, and $\delta < |\mathcal{Y}|^{-d}/d$, it holds that any two $(\mathcal{Y}, \delta, d, N)$-states $|\psi\rangle$ and $|\phi\rangle$ are compatible.*

4.2 Attacking Quantum-Alice Quantum-Bob Protocols

Now we are ready to state our main result in this section, which states that if Conjecture 4.3 holds for parameter δ, then any QCCC key agreement protocols can be broken in roughly $1/\operatorname{poly}(\delta)$ queries. Additionally, by applying Theorem 4.4, we obtain an attack by using exponentially-many queries without resorting to any conjecture. Out results are formulated as the following two theorems.

Theorem 4.5 (Polynomial-query attacks). *Let* (A, B) *be a two-party QCCC protocol where Alice and Bob asks at most* d *queries to a random oracle* h *whose range is* $\mathcal{Y}$*. If Conjecture 4.3 is true, then, there exists an attacker that breaks* (A, B) *by asking* $\mathrm{poly}(d, \log|\mathcal{Y}|)$ *many* classical *queries to* h *and finds the key with probability* ≥ 0.8*.*

Theorem 4.6 (Exponential-query attacks). *Let* (A, B) *be a two-party QCCC protocol with a total of* d *queries to a random oracle* h *whose range is* $\mathcal{Y}$*. Then, there is an attacker who asks an expected number of* $|\mathcal{Y}|^d d^2/\lambda$ *classical queries to* h *and finds the key with probability at least* $1-\lambda$*.*

The rest of this section dedicates to proving Theorem 4.5 and Theorem 4.6. In a nutshell, the proof consists of the following steps.

- In Lemma 4.7, we show that once any two $(\mathcal{Y}, \delta = \varepsilon/\lambda, d, N)$-states are compatible, then any QCCC key agreement protocols can be broken in roughly $1/\mathrm{poly}(\delta)$ queries. The exponential-query attack follows from Theorem 4.4 and Lemma 4.7.
- In Lemma 4.8, we show that if there exists a group $\mathcal{Y}$ such that any key agreement using the oracle with the range $\mathcal{Y}$ is broken by polynomial-query attacks, then any key agreements with a different group $\mathcal{Y}'$ can also be broken by such attacks.

In this section, Alice and Bob always output the same key $k \in \{0,1\}$ with probability 1. Notice that assuming the output is just a bit only makes our impossibility stronger. Besides, we say a key agreement protocol $(\mathsf{A}^h, \mathsf{B}^h)$ using the random oracle h is (τ, s)-broken, if there exists an attacker that finds the key in $(\mathsf{A}^h, \mathsf{B}^h)$ with probability at least τ after asking s many queries to h in expectation. We call the scheme (τ, s)-classically broken, if the same thing holds using only classical queries in the attack.

Lemma 4.7 ((Conditionally) breaking QCCC protocols in the QROM). *Let* $\mathcal{Y}$ *be any finite Abelian group. Let* (A, B) *be a key agreement protocol with at most* d *quantum queries to the random oracle* h *whose range is* $\mathcal{Y}$*. If it holds that any two* $(\mathcal{Y}, \delta = \varepsilon/\lambda, d, N)$*-states are compatible, then* (A, B) *is* $(1-\lambda, d/\varepsilon)$*-classically broken.*

The proof of Lemma 4.7 is given in Section 4.3.

Lemma 4.8 (Group equivalence). *Suppose there exists a finite Abelian group* $\mathcal{Y}$*, a constant* $\tau > 0$ *and a function* $s(\cdot)$ *such that for all* $d \in \mathbb{N}$ *and any single-bit key agreement protocol* $(\mathsf{A}_1^{h_1}, \mathsf{B}_1^{h_1})$ *where Alice and Bob asks* d *queries to random oracles* h_1 *whose range is* $\mathcal{Y}$*, it holds that* $(\mathsf{A}_1^{h_1}, \mathsf{B}_1^{h_1})$ *is* $(\tau, s(d))$*-broken. Then, for any finite Abelian group* $\mathcal{Y}'$*, any* $d' \in \mathbb{N}$*,* $\delta > 0$ *and any single-bit key agreement protocol* $(\mathsf{A}'^{h'}, \mathsf{B}'^{h'})$ *where Alice and Bob asks* d' *queries to random oracles* h' *whose range is* $\mathcal{Y}'$*,* $(\mathsf{A}'^{h'}, \mathsf{B}'^{h'})$ *can be* $(\tau - \delta, 4s(md'))$*-broken, where* $m = \lceil \log_{|\mathcal{Y}|}(d'^3|\mathcal{Y}'|/4\delta^2) \rceil$*.*

The proof of Lemma 4.8 is given in Section 8.2 of the full version [ACC+22].

Proof of Theorem 4.5. Because Conjecture 4.3 is true, there exists a finite Abelian group $\mathcal{Y}$ such that for any $d, N \in \mathbb{N}$, any sufficiently small $\delta = 1/\mathrm{poly}(d)$, it holds that any two $(\mathcal{Y}, \delta, d, N)$-states $|\psi\rangle$ and $|\phi\rangle$ are compatible. Then, Lemma 4.7 guarantees that for any key agreement protocol (A, B) where Alice and Bob asks at most d queries to an oracle h whose range is $\mathcal{Y}$, there exists an attacker that breaks (A, B) by asking $\mathrm{poly}(d)$ many queries to h in expectation and finds the key with probability at least 0.9.

Next, by Lemma 4.8, for any finite Abelian group $\mathcal{Y}'$, $d' \in \mathbb{N}$, $\delta > 0$ and single-bit key agreement $(\mathsf{A}'^{h'}, \mathsf{B}'^{h'})$ where Alice and Bob asks d' queries to random oracles h' with range $\mathcal{Y}'$, $(\mathsf{A}'^{h'}, \mathsf{B}'^{h'})$ can be $(0.9 - \delta, \mathrm{poly}(md'))$-classically broken, where

$$m = \lceil \log_{|\mathcal{Y}|}(^{d'^3|\mathcal{Y}'|}/_{4\delta^2}) \rceil.$$

Choosing $\delta = 0.1$, we obtain a $\mathrm{poly}(d', |\mathcal{Y}'|)$-query attack which finds the key with probability 0.8. Moreover, since $d', \log|\mathcal{Y}'|$ are both at most $\mathrm{poly}(\kappa)$, where κ is the security parameter (as Alice and Bob both run in time $\mathrm{poly}(\kappa)$), this would lead to a $\mathrm{poly}(\kappa)$-query attack. □

Proof of Theorem 4.6. The proof follows from Theorem 4.4 and Lemma 4.7 with $\varepsilon/\lambda = \delta = |\mathcal{Y}|^{-d}/d$. □

4.3 Proof of Lemma 4.7

The rest of this section will be dedicated to proving Lemma 4.7.

Definition 4.9 (Quantum ε-heavy queries). *For $x \in \mathcal{X}$, let $\Pi_x := \sum_{\hat{y} \in \hat{\mathcal{Y}} \setminus \{\hat{0}\}} |\hat{y}\rangle\langle\hat{y}|_{H_x}$. Given a quantum state $|\psi\rangle_{W_\mathsf{A} W_\mathsf{B} H}$, the weight of any $x \in \mathcal{X}$ is defined as*

$$w(x) := \|\Pi_x |\psi\rangle\|_2^2.$$

We call $x \in \mathcal{X}$ a quantum ε-heavy query if $w(x) \geq \varepsilon$.

Construction 4.10 (Attack) *Suppose (A, B) is a quantum-Alice quantum-Bob key agreement protocol using the random oracle h. Given the transcript t, attacking algorithm E' is parameterized by ε and works as follows.*

1. *Prepare $L = \emptyset$ and the classical description of the state*

$$|\psi\rangle_{W'_\mathsf{A} W'_\mathsf{B} H'} = |0\rangle_{W'_\mathsf{A}} |0\rangle_{W'_\mathsf{B}} |\Phi_0\rangle_{H'},$$

   where $W'_\mathsf{A}, W'_\mathsf{B}$ and H' are the simulated registers for Alice, Bob and the oracle prepared by E'.[14]
2. *Simulate the state evolution during the protocol. Concretely, E' calculates the state in $W'_\mathsf{A} W'_\mathsf{B} H'$ after each round in the protocol. Whenever E' encounters the moments in which Alice (Bob) send their messages, E' calculates the post-measurement state that is consistent with t.*
3. *While there is any query $x \notin L$ that is quantum ε-heavy conditioned on (t, L), do the following.*
   (a) *Ask the lexicographically first quantum ε-heavy query x from the real oracle h.*
   (b) *Update the state in $W'_\mathsf{A} W'_\mathsf{B} H'$ to the post-measurement state that is consistent with $(x, h(x))$.*
   (c) *Update L by adding $(x, h(x))$ to L.*
4. *When there is no quantum ε-heavy query left to ask, E' obtains distributions of Alice's and Bob's final keys conditioned on (L, t), and it outputs the key $k \in \{0, 1\}$ that has the highest probability of being Alice's key in this distribution.*

[14] Recall that $|\Phi_0\rangle$ is a uniform superposition over all $h \in \mathcal{H}$, defined as Eq.(1).

Remark 4.11. The attacking algorithm E' is purely classical. It does not need to actually prepare quantum states and apply quantum operation to them. Instead, at each round, the entire protocol, including the sampling of the oracle, can be represented as a pure quantum state. The classical algorithm E' only needs to query the real oracle h classically and simulate how that pure state evolves conditioned on the classical information (t, L) that E' has so far, and all of that is done in Eve's head.

Lemma 4.12 (Efficiency). *Let L be the final list of Eve's algorithm in Construction 4.10. Then $\mathbb{E}[|L|] \leq d/\varepsilon$, where the probability is over the measurement outcomes.*

Proof. By asking queries, Eve gradually gathers a set of query-answer pairs. It naturally introduces a tree where each node corresponds to an intermediate state of L during the procedure. At each node, Eve deterministically chooses the next query q based on t and L and each of its children corresponds to different possible $h(q)$ answered by the oracle. Similar to the proof of Lemma 3.7, in the purified view we denote the state conditioned on t and L by $|\Psi_{t,L}\rangle$. Formally, each node v of the tree consists of the following:

- A label (t, L).
- A quantum state $|\Psi_v\rangle_{W'_{\mathsf{A}}W'_{\mathsf{B}}H'} \coloneqq |\Psi_{t,L}\rangle_{W'_{\mathsf{A}}W'_{\mathsf{B}}H'}$.
- A non-negative real number *total weight* $\mathbf{W}(v)$ defined as

$$\mathbf{W}(v) \coloneqq \sum_{x \in \mathcal{X} \setminus Q_{L'}} \|\Pi_x |\Psi_{t,L'}\rangle\|_2^2.$$

- A Boolean feature $stop(v) \in \{0, 1\}$. If there is no quantum ε-heavy query, then $stop(v) = 1$. In particular, $\mathbf{W}(v) < \varepsilon$ implies $stop(v) = 1$.

The random walk on this tree can start from any node. Whenever $stop(v) = 0$, it moves to of one of its children u according the distribution of measuring the register H_q of $|\Psi_v\rangle$ in the computational basis, where q is Eve's next query at v. Actually, this distribution, denoted by $\Gamma(v)$, is equivalent to the distribution of Eve's query-answer from h conditioned on t and L. By $u \leftarrow \Gamma(v)$ we denote the step from v to its child u. Observes that the depth of the tree is finite since $|L|$ is at most $|\mathcal{X}|$.

For any v and its children u, we have the following property

$$\begin{aligned}
&\mathop{\mathbb{E}}_{u \leftarrow \Gamma(v)}[\mathbf{W}(u)] = \sum_{x \in \mathcal{X} \setminus Q_{L'}} \sum_{y \in \mathcal{Y}} \|\Pi_x |y\rangle\langle y|_{H'_q} |\Psi_v\rangle\|_2^2 \\
&= \sum_{x \in \mathcal{X} \setminus Q_{L'}} \sum_{y \in \mathcal{Y}} \||y\rangle\langle y|_{H'_q} \Pi_x |\Psi_v\rangle\|_2^2 = \sum_{x \in \mathcal{X} \setminus Q_{L'}} \|\Pi_x |\Psi_v\rangle\|_2^2 \\
&= \sum_{x \in \mathcal{X} \setminus Q_L} \|\Pi_x |\Psi_v\rangle\|_2^2 - \|\Pi_q |\Psi_v\rangle\|_2^2 \leq \mathbf{W}(v) - \varepsilon,
\end{aligned} \tag{3}$$

where q is Eve's next query at v, L is the partial oracle of v, and $Q_{L'} \coloneqq Q_L \cup \{q\}$. The second equality holds since $|y\rangle\langle y|_{H'_q}$ commutes with Π_x for all $x \in \mathcal{X} \setminus Q_{L'}$, and the inequality is due to the heaviness of q.

We claim the following inequality holds for every v

$$\mathbb{E}[|S(v)|] \leq \frac{\mathbf{W}(v)}{\varepsilon}, \tag{4}$$

where by $S(v)$ we denote the total number of steps that the random walk takes when starting form v. We prove it by induction on the depth of the starting node. By D we denote the depth of the tree. For v in depth D we shall have $stop(v) = 1$, in which case $|S(v)| = 0 \leq \mathbf{W}(v)/\varepsilon$, and so the claim follows. Now suppose the inequality holds for depth i nodes and we move to v in depth $i-1$. If $stop(v) = 0$, again we have $|S(v)| = 0 \leq \mathbf{W}(v)/\varepsilon$ which is what we need. Otherwise, by induction and the linearity of expectation,

$$\begin{aligned}
\mathbb{E}[|S(v)|] &= 1 + \mathop{\mathbb{E}}_{u \leftarrow \Gamma(v)}[\mathbb{E}[|S(u)|]] \\
&\leq 1 + \mathop{\mathbb{E}}_{u \leftarrow \Gamma(v)}[\mathbf{W}(u)/\varepsilon] \\
&= 1 + \frac{\mathbb{E}_{u \leftarrow \Gamma(v)}[\mathbf{W}(u)]}{\varepsilon} \\
&\leq 1 + \frac{\mathbf{W}(v) - \varepsilon}{\varepsilon} = \frac{\mathbf{W}(v)}{\varepsilon},
\end{aligned}$$

where the first inequality is due to induction hypothesis and the second inequality follows by Eq. 3. By Lemma 2.4, the total weight of the root R (where the state is $|\Psi_t\rangle$ in the purified view) is at most d since

$$\mathbf{W}(R) = \sum_{x \in \mathcal{X}} \left\| \sum_{\hat{h} \in \hat{\mathcal{H}}} \alpha_{\hat{h}} |\psi_{\hat{h}}\rangle_{W'_{\mathsf{A}} W'_{\mathsf{B}}} \Pi_x |\hat{h}\rangle_{H'} \right\|_2^2 = \sum_{\hat{h} \in \hat{\mathcal{H}}} |\hat{h}| \cdot |\alpha_{\hat{h}}|^2 \leq d \cdot \sum_{\hat{h} \in \hat{\mathcal{H}}} |\alpha_{\hat{h}}|^2 = d,$$

where we represent the attached state as $|\Psi_t\rangle_{W'_{\mathsf{A}} W'_{\mathsf{B}} H'} = \sum_{\hat{h}} \alpha_{\hat{h}} |\psi_{\hat{h}}\rangle_{W'_{\mathsf{A}} W'_{\mathsf{B}}} |\hat{h}\rangle_{H'}$. Therefore, starting from the root we have $\mathbb{E}[|L|] \leq d/\varepsilon$ by Eq. 4. □

Lemma 4.13 (Success). *Suppose that Alice and Bob ask a total of d quantum queries. If any two $(|\mathcal{Y}|, \delta = \varepsilon/\lambda, d, N)$-states are compatible, then there is an eavesdropper E who finds the key k with probability at least $1 - \lambda$.*

Proof. Consider the purified version of the protocol. Let $|\Psi_t\rangle_{WH}$ be the joint state after the protocol finishes, conditioned on the transcript t. By Lemma 2.4 it holds that $|\hat{h}^H_{\max}(|\Psi_t\rangle)| \leq d$. After E' learns the heavy queries, the resulting state becomes $|\Psi_{t,L}\rangle$ conditioned on L. Similarly, by Lemma 2.7 it holds that $|\hat{h}^{H'}_{\max}(|\Psi_{t,L}\rangle)| \leq d$. Since the oracle registers corresponding to Q_L are now measured, we can consider the "truncated" version of $|\Psi_{t,L}\rangle_{WH}$ by discarding those registers. Let $H' = \{H_x\}_{x \in \mathcal{X} \setminus Q_L}$ be the set of remaining registers. By $|\Psi_{t,L}\rangle_{WH'}$ we denote the truncated $|\Psi_{t,L}\rangle_{WH}$.

Now, set the register H in Definition 4.1 to be H'. The state $|\Psi_{t,L}\rangle$ is d-sparse and ε-light by definition, so $|\Psi_{t,L}\rangle$ is a $(|\mathcal{Y}|, \varepsilon, d)$-state. Recall that at the end of the attack, E' learns all the heavy queries, calculates the key distribution of $|\Psi_{t,L}\rangle$ among the remaining oracles and outputs the key with the highest probability to be outputted. We are going to show that there exist a key $k = b \in \{0, 1\}$ such that the probability of the key b in the key distribution of $|\Psi_{t,L}\rangle$, denoted by $\Pr[k = b \text{ in } |\Psi_{t,L}\rangle]$, is larger than $1 - \lambda$. We will prove this by contradiction. Namely, in the following, suppose $\Pr[k = b \text{ in } |\Psi_{t,L}\rangle] \geq \lambda$ for both $b = 0$ and $b = 1$.

Let $|\Psi_{t,L,k=b}\rangle$ be the residual state of $|\Psi_{t,L}\rangle$ conditioned on $k = b$. Observe that $|\Psi_{t,L,k=b}\rangle$ is a $(C, \varepsilon/\lambda, d)$-state for both $k \in \{0, 1\}$. In addition, $|\Psi_{t,L,k=b}\rangle$ is d-sparse

since $|\Psi_{t,L}\rangle$ is d-sparse and conditioning on k is a process acting on A and B's registers and will not affect the sparsity of the oracle. $|\Psi_{t,L,k=b}\rangle$ is ε/λ-light because $|\Psi_{t,L}\rangle$ is ε-light and $\Pr[k=b \text{ in } |\Psi_{t,L}\rangle] \geq \lambda$. By the premise in the lemma statement, $|\Psi_{t,L,k=0}\rangle$ and $|\Psi_{t,L,k=1}\rangle$ are compatible, which means that there exists an oracle h, a state $w_{\mathsf{A}} \in W_{\mathsf{A}}$ which outputs the key $k=0$, and a state $w_{\mathsf{B}} \in W_{\mathsf{B}}$ outputs the key $k=1$ such that h is consistent with both w_{A} and w_{B} with nonzero probability, that is, there is a nonzero chance that in a real execution of the protocol, A outputs the key 0 and B outputs the key 1, which violates the perfect completeness of the protocol. □

Proof of Theorem 4.7. We use the Eve of Construction 4.10 with parameter ε. Then, by Lemma 4.12, the expected number of queries of Eve is at most d/ε, and by Lemma 4.13, it finds the key with probability $1-\lambda$. □

5 Case of Exponentially Small Influences: Proving Theorem 4.4

Before proving Theorem 4.4, we describe a connection between $(|\mathcal{Y}|, \delta, d, N)$-states and distributions of polynomials with bounded degree and influence, giving an alternative formulation of Conjecture 4.3.

5.1 The Polynomial Formulation

As in the rest of the paper, we let $\mathcal{Y}$ be an Abelian group of order $|\mathcal{Y}|$ and $\hat{\mathcal{Y}}$ be its dual group having $\hat{0}$ as the identity element. Recall that we are working with quantum states over a register H whose basis states are all functions $h : \mathcal{X} \to \mathcal{Y}$ for some $|\mathcal{X}| = N$. To keep the notation clean in this section, we identify $\mathcal{X}$ with $[N]$ and view functions $h : \mathcal{X} \to \mathcal{Y}$ as vectors in $\mathcal{Y}^N$ (i.e., we write h_i rather than $h(x)$ for a typical value).

We recall that any $f : \mathcal{Y}^N \to \mathbb{C}$ can be written in terms of its Fourier transform

$$f(\mathbf{x}) = \sum_{\chi \in \hat{\mathcal{Y}}^N} \hat{f}(\chi) \prod_{i=1}^{N} \chi_i(\mathbf{x}_i)$$

The *degree* of a character $\chi \in \hat{\mathcal{Y}}^N$ is $\deg(\chi) = |\{i \,|\, \chi_i \neq \hat{0}\}|$, and the degree of f is $\deg(f) = \max\{\deg(\chi) \,|\, \hat{f}(\chi) \neq 0\}$. The *influence* of variable i on f is $\mathrm{Inf}_i(f) = \sum_{\substack{\chi \in \hat{\mathcal{Y}}^N \\ \chi_i \neq \hat{0}}} |\hat{f}(\chi)|^2$. We denote by $\max \mathrm{Inf}(f) = \max_{i=1\ldots N} \mathrm{Inf}_i(f)$ the maximum influence of f.

Definition 5.1 (State polynomial). *For a quantum state $|\psi\rangle$ over the register H, the* state polynomial *of $|\psi\rangle$ is the function $f_\psi : \mathcal{Y}^N \to \mathbb{C}$ defined by*

$$f_\psi(h) = |\mathcal{Y}|^{N/2} \cdot \langle \psi | h \rangle = \sum_{\chi \in \hat{\mathcal{Y}}^N} \langle \psi | \chi \rangle \prod_{i=1}^{N} \chi_i(h_i). \tag{5}$$

Lemma 5.2 (Sparsity vs. degree, heaviness vs. influence). *For a quantum state $|\psi\rangle$ over register H, f_ψ has the following properties.*

1. *f_ψ has ℓ_2-norm equal to 1, i.e.,* $\mathbb{E}_{\mathbf{x}\leftarrow\mathcal{Y}^N}|f_\psi(\mathbf{x})|^2 = 1$.
2. *$|\psi\rangle$ is d-sparse if and only if* $\deg(f_\psi) \le d$.
3. *$|\psi\rangle$ has no δ-heavy queries if and only if* $\max \mathrm{Inf}(f_\psi) \le \delta$.

Proof. For Item 1, we have by definition $\mathbb{E}_{\mathbf{x}\leftarrow\mathcal{H}}[|f_\psi(\mathbf{x})|^2 = \sum_h |\langle\psi|h\rangle|^2 = 1$ (since the set of h form a basis for the space). For Item 2, recall from Definition 4.1 that $|\psi\rangle$ is d-sparse if and only if $|\hat{h}^H_{\max}(|\psi\rangle)| \le d$, i.e., if for $\chi \in \hat{\mathcal{Y}}$, we have $\langle\psi|\chi\rangle \neq 0$ only if $d \ge |\{i \,|\, \chi_i \neq \hat{0}\}| = \deg(\chi)$. Equivalently, the non-zero terms in the right hand side of (5) are those where $\deg(\chi) \le d$, i.e., $\deg(f_\psi) \le d$. Finally, for Item 3, recall from Definition 4.9 that $|\psi\rangle$ has no δ-heavy queries if and only if $\|\Pi_i|\psi\rangle\|_2^2 \le \delta$ for all $i \in [N]$, where $\Pi_i = \sum_{\chi_i\in\hat{Y}\setminus\hat{0}} |\chi_i\rangle\langle\chi_i|_{H_i}$. Expanding, we see that

$$\|\Pi_i|\psi\rangle\|_2^2 = \sum_{\substack{\chi\in\hat{Y}^N \\ \chi_i\neq\hat{0}}} |\langle\psi|\chi\rangle|^2 = \mathrm{Inf}_i(f_\psi).$$

□

Definition 5.3 (State polynomial distribution). *For a quantum state $|\psi\rangle$ over registers W, H, the* state polynomial distribution *of $|\psi\rangle$ is the distribution F_ψ over polynomials $f : \mathcal{Y} \to \mathbb{C}$ which is sampled by measuring W in some fixed basis and taking the resulting state polynomial for H.*

Observation 5.4 *Two quantum states $|\psi\rangle$ and $|\phi\rangle$ over registers W, H are compatible if and only if there exist $f \in \mathrm{supp}(F_\psi)$, $g \in \mathrm{supp}(F_\phi)$ and an $\mathbf{x} \in \mathcal{Y}^N$ such that $f(\mathbf{x}) \cdot g(\mathbf{x}) \neq 0$.*

The observations above motivate us to formulate our main conjecture in terms of polynomials. Notice that, in the following formulation, we focus on the distributions of functions whose range is $\mathbb{R}$ instead of $\mathbb{C}$. Later on, in Theorem 5.6, we will show that it suffices to consider real functions.

Conjecture 5.5 *There exists a finite Abelian group $\mathcal{Y}$ and a function $\delta(d) = 1/\mathrm{poly}(\cdot)$ such that the following holds for all d. Let F and G be two distributions of functions from $\mathcal{Y}^N$ to $\mathbb{R}$ such that the following holds for all $f \in \mathrm{supp}(F)$ and $g \in \mathrm{supp}(G)$.*

- **Unit ℓ_2 norm:** *f and g have ℓ_2-norm 1.*
- ***d*-degrees:** *$\deg(f) \le d$ and $\deg(g) \le d$.*
- ***δ*-influences on average:** *For all $i \in [N]$, we have $\mathbb{E}_{f\leftarrow F}[\mathrm{Inf}_i(f)] \le \delta$ and $\mathbb{E}_{g\leftarrow G}[\mathrm{Inf}_i(g)] \le \delta$, where $\delta = \delta(d)$.*

Then, there is an $f \in \mathrm{supp}(F)$, $g \in \mathrm{supp}(G)$, and $\mathbf{x} \in \mathcal{Y}^N$ such that $f(\mathbf{x}) \cdot g(\mathbf{x}) \neq 0$.

Theorem 5.6. *Conjecture 5.5 is true if and only if Conjecture 4.3 is true.*

The proof is given in Appendix A of the full version [ACC+22].

5.2 Proving Theorem 4.4

In this subsection, we prove Theorem 4.4, using the polynomial formulation explained in the previous subsection. In other words, we prove a weaker version of Conjecture 5.5 where we set $\delta < |\mathcal{Y}|^{-d}/d$. Interestingly, the theorem holds without any influence condition on F, and without any degree restriction on G. I.e., we only use that there is an $f \in \mathrm{supp}(F)$ of degree $\leq d$, and that $\mathbb{E}_{g \leftarrow G}[\mathrm{Inf}_i(g)] \leq \delta$ for all $i \in [N]$.

For any $f \in \mathrm{supp}(F)$, let $f(\mathbf{x}) = \sum_{\chi \in \hat{\mathcal{Y}}^{\mathcal{X}}} \hat{f}(\chi)\chi(\mathbf{x})$ and $\chi^* \in \hat{\mathcal{Y}}^N$ be a character for which $\hat{f}(\chi) \neq 0$ and $\deg(\chi) = \deg(f)$. Since $\deg(f) \leq d$ we can without loss of generality assume that $\chi_i^* = \hat{0}$ for $i = d+1, \ldots, N$ by reordering the coordinates.

Note that for any partial assignment $\mathbf{x}_{>d} = (x_{d+1}, \ldots, x_N)$, the restricted function $f|_{\mathbf{x}_{>d}}$ is non-constant and in particular there exists a $\mathbf{x}_{\leq d}$ such that $f(\mathbf{x}_{\leq d}, \mathbf{x}_{>d}) \neq 0$.

For any function $g : \mathcal{Y}^N \to \mathbb{C}$, decompose it as

$$g(\mathbf{x}) = \sum_{\chi \in \hat{\mathcal{Y}}^d} g_\chi(\mathbf{x}_{>d})\chi(\mathbf{x}_{\leq d})$$

for $|\mathcal{Y}|^d$ functions $\{g_\chi\}_{\chi \in \hat{\mathcal{Y}}^d}$ on $\mathbf{x}_{>d}$. Writing $\hat{\mathbf{0}} = (\hat{0}, \ldots, \hat{0}) \in \mathcal{Y}^d$ we then have

$$\sum_{\chi \neq \hat{\mathbf{0}}} \mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[|g_\chi(\mathbf{x}_{>d})|^2\right] \leq \sum_{i=1}^{d} \sum_{\chi_i \neq \hat{0}} \mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[|g_\chi(\mathbf{x}_{>d})|^2\right] = \sum_{i=1}^{d} \mathrm{Inf}_i(g)$$

and $\mathbb{E}_{\mathbf{x}_{>d}} \left[|g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})|^2\right] \geq \|g\|_2^2 - \sum_{i=1}^{d} \mathrm{Inf}_i(g)$. Thus, we have

$$\mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[|g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})|^2 - (|\mathcal{Y}|^d - 1) \sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})|^2 \right] \geq \|g\|_2^2 - |\mathcal{Y}|^d \sum_{i=1}^{d} \mathrm{Inf}_i(g)$$

Taking the expectation over $g \leftarrow G$ and using the condition $\mathbb{E}_{g \leftarrow G}[\mathrm{Inf}_i(g)] \leq \delta < |\mathcal{Y}|^{-d}/d$ on the influences of G we thus conclude

$$\mathop{\mathbb{E}}_{g \leftarrow G} \mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[|g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})|^2 - (|\mathcal{Y}|^d - 1) \sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})|^2 \right] > 0.$$

In particular there exists a $g \in \mathrm{supp}(G)$ such that

$$\mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[|g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})|^2\right] > \mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[(|\mathcal{Y}|^d - 1) \sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})|^2 \right] \geq \mathop{\mathbb{E}}_{\mathbf{x}_{>d}} \left[\left(\sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})| \right)^2 \right],$$

where the second inequality is Cauchy-Schwarz. It follows that there is $\mathbf{x}_{>d}$ such that

$$|g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})| > \sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})|.$$

As observed above, for this $\mathbf{x}_{>d}$ there must exist some $\mathbf{x}_{\leq d}$ such that $f(\mathbf{x}_{\leq d}, \mathbf{x}_{>d}) \neq 0$. But, that means we obtain the following as desired.

$$|g(\mathbf{x}_{\leq d}, \mathbf{x}_{>d})| = \left| \sum_{\chi \in \hat{\mathcal{Y}}^d} g_\chi(\mathbf{x}_{>d}) \chi(\mathbf{x}_{\leq d}) \right| \geq |g_{\hat{\mathbf{0}}}(\mathbf{x}_{>d})| - \sum_{\chi \neq \hat{\mathbf{0}}} |g_\chi(\mathbf{x}_{>d})| > 0.$$

References

AA09. Scott Aaronson and Andris Ambainis. The need for structure in quantum speedups. *arXiv preprint arXiv:0911.0996*, 2009. 6, 11

Aar05. Scott Aaronson. Quantum computing, postselection, and probabilistic polynomial-time. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2063):3473–3482, 2005. 10

ABG+21. Amit Agarwal, James Bartusek, Vipul Goyal, Dakshita Khurana, and Giulio Malavolta. Post-quantum multi-party computation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 435–464. Springer, 2021. 11

ACC+22. Per Austrin, Hao Chung, Kai-Min Chung, Shiuan Fu, Yao-Ting Lin, and Mohammad Mahmoody. On the impossibility of key agreements from quantum random oracles. Cryptology ePrint Archive, Paper 2022/218, 2022. https://eprint.iacr.org/2022/218. 4, 5, 6, 10, 17, 21, 26

ACGH20. Gorjan Alagic, Andrew M Childs, Alex B Grilo, and Shih-Han Hung. Non-interactive classical verification of quantum computation. In *Theory of Cryptography Conference*, pages 153–180. Springer, 2020. 11

ACP21. Prabhanjan Ananth, Kai-Min Chung, and Rolando L La Placa. On the concurrent composition of quantum zero-knowledge. In *Annual International Cryptology Conference*, pages 346–374. Springer, 2021. 11

Bar17. Boaz Barak. The complexity of public-key cryptography. In *Tutorials on the Foundations of Cryptography*, pages 45–77. Springer, 2017. 2

Bar21. James Bartusek. Secure quantum computation with classical communication. Cryptology ePrint Archive, Report 2021/964, 2021. https://ia.cr/2021/964. 11

BB84. CH Bennett and G Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proc. of IEEE Int. Conf. on Computers, Systems, and Signal Processing*, pages 175–179, 1984. 2

BGI08. Eli Biham, Yaron J Goren, and Yuval Ishai. Basing weak public-key cryptography on strong one-way functions. In *Theory of Cryptography Conference*, pages 55–72. Springer, 2008. 2

BGP00. Mihir Bellare, Oded Goldreich, and Erez Petrank. Uniform generation of NP-witnesses using an NP-oracle. *Information and Computation*, 163(2):510–526, 2000. 10

BHK+15. Gilles Brassard, Peter Hoyer, Kassem Kalach, Marc Kaplan, Sophie Laplante, and Louis Salvail. Key establishment à la merkle in a quantum world, 2015. 3

BKS21. Nir Bitansky, Michael Kellner, and Omri Shmueli. Post-quantum resettably-sound zero knowledge. In *Theory of Cryptography Conference*, pages 62–89. Springer, 2021. 11

BKSY11. Zvika Brakerski, Jonathan Katz, Gil Segev, and Arkady Yerukhimovich. Limits on the power of zero-knowledge proofs in cryptographic constructions. In *Theory of Cryptography Conference*, pages 559–578. Springer, 2011. 6

BKVV20. Zvika Brakerski, Venkata Koppula, Umesh Vazirani, and Thomas Vidick. Simpler proofs of quantumness. *arXiv preprint arXiv:2005.04826*, 2020. 11

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), 2017. 2, 5, 6, 7, 18

BMG09a. Boaz Barak and Mohammad Mahmoody-Ghidary. Merkle puzzles are optimal - an O(n^2)-query attack on any key exchange from a random oracle. In Shai Halevi, editor, *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, volume 5677 of *Lecture Notes in Computer Science*, pages 374–390. Springer, 2009. 5

BMG09b. Boaz Barak and Mohammad Mahmoody-Ghidary. Merkle puzzles are optimal - an $O(n^2)$-query attack on any key exchange from a random oracle. In Shai Halevi, editor, *Advances in Cryptology – CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 374–390, Santa Barbara, CA, USA, August 16–20, 2009. Springer, Heidelberg, Germany. 10

BS08. Gilles Brassard and Louis Salvail. Quantum merkle puzzles. In *International Conference on Quantum, Nano and Micro Technologies (ICQNM)*, pages 76–79. IEEE Computer Society, 2008. 2, 3

BS20. Nir Bitansky and Omri Shmueli. Post-quantum zero knowledge in constant rounds. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 269–279, 2020. 11

CCY20. Nai-Hui Chia, Kai-Min Chung, and Takashi Yamakawa. Classical verification of quantum computations with efficient verifier. In *Theory of Cryptography Conference*, pages 181–206. Springer, 2020. 11

CFHL21. Kai-Min Chung, Serge Fehr, Yu-Hsuan Huang, and Tai-Ning Liao. On the compressed-oracle technique, and post-quantum security of proofs of sequential work. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 598–629. Springer, 2021. 12

CX21. Shujiao Cao and Rui Xue. Being a permutation is also orthogonal to one-wayness in quantum world: Impossibilities of quantum one-way permutations from one-wayness primitives. *Theoretical Computer Science*, 855:16–42, 2021. 10

DFKO06. Irit Dinur, Ehud Friedgut, Guy Kindler, and Ryan O'Donnell. On the fourier tails of bounded functions over the discrete cube. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 437–446, 2006. 11

DH76. Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Transactions on Information Theory*, 22(6):644–654, 1976. 2

GKM$^+$00. Yael Gertner, Sampath Kannan, Tal Malkin, Omer Reingold, and Mahesh Viswanathan. The relationship between public key encryption and oblivious transfer. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 325–335. IEEE, 2000. 5

Gro96. Lov K Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*, pages 212–219, 1996. 3

HHRS07. Iftach Haitner, Jonathan J. Hoch, Omer Reingold, and Gil Segev. Finding collisions in interactive protocols - A tight lower bound on the round complexity of statistically-hiding commitments. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007), October 20-23, 2007, Providence, RI, USA, Proceedings*, pages 669–679, 2007. 10

HMO$^+$21. Iftach Haitner, Noam Mazor, Rotem Oshman, Omer Reingold, and Amir Yehudayoff. On the communication complexity of key-agreement protocols. *arXiv preprint arXiv:2105.01958*, 2021. 11

HMST21. Iftach Haitner, Noam Mazor, Jad Silbak, and Eliad Tsfadia. On the complexity of two-party differential privacy, 2021. 10

Hol05. Thomas Holenstein. Key agreement from weak bit agreement. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 664–673, 2005. 10

HOZ16. Iftach Haitner, Eran Omri, and Hila Zarosim. Limits on the usefulness of random oracles. *Journal of Cryptology*, 29(2):283–335, 2016. 11

HY18. Akinori Hosoyamada and Takashi Yamakawa. Finding collisions in a quantum world: Quantum black-box separation of collision-resistance and one-wayness. Cryptology ePrint Archive, Report 2018/1066, 2018. https://ia.cr/2018/1066. 3

HY20. Akinori Hosoyamada and Takashi Yamakawa. Finding collisions in a quantum world: quantum black-box separation of collision-resistance and one-wayness. In *International Conference on the Theory and Application of Cryptology and Information Security*, pages 3–32. Springer, 2020. 3, 5, 10

IR89. Russell Impagliazzo and Steven Rudich. Limits on the provable consequences of one-way permutations. In *Proceedings of the 21st Annual ACM Symposium on Theory of Computing (STOC)*, pages 44–61. ACM Press, 1989. 2, 5, 6, 7, 10

KSS00. Jeff Kahn, Michael Saks, and Cliff Smyth. A dual version of reimer's inequality and a proof of rudich's conjecture. In *Proceedings 15th Annual IEEE Conference on Computational Complexity*, pages 98–103. IEEE, 2000. 11

Mah18. Urmila Mahadev. Classical verification of quantum computations. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 259–267. IEEE, 2018. 11

Mer74. R. Merkle. C.s. 244 project proposal. In *Facsimile available at http://www.merkle.com/1974.*, 1974. 2

MMP14. Mohammad Mahmoody, Hemanta K Maji, and Manoj Prabhakaran. Limits of random oracles in secure computation. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 23–34, 2014. 11

OSSS05. Ryan O'Donnell, Michael Saks, Oded Schramm, and Rocco A Servedio. Every decision tree has an influential variable. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 31–39. IEEE, 2005. 11

RSA78. Ronald L. Rivest, Adi Shamir, and Leonard M. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, Feb 1978. 2

RTV04. Omer Reingold, Luca Trevisan, and Salil P. Vadhan. Notions of reducibility between cryptographic primitives. In *Theory of Cryptography, First Theory of Cryptography Conference, TCC 2004*, volume 2951 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2004. 10

Rud88. Steven Rudich. Limits on the provable consequences of one-way functions. *Ph. D. Thesis, University of California*, 1988. 10

Zha19. Mark Zhandry. How to record quantum queries, and applications to quantum indifferentiability. In *Annual International Cryptology Conference*, pages 239–268. Springer, 2019. 7, 12

Zha21. Jiayu Zhang. Succinct blind quantum computation using a random oracle. In *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 1370–1383, 2021. 11