

# Succinct Arguments for QMA from Collapsing Hash Functions

James Bartusek\*

Giulio Malavolta†

## Abstract

We prove the existence of succinct arguments for QMA, assuming only the existence of collapsing hash functions. This is the first scheme that relies only on unstructured “Minicrypt” assumptions, which are not known to imply public-key encryption.

Our main technical contribution is a quantum-succinct *claw-state generation* protocol that allows us to bootstrap a small number of quantum correlations into an arbitrarily large number of claw-state correlations, using classical communication only. This improves upon the work of [Zhang, STOC 2021], having better round complexity, a proof in the standard model, and being overall much simpler. This yields a quantum-succinct blind delegation of quantum computation protocol from one-way functions, which we plug into the communication-compression compiler of [Bartusek, Liu, and Malavolta, EUROCRYPT 2026] to obtain succinct arguments for QMA.

## Contents

|          |                                       |           |
|----------|---------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                   | <b>2</b>  |
| 1.1      | Results . . . . .                     | 2         |
| 1.2      | Technical Outline . . . . .           | 3         |
| 1.3      | AI Disclosure . . . . .               | 7         |
| <b>2</b> | <b>Preliminaries</b>                  | <b>7</b>  |
| 2.1      | Succinct Arguments for QMA . . . . .  | 8         |
| 2.2      | Claw States . . . . .                 | 8         |
| 2.3      | Pseudorandom Functions . . . . .      | 9         |
| <b>3</b> | <b>Claw-State Generation Protocol</b> | <b>10</b> |
| 3.1      | Protocol Description . . . . .        | 10        |
| 3.2      | Analysis . . . . .                    | 11        |
| <b>4</b> | <b>Sequential Repetition</b>          | <b>15</b> |
| <b>5</b> | <b>Succinct Arguments for QMA</b>     | <b>18</b> |
| 5.1      | Blind Delegation . . . . .            | 20        |
| 5.2      | Putting Things Together . . . . .     | 22        |
|          | <b>References</b>                     | <b>23</b> |

---

\*Columbia University [bartusek.james@gmail.com](mailto:bartusek.james@gmail.com)

†Bocconi University [giulio.malavolta@unibocconi.it](mailto:giulio.malavolta@unibocconi.it)

# 1 Introduction

Succinct arguments enable the verification of mathematical statements using significantly fewer resources than required to process a complete proof. Pioneered by Kilian [Kil92], succinct arguments have become a pillar of study in both foundational and applied cryptography. In his work, Kilian showed how to compile any probabilistically checkable proof system (PCP) [ALM<sup>+</sup>98, AS98] for NP into a succinct argument for NP using only collision-resistant hash functions.

Recent years have witnessed a surge of interest in quantum information-processing, giving rise to a natural follow-up question: Under what cryptographic assumptions do there exist succinct arguments for all of QMA, namely, what type of cryptographic structure enables the extremely efficient verification of statements with *quantum* proofs? A series of recent works [BTL<sup>+</sup>22, MNZ24, GTNV25, BLM26] has culminated in a construction of succinct arguments for QMA from the following two cryptographic ingredients:

- Collapsing hash functions: the post-quantum analogue of collision-resistance.
- Oblivious state preparation: a generic “public-key” style assumption that can be instantiated from LWE, LIP, or assumptions on cryptographic group actions [BK25, BMM26].

The state of the art thus highlights a significant gap from the classical setting: While succinct arguments for NP only require unstructured cryptography in the form of collision-resistant hash functions (placing them in “Minicrypt”), succinct arguments for QMA are only known from public-key style assumptions, positioning them in “Cryptomania”.

In this work, we establish that succinct arguments for QMA do in fact live in Minicrypt by proving their existence assuming only collapsing hash functions [Unr16]. As collapsing is the post-quantum analogue of collision-resistance, this exactly matches the weakest assumption under which Kilian’s succinct arguments for NP are known to be post-quantum secure [CMSZ22].

## 1.1 Results

Our main result is stated below, where  $\lambda$  denotes the security parameter.

**Theorem 1.1.** *There exists a fixed polynomial  $\text{poly}$  such that, assuming collapsing hash functions, there exists an argument system for QMA with completeness  $1 - \text{negl}(\lambda)$ , soundness error  $\text{negl}(\lambda)$ , total communication bounded by  $\text{poly}(\lambda)$ , and verifier runtime  $|x| \cdot \text{poly}(\lambda)$ .*

*The protocol uses quantum communication (and thus a quantum verifier) and makes non-black-box use of the collapsing hash function.*

Above,  $x$  refers to the QMA statement, and  $|x|$  is its bit length. In particular,  $|x| \cdot \text{poly}(\lambda)$  has no dependence on the witness length or the time required to run the original QMA verifier on  $x$  and a witness state  $|\psi\rangle$ . Note that, because of the non-black-box use of the collapsing hash function, we do not automatically obtain a succinct argument for QMA in the quantum random oracle model.

**Remark 1.2** (Prepare-and-send verifier). *One potentially desirable aspect of our protocol is that it does not require the verifier to keep any quantum memory between rounds. In particular, each of the verifier’s quantum messages can be prepared from its classical state immediately before sending, and no auxiliary quantum registers are kept by the verifier after the message is sent.*

Our main technical ingredient is a new construction of blind delegation of quantum computation, which asks whether a resource-constrained client can delegate a quantum computation to a server in such a manner that the server learns nothing about the client’s computation. One natural resource constraint is the size of the quantum circuit that the client runs during the course of the protocol. If the size is 0, we refer to this as *classical* blind delegation of quantum computation, and if the size is bounded by a fixed polynomial in the security parameter, *independent* of the size of the computation being delegated, we refer to this as *quantum-succinct* blind delegation of quantum computation.

It was shown by [Zha21] that quantum-succinct blind delegation of quantum computation exists in the quantum random oracle model. In this work, we design and prove the security of a vastly simpler protocol that only requires one-way functions, yielding the following theorem.

**Theorem 1.3.** *Assuming one-way functions, there exists quantum-succinct blind delegation of quantum computation.*

**Discussion.** Recall that Kilian’s result used collision-resistant hash functions to compile a PCP for NP into a succinct argument for NP. Recently, a quantum analogue of Kilian’s compiler was worked out by [GJMZ23], establishing that PCPs for QMA can be compiled into succinct arguments for QMA, utilizing only unstructured cryptography. In fact, they assume only the existence of pseudorandom unitaries, a “Microcrypt” assumption that is weaker than even one-way functions. Unfortunately, it remains a major unresolved problem whether PCPs for QMA actually exist.

Another point worth making is that [GJMZ23] and our protocol both make use of quantum communication and thus a quantum verifier. One could ask whether the verifier can be made completely classical, while remaining succinct. This is indeed known, but under public-key assumptions such as LWE. However, improving the result to classical, even *non-succinct*, verification of QMA in Minicrypt would be considered a major breakthrough, so we regard the quantum communication component of our protocol as a crucial relaxation given current techniques.

Therefore, we can summarize the state of succinct arguments for QMA as follows.

- [GJMZ23]: Assuming the quantum PCP conjecture, there exist succinct arguments for QMA from pseudorandom unitaries. That is, assuming the quantum PCP conjecture, succinct arguments for QMA exist in Microcrypt.
- Our result: Succinct arguments for QMA exist in Minicrypt, without assuming the quantum PCP conjecture. In particular, succinct arguments for QMA exist assuming only collapsing hash functions.
- [MNZ24, GTNV25, BLM26]: Succinct arguments for QMA *with classical verification* exist in Cryptomania, without assuming the quantum PCP conjecture. In particular, succinct classically-verifiable arguments for QMA exist assuming a standard public-key assumption such as LWE.

## 1.2 Technical Outline

Before we explain our approach to the problem, let us recall some necessary background information to motivate our design choices. Our starting point is the recent work of [BLM26], who proposed a compiler that, assuming collapsing hash functions, compresses the communication complexity of

any classical-verifier interactive protocol  $\Pi$  with  $r(\lambda)$  rounds. For any  $\varepsilon = 1/\text{poly}(\lambda)$ , the compiled protocol  $\tilde{\Pi}$  has negligible completeness loss, soundness loss  $O(\varepsilon) + \text{negl}(\lambda)$ , and satisfies

$$\text{CC}(\tilde{\Pi}) \leq r(\lambda) \text{poly}(\lambda, 1/\varepsilon), \quad \text{verifier time} \leq |x| r(\lambda) \text{poly}(\lambda, 1/\varepsilon),$$

where  $x$  is the common input and the polynomial is fixed independently of the original message lengths. Thus, for our purposes, the remaining task is to construct a QMA argument with a fixed  $\text{poly}(\lambda)$  number of rounds but otherwise arbitrary classical communication, which can be fed into the above compiler.

Prior work [BLM26] constructs such a round-efficient argument using the compiled non-local games framework [KLVY23, NZ23, BK25, BKM<sup>+</sup>25]. The approach here is to start from an information-theoretically secure *two-prover* protocol and compile it into a single-prover protocol using cryptography. The cryptographic ingredient is *blind delegation*, which is an interactive protocol between a server and a client, where the client holds a private classical input  $x$ , the server has a quantum input  $\rho$ , and they both have the public description of a quantum circuit  $Q$ . In the ideal honest execution, the client obtains Pauli keys  $(r, s)$  and the server obtains  $\sigma$  such that

$$Z^s X^r \sigma X^r Z^s = Q(x, \rho).$$

Correctness allows negligible error and includes arbitrary reference systems, as specified in [Definition 5.4](#). We require that, for every QPT server  $\mathcal{A}$  and every pair of equal-length inputs  $(x_0, x_1)$ , the views of the server are computationally indistinguishable:

$$\text{View}_{\mathcal{A}}(\rho, x_0) \approx_c \text{View}_{\mathcal{A}}(\rho, x_1).$$

[BLM26] shows how to combine a blind delegation protocol from [BK25, BKM<sup>+</sup>25] with a two-prover game from [MNZ24] to obtain the desired round-efficient argument. These protocols however require the existence of public-key assumptions such as oblivious state preparation or trapdoor claw-free functions. Constructing blind delegation from collapsing hashes alone appears unlikely (or at least very difficult given known techniques), since blind delegation implies (classical-communication) key agreement [BK25].

**Idea I: Relaxing the Model.** Our first simple, but crucial, observation is that the client does not need to be completely classical in the above interaction. For instance, it could be acceptable to construct a blind delegation protocol where the client performs a *fixed polynomial amount* of quantum operations, so long as we can still apply the aforementioned communication-compression compiler to the (potentially long) classical messages. Henceforth, we refer to such a protocol as *quantum succinct*. With this observation in mind, let us try to simplify the problem even further: In [BK25] it is shown that a blind delegation protocol can be generically derived from *claw-state generation*, an interactive protocol where at the end, the server obtains multiple claw states

$$\frac{1}{\sqrt{2}}(|0, x_0\rangle + |1, x_1\rangle)$$

and the client obtains a classical description of their labels  $(x_0, x_1)$ . The only security guarantee that we require is that the server should not be able to guess *both*  $x_0$  and  $x_1$ . Thus, our problem further reduces to designing a round-efficient and quantum-succinct claw-state generation protocol: Allowing the client to use  $\text{poly}(\lambda, \log T)$  quantum gates and a (possibly large) amount of classical communication, we want an interactive protocol where the server obtains  $T$  such claw states. Climbing up the chain of implications explained above, this will suffice for our main result.

In fact, a very similar problem was considered by Zhang [Zha21], precisely motivated by the construction of a blind delegation protocol with a quantum-succinct client. He shows how to *bootstrap* a small initial quantum communication into an arbitrary polynomial number of claw states, using classical communication only. However his approach has two limitations: First, the round complexity of his protocol is polynomial in  $T$ , rendering it unusable for us since we require a round-efficient protocol. Second, the security is proved in the quantum random-oracle model. This is also problematic, since the communication-compression compiler makes non-black-box use of the underlying protocol (in this case, it would require the circuit description of the underlying hash function). This would render the security claim of the underlying protocol vacuous.

Instead, we propose a new approach to quantum-succinct claw-state generation, which relies only on (quantum secure) one-way functions and has the added benefit of being dramatically simpler than Zhang’s protocol. Next, we give a more detailed overview of our protocol.

**Idea II: Quantum-Succinct Claw-Generation from One-Way Functions.** Assume that  $T$  is a power of 2 and set  $n = \lambda$ . Prior to the interaction, the client samples several PRF keys denoted by  $k \in \{0, 1\}^\lambda$  and  $k_{i,b} \in \{0, 1\}^\lambda$ , along with random strings  $x_{i,b} \in \{0, 1\}^n$ , for  $i \in [\log T]$  and  $b \in \{0, 1\}$ . For any index  $t \in [T] \cong \{0, 1\}^{\log T}$  we will think of  $(y_{t,0}, y_{t,1}) = f_k(t)$  as the strings that determine the  $t$ -th output claw. The client prepares and sends the quantum state

$$\frac{1}{\sqrt{T}} \sum_{t \in \{0,1\}^{\log T}} \left( \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) \frac{|0, y_{t,0}\rangle + |1, y_{t,1}\rangle}{\sqrt{2}} \quad (1)$$

to the server, which can be prepared efficiently (that is, in time  $\text{poly}(\lambda, \log T)$ ) with the knowledge of the client’s randomness. In addition the client also sends the large, but classical, table

$$c_{t,i} = f_k(t) \oplus f_{k_{i,1-t_i}}(t) \quad \text{for } t \in \{0, 1\}^{\log T}, i \in [\log T]. \quad (2)$$

To get some intuition, it is useful to think about each of the  $T$  states in the superposition of Equation (1): In the “ $t$ -subspace”, the server can compute  $(y_{t',0}, y_{t',1})$  for every  $t'$  *except for*  $t$ . Indeed, for  $t' \neq t$ , choose any  $i$  with  $t'_i \neq t_i$ . Then  $c_{t',i} \oplus f_{k_{i,t'_i}}(t') = f_k(t') = (y_{t',0}, y_{t',1})$ . We also mention that while the random  $x_{i,t_i}$  are not used in this derivation, their presence will be useful for arguing security.

Coherently preparing the missing claws and placing every output at its designated index, the server can then prepare the state

$$\left( \bigotimes_{i \in [\log T]} \frac{|0, x_{i,0}, k_{i,0}\rangle + |1, x_{i,1}, k_{i,1}\rangle}{\sqrt{2}} \right) \otimes \bigotimes_{t \in \{0,1\}^{\log T}} \frac{|0, y_{t,0}\rangle + |1, y_{t,1}\rangle}{\sqrt{2}}. \quad (3)$$

To complete the protocol, the server measures the first registers in the Hadamard basis and returns  $d = (d_1, \dots, d_{\log T})$ , parsed as  $d_i = (d_i^{(0)}, d_i^{(1)}, d_i^{(2)}) \in \{0, 1\} \times \{0, 1\}^n \times \{0, 1\}^\lambda$ . The client accepts exactly when, for every  $i$ ,

$$d_i^{(1)} \neq 0, \quad \langle d_i, (1, x_{i,0} \oplus x_{i,1}, k_{i,0} \oplus k_{i,1}) \rangle = 0. \quad (4)$$

Honest Hadamard measurements satisfy the parity constraint, and  $d_i^{(1)} = 0$  occurs with negligible probability. The honest server thereby deletes the first registers and retains the output claw states. Moreover, notice that the round complexity of this protocol is constant, and the client’s quantum

operations are bounded by  $\text{poly}(\log T, n, \lambda) = \text{poly}(\lambda)$ , whereas the client’s classical runtime and the server’s total runtime are  $\text{poly}(T, \lambda)$ .

For security, we will only be able to prove that, for a *random* choice of index  $t^* \leftarrow [T]$  sampled at the conclusion of the protocol, the adversary cannot both pass Equation (4) and output  $f_k(t^*) = (y_{t^*,0}, y_{t^*,1})$  with better than *constant* probability. We refer the reader to the technical sections for a precise analysis, and here we just say that the proof will boil down to splitting the state of the adversary into two components:

- The first component will be bounded by the success probability of an experiment where the adversary is given *only* the part of the state indexed by  $t^*$ , instead of the superposition in Equation (1), which allows us to appeal to the security of the PRF and replace the table entries  $c_{t^*,i}$  by uniform strings and  $f_k(t^*)$  by a uniform pair. In this hybrid experiment, the success probability of outputting  $f_k(t^*)$  can therefore be bounded by negligible.
- The second component is a residual “unbalanced” state derived by subtracting off the (appropriately weighted)  $t^*$  branch, and whose averaged squared norm after the deletion check can only be at most  $\frac{1}{4}$ : each individual branch passes the checks with probability at most  $\frac{1}{T}$ , and the imbalance in the amplitudes of the branches can then be used to derive the  $\frac{1}{4}$  bound.

Overall, we combine the two bounds using standard inequalities, and obtain a total bound on passing Equation (4) and guessing the randomly selected output pair by  $\frac{1}{4}$  (plus an irrelevant negligible term). The next goal is therefore to *amplify* the hardness of these output claw states.

**Hardness Amplification.** To amplify, we use an idea from Zhang [Zha21], which consists in running  $\lambda$  independent copies of the expansion protocol sequentially and then combining  $\lambda$  weakly-secure claw states into one strongly-secure claw state. Only after all copies terminate, the client chooses independent uniform permutations  $\pi_i \in S_T$  for each repetition  $i \in [\lambda]$  and sends them to the server. For an output index  $t$ , the server combines the  $\lambda$  claws indexed by  $\{\pi_i(t)\}_{i \in [\lambda]}$  by measuring the XOR of the first claw’s leading qubit with each other leading qubit, obtaining the outcomes  $e_{2,t}, \dots, e_{\lambda,t}$ . This effectively *glues* together independent claw states into a single claw  $\frac{|0, Y_{t,0}\rangle + |1, Y_{t,1}\rangle}{\sqrt{2}}$ , where

$$\begin{aligned} Y_{t,0} &= y_{\pi_1(t),0}^{(1)} \parallel y_{\pi_2(t),e_{2,t}}^{(2)} \parallel \dots \parallel y_{\pi_\lambda(t),e_{\lambda,t}}^{(\lambda)}, \\ Y_{t,1} &= y_{\pi_1(t),1}^{(1)} \parallel y_{\pi_2(t),1 \oplus e_{2,t}}^{(2)} \parallel \dots \parallel y_{\pi_\lambda(t),1 \oplus e_{\lambda,t}}^{(\lambda)}. \end{aligned}$$

The important observation is that guessing  $Y_{t,0}, Y_{t,1}$  implies also guessing *all*  $\lambda$  constituents simultaneously. We can therefore hope to conclude that, if each individual claw is  $\frac{1}{4}$  secure, then the probability of guessing the combined claw state must be bounded by  $4^{-\lambda} + \text{negl}(\lambda)$ .

**Idea III: Sequential Repetition for Quantum Search Games.** To make this intuition formal, we prove a general *sequential repetition theorem for quantum search games*: A quantum search game is an interactive protocol between a prover and a verifier, where the goal of the prover is to guess some secret  $x$ , and the verifier can either accept or reject. In  $\ell$  sequential copies, the verifier uses independent fresh randomness for each, the adversary retains arbitrary quantum memory, and it may postpone *all* predictions until every interaction has finished. We show that the combined success probability degrades exponentially with  $\ell$ . As usual in this context, the difficulty is that

conditioning on success in other copies can disturb the quantum memory, and the verifier’s secret target is unavailable to a reduction.

For any round  $j$ , we denote by  $q_j$  the maximum probability, over arbitrary starting states just before copy  $j$ , of guessing all targets from  $j$  through  $\ell$  using the remaining strategy. Our goal will be to prove that  $q_j \lesssim \varepsilon q_{j+1}$  where  $\varepsilon$  is the bound on the single search game, with a reduction to the bound on the single-copy search game. We can imagine giving the reduction the best-possible quantum state as non-uniform advice, then let the reduction interact with the challenger for the  $j$ -th copy and simulate the rest of the interactions locally. If  $q_{j+1}$  is sufficiently close to one, simply running that suffix already contradicts the one-copy bound. Otherwise, we need to amplify the successful part of the suffix *coherently*. We do this using the quantum singular value transformation (QSVT) [GSLW19], which is a powerful *quantum rewinding* technique. In particular, rewinding enables us to effectively condition on the adversary successfully answering targets  $j + 1$  through  $\ell$  (note that the reduction knows what these targets are, the only missing information is target  $j$ ), and only then output its guess for target  $j$ . This increases its overall probability of success by a factor of roughly  $1/q_{j+1}$ , so if we had assumed for contradiction that  $q_j > \varepsilon q_{j+1}$ , then our reduction outputs the  $j$ -th target with probability  $> \varepsilon$ , a contradiction. We refer the reader to the technical sections for more details.

**Putting Things Together.** Taking a step back, we have therefore obtained, from quantum-secure one-way functions, a  $\text{poly}(\lambda)$ -round claw-state generation protocol for any polynomial  $T$ : At the end of the interaction, the honest server holds  $\bigotimes_t \frac{|0, Y_{t,0}\rangle + |1, Y_{t,1}\rangle}{\sqrt{2}}$ , and outputting any classical pair is computationally hard. Quantum communication is only client-to-server, with  $\text{poly}(\lambda, \log T)$  preparation cost and circuits fixed before interaction. The [BK25] transformation converts these correlations into a blind delegation protocol with the same efficiency properties, which supplies the missing ingredient for the communication-compression compiler of [BLM26]. Combining all the steps, we obtain a succinct argument for QMA from collapsing hash functions.

### 1.3 AI Disclosure

The overall architecture and general proof strategy were developed without AI assistance. ChatGPT was used to refine the protocol, assist with technical proofs, revise the exposition, and check references. The authors take full responsibility for the content.

## 2 Preliminaries

Let  $\lambda$  be the security parameter. We write  $[n] = \{1, \dots, n\}$  and use logarithms to base 2. For a finite set  $S$ , we write  $s \sim S$  when  $s$  is sampled uniformly from  $S$ . All inner products between bit strings are over  $\mathbb{F}_2$ .

All registers are finite-dimensional Hilbert spaces. We consider security against non-uniform QPT adversaries that may initialize a polynomial-size register to an arbitrary non-uniform state independent of fresh honest randomness. All cryptographic assumptions and computational indistinguishability statements quantify over adversaries with such auxiliary input; negligible terms may depend on the adversary. Instance and circuit sizes are polynomially bounded in the security parameter, while the polynomials in succinctness bounds are fixed independently of the original verification time and witness size.

We use sans-serif letters such as  $A, B, R$  for registers,  $\mathcal{H}_A$  for the Hilbert space of register  $A$ , and  $I_A$  for its identity operator. Register labels on states and operators indicate where they act; omitted registers are left unchanged. Braces with subscripts denote indexed lists. We write  $\parallel$  for concatenation of bit strings,  $\|\cdot\|$  for vector or operator norm, and

$$\|A\|_1 := \text{Tr} \sqrt{A^\dagger A}$$

for the trace norm. The trace distance of states  $\rho, \sigma$  is  $\frac{1}{2}\|\rho - \sigma\|_1$ . A POVM is a family of positive operators summing to the identity. We write  $\approx_c$  for computational indistinguishability and  $\text{negl}(\lambda)$  for a function smaller than every inverse polynomial.

We use compressing collapsing hash functions in the sense of [Unr16]: After a hash is evaluated coherently and its output is measured, additionally measuring the input is computationally undetectable, even given the remaining workspace. We use this property through the compiler of [BLM26]. All one-way functions and pseudorandom functions are assumed secure against the quantum adversaries specified above.

## 2.1 Succinct Arguments for QMA

Fix a QMA promise problem  $L = (L_{\text{yes}}, L_{\text{no}})$ , with a polynomial-time uniform quantum verifier  $V = \{V_x\}$ . For  $x \in L_{\text{yes}}$ , there exists a polynomial-size quantum witness  $\rho$  such that  $V_x(\rho)$  accepts with probability at least  $2/3$ ; for  $x \in L_{\text{no}}$ , no witness causes  $V_x$  to accept with probability greater than  $1/3$ . We recall the definition of succinct arguments for QMA below [BTL<sup>+</sup>22].

**Definition 2.1** (Succinct Argument for QMA). *A succinct argument for the QMA promise problem  $L$  is an interactive protocol  $\Pi$  between a prover and a verifier where the common input is  $(1^\lambda, x)$  and the prover additionally receives a quantum state  $\rho$ . The protocol satisfies the following properties.*

- *Correctness: For  $x \in L_{\text{yes}}$ , there exists a QPT prover that, on input a valid witness  $\rho$ , causes the verifier to accept with probability at least  $\frac{2}{3}$ .*
- *Soundness: For  $x \in L_{\text{no}}$  and all QPT provers on input an arbitrary state  $\rho$ , the probability that the verifier accepts is bounded by  $\frac{1}{3}$ .*
- *Succinctness: The communication complexity of the protocol is bounded by  $\text{poly}(\lambda)$  and the verifier's complexity is bounded by  $|x| \cdot \text{poly}(\lambda)$ .*

## 2.2 Claw States

We refer to states of the form

$$|\psi_{x_0, x_1}\rangle := \frac{|0, x_0\rangle + |1, x_1\rangle}{\sqrt{2}}$$

for  $x_0, x_1 \in \{0, 1\}^n$ , as *claw states*. Let  $\rho_{x_0, x_1} := |\psi_{x_0, x_1}\rangle\langle\psi_{x_0, x_1}|$ , it is easy to see that  $\rho_{x_0, x_1}$  can be prepared by a polynomial-size quantum circuit, given  $(x_0, x_1)$ . The following lemma shows that a uniformly sampled claw state is hard to guess.

**Lemma 2.2.** *Consider the following experiment:*

- *Sample  $x_0, x_1 \sim \{0, 1\}^n$ , then prepare the state  $\rho_{x_0, x_1}$  and send it to the adversary.*
- *The adversary returns  $(x_0^*, x_1^*)$ .*

- The adversary wins if  $(x_0, x_1) = (x_0^*, x_1^*)$ .

The probability that any adversary succeeds in this experiment is at most  $2^{1-n}$ .

*Proof.* Let  $\mathcal{C}$  be the claw register, with  $\mathcal{H}_{\mathcal{C}} \cong \mathbb{C}^{2^{n+1}}$ , and let  $\mathcal{A}$  be the adversary's auxiliary register. We may assume without loss of generality that its initial state is pure, say  $|\alpha\rangle_{\mathcal{A}}$ : a mixed state can be purified by enlarging  $\mathcal{A}$ . This state is independent of  $(x_0, x_1)$ .

Let  $\{M_{x_0, x_1}\}_{x_0, x_1}$  be the POVM on  $\mathcal{C}\mathcal{A}$  describing the adversary's guess. To obtain a measurement on the claw register alone, define

$$N_{x_0, x_1} := (I_{\mathcal{C}} \otimes \langle \alpha |_{\mathcal{A}}) M_{x_0, x_1} (I_{\mathcal{C}} \otimes |\alpha\rangle_{\mathcal{A}}).$$

Each  $N_{x_0, x_1}$  is positive, and  $\sum_{x_0, x_1} N_{x_0, x_1} = I_{\mathcal{C}}$ , thus these operators form a POVM. They give exactly the same guessing probabilities because

$$\text{Tr}(M_{x_0, x_1}(\rho_{x_0, x_1} \otimes |\alpha\rangle\langle\alpha|)) = \text{Tr}(N_{x_0, x_1} \rho_{x_0, x_1}).$$

The adversary's success probability is therefore

$$\begin{aligned} \frac{1}{2^{2n}} \sum_{x_0, x_1} \text{Tr}(N_{x_0, x_1} \rho_{x_0, x_1}) &\leq \frac{1}{2^{2n}} \sum_{x_0, x_1} \text{Tr}(N_{x_0, x_1}) \\ &= \frac{1}{2^{2n}} \text{Tr}\left(\sum_{x_0, x_1} N_{x_0, x_1}\right) \\ &= \frac{1}{2^{2n}} \text{Tr}(I_{\mathcal{C}}) \\ &= 2^{1-n}, \end{aligned}$$

where the inequality uses  $\rho_{x_0, x_1} \preceq I_{\mathcal{C}}$ . □

## 2.3 Pseudorandom Functions

Let  $T$  be a power of 2 and  $f_k : \{0, 1\}^{\log T} \rightarrow \{0, 1\}^{2n}$  be a keyed function, for  $k \in \{0, 1\}^{\lambda}$ . We say that  $f_k$  is a pseudorandom function if its table of outputs is computationally indistinguishable from a uniformly random table. In this work we consider the special case where  $T = \text{poly}(\lambda)$  and we assume that the function satisfies the following notion of pseudorandomness, which is implied by the regular definition of a pseudorandom function.

**Definition 2.3** (Pseudorandomness). *A function  $f_k : \{0, 1\}^{\log T} \rightarrow \{0, 1\}^{2n}$  is pseudorandom if, for all  $x^* \in \{0, 1\}^{\log T}$ , the following distributions are computationally indistinguishable*

$$(\{f_k(x)\}_{x \neq x^*}, f_k(x^*)) \approx_c (\{f_k(x)\}_{x \neq x^*}, y^*).$$

Here  $k \sim \{0, 1\}^{\lambda}$  and  $y^* \sim \{0, 1\}^{2n}$  are independent, and the non-target entries are listed in the same fixed order on both sides.

Such pseudorandom functions can be constructed assuming the existence of (quantum-secure) one-way functions [Zha12].

### 3 Claw-State Generation Protocol

#### 3.1 Protocol Description

Set  $n = \lambda$  and let  $T = \text{poly}(\lambda)$  be a power of 2. We consider the following interactive protocol between a client and a server.

- (i) The client samples  $k \sim \{0, 1\}^\lambda$ , then for all  $i \in [\log T]$  and  $b \in \{0, 1\}$  it samples  $x_{i,b} \sim \{0, 1\}^n$  and  $k_{i,b} \sim \{0, 1\}^\lambda$ .
- (ii) For all  $t \in \{0, 1\}^{\log T}$  and  $i \in [\log T]$ , the client sends to the server

$$c_{t,i} := f_k(t) \oplus f_{k_{i,1-t_i}}(t) \in \{0, 1\}^{2n}.$$

- (iii) The client computes the state

$$\bigotimes_{i \in [\log T]} \left( \frac{|0, x_{i,0}, k_{i,0}\rangle + |1, x_{i,1}, k_{i,1}\rangle}{\sqrt{2}} \right) = \frac{1}{\sqrt{T}} \sum_{t \in \{0,1\}^{\log T}} \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle.$$

Let  $(y_{t,0}, y_{t,1}) = f_k(t)$ . The client applies the isometry

$$|t\rangle |0\rangle \mapsto |t\rangle |\psi_{y_{t,0}, y_{t,1}}\rangle$$

to the leading qubits of the above state and a fresh output register to obtain

$$\frac{1}{\sqrt{T}} \sum_{t \in \{0,1\}^{\log T}} \left( \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) \otimes |\psi_{y_{t,0}, y_{t,1}}\rangle.$$

The client returns its auxiliary qubits to zero and sends the resulting state to the server.

- (iv) The server maps the state received from the client to

$$\left( \frac{1}{\sqrt{T}} \sum_{t \in \{0,1\}^{\log T}} \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) \bigotimes_{t \in \{0,1\}^{\log T}} |\psi_{y_{t,0}, y_{t,1}}\rangle.$$

Note that this isometry can be implemented efficiently, given  $\{c_{t,i}\}_{t,i}$ . Indeed, for any fixed  $t$  and any  $t' \neq t$ , the server can recover  $f_k(t')$  by selecting an index  $i$  such that  $t'_i \neq t_i$ , so that  $k_{i,1-t'_i} = k_{i,t_i}$ , and computing  $c_{t',i} \oplus f_{k_{i,t_i}}(t') = f_k(t') = (y_{t',0}, y_{t',1})$ . Then the server can compute the claw state  $|\psi_{y_{t',0}, y_{t',1}}\rangle$  by evaluating the claw-preparation circuit. Running this algorithm coherently and reordering the registers leads to the state as described above.

- (v) The server measures its first registers in the Hadamard basis to obtain  $d = (d_1, \dots, d_{\log T})$ , where  $d_i \in \{0, 1\} \times \{0, 1\}^n \times \{0, 1\}^\lambda$ , which are sent to the client.
- (vi) The client parses  $d_i = (d_i^{(0)}, d_i^{(1)}, d_i^{(2)})$ , and for all  $i \in [\log T]$  it checks that  $d_i^{(1)} \neq 0$  and that

$$\langle d_i, (1, x_{i,0} \oplus x_{i,1}, k_{i,0} \oplus k_{i,1}) \rangle = 0.$$

We say that the server *passes* the protocol if all checks of the client succeed. If this is the case, the client returns  $\{y_{t,0}, y_{t,1}\}_{t \in \{0,1\}^{\log T}}$  as its private local output.

Note that the client's quantum operations are confined in step (iii) and their runtime is bounded by some  $\text{poly}(\log T, n, \lambda)$ . On the other hand, the client's classical runtime as well as the server's total runtime is bounded by  $\text{poly}(T, n, \lambda)$ .

Moreover, it is easy to see that an honest server passes the protocol except with probability at most  $(\log T)2^{-n}$ , since the probability that  $d_i^{(1)} = 0$  is exactly  $2^{-n}$ .

### 3.2 Analysis

Next, we show that the output claw states are hard to guess. Assuming  $n = \Omega(\lambda)$ , we show the following.

**Theorem 3.1.** *Consider the following experiment:*

- Run the client-server protocol from [Section 3.1](#).
- After the server sends  $d$ , the client samples  $t^* \sim \{0, 1\}^{\log T}$  and sends this index and the strings  $\{y_{t,0}, y_{t,1}\}_{t \neq t^*}$  to the server.
- The server replies with  $(y_0^*, y_1^*)$  and wins if  $(y_0^*, y_1^*) = (y_{t^*,0}, y_{t^*,1})$  and  $d$  passes the client's checks.

Assuming quantum-secure one-way functions, for every QPT server there is a negligible function  $\mu$  such that its success probability is at most  $1/4 + \mu(\lambda)$ .

*Proof.* Let  $X$  contain the client's random choices of  $k$  and  $\{x_{i,b}, k_{i,b}\}_{i \in [\log T], b \in \{0,1\}}$ , and let  $\text{Valid}_X(d)$  be its verification predicate. We use the following registers throughout this proof and the two lemmas below:  $\mathbf{C}$  contains the client's quantum message,  $\mathbf{A}$  contains the server's pure auxiliary input  $|\alpha\rangle$ ,  $\mathbf{D}$  contains the client's copy of  $d$ ,  $\mathbf{B}$  contains all registers retained by the server after sending  $d$ , and  $\mathbf{L}$  contains the client's final classical message.

Let  $\mathcal{A} : \mathbf{CA} \rightarrow \mathbf{DB}$  be a unitary dilation of the server's first stage. Its dependence on the classical table  $\{c_{t,i}\}_{t,i}$  is implicit. The server may retain its own copy of  $d$  in  $\mathbf{B}$ , but all its subsequent operations act trivially on  $\mathbf{D}$ . Define

$$\Pi_X := \sum_{d: \text{Valid}_X(d)=1} |d\rangle\langle d|_{\mathbf{D}} \otimes I_{\mathbf{B}}.$$

For fixed  $X$ , the client sends its final message by applying the isometry

$$V_t |\varphi\rangle_{\mathbf{DB}} = |\varphi\rangle_{\mathbf{DB}} \otimes |t, \{y_{u,0}, y_{u,1}\}_{u \neq t}\rangle_{\mathbf{L}}.$$

The dependence of  $V_t$  on  $X$  is implicit. Let  $\{M_{t,y_0,y_1}\}_{y_0,y_1 \in \{0,1\}^n}$  be the POVM on  $\mathbf{BL}$  describing the server's second message, extended by the identity on  $\mathbf{D}$ .

We omit register subscripts in the remaining state formulas. The subnormalized state for which the client accepts, right before the final measurement, is

$$|\xi_{X,t^*}\rangle := \frac{1}{\sqrt{T}} V_{t^*} \sum_t \Pi_X \mathcal{A} \left( \left( \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) |\psi_{y_{t,0}, y_{t,1}}\rangle |\alpha\rangle \right).$$

Define also the possibly subnormalized state

$$|\zeta_{X,t^*}\rangle := V_{t^*} \Pi_X \mathcal{A} \left( \left( \bigotimes_{i \in [\log T]} |t_i^*, x_{i,t_i^*}, k_{i,t_i^*}\rangle \right) |\psi_{y_{t^*,0}, y_{t^*,1}}\rangle |\alpha\rangle \right).$$

Both states are on DBL. We use the decomposition

$$|\xi_{X,t^*}\rangle = \left( |\xi_{X,t^*}\rangle - \frac{\sqrt{T}}{2} |\zeta_{X,t^*}\rangle \right) + \frac{\sqrt{T}}{2} |\zeta_{X,t^*}\rangle.$$

Intuitively, the second summand isolates the contribution obtained when the server is given the single claw state  $|\psi_{y_{t^*,0}, y_{t^*,1}}\rangle$ , which we will bound by the hardness of recovering both strings from one claw state. The residual term can be bounded using only the verification constraints. The coefficient  $\sqrt{T}/2$  is chosen so that, after averaging over  $t^*$ , the cross term in the expansion of the residual squared norm cancels its  $\| |\xi_{X,t^*}\rangle \|^2$  term.

By Cauchy–Schwarz and using  $0 \preceq M_{t^*, y_{t^*,0}, y_{t^*,1}} \preceq I$ , the square root of the prover’s total success probability satisfies

$$\begin{aligned} & \sqrt{\frac{1}{T} \sum_{t^*} \mathbb{E}_X \langle \xi_{X,t^*} | M_{t^*, y_{t^*,0}, y_{t^*,1}} | \xi_{X,t^*} \rangle} \\ & \leq \sqrt{\frac{1}{T} \sum_{t^*} \mathbb{E}_X \left\| |\xi_{X,t^*}\rangle - \frac{\sqrt{T}}{2} |\zeta_{X,t^*}\rangle \right\|^2} + \frac{1}{2} \sqrt{\sum_{t^*} \mathbb{E}_X \langle \zeta_{X,t^*} | M_{t^*, y_{t^*,0}, y_{t^*,1}} | \zeta_{X,t^*} \rangle}. \end{aligned} \tag{5}$$

For fixed  $X$ , the definitions and the isometry property give

$$V_t^\dagger |\xi_{X,t}\rangle = \frac{1}{\sqrt{T}} \sum_u V_u^\dagger |\zeta_{X,u}\rangle.$$

In particular, the left-hand side is independent of  $t$ . Thus

$$\sum_t \operatorname{Re} \langle \xi_{X,t} | \zeta_{X,t} \rangle = \sqrt{T} \| |\xi_{X,t^*}\rangle \|^2 \quad \text{for every } t^*.$$

Expanding the squared norm on the right-hand side of Equation (5) now gives

$$\begin{aligned} \frac{1}{T} \sum_t \left\| |\xi_{X,t}\rangle - \frac{\sqrt{T}}{2} |\zeta_{X,t}\rangle \right\|^2 &= \frac{1}{T} \sum_t \| |\xi_{X,t}\rangle \|^2 + \frac{1}{4} \sum_t \| |\zeta_{X,t}\rangle \|^2 - \frac{1}{\sqrt{T}} \sum_t \operatorname{Re} \langle \xi_{X,t} | \zeta_{X,t} \rangle \\ &= \frac{1}{4} \sum_t \| |\zeta_{X,t}\rangle \|^2. \end{aligned}$$

Averaging over  $X$  and applying Lemma 3.2 (proven below) bounds the expression under the first square root on the right-hand side of Equation (5) by  $1/4$ . For the second term, Lemma 3.3 (proven below) gives

$$\frac{1}{4} \sum_{t^*} \mathbb{E}_X \langle \zeta_{X,t^*} | M_{t^*, y_{t^*,0}, y_{t^*,1}} | \zeta_{X,t^*} \rangle \leq \frac{2^{1-n} + \mu(\lambda)}{4}.$$

Combining these bounds, the success probability is at most

$$\left( \frac{1}{2} + \frac{1}{2} \sqrt{2^{1-n} + \mu(\lambda)} \right)^2 = \frac{1}{4} + \operatorname{negl}(\lambda),$$

since  $n = \Omega(\lambda)$ . □

Using the notation introduced above, we prove the two remaining technical statements in the following.

**Lemma 3.2.**

$$\sum_{t \in \{0,1\}^{\log T}} \mathbb{E}_X \|\zeta_{X,t}\|^2 \leq 1.$$

*Proof.* Fix  $t \in \{0,1\}^{\log T}$  and condition on all client randomness except  $\{x_{i,1-t_i}\}_{i \in [\log T]}$ . Write

$$\mathcal{A} \left( \left( \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) |\psi_{y_{t,0}, y_{t,1}}\rangle |\alpha\rangle \right) = \sum_d |d\rangle |\phi_{t,d}\rangle$$

for possibly subnormalized states  $|\phi_{t,d}\rangle$  on  $\mathcal{B}$ , with  $|d\rangle$  on  $\mathcal{D}$ . These states are fixed under the conditioning: the input state contains only  $\{x_{i,t_i}\}_{i \in [\log T]}$ , and the classical table  $\{c_{t,i}\}_{t,i}$  depends only on the keys. Since  $V_t$  is an isometry,

$$\begin{aligned} \mathbb{E}_{\{x_{i,1-t_i}\}_i} \|\zeta_{X,t}\|^2 &= \mathbb{E}_{\{x_{i,1-t_i}\}_i} \left\| \Pi_X \sum_d |d\rangle |\phi_{t,d}\rangle \right\|^2 \\ &= \sum_d \mathbb{E}_{\{x_{i,1-t_i}\}_i} \mathbf{1}_{\text{Valid}_X(d)=1} \|\phi_{t,d}\|^2 \\ &\leq \frac{1}{T} \sum_d \|\phi_{t,d}\|^2 = \frac{1}{T}. \end{aligned}$$

The second equality uses the orthogonality of the states  $|d\rangle$ , and the inequality follows from

$$\mathbb{E}_{\{x_{i,1-t_i}\}_i} \mathbf{1}_{\text{Valid}_X(d)=1} = \frac{1}{T} \mathbf{1}_{\forall i \in [\log T], d_i^{(1)} \neq 0}. \quad (6)$$

Indeed, if  $d_i^{(1)} = 0$  for some  $i$ , both sides vanish. Otherwise, each verification check is a nonconstant affine equation in the independent uniform string  $x_{i,1-t_i}$ , and hence holds with probability exactly  $1/2$ . There are  $\log T$  independent checks, so their joint probability is  $2^{-\log T} = 1/T$ .

Averaging over the remaining client randomness and summing over  $t$  proves the claim.  $\square$

**Lemma 3.3.** *Assuming quantum-secure one-way functions, for every QPT server there is a negligible function  $\mu$  such that*

$$\sum_{t \in \{0,1\}^{\log T}} \mathbb{E}_X \langle \zeta_{X,t} | M_{t,y_{t,0},y_{t,1}} | \zeta_{X,t} \rangle \leq 2^{1-n} + \mu(\lambda).$$

*Proof.* Consider the following modified experiment:

- Sample  $t^* \sim \{0,1\}^{\log T}$ , then generate the client's complete setup and classical table  $\{c_{t,i}\}_{t,i}$  as in [Section 3.1](#), except that the client sends the state

$$\left( \bigotimes_{i=1}^{\log T} |t_i^*, x_{i,t_i^*}, k_{i,t_i^*}\rangle \right) \otimes |\psi_{y_{t^*,0}, y_{t^*,1}}\rangle$$

to the server.

- After the server sends  $d$ , the client always sends  $t^*$  and  $\{y_{t,0}, y_{t,1}\}_{t \neq t^*}$  to the server.
- The server replies with  $(y_0^*, y_1^*)$ . It wins if  $d_i^{(1)} \neq 0$  for all  $i \in [\log T]$  and

$$(y_0^*, y_1^*) = (y_{t^*,0}, y_{t^*,1}).$$

We first show that the success probability in this modified experiment equals the sum in the statement. Fix  $t$  and condition on all client randomness except  $\{x_{i,1-t_i}\}_{i \in [\log T]}$ . As in [Lemma 3.2](#), write

$$\mathcal{A} \left( \left( \bigotimes_{i \in [\log T]} |t_i, x_{i,t_i}, k_{i,t_i}\rangle \right) |\psi_{y_{t,0}, y_{t,1}}\rangle |\alpha\rangle \right) = \sum_d |d\rangle |\phi_{t,d}\rangle.$$

The states  $|\phi_{t,d}\rangle$ , the isometry  $V_t$ , and the operator  $M_{t,y_{t,0}, y_{t,1}}$  are fixed under this conditioning. Since  $V_t$  and  $M_{t,y_{t,0}, y_{t,1}}$  act trivially on  $D$  and  $\langle d' | d \rangle = 0$  for  $d' \neq d$ , all terms with different values of  $d$  vanish. Therefore,

$$\langle \zeta_{X,t} | M_{t,y_{t,0}, y_{t,1}} | \zeta_{X,t} \rangle = \sum_d \mathbf{1}_{\text{Valid}_X(d)=1} \cdot (\langle d | \langle \phi_{t,d} |) V_t^\dagger M_{t,y_{t,0}, y_{t,1}} V_t (|d\rangle |\phi_{t,d}\rangle).$$

Every quadratic form on the right is nonnegative and independent of the hidden strings. Applying [Equation \(6\)](#) therefore gives

$$\mathbb{E}_{\{x_{i,1-t_i}\}_i} \langle \zeta_{X,t} | M_{t,y_{t,0}, y_{t,1}} | \zeta_{X,t} \rangle = \frac{1}{T} \sum_{d: \forall i, d_i^{(1)} \neq 0} (\langle d | \langle \phi_{t,d} |) V_t^\dagger M_{t,y_{t,0}, y_{t,1}} V_t (|d\rangle |\phi_{t,d}\rangle).$$

Averaging over the remaining randomness and summing over  $t$  yields

$$\sum_t \mathbb{E}_X \langle \zeta_{X,t} | M_{t,y_{t,0}, y_{t,1}} | \zeta_{X,t} \rangle = \frac{1}{T} \sum_t \mathbb{E}_X \sum_{d: \forall i, d_i^{(1)} \neq 0} (\langle d | \langle \phi_{t,d} |) V_t^\dagger M_{t,y_{t,0}, y_{t,1}} V_t (|d\rangle |\phi_{t,d}\rangle).$$

The right-hand side is exactly the success probability of the modified experiment: the factor  $1/T$  samples  $t^*$  uniformly, the restriction on  $d$  enforces the nonzero checks, and the POVM element enforces the correct output.

It remains to bound this probability. Fix a target  $t^* \in \{0, 1\}^{\log T}$  and consider the following hybrids:

- **Hybrid 0:** This is the modified experiment above, with target  $t^*$ .
- **Hybrid 1:** Replace the entries  $\{c_{t^*,i}\}_{i \in [\log T]}$  by independent uniform strings.

We make these replacements one at a time. For the  $i$ th replacement, the masking key  $k_{i,1-t_i^*}$  does not occur in the state sent to the server. A reduction uses the PRF challenge table for this key to construct the entire table, changing only its value at  $t^*$ . It samples the other keys itself and simulates all remaining messages. Because the parity checks have been omitted, the simulation never needs the key  $k_{i,1-t_i^*}$  itself. Thus, [Definition 2.3](#) makes each replacement computationally indistinguishable.

- **Hybrid 2:** Replace  $(y_{t^*,0}, y_{t^*,1}) = f_k(t^*)$  by an independent uniform pair.

The  $t^*$ -th row of the classical table is now uniform and independent of this pair. A reduction can therefore use its PRF challenge table for the master key  $k$  to prepare the target claw and all remaining messages. Again, only the value at  $t^*$  changes, so [Definition 2.3](#) makes this hybrid computationally indistinguishable from Hybrid 1.

Note that these simulations are efficient because the PRF domain has size  $T = \text{poly}(\lambda)$ .

In Hybrid 2, the target pair is uniform, and the server receives only its claw state together with data independent of that pair. Even after dropping the nonzero checks, its success probability is at most  $2^{1-n}$ , by [Lemma 2.2](#). The  $\log T + 1$  replacements change this probability by a negligible amount. Averaging over  $t^*$  proves the statement; non-uniform security makes the negligible bound uniform over the polynomially many targets.  $\square$

## 4 Sequential Repetition

In this section, we give a generic sequential repetition theorem for any quantum interactive search game.

**Definition 4.1** (Quantum interactive search game). *A quantum interactive search game  $\Gamma$  is specified by a polynomial-size quantum interactive circuit family  $V = \{V_\lambda\}_\lambda$ . At the end of an execution it outputs a classical string*

$$x \in \{0, 1\}^{m(\lambda)} \cup \{\perp\}$$

*in a designated classical register  $\mathsf{X}$ .*

*For any polynomial-size interactive strategy  $B$  with input register  $\mathsf{M}$  and output register  $\mathsf{M}'$ , interaction with  $V_\lambda$  defines the map*

$$\text{Exec}_{\Gamma, \lambda}[B] : \mathsf{M} \longrightarrow \mathsf{M}'\mathsf{X}. \quad (7)$$

**Definition 4.2** (Success probability). *A QPT adversary  $\mathcal{B} = (|\psi\rangle, B, B_{\text{out}})$  against a quantum interactive search game  $\Gamma$  consists of a polynomial-size advice family  $\{|\psi_\lambda\rangle\}_\lambda$ , a polynomial-size interactive strategy  $B$ , and a polynomial-size predictor circuit  $B_{\text{out}}$  that takes register  $\mathsf{M}'$  as input and produces a classical string  $y \in \{0, 1\}^{m(\lambda)}$  in a designated classical register  $\mathsf{Y}$ . The success probability  $p_{\Gamma, \lambda}(\mathcal{B})$  of the adversary is defined as*

$$p_{\Gamma, \lambda}(\mathcal{B}) := \text{Tr}[\Pi_{\text{win}}(B_{\text{out}} \otimes I_{\mathsf{X}})\text{Exec}_{\Gamma, \lambda}[B](|\psi_\lambda\rangle\langle\psi_\lambda|)],$$

where

$$\Pi_{\text{win}} := \sum_{x \in \{0, 1\}^{m(\lambda)}} |x\rangle\langle x|_{\mathsf{X}} \otimes |x\rangle\langle x|_{\mathsf{Y}}.$$

*The winning projector acts as the identity on unmentioned registers.*

**Definition 4.3** (Sequential success probability). *Given a repetition parameter  $\ell = \ell(\lambda)$ , a QPT adversary  $\mathcal{A} = (|\psi\rangle, A_1, \dots, A_\ell, A_{\text{out}})$  against the  $\ell$ -sequentially-repeated quantum interactive search game  $\Gamma$  has success probability  $p_{\Gamma, \lambda}^\ell(\mathcal{A})$  defined as follows. Each copy uses fresh verifier randomness and workspace.*

- Set  $\sigma_0 := |\psi_\lambda\rangle\langle\psi_\lambda|$  on register  $\mathsf{M}_0$ .
- For  $i \in [\ell]$ , set  $\sigma_i := (\text{Exec}_{\Gamma, \lambda}[A_i]_{\mathsf{M}_{i-1} \rightarrow \mathsf{M}_i} \otimes I_{\mathsf{X}_1, \dots, \mathsf{X}_{i-1}})(\sigma_{i-1})$ .
- $p_{\Gamma, \lambda}^\ell(\mathcal{A}) := \text{Tr}[\Pi_{\text{win}}^\ell(A_{\text{out}} \otimes I_{\mathsf{X}_1, \dots, \mathsf{X}_\ell})(\sigma_\ell)],$

where

$$\Pi_{\text{win}}^\ell := \sum_{x_1, \dots, x_\ell \in \{0, 1\}^{m(\lambda)}} |x_1, \dots, x_\ell\rangle\langle x_1, \dots, x_\ell|_{\mathsf{X}_1, \dots, \mathsf{X}_\ell} \otimes |x_1, \dots, x_\ell\rangle\langle x_1, \dots, x_\ell|_{\mathsf{Y}_1, \dots, \mathsf{Y}_\ell}.$$

**Theorem 4.4.** *For any quantum interactive search game  $\Gamma$ , any polynomial  $\ell = \ell(\lambda) \geq 1$ , and any  $\varepsilon(\lambda) \in [0, 1]$ , if for all QPT adversaries  $\mathcal{B}$ ,  $p_{\Gamma, \lambda}(\mathcal{B}) \leq \varepsilon(\lambda) + \text{negl}(\lambda)$ , then for all QPT adversaries  $\mathcal{A}$  against the  $\ell$ -sequentially-repeated game,  $p_{\Gamma, \lambda}^\ell(\mathcal{A}) \leq \varepsilon(\lambda)^\ell + \text{negl}(\lambda)$ .*

The main quantum information tool that we use to prove [Theorem 4.4](#) is the following uniform singular value amplification result from [\[GSLW19\]](#).

**Lemma 4.5** ([\[GSLW19\]](#), Theorem 30 and the paragraph immediately following it). *Let  $U$  be a unitary and  $\Pi_{\text{in}}, \Pi_{\text{out}}$  be two projectors such that  $U, U^\dagger$  and coherent tests of the projectors are implementable by quantum circuits of size at most  $s$ . Let  $\gamma > 1, \delta, \zeta \in (0, 1/2)$ , define  $K := \Pi_{\text{out}} U \Pi_{\text{in}}$ , and suppose that  $\|K\| \leq \frac{1-\delta}{\gamma}$ . Then there exists a unitary  $W$ , using one additional qubit  $\mathcal{C}$  and implementable by a circuit of size at most  $\text{poly}(s, \gamma, 1/\delta, \log(1/\zeta))$ , such that*

$$\|(\langle 0|_{\mathcal{C}} \otimes \Pi_{\text{out}}) W (|0\rangle_{\mathcal{C}} \otimes \Pi_{\text{in}}) - \gamma K\| < \zeta.$$

Next we give a direct corollary stated in a manner that will be convenient for our proof of [Theorem 4.4](#).

**Corollary 4.6.** *Let  $U, \Pi_{\text{in}}, \Pi_{\text{out}}, K, \delta, \zeta$  be as above and suppose that  $c > 0$  and  $\|K\| \leq \sqrt{c} < 1 - \delta$ . Then there exists a unitary  $W$  implementable by a circuit of size at most  $\text{poly}(s, 1/c, 1/\delta, \log(1/\zeta))$  such that the following holds. Define*

$$L := (\langle 0|_{\mathcal{C}} \otimes \Pi_{\text{out}}) W (|0\rangle_{\mathcal{C}} \otimes \Pi_{\text{in}}).$$

*For any register  $\mathcal{R}$ , normalized state  $|\psi\rangle$  in the image of  $\Pi_{\text{in}} \otimes I_{\mathcal{R}}$ , and any projector  $\Pi$ ,*

$$\|\Pi(L \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 \geq \frac{(1-\delta)^2}{c} \|\Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 - \zeta.$$

*Proof.* Set  $\gamma' = \frac{1-\delta}{\sqrt{c}}, \zeta' = \zeta/2$ , and apply [Lemma 4.5](#) with parameters  $\gamma', \delta, \zeta'$  to obtain  $L$  as defined in the corollary statement such that  $\|L - \gamma' K\| < \frac{\zeta}{2}$ . Let

$$|\nu\rangle := ((L - \gamma' K) \otimes I_{\mathcal{R}}) |\psi\rangle$$

and note that  $\|\nu\| < \zeta/2$ . We have that

$$\begin{aligned} \|\Pi(L \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 &= \|\gamma' \Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle + \Pi |\nu\rangle\|^2 \\ &= \|\gamma' \Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 + 2\gamma' \text{Re} \langle \psi | (K^\dagger \otimes I_{\mathcal{R}}) \Pi |\nu\rangle + \|\Pi |\nu\rangle\|^2 \\ &\geq \|\gamma' \Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 - 2\|\gamma' \Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle\| \|\Pi |\nu\rangle\| \\ &> \frac{(1-\delta)^2}{c} \|\Pi(K \otimes I_{\mathcal{R}}) |\psi\rangle\|^2 - \zeta. \end{aligned}$$

Here the last step uses  $\|\gamma' K\| \leq 1 - \delta < 1$  and  $\|\Pi |\nu\rangle\| < \zeta/2$ . □

Finally, we prove [Theorem 4.4](#).

*Proof.* (of [Theorem 4.4](#)) Fix a search game  $\Gamma$ , polynomial  $\ell = \ell(\lambda)$ ,  $\varepsilon = \varepsilon(\lambda)$  and any adversary  $\mathcal{A} = (|\psi\rangle, A_1, \dots, A_\ell, A_{\text{out}})$ . For  $j \in [\ell]$ , define  $q_j$  as follows. Start immediately before copy  $j$  and initialize  $A_j$  with an arbitrary state. Run  $A_j$  through  $A_\ell$  interacting with  $V$ , run  $A_{\text{out}}$  to produce guesses  $y_j, \dots, y_\ell$ , and accept if  $y_j = x_j, \dots, y_\ell = x_\ell$ . Then  $q_j$  is the maximum over all initial states of the probability of accepting.

Define  $q_{\ell+1} = 1$ , and note that  $p_{\Gamma,\lambda}^\ell(\mathcal{A}) \leq q_1 \leq q_2 \leq \dots \leq q_{\ell+1}$ . The theorem follows if  $\max_{j \in [\ell]} (q_j - \varepsilon q_{j+1})_+$  is negligible, where  $(u)_+ := \max\{u, 0\}$ : iterating this uniform recurrence gives  $q_1 \leq \varepsilon^\ell + \ell \cdot \text{negl}(\lambda)$ . Towards contradiction, suppose that there exist a polynomial  $r(\lambda) \geq 1$ , an infinite set of security parameters, and an index  $j = j(\lambda) \in [\ell]$  such that, on this infinite set,  $q_j > \varepsilon q_{j+1} + \eta$ , where  $\eta := 1/r(\lambda)$ . Include  $j$  in the non-uniform advice. Note that  $q_{j+1} \geq q_j \geq \eta$ , and consider the following two cases.

First suppose that  $q_{j+1} \geq 1 - \eta/16$ . Define an adversary  $\mathcal{B} = (|\phi\rangle, B, B_{\text{out}})$  against the  $j$ 'th sequential game as follows. Let  $|\phi\rangle$  be a pure state that attains  $q_j$ ,  $B = A_j$ , and  $B_{\text{out}}$  run the remainder of  $A_{j+1}, \dots, A_\ell, A_{\text{out}}$  while internally simulating  $V$  in each game, and outputting the value  $y_j$  output by  $A_{\text{out}}$ . Then the probability that  $y_j = x_j$  is at least

$$\varepsilon q_{j+1} + \eta \geq \varepsilon - \frac{\varepsilon \eta}{16} + \eta \geq \varepsilon + \frac{15\eta}{16},$$

a contradiction.

Next suppose that  $q_{j+1} < 1 - \eta/16$ . Let  $U$  be a purification, retaining all measurement outcomes and discarded registers, of the procedure that runs  $A_{j+1}, \dots, A_\ell, A_{\text{out}}$  while internally simulating  $V$  in each game. Let  $M$  be the register passed to  $A_{j+1}$  and  $Z$  be the auxiliary register needed to define the purification  $U$ , initialized to  $|0\rangle_Z$ . Define  $\Pi_{\text{in}} = I_M \otimes |0\rangle\langle 0|_Z$  and define  $\Pi_{\text{out}}$  to be the projector applied to the output of  $U$  that checks that  $y_i = x_i$  for all  $i \in \{j+1, \dots, \ell\}$ . Then defining

$$K = \Pi_{\text{out}} U \Pi_{\text{in}},$$

we have that  $\|K\|^2 = q_{j+1}$ . The circuits for  $U, U^\dagger$  and coherent tests of both projectors have polynomial size.

Set  $b = \lceil \log_2(32/\eta) \rceil$  and  $c = 2^{-b} \lceil 2^b q_{j+1} \rceil$ , included in the non-uniform advice, so that  $q_{j+1} \leq c \leq q_{j+1} + \eta/32 < 1$ , and apply [Corollary 4.6](#) with  $\delta, \zeta = \eta/128$ . Note that the hypothesis holds since  $c < 1 - \eta/32 < (1 - \eta/128)^2 = (1 - \delta)^2$ . Since  $c \geq \eta$ , we obtain a polynomial-size circuit  $W$  which we will use to define an adversary  $\mathcal{B} = (|\phi\rangle, B, B_{\text{out}})$  as follows. Let  $|\phi\rangle$  be a pure state that attains  $q_j$  and  $B = A_j$ . Let  $B_{\text{out}}$  prepare  $|0\rangle_Z$  and an additional single-qubit register  $|0\rangle_C$ , run  $W$ , and attempt to project the output onto  $|0\rangle\langle 0|_C \otimes \Pi_{\text{out}}$ . If successful it measures and outputs  $y_j$ . Otherwise, it outputs an arbitrary string.

Let  $|\phi'\rangle_{M,Z,X_j,E}$  be a purification of the output of  $B$  interacting with  $V$ , tensored with  $|0\rangle_Z$ , where  $X_j$  holds  $V$ 's output and  $E$  purifies the interaction. Then  $B_{\text{out}}$  acts on  $M, Z, C$  and we take  $R = X_j E$  to be the register that it does not touch. Let  $\Pi$  be the projector checking that  $x_j = y_j$  and let  $L$  be as defined in the statement of [Corollary 4.6](#). Then the probability that  $\mathcal{B}$  succeeds is at least

$$\begin{aligned} \|\Pi(L \otimes I_R) |\phi'\rangle\|^2 &\geq \frac{(1-\delta)^2}{c} \|\Pi(K \otimes I_R) |\phi'\rangle\|^2 - \zeta \\ &= (1-\zeta)^2 \frac{q_j}{c} - \zeta \geq \frac{q_j}{c} - 3\zeta \\ &\geq \frac{\varepsilon q_{j+1} + \eta}{c} - 3\zeta = \varepsilon + \frac{\eta - \varepsilon(c - q_{j+1})}{c} - 3\zeta \\ &\geq \varepsilon + \frac{\eta - \eta/32}{c} - 3\zeta \geq \varepsilon + \frac{31\eta}{32} - 3\zeta \\ &= \varepsilon + \frac{121\eta}{128}, \end{aligned}$$

a contradiction. □

## 5 Succinct Arguments for QMA

Let  $T = T(\lambda)$  be a polynomial in the security parameter and a power of 2, and identify  $[T]$  with  $\{0, 1\}^{\log T}$  using its canonical encoding. Consider the following protocol between a client and a (possibly malicious) server.

- (i) Run  $\lambda$  independent copies of the protocol in [Section 3.1](#) in sequence. Denote the output state held by the (honest) server in the  $i$ -th copy by

$$\bigotimes_{t \in \{0,1\}^{\log T}} \left| \psi_{y_{t,0}, y_{t,1}}^{(i)} \right\rangle$$

up to an irrelevant global phase.

- (ii) After the completion of all the protocols, the client samples independent uniform permutations  $\pi_i \sim S_T$  and sends them to the server.
- (iii) For every  $t \in \{0, 1\}^{\log T}$ , the server takes the states indexed by  $\pi_i(t)$ . For each  $i \in \{2, \dots, \lambda\}$ , it applies a CNOT from the first leading qubit to the  $i$ -th leading qubit, measures the latter in the computational basis, and records the outcome  $e_{i,t} \in \{0, 1\}$ . After discarding the measured leading qubits, the remaining state is  $|\psi_{Y_{t,0}, Y_{t,1}}\rangle$ , where

$$\begin{aligned} Y_{t,0} &:= y_{\pi_1(t),0}^{(1)} \| y_{\pi_2(t),e_{2,t}}^{(2)} \| \dots \| y_{\pi_\lambda(t),e_{\lambda,t}}^{(\lambda)}, \\ Y_{t,1} &:= y_{\pi_1(t),1}^{(1)} \| y_{\pi_2(t),1 \oplus e_{2,t}}^{(2)} \| \dots \| y_{\pi_\lambda(t),1 \oplus e_{\lambda,t}}^{(\lambda)}. \end{aligned} \tag{8}$$

The server sends the outcomes  $\{e_{i,t}\}_{i,t}$  to the client.

- (iv) The client aborts if the verification in any of the copies fails and otherwise it locally computes  $\{Y_{t,0}, Y_{t,1}\}_t$  as specified above.

Clearly the only quantum communication between the client and the server is the one happening in the protocol from [Section 3.1](#). Therefore, quantum states are only sent from the client to the server. Moreover, before any message is exchanged, the client can sample the classical description of a circuit  $C_i$  of size  $\text{poly}(\log T, \lambda) = \text{poly}(\lambda)$  and compute

$$C_i |0\rangle_{\mathbf{Q}} |0\rangle_{\mathbf{W}} = |\psi_i\rangle_{\mathbf{Q}} |0\rangle_{\mathbf{W}}$$

where  $|\psi_i\rangle$  is the state sent from the client to the server in the  $i$ -th quantum communication round. In particular, the state sent in the  $i$ -th round is independent of the protocol transcript (but might depend on the client's internal randomness). We refer to an interactive protocol that satisfies such properties as being *quantum-succinct and history-independent*.

**Definition 5.1** (Quantum-Succinctness and History-Independence). *A protocol with public input  $x$  is quantum-succinct and history-independent if the client: (i) samples a uniform classical seed  $s$  of fixed  $\text{poly}(\lambda)$  length before interaction; (ii) computes from  $(s, x)$ , in time  $|x| \text{poly}(\lambda)$ , descriptions of circuits  $C_i$  of fixed total size  $\text{poly}(\lambda)$ , each preparing a message in fresh registers*

$$C_i |0\rangle_{\mathbf{Q}_i} |0\rangle_{\mathbf{W}_i} = |\psi_i\rangle_{\mathbf{Q}_i} |0\rangle_{\mathbf{W}_i},$$

*sends  $\mathbf{Q}_i$  at a publicly pre-determined round, and discards  $\mathbf{W}_i$ ; and (iii) is otherwise classical, with messages and acceptance determined by  $s, x$ , any private classical input, and the classical transcript.*

The protocol above satisfies this definition: Its preparation circuits use only the short sampled key and string lists and efficient PRF evaluation. These lists may be sampled directly, with only the remaining classical coins expanded from a PRG seed. Replacing those coins by PRG output changes QPT views only negligibly; the proofs below use uniform coins. Moreover, we say that a protocol is a *claw-state correlation* protocol if it satisfies Lemma 5.2 and Lemma 5.3, which we prove in the following.

**Lemma 5.2.** *When both parties are honest, except with negligible probability, at the end of the protocol execution the server holds the state*

$$\bigotimes_{t \in \{0,1\}^{\log T}} |\psi_{Y_{t,0}, Y_{t,1}}\rangle$$

and the client holds  $\{Y_{t,0}, Y_{t,1}\}_{t \in \{0,1\}^{\log T}}$ .

*Proof.* Fix  $t$ . By correctness of the weak protocol, before gluing the server holds the tensor product of the  $\lambda$  claws indexed by  $\pi_i(t)$ . After the CNOTs this state is

$$2^{-\lambda/2} \sum_{b_1, \dots, b_\lambda \in \{0,1\}} |b_1, b_1 \oplus b_2, \dots, b_1 \oplus b_\lambda\rangle \bigotimes_{i=1}^{\lambda} |y_{\pi_i(t), b_i}^{(i)}\rangle.$$

Measuring the last  $\lambda - 1$  leading qubits fixes  $b_i = b_1 \oplus e_{i,t}$  for  $i \geq 2$ , leaving two equally weighted terms. After normalization and discarding the measured qubits, the remaining state is therefore

$$\frac{|0, Y_{t,0}\rangle + |1, Y_{t,1}\rangle}{\sqrt{2}} = |\psi_{Y_{t,0}, Y_{t,1}}\rangle,$$

with strings given by Equation (8). The gluing operations act separately for each  $t$ , so the resulting states form the claimed tensor product, and the communicated measurement outcomes let the client compute the same strings. Finally, each weak copy rejects an honest server with probability at most  $(\log T)2^{-n}$ . By a union bound over the  $\lambda$  copies, the abort probability is at most  $\lambda(\log T)2^{-n} = \text{negl}(\lambda)$ .  $\square$

The following lemma establishes the security of the protocol.

**Lemma 5.3.** *For every non-uniform QPT server and every  $t^* \in \{0,1\}^{\log T}$ , consider the experiment that runs the protocol, gives the server  $\{Y_{t,0}, Y_{t,1}\}_{t \neq t^*}$  after the client produces its output, and asks it to return  $(Y_{t^*,0}, Y_{t^*,1})$ . Assuming the existence of quantum-secure one-way functions, the probability that the client does not abort and the server returns the correct pair is negligible.*

*Proof.* Fix any  $t^* = t^*(\lambda) \in \{0,1\}^{\log T}$ . We reduce recovery of its pair of strings to winning  $\lambda$  sequential copies of the search game in Theorem 3.1. The private output of each game is its target pair if the check passes, and  $\perp$  otherwise. The reduction forwards each weak-protocol interaction to the server, but stores the disclosed target index  $u_i$  and strings for the other claws without revealing them.

After all copies have finished, the reduction samples independent permutations uniformly subject to  $\pi_i(t^*) = u_i$  and sends them to the server. Since the  $u_i$  are independent uniform indices and were withheld during the weak interactions, this has exactly the joint distribution of the original protocol. Once the server sends its gluing bits, the stored strings determine every combined pair at  $t \neq t^*$ , so the reduction can supply all the required leakage.

A correct guess of  $(Y_{t^*,0}, Y_{t^*,1})$  gives every raw target pair by splitting the strings into blocks and undoing the swaps specified in Equation (8). Failed checks can be followed by dummy continuations, since those branches cannot win. Thus Theorems 3.1 and 4.4 give

$$\Pr[\text{no abort and a correct target pair}] \leq 4^{-\lambda} + \text{negl}(\lambda) = \text{negl}(\lambda).$$

Non-uniformity makes this bound uniform over target sequences. A union bound over the polynomially many targets also gives hardness of recovering any pair without being given the other strings.  $\square$

## 5.1 Blind Delegation

We recall the definition of a blind delegation protocol.

**Definition 5.4** (Blind Delegation). *A blind delegation protocol for a public quantum circuit, viewed as a map  $Q : \mathbf{XB} \rightarrow \mathbf{O}$ , takes a private classical input  $x$  from the client and a quantum input on  $\mathbf{B}$  from the server. In an honest execution, the client obtains one-time-pad keys  $(r, s)$ . For every input state  $\rho_{\mathbf{BR}}$  with arbitrary reference  $\mathbf{R}$ , let  $\sigma_{\mathbf{OR}}$  be the server's output and reference after applying  $Z^s X^r$  on  $\mathbf{O}$  whenever the client does not abort, averaged over all randomness and measurement outcomes. Represent abort by an orthogonal flag in  $\mathbf{O}$ , and embed the ideal output in the non-abort subspace. Correctness requires*

$$\frac{1}{2} \|\sigma_{\mathbf{OR}} - (Q \otimes I_{\mathbf{R}})(|x\rangle\langle x|_{\mathbf{X}} \otimes \rho_{\mathbf{BR}})\|_1 \leq \text{negl}(\lambda).$$

*For every QPT server  $\mathcal{A}$ , every  $\rho_{\mathbf{BR}}$ , and all equal-length inputs  $x_0, x_1$ , blindness requires*

$$\text{View}_{\mathcal{A}}(\rho_{\mathbf{BR}}, x_0) \approx_c \text{View}_{\mathcal{A}}(\rho_{\mathbf{BR}}, x_1),$$

*where the view includes the server's entire output,  $\mathbf{R}$ , and the abort flag, with the same public circuit  $Q$  in both experiments. The factor  $I_{\mathbf{R}}$  leaves the reference untouched.*

Using the protocol defined above and a result from [BK25], we can establish the existence of a quantum-succinct and history-independent blind delegation protocol.

**Lemma 5.5.** *Assuming the existence of quantum-secure one-way functions, for any polynomial-size quantum circuit  $Q$ , there exists a quantum-succinct and history-independent (Definition 5.1) blind delegation protocol for  $Q$ . Moreover, its round complexity is bounded by  $\text{poly}(\lambda, \delta)$ , where  $\delta$  is the  $\mathbf{T}$ -depth of  $Q$ .*

*Proof.* Use a Clifford+ $\mathbf{T}$  realization of  $Q$ , including any negligible approximation error in correctness. Let  $T = \text{poly}(\lambda)$  be a power-of-two upper bound on its number of  $\mathbf{T}$  gates and let  $\Pi_{\text{claw}}$  be a quantum-succinct and history-independent claw-state correlation protocol with  $T$  output claw states and round complexity  $\text{poly}(\lambda)$ , i.e. it satisfies Definition 5.1, Lemma 5.2, and Lemma 5.3. Above we showed that  $\Pi_{\text{claw}}$  exists assuming quantum-secure one-way functions. Unused correlations may be discarded; if there are no  $\mathbf{T}$  gates, omit this setup.

We will use  $\Pi_{\text{claw}}$  to construct a blind delegation protocol  $\Pi_{\text{blind}}$  for  $Q$  by applying two transformations from [BK25]: (1) BB84 correlations from claw-state correlations, and (2) blind delegation from BB84 correlations.

First, we specify the definition of a BB84 correlation protocol  $\Pi_{\text{BB84}}$ . Let  $T = \text{poly}(\lambda)$  be the number of output correlations.

- **Correctness:** When both parties are honest, except with negligible probability, at the end of the protocol execution the server holds the state

$$\bigotimes_{i \in [T]} H^{\theta_i} |x_i\rangle,$$

and the client holds  $\{\theta_i, x_i\}_{i \in [T]}$ .

- **Security:** For every non-uniform QPT server and every  $i^* = i^*(\lambda) \in [T]$ , let  $\text{ok}$  denote that the client does not abort. On this branch give the server  $\{\theta_j, x_j\}_{j \neq i^*}$  and let  $\hat{\theta}_{i^*}$  be its guess. We require

$$\left| \Pr[\text{ok} \wedge \hat{\theta}_{i^*} = \theta_{i^*}] - \frac{1}{2} \Pr[\text{ok}] \right| \leq \text{negl}(\lambda). \quad (9)$$

The view includes the abort flag and the server's full workspace. On abort the verifier samples the basis bits  $\theta_i$  independently and uniformly at random.

For the first transformation, [BK25, Theorem 4.7] converts a claw-state generator into an oblivious state preparation (OSP) protocol. Apply this conversion in parallel to each of our  $T$  claws. To meet its requirement that the two strings be distinct, set  $L_{i,b} := b||Y_{i,b}$  and have the server copy its leading qubit into a fresh data qubit, obtaining  $\frac{1}{\sqrt{2}}(|0, L_{i,0}\rangle + |1, L_{i,1}\rangle)$ . Recovering both strings recovers the original pair. For security of coordinate  $i$ , the reduction receives the other pairs of strings, simulates the other conversion clients, and computes their BB84 labels. The Goldreich–Levin reduction in that theorem then applies jointly with this leakage and the server's workspace: a noticeable advantage in the unconditional bound above gives a noticeable probability of recovering the target pair without abort, contradicting Lemma 5.3. Correctness follows from the cited conversion and Lemma 5.2. The transformation adds only classical communication and polynomially many rounds, so the resulting  $\Pi_{\text{BB84}}$  remains quantum-succinct and history-independent.

Generate this entire batch before sending any message depending on the client's input. For the second transformation, [BK25, Lemma 6.10 and Theorem 6.11] gives a protocol for blind delegation from any OSP. The BK25 protocol runs one OSP for each T-gate, and here we will use the  $i$ -th OSP correlation  $H^{\theta_i} |x_i\rangle, (\theta_i, x_i)$  in place of the OSP protocol for the  $i$ -th T-gate. In [BK25], the protocol for the  $i$ -th T-gate requires the client to input a *chosen* basis  $\theta'_i$ , while our OSP correlation yields a basis  $\theta_i$  that is not controlled by the client. To remedy this, we simply have the client first send the bit  $b_i = \theta_i \oplus \theta'_i$  to the server, who applies  $H^{b_i}$  to its state. After this, correctness follows directly from the arguments in [BK25]. Further, note that the security of the OSP correlation protocol implies that the server's view on client input  $\theta' = 0$  and  $\theta' = 1$  are indistinguishable even given  $\{\theta_j, x_j\}_{j \neq i}$ .

For blindness, switch the chosen bases to 0 in reverse gate order, as in [BK25, Theorem 6.11]. At gate  $k$ , the reduction knows all labels except  $(\theta_k, x_k)$ : the earlier labels and transcript determine  $\theta'_k$ , while all later chosen bases have already been replaced by 0. The BB84 guarantee makes  $\theta_k$  indistinguishable from a fresh uniform bit jointly with the known labels and the server's state, so sending  $\theta_k \oplus \theta'_k$  or  $\theta_k$  gives indistinguishable views. Subsequent visible messages use only known labels; the unknown  $x_k$  affects only private key updates. After polynomially many switches, the only input-dependent message is protected by the initial classical one-time pad, giving statistical blindness.

Finally, as also observed in [BLM26], the client and server can participate in all OSPs for a given layer of T-gates in parallel, with that layer's chosen bases fixed before its messages are sent. Thus, the round complexity of the blind delegation protocol is bounded by  $\text{poly}(\lambda, \delta)$ . Moreover, since

this transformation requires only classical communication, the blind delegation protocol remains quantum-succinct and history-independent. □

## 5.2 Putting Things Together

In [BLM26, Theorem 54] it is shown that a blind delegation protocol implies an argument for QMA.

**Lemma 5.6.** *Assuming a blind delegation protocol with round complexity  $\text{poly}(\lambda, \delta)$  with  $\delta$  the T-depth of the computed quantum circuit, there is an argument for QMA with completeness  $1 - \text{negl}(\lambda)$ , soundness at most a constant  $s < 1$ , and a fixed  $\text{poly}(\lambda)$  number of rounds.*

*Moreover, if the blind delegation protocol satisfies Definition 5.1, then so does the resulting interactive argument.*

The moreover part of the above statement is not explicitly stated in [BLM26], but it is implicit in their proof because the protocol uses the blind delegation (see also [BLM26, Theorem 40]) as a black-box and everything else is classical communication. Next, we invoke the communication compression compiler from [BLM26].

**Lemma 5.7.** *Let  $\Pi$  be an  $r(\lambda)$ -round interactive protocol satisfying Definition 5.1, with no private client input besides  $s$ , where  $r$  is bounded by a fixed polynomial and the public input  $x$  specifies the classical next-message and decision algorithms with description size  $O(|x|)$ . Assuming collapsing hash functions, for every  $\varepsilon = 1/\text{poly}(\lambda)$  there is a compiled protocol  $\tilde{\Pi}$  with negligible honest-completeness loss such that, for every QPT compiled prover  $\tilde{P}^*$ , there is a QPT original prover  $P^*$  with*

$$\Pr[\tilde{V}^{\tilde{P}^*} \text{ accepts}] \leq \Pr[V^{P^*} \text{ accepts}] + O(\varepsilon) + \text{negl}(\lambda).$$

*Its communication is at most  $r(\lambda) \text{poly}(\lambda, 1/\varepsilon)$  and its verifier time is at most  $|x|r(\lambda) \text{poly}(\lambda, 1/\varepsilon)$ .*

*Proof.* [BLM26, Theorem 18] states the same theorem for a protocol with a completely classical verifier, but we argue that the same holds for any interactive protocol satisfying Definition 5.1. First, recall that we can assume without loss of generality that the verifier's classical computation is deterministic once it samples a random seed  $s \sim \{0, 1\}^{\text{poly}(\lambda)}$  prior to any interaction. Keep the short randomness for quantum preparation explicitly in  $s$  and derive the remaining classical coins using a quantum-secure PRG. Note that by history-independence (Definition 5.1), this also means that the quantum states sent by the verifier are fully determined prior to any interaction.

Let us now recall the compilation procedure for a classical verifier from [BLM26]. For any round of interaction:

- The prover computes locally its next message of the protocol and sends a succinct commitment to it, together with a state-preserving succinct argument of knowledge for its pre-image.
- The prover and the verifier engage in a chosen-input committed secure function sampling (SFS) protocol, where the verifier's input is  $s$  and the prover's inputs are the committed messages. At the end of the interaction, the prover receives the value of the verifier's next-message function.

At the end of the interaction, the verifier reveals  $s$ , and the prover supplies a succinct argument of knowledge proving that the committed messages form an accepting transcript.

The verifier also sends the original quantum states at their prescribed positions. For soundness, keep  $s$  and all external registers as an untouched reference. Following [BLM26, Definitions 11 and 19, Lemma 20, and Section 5.4], extract each committed message using the state-preserving extractor and simulate the corresponding chosen-input SFS from the next-message value on the extracted transcript. These guarantees preserve the joint state up to the prescribed distinguishing error, including correlations between the prover’s quantum registers and  $s$ ; the chosen-input transformation uses fresh garbling randomness. At quantum-message positions, the reduction forwards the original verifier’s register once, without knowing  $s$  or cloning the state. Choose the inverse-polynomial per-call accuracies so that all extraction and simulation errors sum to  $O(\varepsilon)$ ; the fixed polynomial bound on  $r$  absorbs the resulting overhead.

After the last commitment, extract the final accepting-transcript witness. Collision resistance forces its messages to equal those already extracted, except with negligible probability. Hence compiled acceptance implies that the original verifier’s classical predicate accepts the extracted transcript, up to the stated errors. The final seed reveal and proof cannot change those fixed messages and can be omitted for this upper bound. Thus an original prover forwards the extracted classical messages, uses the verifier’s classical replies as SFS-simulator inputs, and forwards its quantum replies as above. Honest correctness of the subprotocols separately gives negligible completeness loss. Their efficiency and Definition 5.1 give the claimed bounds.  $\square$

Combining Lemmas 5.3 and 5.5 to 5.7, along with the fact that collapsing hashes imply the existence of one-way functions, we obtain a succinct argument for QMA with constant soundness. Sequential repetition of Theorem 5.8 yields our main result Theorem 1.1.

**Theorem 5.8.** *Assuming collapsing hash functions, there exists a constant  $0 < s < 1$  and a succinct argument for QMA with completeness  $1 - \text{negl}(\lambda)$ , soundness error  $s$ , total communication bounded by  $\text{poly}(\lambda)$ , and verifier runtime  $|x| \cdot \text{poly}(\lambda)$ .*

*Proof.* Let  $s_0 < 1$  be the soundness in Lemma 5.6. Choose the compilation accuracy so that Lemma 5.7 gives soundness at most  $s = (1 + s_0)/2 < 1$  for sufficiently large  $\lambda$ , while honest completeness remains  $1 - \text{negl}(\lambda)$ . The fixed round bound gives the claimed communication and verifier time. For Theorem 1.1, repeat sequentially  $\lambda$  times with fresh verifier randomness, accepting only if every copy accepts, and give the honest prover fresh witness copies. Completeness loss is still negligible by a union bound. Apply Theorem 4.4 to the search game whose private target is 0 on acceptance and  $\perp$  otherwise, soundness is at most  $s^\lambda + \text{negl}(\lambda)$ .  $\square$

## Acknowledgments

JB is supported by the Air Force Office of Scientific Research under agreement number FA9550261B239. GM is supported by the European Research Council through an ERC Starting Grant (Grant agreement No. 101077455, ObfusQation) and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy - EXC 2092 CASA – 390781972.

## References

- [ALM<sup>+</sup>98] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998. <https://doi.org/10.1145/278298.278306>. 2
- [AS98] Sanjeev Arora and Shmuel Safra. Probabilistic checking of proofs: A new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998. <https://doi.org/10.1145/273865.273901>. 2
- [BK25] James Bartusek and Dakshita Khurana. On the power of oblivious state preparation. In *Advances in Cryptology—CRYPTO 2025*, volume 16001 of *Lecture Notes in Computer Science*, pages 575–607. Springer, 2025. Full version: <https://arxiv.org/abs/2411.04234v1>. 2, 4, 7, 20, 21
- [BKM<sup>+</sup>25] Kaniuar Bacho, Alexander Kulpe, Giulio Malavolta, Simon Schmidt, and Michael Walter. Compiled nonlocal games from any trapdoor claw-free function. In *Advances in Cryptology—CRYPTO 2025*, volume 16001 of *Lecture Notes in Computer Science*, pages 642–673. Springer, 2025. <https://eprint.iacr.org/2024/1829>. 4
- [BLM26] James Bartusek, Jiahui Liu, and Giulio Malavolta. A modular approach to succinct arguments for QMA. In *Advances in Cryptology—EUROCRYPT 2026*, volume 16547 of *Lecture Notes in Computer Science*, pages 446–474. Springer, 2026. Full version: <https://arxiv.org/abs/2606.10408v1>. 2, 3, 4, 7, 8, 21, 22, 23
- [BMM26] Pedro Branco, Giulio Malavolta, and Zayd Maradni. Fully-homomorphic encryption from lattice isomorphism. In *Theory of Cryptography—TCC 2025*, volume 16268 of *Lecture Notes in Computer Science*, pages 220–252. Springer, 2026. First published online in December 2025. <https://eprint.iacr.org/2025/993>. 2
- [BTL<sup>+</sup>22] James Bartusek, Yael Tauman Kalai, Alex Lombardi, Fermi Ma, Giulio Malavolta, Vinod Vaikuntanathan, Thomas Vidick, and Lisa Yang. Succinct classical verification of quantum computation. In *Advances in Cryptology—CRYPTO 2022*, volume 13508 of *Lecture Notes in Computer Science*, pages 195–211. Springer, 2022. <https://arxiv.org/abs/2206.14929>. 2, 8
- [CMSZ22] Alessandro Chiesa, Fermi Ma, Nicholas Spooner, and Mark Zhandry. Post-quantum succinct arguments: Breaking the quantum rewinding barrier. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 49–58. IEEE, 2022. <https://arxiv.org/abs/2103.08140>. 2
- [GJMZ23] Sam Gunn, Nathan Ju, Fermi Ma, and Mark Zhandry. Commitments to quantum states. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1579–1588. ACM, 2023. <https://arxiv.org/abs/2210.05138>. 3
- [GSLW19] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: Exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 193–204. ACM, 2019. Full version: <https://arxiv.org/abs/1806.01838v1>. 7, 16

- [GTNV25] Sam Gunn, Yael Tauman Kalai, Anand Natarajan, and Ági Villányi. Classical commitments to quantum states. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC)*, pages 234–244. ACM, 2025. <https://arxiv.org/abs/2404.14438>. 2, 3
- [Kil92] Joe Kilian. A note on efficient zero-knowledge proofs and arguments (extended abstract). In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing (STOC)*, pages 723–732. ACM, 1992. <https://doi.org/10.1145/129712.129782>. 2
- [KLVY23] Yael Kalai, Alex Lombardi, Vinod Vaikuntanathan, and Lisa Yang. Quantum advantage from any non-local game. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing (STOC)*, pages 1617–1628. ACM, 2023. <https://arxiv.org/abs/2203.15877>. 4
- [MNZ24] Tony Metger, Anand Natarajan, and Tina Zhang. Succinct arguments for QMA from standard assumptions via compiled nonlocal games. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1193–1201. IEEE, 2024. <https://arxiv.org/abs/2404.19754>. 2, 3, 4
- [NZ23] Anand Natarajan and Tina Zhang. Bounding the quantum value of compiled nonlocal games: From CHSH to BQP verification. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1342–1348. IEEE, 2023. <https://arxiv.org/abs/2303.01545v2>. 4
- [Unr16] Dominique Unruh. Computationally binding quantum commitments. In *Advances in Cryptology—EUROCRYPT 2016*, volume 9666 of *Lecture Notes in Computer Science*, pages 497–527. Springer, 2016. <https://eprint.iacr.org/2015/361>. 2, 8
- [Zha12] Mark Zhandry. How to construct quantum random functions. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 679–687. IEEE, 2012. <https://eprint.iacr.org/2012/182>. 9
- [Zha21] Jiayu Zhang. Succinct blind quantum computation using a random oracle. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 1370–1383. ACM, 2021. <https://arxiv.org/abs/2004.12621>. 3, 5, 6