

MIT Open Access Articles

Classical Commitments to Quantum States

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Sam Gunn, Yael Tauman Kalai, Anand Natarajan, and Ági Villányi. 2025. Classical Commitments to Quantum States. In Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC '25). Association for Computing Machinery, New York, NY, USA, 234–244.

As Published: <https://doi.org/10.1145/3717823.3718264>

Publisher: ACM|Proceedings of the 57th Annual ACM Symposium on Theory of Computing

Persistent URL: <https://hdl.handle.net/1721.1/164618>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution-Noncommercial



Classical Commitments to Quantum States

Sam Gunn
gunn@berkeley.edu
UC Berkeley
Berkeley, CA, USA

Anand Natarajan
anandn@mit.edu
MIT
Cambridge, MA, USA

Yael Tauman Kalai
tauman@mit.edu
MIT
Cambridge, MA, USA

Ági Villányi
agivilla@mit.edu
MIT
Cambridge, MA, USA

Abstract

We define the notion of a classical commitment scheme to quantum states, which allows a quantum prover to compute a classical commitment to a quantum state, and later open each qubit of the state in either the standard or the Hadamard basis. Our notion is a strengthening of the measurement protocol from Mahadev (STOC 2018). We construct such a commitment scheme from the post-quantum Learning With Errors (LWE) assumption, and more generally from any noisy trapdoor claw-free function family that has the distributional strong adaptive hardcore bit property (a property that we define in this work).

Our scheme is *succinct* in the sense that the running time of the verifier in the commitment phase depends only on the security parameter (independent of the size of the committed state), and its running time in the opening phase grows only with the number of qubits that are being opened (and the security parameter). As a corollary we obtain a classical succinct argument system for **QMA** under the post-quantum LWE assumption. Previously, this was only known assuming post-quantum secure indistinguishability obfuscation. As an additional corollary we obtain a generic way of converting any X/Z quantum PCP into a succinct argument system under the quantum hardness of LWE.

CCS Concepts

• **Theory of computation** → **Interactive proof systems; Cryptographic protocols; Quantum complexity theory.**

Keywords

Succinct arguments, quantum interactive proofs, quantum Merlin-Arthur proofs, quantum commitment protocols

ACM Reference Format:

Sam Gunn, Yael Tauman Kalai, Anand Natarajan, and Ági Villányi. 2025. Classical Commitments to Quantum States. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC '25)*, June 23–27, 2025, Prague, Czechia. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3717823.3718264>



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

STOC '25, Prague, Czechia

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1510-5/25/06

<https://doi.org/10.1145/3717823.3718264>

1 Introduction

A commitment scheme is one of the most basic primitives in classical cryptography, with far reaching applications ranging from zero-knowledge proofs [5, 11], identification schemes and signature schemes [9], secure multi-party computation protocols [6, 12], and succinct arguments [19]. There is a long history of studying commitments to *classical* information, both in the classical and post-quantum worlds, and recently, Gunn et al. [13] systematically explored commitments to *quantum* states, using quantum messages. In this work, we initiate the formal study of commitments to quantum states using *classical* messages. Specifically, we study the existence of *classical* commitments to quantum states, where all messages (the commitment and the opening) are classical, and the receiver is a classical machine. This setting models a likely future where classical devices will have access to powerful (possibly untrusted) quantum devices. The goal of this work is to provide the foundations needed for these classical devices to use the (untrusted) quantum devices effectively.

Our major contributions are a definition of a classical commitment to quantum states, including a sensible notion of a classical opening of a committed quantum state; a construction based on the post-quantum Learning With Errors (LWE) assumption; and a construction of a *succinct* commitment to quantum states (analogous to Merkle hashing in the classical setting [17]), also under post-quantum LWE.¹ As an immediate application, we obtain a succinct classical argument system for **QMA** based only on post-quantum hardness of LWE, improving on previous work which required indistinguishability obfuscation [1]. To our knowledge, our work constitutes the first work to define a notion of a binding classical commitment to quantum states, and to give a construction that achieves it.

Our construction builds directly on the seminal *measurement protocol* of Mahadev [16], which was used by her to construct the first classical argument system for **QMA**. Loosely speaking, a measurement protocol is a way for a classical verifier to request a quantum prover to measure each qubit of a quantum state (of the prover's choice) in the X or Z basis, with the guarantee that the prover's opening must be "consistent with a quantum state." This motivates our definition of a classical *opening* of a quantum state: the receiver should be able to request the sender to open each qubit of the committed state in either the X or Z basis. (One

¹More generally, our constructions are based on the existence of a (noisy) claw-free trapdoor function family with a distributional strong adaptive hard-core bit property, which in particular can be instantiated under the LWE assumption.

could imagine asking for openings in more general bases, but these two seem to be a desirable minimum.) However, a measurement protocol does not automatically give rise to a commitment, for several reasons. First, there is a major structural difference: in a measurement protocol, all phases of the protocol—even the keys chosen in the initial setup—may depend on the choice of opening basis! (Indeed, in Mahadev’s protocol, the keys consist of either “2-to-1” or “injective” claw-free functions depending on the basis to be measured.) This is far from what we would like in a commitment: the initial “commitment” phase should be *completely* independent of the basis in which the receiver ultimately chooses to request an opening.

Thus, the first step to building our construction is to convert Mahadev’s measurement protocol into something having the syntax of a commitment², and henceforth we refer to this modified protocol as Mahadev’s “weak” commitment³. In the most basic version of this protocol, a quantum sender holding a qubit in state $|\psi\rangle$ interacts with a classical receiver, sending a classical message that commits to $|\psi\rangle$. Later, the sender is requested by the receiver to “open” the committed qubit in either the standard or the Hadamard basis. To open, the sender performs an appropriate measurement and returns the outcome, which can be *decoded* by the receiver (using a cryptographic trapdoor), to obtain an outcome from measuring $|\psi\rangle$ in the appropriate basis.

A commitment scheme must be *binding*, meaning that the sender cannot change their mind about the committed state once the commitment has been sent. It turns out that the modified Mahadev scheme is a “weak” commitment because it partially satisfies the binding property: it is binding in the standard basis, but *not at all* binding in the Hadamard basis. In fact, the sender, after committing to $|+\rangle$, can always freely change the committed state to $|-\rangle$ without ever being detected! Relatedly, in the modified Mahadev scheme, the receiver performs a test on the opening in the standard basis case, and only accepts the opening if it is valid, but performs no test in the Hadamard case.

Motivated by this observation, we show that a simple twist on Mahadev’s weak commitment is truly binding (in a rigorous sense which we define) in both bases. We elaborate on our binding definition in Sections 1.1 and 2, and on our construction in Sections 1.2 and 2, and below only give a teaser. In our construction, the sender first commits to $|\psi\rangle$ under Mahadev’s weak commitment, generating a commitment string y_0 and a (multi-qubit) post-commitment state $|\psi_1\rangle$. It then coherently *opens* this state in the Hadamard basis—that is, it executes a unitary version of the opening algorithm, but does not perform the final measurement, instead producing a quantum state $|\psi_1\rangle$. Finally, the sender applies Mahadev’s weak commitment *again* to the state $|\psi_1\rangle$, qubit-by-qubit, obtaining a vector of commitment strings \vec{y} and a post-commitment state $|\psi_2\rangle$. The strings (y_0, \vec{y}) now constitute a classical commitment to the state $|\psi\rangle$. To open this commitment in the Hadamard basis, the sender simply applies the standard basis opening procedure for the second Mahadev commitment, yielding a string z which the receiver will

test and decode using the commitment vector \vec{y} . By the standard-basis binding of Mahadev’s commitment, we are guaranteed that the decoded outcome from z —assuming the test passes—yields the same result as measuring $|\psi_1\rangle$ in the standard basis, and by construction, this gives a Hadamard-basis opening of $|\psi\rangle$, which it can then decode using the commitment string y_0 . But how do we open the commitment in the standard basis? It is far from obvious that this is even possible! For this we exploit specific features of the Mahadev scheme—in particular, the fact that the opening procedure is “native”: Opening in the standard basis constitutes measuring the registers in the standard basis, and opening in the Hadamard basis constitutes measuring the registers in the Hadamard basis. This fact is useful both to argue that the opening is correct and to prove that the binding property is achieved. We note that in our new scheme the verifier tests the validity of both the standard basis opening and the Hadamard basis opening, and decodes both openings using the cryptographic trapdoor.⁴

Our basic construction for a single qubit can be extended to states with any number of qubits to get a *non-succinct* commitment to a quantum state. We next ask whether our commitment scheme can be made *succinct*: can the sender commit to an ℓ qubit state, and open to a small number of these qubits, by exchanging much fewer than ℓ bits with the receiver? Here, already in the case of “weak” commitments, there is a significant technical obstacle with just the *very first message* from the receiver to the sender: openings in Mahadev’s scheme can leak information about the secret key, so each committed qubit must use a fresh secret key to maintain any security at all. This means that, already in the initial key-exchange phase, the receiver must send the sender $\geq \ell$ bits. We show that, surprisingly, the “strong” binding property of our commitment, together with specific properties of the underlying (noisy) trapdoor claw-free family, allows us to overcome this barrier. Namely, we show that strong binding, together with specific properties of the underlying (noisy) trapdoor claw-free family, implies that the openings do not leak information about the key in our scheme, allowing us to use the same key for all committed qubits. We emphasize that, even to obtain a succinct “weak” commitment, or a succinct measurement protocol, the only route we know of using standard (post-quantum) cryptographic assumptions is through our strongly binding commitments! We view this as an interesting indication of the possible usefulness of our strong binding property in further applications.

As a teaser for how exactly the leakage occurs, and how we avoid it, for now we remark that in the Mahadev weak commitment, the adversary can cause the receiver to generate outputs of the form $d' \cdot s$, for known vectors d' of its choice, where s is the secret. This means that the output for sufficiently many qubits may leak the secret s . For an honest sender, this would not be an issue because the vectors d' would be obtained by a quantum measurement with unpredictable answers, and thus have high min-entropy. We show that in our scheme, even *dishonest* senders are forced to produce d' with (sufficient) min-entropy, because of the additional tests done

²Technically, we do this by always using the “2-to-1” mode of the claw-free function. Moreover, we do not even rely on the existence of a dual-mode (as was done by Mahadev [16]), and simply use a “2-to-1” claw-free family.

³We refer to it as a weak commitment since (as we elaborate on below) it does not have the desired binding property.

⁴We mention that in Mahadev’s scheme, the verifier only tests the validity of the standard basis opening, and this test, as well as the decoding, is done publicly (without the trapdoor). The verifier uses the trapdoor only to decode the Hadamard basis opening, which it did not test.

in our opening procedure. This is what prevents the outcomes from leaking information about s .

Reusing the key directly only gives us a short first message, which yields a “semi-succinct” commitment, in which messages from the receiver are short, but messages from the sender are long. In fact, this already yields an application of our results: a *fully-succinct* classical argument system for **QMA** which is secure assuming post quantum security of LWE. We obtain this by following the template of Bartusek et al. [1], but replacing their use of Mahadev’s measurement protocol with our succinct commitment.

THEOREM 1.1 (INFORMAL). *There exists a (classical) succinct interactive argument for **QMA** under the post-quantum Learning With Errors (LWE) assumption.*⁵

This improves on the result of [1] in terms of cryptographic assumptions: they required the assumption of post-quantum indistinguishability obfuscation (iO) to succinctly generate ℓ keys for Mahadev’s protocol, whereas our protocol only requires the post-quantum security of LWE. It is currently not known how to deduce post-quantum iO from *any* standard cryptographic assumptions, whereas LWE is the “paradigmatic” post-quantum cryptographic assumption.

To construct a succinct argument system for **QMA**, the approach we and [1] both follow is to construct a semi-succinct argument system, and then make it fully succinct by composing with (state-preserving) post-quantum interactive arguments of knowledge [8, 15]. It turns out that the same tools let us construct outright a fully succinct commitment scheme: for this to be meaningful, we imagine that the sender only opens to a small number of qubits chosen by the receiver, rather than to all of the qubits. In classical cryptography, succinct commitments are natural partners of PCPs, as they enable a verifier to delegate the task of checking a PCP to the prover. While quantum PCPs do not currently exist, we hope that our succinct commitment can be paired with a suitable future PCP to design interesting protocols.

1.1 The Definition

Defining a non-succinct commitment scheme. Our definition of a (non-succinct) commitment scheme is a natural extension of the classical counterpart. It consists of a key generation algorithm Gen that takes as input the security parameter 1^λ and a length parameter 1^ℓ and outputs a pair of public and secret keys (pk, sk) ; a commit algorithm Commit that takes as input a public key pk and an ℓ -qubit quantum state σ and outputs a classical string \mathbf{y} and a post-commitment state ρ , where \mathbf{y} is the commitment to the quantum state σ ;⁶ an open algorithm Open that takes as input the post-commitment state ρ and a basis choice $\mathbf{b} = (b_1, \dots, b_\ell) \in \{0, 1\}^\ell$, where $b_i = 0$ corresponds to opening the i ’th qubit in the standard basis and $b_i = 1$ corresponds to opening the i ’th qubit in the Hadamard basis, and outputs an opening $\mathbf{z} \in \{0, 1\}^{\ell \cdot \text{poly}(\lambda)}$; and the final algorithm Out that takes as input a secret key sk , a commitment string \mathbf{y} , a basis choice $\mathbf{b} \in \{0, 1\}^\ell$ and an opening \mathbf{z} ,

⁵More generally, assuming the existence of a (noisy) trapdoor claw free function family with a distributional strong adaptive hard-core bit property, which we elaborate on later on.

⁶We note that both the length of pk and the length of the commitment string \mathbf{y} may grow polynomially with the length ℓ of the committed state σ .

and outputs the measurement result $\mathbf{m} \in \{0, 1\}^\ell$ or \perp if the opening is rejected.⁷

We mention that the above syntax yields a commitment scheme that is *privately verifiable* in the sense that sk is needed to decode the measurement value \mathbf{m} from the opening value \mathbf{z} . While it would be desirable to construct a commitment scheme that is publicly verifiable, where Gen only generates a public key pk , and this public key is used by the opening algorithm to generate the output \mathbf{m} along with an opening \mathbf{z} which can be verified given pk , we believe that this public key variant is impossible to achieve. This impossibility was formalized on the quantum setting (i.e., where the commitment is a quantum state) by [13], and we leave it as an open problem to prove the impossibility in the classical setting.

We require two properties from our commitment scheme: completeness and binding. We note that for commitments to classical strings it is common to require a *hiding* property. We do not require it since one can easily obtain hiding by committing to the commitment string \mathbf{y} using a classical commitment scheme (that is binding and hiding).

- **Correctness.** The correctness property asserts that if an honest committer commits to an ℓ -qubit state σ then for any basis choice $\mathbf{b} \in \{0, 1\}^\ell$, the algorithm Out , applied to the opening string \mathbf{z} generated by Open , yields an output \mathbf{m} whose distribution is statistically close to the distribution obtained by simply measuring σ in the basis \mathbf{b} .
- **Binding.** Loosely speaking, the binding property asserts that for any (possibly malicious) QPT algorithm Commit^* that commits to an ℓ -qubit quantum state, there is a *single* extracted quantum state τ such that for *any* QPT algorithm Open^* and *any* basis (b_1, \dots, b_ℓ) , where $b_i = 0$ corresponds to measuring the i ’th qubit in the standard basis and $b_i = 1$ corresponds to measuring it in the Hadamard basis, the output obtained by $\text{Open}^*(b_1, \dots, b_\ell)$ is computationally indistinguishable from measuring τ in basis (b_1, \dots, b_ℓ) , assuming Open^* is always accepted. We relax the requirement that Open^* is always accepted, and allow Open^* to be rejected with probability δ at the price of the two distributions being $O(\sqrt{\delta})$ -computationally indistinguishable. We elaborate on the binding property in Section 2. We note that our definition of binding is nontrivial only for senders that are accepted with a high success probability. By repeating the protocol sequentially $O(1/\delta \cdot \log(1/\delta))$ times we can ensure that if all the openings are accepted with probability $\geq \delta$ then a random one of these openings is accepted with probability $1 - \delta$. While this is weaker than classical notions of binding commitments (which apply to any sender that is accepted with non-negligible probability), it is sufficient for constructing a succinct argument system for **QMA**.

Comparison with Mahadev’s measurement protocol. Our commitment scheme is stronger than that of a *measurement protocol*, originally considered in [16] and formally defined in [1]. Beyond the syntactic difference, where in a measurement protocol the opening basis must be determined during the key generation phase (and the

⁷We note that in the actual definition we partition this algorithm into two parts: Ver and Out where the former only outputs a bit indicating if the opening is valid or not and the latter outputs the actual opening if valid. This partition is only for convenience.

key generation algorithm takes as input the basis $\mathbf{b} \in \{0, 1\}^\ell$, our binding property is significantly stronger. A measurement protocol guarantees that any (possibly malicious) QPT algorithm Open^* must be consistent with an ℓ -qubit state, but different opening algorithms can be consistent with different quantum states.

Defining a succinct commitment scheme. The syntax for a succinct commitment differs quite substantially from the syntax of a non-succinct commitment described above. First, Gen only takes as input the security parameter 1^λ (and does not take as input the length parameter 1^ℓ); in addition, Commit is required to output a succinct commitment of size $\text{poly}(\lambda)$. However, there is a more substantial difference which stems from the fact that, similarly to the non-succinct variant, we require a succinct commitment to have a binding property that asserts that one can extract an ℓ -qubit quantum state τ such that the output distribution of any successful opening is indistinguishable from measuring τ . Since in this setting we consider opening algorithms that only open a few of the qubits, there is no way we can extract an ℓ -qubit state from such algorithms. As a remedy, we add an *interactive test phase*. This test phase is executed with probability $1/2$, and if executed then at the end of it the verifier outputs 0 or 1, indicating accept or reject, and the protocol terminates without further executing the opening phase, since the test protocol destroys the state. We note that Mahadev’s measurement protocol has a non-interactive test phase which is executed with probability $1/2$. In our setting this test phase is *interactive*. It is this interactive nature that allows us to extract a large state from a succinct protocol.

1.2 The Construction

Our construction: the single qubit case. We construct the commitment scheme in stages. We first construct a *single-qubit* commitment scheme; this scheme is inspired by the construction from Mahadev [16]. We elaborate on it in Section 2, but give a very high-level description here. First, let us recall Mahadev’s weak commitment for a single qubit. In this scheme, the sender receives a public key that enable it to evaluate a *two-to-one trapdoor claw-free* (TCF) function $f : \{0, 1\} \times \mathcal{X} \rightarrow \mathcal{Y}$.⁸ For every image $y \in \mathcal{Y}$, there are exactly two preimages, which have the form $(0, x_0)$ and $(1, x_1)$, where $x_0, x_1 \in \{0, 1\}^n$, but any such pair (called a “claw”) is cryptographically hard to find. In Mahadev’s scheme, to commit to a qubit in state $|\psi\rangle = \sum_{b \in \{0, 1\}} \alpha_b |b\rangle$, the sender first prepares

$$\sum_{b \in \{0, 1\}} \sum_{x \in \mathcal{X}} \alpha_b |b\rangle |x\rangle |f(b, x)\rangle,$$

and then measures the last register to obtain a random outcome y . The resulting state is the $(n + 1)$ -qubit state:

$$\sum_{b \in \{0, 1\}} \alpha_b |b\rangle |x_b\rangle$$

To open this in the standard basis, the honest sender measures in the standard basis and returns (b, x_b) ; the receiver checks that $f(b, x_b) = y$, and if so, records a measurement outcome of b . Intuitively, this constitutes a “binding” commitment in the standard basis because it is impossible for the sender to know both x_0 and x_1 ,

⁸We mention that under the LWE assumption we only have a “noisy” TCF function family, which was constructed in [4]. We do not go into this technicality in the introduction and overview sections.

and thus impossible to flip between them. To open in the Hadamard basis, the honest sender measures in the *Hadamard* basis; a short calculation shows that the outcome is a random string $d \in \{0, 1\}^{n+1}$, where the probability that $d \cdot (1, x_0 \oplus x_1) \equiv 0 \pmod{2}$ is exactly equal to $|\alpha_0 + \alpha_1|^2/2$, the probability that a Hadamard basis measurement on the *original* state $|\psi\rangle$ would have yielded $+$. The receiver uses the cryptographic trapdoor to compute $d \cdot (1, x_0 \oplus x_1) \pmod{2}$ as the measurement outcome of the opening, and performs *no* test. This is not at all a binding commitment: indeed, the “commitments” to a Hadamard basis states $|\pm\rangle$ look like

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

and one can easily map from one state to the other by applying a Pauli Z operator to the first qubit.

We now describe our modification to convert this weak commitment (denoted commit_W) into a binding commitment: simply apply a Hadamard transform to the post-commitment state, and then weakly commit again to the resulting n -qubit state, applying the Mahadev scheme qubit by qubit, with a new TCF function f_i for each qubit.

$$\begin{aligned} & \sum_b \alpha_b |b\rangle \\ & \mapsto^{\text{commit}_W \rightarrow y_0} \sum_b \alpha_b |b, x_b\rangle \\ & \mapsto^{H^{\otimes (n+1)}} \sum_{d \in \{0, 1\}^{n+1}} \beta_d |d\rangle \\ & \mapsto^{\text{commit}_W \rightarrow y_1, \dots, y_{n+1}} \sum_d \beta_d |d_1, x'_{1, d_1}\rangle \dots |d_{n+1}, x'_{n+1, d_{n+1}}\rangle. \end{aligned}$$

Here, d_j denotes the j th bit of d , and $x'_{j, b}$ denotes the corresponding preimage of y_j under the TCF function f_j (so $f_j(b, x'_{j, b}) = y_j$).

Let us see how to open this commitment. It will be easier to start with the Hadamard basis: to open in this basis, the sender measures their state in the *standard* basis, and returns the string $(d_1, x_1, \dots, d_{n+1}, x_{n+1})$. The receiver checks that each (d_i, x_i) is a preimage of the corresponding y_i , and then records the measurement outcome as $(d_1, \dots, d_{n+1}) \cdot (1, x_0 \oplus x_1)$. To open in the standard basis, the sender measures their state in the *Hadamard* basis, obtaining a (long) string z , and the receiver converts this into a measurement outcome by applying the Mahadev procedure for the *Hadamard* basis. Specifically, it first splits z into equal blocks of size $n + 1$, and applies the Mahadev Hadamard procedure on each block, to get $n + 1$ bits m_1, \dots, m_{n+1} .

$$\begin{aligned} z &= (z_1, \dots, z_{n+1}) \\ & \mapsto (m_1 = z_1 \cdot (1, x'_{1, 0} \oplus x'_{1, 1}), \dots, \\ & \quad m_{n+1} = z_{n+1} \cdot (1, x'_{n+1, 0} \oplus x'_{n+1, 1})) \end{aligned}$$

Now, this corresponds to the outcome of opening the weak commitment of $\sum_d \beta_d |d\rangle$ in the Hadamard basis. But this state in turn was equal to the Hadamard transform of $\sum_b \alpha_b |b, x_b\rangle$. Thus, the outcomes m_1, \dots, m_{n+1} should look like the outcome of measuring $\sum_b \alpha_b |b, x_b\rangle$ in the standard basis: that is, like a preimage of y_0 under the TCF function f ! Thus, the receiver tests the outcomes by

checking that

$$f(m_1, \dots, m_{n+1}) = y_0,$$

and if this passes, it records m_1 as the measurement outcome.

At an intuitive level, what makes this commitment scheme binding is that the receiver performs a test in *both* bases. More formally, we show binding in two parts: (1) there exists a qubit state consistent with the openings reported by the sender, and (2) for any two opening algorithms, the openings they generate are statistically indistinguishable. The proof of (1) uses standard techniques from the analysis of Mahadev’s protocol—in particular, the “swap isometry” as presented in [20], but the proof of (2) is new to our work. Our arguments are based on the *collapsing* property of the TCF functions used to generate y_1, \dots, y_n (in the Hadamard basis case), and y_0 (in the standard basis case). (Jumping ahead, we note that in the succinct setting the situation is reversed. We can obtain (2) basically “for free” from the non-succinct setting, whereas the proof of (1) incurs most of the technical burden in this work.)

Our construction: multiple qubits, and succinctness. From the single-qubit scheme described above, we construct a *non-succinct multi-qubit* commitment scheme, by committing qubit-by-qubit, and thus repeating the single-qubit construction ℓ -times, where ℓ is the number of qubits we wish to commit to. This transformation is generic and can be used to convert *any* single-qubit commitment scheme into a *non-succinct* multi-qubit one. We emphasize that in the resulting ℓ -qubit scheme, both the public key and the commitment string grow with ℓ , since the former consists of ℓ public-keys and the latter consists of ℓ commitment strings, where each corresponds to the underlying single-qubit scheme. We then convert this scheme into a succinct commitment scheme. This is done in two stages:

- (1) **Stage 1:** Reuse the same public key, as opposed to choosing ℓ independent ones. Namely, the public key consists of a single public key pk corresponding the underlying single-qubit commitment scheme. To commit to an ℓ -qubit state, commit qubit-by-qubit while using the same public key pk . We refer to such a commitment scheme as *semi-succinct* since the public key is succinct but the commitment is not. We note that while this construction is generic, the analysis is not. In general, reusing the same public-key may break the binding property. We prove that if we start with our specific single-qubit commitment scheme then the resulting semi-succinct multi-qubit scheme remains sound. We recall, that as mentioned above, if we start with Mahadev’s single qubit weak commitment protocol and convert it into a multi-qubit weak commitment while reusing the same public key, then the resulting measurement protocol becomes insecure. The reason is that a malicious sender may generate openings d in the Hadamard basis that cause the receiver’s “decoding” outcomes $d \cdot (1, x_0 \oplus x_1)$ to leak bits of sk —recall that the receiver must use the secret key to decode, as x_0 and x_1 cannot be computed efficiently without it. Indeed, the TCF function family that we (and Mahadev) use is the LWE-based construction due to [4], which has the property

that $d \cdot (1, x_0 \oplus x_1) = d' \cdot s$, where s is a secret key⁹ and d' can be efficiently computed from d and x_0 . Once enough information about the secret s has been revealed, the scheme is no longer a secure measurement protocol, let alone a secure commitment: with knowledge of s , it becomes easy to distinguish the outcomes of the commitment from outcomes of measuring a true quantum state! Thus, to argue the security of our semi-succinct scheme, we must exploit specific properties of our single-qubit scheme. Indeed, we crucially use the *binding* property of our scheme to show that the openings z reported by a successful sender must always have high min-entropy, which in our construction implies that d' has min-entropy. We then use a specific property of the underlying TCF function family from [4], which we call the “distributional strong adaptive hardcore bit” property. Roughly, this property ensures that if the opening d has min-entropy then $d \cdot (1, x_0 \oplus x_1)$ (which in their construction is equal to $d' \cdot s$) does not reveal information about sk .

- (2) **Stage 2:** Convert any semi-succinct commitment scheme into a succinct one. This part is generic and shows how to convert *any* semi-succinct commitment scheme into a succinct one. Our transformation is almost identical to that from [1], who showed how to convert any semi-succinct interactive argument (which is one where only the verifier’s communication is succinct, and where the prover’s communication can be long) into a fully succinct one. We elaborate on the high-level idea behind this transformation in Section 2.

1.3 Applications

We show how to use our succinct commitment scheme to construct succinct interactive argument for **QMA**. As a simpler bonus, we also use it show how to compile a hypothetical quantum PCP in “ X/Z form” into a succinct interactive argument. For the X/Z PCP compiler the idea is simple: In the succinct interactive argument the prover first succinctly commits to the X/Z PCP, then the verifier sends its X/Z queries and finally the prover opens the relevant qubits in the desired basis. The succinct interactive argument for **QMA** is more complicated, and follows the blueprint from [1, 16]. We elaborate on this in Section 2.1.

1.4 Related Works

Our work is inspired by the measurement protocol of Mahadev [16], which has the same correctness guarantee as our commitment scheme. However, a measurement protocol (as was formally defined in [1]) does not require binding to hold; rather it only requires that an opening is consistent with a qubit. This qubit may be different for different opening algorithms. Indeed, the measurement protocol of Mahadev, as well as the ones from followup works, are not binding in the Hadamard basis. Mahadev uses this measurement protocol to construct classical interactive arguments for **QMA**. Mahadev’s measurement protocol, which was proven to be secure under the post-quantum LWE assumption, is a key ingredient in our construction.

⁹In their construction the public key is an LWE tuple $(A, As + e)$. The secret key is actually a trapdoor of the matrix A but revealing the secret s is sufficient to break security.

Mahadev’s measurement protocol is not succinct. In a followup work, Bartusek et al. [1] constructed a succinct measurement protocol, by using Mahadev’s measurement protocol as a key ingredient, and thus obtaining a succinct classical interactive arguments for **QMA**. However the security of their protocol, and thus the soundness of the resulting **QMA** argument, relies on the existence of a post-quantum secure indistinguishable obfuscation scheme (in addition the post-quantum LWE assumption). We mention that Chia, Chung and Yamakawa [7] also construct a succinct measurement protocol, which they use to obtain a succinct 2-message argument for **QMA**. However, in their scheme the prover and verifier share a polynomial-sized structured reference string (which requires a trusted setup to instantiate), and their security is heuristic.¹⁰

We improve upon these works by constructing a succinct classical commitment scheme for quantum states that guarantees binding (which is a stronger security condition than the one offered by a measurement protocol), based only on the post-quantum LWE assumption. As a result, we obtain a succinct classical interactive arguments for **QMA**, under the post-quantum LWE assumption. Our analysis makes use of techniques developed in [1, 16, 20], in addition to several new ideas that are needed to obtain our results.

We mention that our work, as well as all prior works mentioned above, require the receiver (a.k.a the verifier) to hold a secret key sk which is needed to decode the prover’s message and obtain the measurement output. We mention that the recent work of Bartusek et al. [2] considers the public-verifiable setting, where decoding can be done publicly. They construct a publicly verifiable measurement protocol in an oracle model, which is used as a building block in their obfuscation of pseudo-deterministic quantum circuits.

So far we only focused on prior work where the verifier (and hence the communication) is classical. We mention that recently Gunn et al. [13] defined and constructed a *quantum* commitment scheme to quantum states, where *both* parties are quantum. In their setting, the quantum committer sends a quantum commitment to the receiver, and later opens by sending a quantum opening. The receiver then applies some unitary operation to recover the committed quantum state. This is in contrast to the classical setting where the receiver is classical and cannot hope to recover the committed quantum state, and instead only obtains an opening in a particular basis (standard or Hadamard). We mention that the quantum commitment scheme from [13] relies on very weak cryptographic assumptions, and in particular, ones that are implied by the existence of one-way functions.

Finally, simultaneously and using different techniques from this work, a succinct argument system for **QMA** based on the assumption of quantum Fully Homomorphic Encryption (qFHE) was achieved by [18]. While both papers use common techniques from [1] to go from semi-succinctness to full succinctness, the core techniques are essentially disjoint. In particular, [18] does not use commitments to quantum states, but instead directly analyzes the soundness of the KLVY [14] compilation of a particular semi-succinct two-prover interactive proof for **QMA**. We leave it as an interesting open question for future work whether their result can yield an alternate construction of our primitive of quantum commitments.

¹⁰More specifically, their scheme uses a hash function h , and it is proved to be secure when h is modeled as a random oracle, but the *protocol description itself* explicitly requires the code of h (i.e. uses h in a non-black-box way).

2 Technical Overview

In this section we describe the ideas behind our commitment schemes and their applications in more depth yet still informally. Our first contribution is defining the notion of a classical commitment scheme to quantum states. Let us start with the non-succinct version, and in particular the single-qubit case. As mentioned in the introduction, such a commitment scheme consists of algorithms

(Gen, Commit, Open, Out)

where Gen is a PPT algorithm that takes as input the security parameter 1^λ and outputs a pair of keys (pk, sk) ; Commit is a QPT algorithm that takes as input a public key pk and a single-qubit quantum state σ and outputs a classical commitment string y and a post-commitment state ρ ; Open is a QPT algorithm that takes as input the post-commitment state ρ and a bit $b \in \{0, 1\}$, where $b = 0$ corresponds to a standard basis opening and $b = 1$ corresponds to a Hadamard basis opening, and outputs a classical opening z ; and Out is a polynomial-time algorithm that takes as input the secret key sk , a commitment string y , a basis $b \in \{0, 1\}$ and an opening z and it outputs an element in $\{0, 1, \perp\}$.

We require the scheme to satisfy a correctness and a binding property. The correctness property is straightforward and was formalized in prior work [1]. It is the binding property that is tricky to formulate and achieve.

Defining Binding: the single qubit setting. In the classical setting, the binding condition asserts that for any poly-size algorithm Commit^* that generates a commitment y (to some classical string), and for any two poly-size algorithms Open_1^* and Open_2^* , the probability that they successfully open to different strings is negligible. In the quantum setting the analogous property is the following: For any QPT algorithm Commit^* that generates a commitment y (to a quantum state), and for QPT algorithms Open_1^* and Open_2^* (that are accepted with probability 1) and every basis choice $b \in \{0, 1\}$, the output distributions of Open_1^* and Open_2^* are statistically close or computationally indistinguishable. This is indeed one of the properties we require.¹¹ But this property on its own is not enough. We also need to ensure that the opening is consistent with some qubit. Namely, we require that there exists a QPT extractor Ext such that for every QPT algorithm Open^* (that is accepted with probability 1), Ext given black-box access to Open^* can extract from Open^* a quantum state τ such that for every basis $b \in \{0, 1\}$ the output of Open^* is computationally indistinguishable from measuring τ in basis b . We mention that this latter condition was formalized in [1] as a security property from a measurement protocol.

We construct a commitment scheme that achieves the above two properties. However, to make this definition meaningful we must consider opening algorithms that are accepted with probability smaller than 1. Indeed, we consider opening algorithms that are accepted with probability $1 - \delta$ and obtain $O(\sqrt{\delta})$ -indistinguishability in both the requirements above. We note that we can assume that

¹¹Jumping ahead, we note that our non-succinct commitment scheme achieves statistical closeness and our succinct commitment scheme achieves computational indistinguishability. We mention that Mahadev’s scheme [16], as well as its successors [1], do not satisfy this property since these schemes offer no binding on the Hadamard basis.

Open* is accepted with probability $1 - \delta$ by repeating the commitment protocol $\Omega(1/\delta)$ times (assuming the committer has many copies of the state they wish to commit to).

The multi-qubit setting. So far we focused on the single-qubit setting. When generalizing the definitions to the multi-qubit setting we distinguish between the non-succinct setting and the succinct setting, starting with the former. The syntax can be generalized to the non-succinct multi-qubit setting in a straightforward way by committing and opening qubit-by-qubit. Generalizing the binding definition to the multi-qubit setting is a bit tricky. In particular, recall that we assumed that Open* is accepted with high probability when opening in both bases. As mentioned, this is a reasonable assumption since we can require the committer to commit to its state many ($\Omega(1/\delta)$) times, then open half of the commitments in the standard basis and half of them in the Hadamard basis. If any of them are rejected then output \perp and otherwise, choose a random one that was opened in the desired basis b and use that as the opening. Generalizing this to the ℓ -qubit setting must be done with care to avoid an exponential blowup in ℓ . Clearly, we do not want to assume that for every basis choice $(b_1, \dots, b_\ell) \in \{0, 1\}^\ell$, Open* successfully opens in this basis with high probability, since we cannot enforce this without incurring an exponential blowup. Yet, in order for our extractor to be successful, we need to ensure that Open* succeeds in opening each qubit in each basis with high probability. To achieve this, without incurring an exponential blowup, we require that Open* succeeds with high probability to open all the qubits in the standard basis (i.e., succeeds with $(b_1, \dots, b_\ell) = (0, \dots, 0)$) and succeeds with high probability to open all the qubits in the Hadamard basis (i.e., succeeds $(b_1, \dots, b_\ell) = (1, \dots, 1)$). This can be achieved via repetitions, as in the single qubit setting. Specifically, in this setting we ask $1/3$ of the repetitions to be opened in the 0^ℓ basis, $1/3$ to be opened in the 1^ℓ basis, and the remaining $1/3$ to be opened in the desired (b_1, \dots, b_ℓ) basis. Jumping ahead, we note that the extractor Ext uses Open* with basis (b, \dots, b) to extract the state τ . We refer the reader to the full version of the paper for the formal definition.

Our construction for the single qubit case. We start by describing our commitment scheme in the single-qubit case. Our starting point is Mahadev’s [16] measurement protocol. Her protocol is binding in the standard basis but offers no binding guarantees, and in fact fails to provide any form of binding, when opening in the Hadamard basis. Moreover, in her protocol the opening basis must be determined ahead of time and the public key pk used to compute the commitment string depends on this basis. Specifically, her protocol uses a family of (noisy) trapdoor claw-free functions, where functions can be generated either in an *injective* mode or in a *two-to-one* mode. The public key of the commitment scheme consists of a public key corresponding to an injective function if the verifier wishes to open in the standard basis, and corresponds to a two-to-one function if the verifier wishes to open in the Hadamard basis.

We first notice that it is not necessary to determine the opening basis in the key generation phase. In fact, we show that one can always use the two-to-one mode, irrespective of the basis we wish to open in. Moreover, we show that this “dual mode” property is not needed altogether. This observation is quite straightforward and was implicitly used in the analysis in prior work [1, 20].

Our first instrumental idea is that we can obtain binding in both bases if we compose Mahadev’s weak commitment twice! Namely, to commit to a state σ , we first apply Mahadev’s measurement protocol, denoted by Commit_W , to obtain

$$(y, \rho) \leftarrow \text{Commit}_W(pk, \sigma).$$

As mentioned, this already guarantees binding when opening in the standard basis, but fails to provide binding when opening in the Hadamard basis. To fix this we make use of the fact that Mahadev’s measurement protocol has the property that the Open algorithm always measures the post-commitment state in either the standard basis or the Hadamard basis. We apply to the post-commitment state ρ the unitary that computes Hadamard opening $\text{Open}(\cdot, 1)$, which is simply the Hadamard unitary $H^{\otimes(n+1)}$, where $n + 1$ is the number of qubits in ρ (n being the security parameter associated with the underlying NTCF family), and we commit to the resulting state. Namely, we compute

$$\rho' \leftarrow H^{\otimes(n+1)}[\rho] \text{ and } (y', \rho'') \leftarrow \text{Commit}_W(pk', \rho'),$$

where pk and pk' are independent keys,¹² and where throughout our paper we use the shorthand

$$U[\rho] = U\rho U^\dagger$$

to denote the application of a unitary U to a mixed state ρ .

To open the commitment in the Hadamard basis, we just need to measure ρ' in the *standard* basis. Binding in the Hadamard basis follows from the fact that ρ' was committed to via the classical string y' , and from the fact that Mahadev’s measurement protocol provides binding in the standard basis. However, it is no longer clear how to open in the standard basis, since the original post-commitment state ρ is no longer available, and has been replaced with ρ'' . Here we use the desired property mentioned above, specifically, that algorithm Open generates a standard basis opening by measuring the state in the standard basis, and generates a Hadamard basis opening by measuring the state in the Hadamard basis. This implies that measuring ρ in the standard basis is equivalent to measuring ρ' in the Hadamard basis.

The reader may be concerned that we may have lost the binding in the standard basis, since opening in the Hadamard basis is not protected. But this is not the case, since it is the commitment string y that binds the standard basis measurement, and the commitment string y' that binds the Hadamard basis measurement.

Multi-qubit commitments. One can use this single qubit commitment scheme to commit to an ℓ -qubit state, by committing qubit-by-qubit. This results with a long commitment string of size $\ell \cdot \text{poly}(\lambda)$ and with a long public key, since the public key consists of ℓ public keys (pk_1, \dots, pk_ℓ) , where each pk_i is generated according to the single qubit scheme. As mentioned in the introduction, our main goal is to construct a succinct commitment scheme. Following the blueprint of [1], we do this in two steps. We first construct a *semi-succinct* commitment scheme where the commitment string is long, but the public-key is succinct. We then show how to convert the semi-succinct scheme into a fully succinct one.

¹²Using different and independent public keys pk and pk' is important in our analysis.

Semi-succinct commitments. In our semi-succinct commitment scheme we generate a single key pair $(pk, sk) \leftarrow \text{Gen}(1^\lambda)$ corresponding to the single-qubit scheme, and simply use pk to commit to each and every one of the qubits. The question is whether this is sound. Let us first describe the main issue that comes up when trying to prove soundness, and then we will show how we overcome it. The issue is that our commitment scheme is privately verifiable, and thus a QPT algorithm Open^* , which produces an opening z , does not know the corresponding output bit $m = \text{Out}(sk, y, b, z)$ since sk is needed to compute m . Therefore, perhaps a malicious QPT algorithm Open^* can generate z in a way such that m leaks information about sk . In particular, perhaps Open^* can generate ℓ openings z_1, \dots, z_ℓ such that their corresponding outputs m_1, \dots, m_ℓ completely leak sk .

Recall that our binding property consists of two parts: The first asserts that for any QPT algorithm Commit^* that commits to an ℓ -qubit state via a classical commitment string y , it holds that for any two QPT opening algorithms Open_1^* and Open_2^* and any basis choice (b_1, \dots, b_ℓ) , the output distributions produced by these two opening algorithms are (computationally or statistically) close. In our construction we get *statistical closeness*, and hence the closeness holds even if sk is leaked. Indeed, the proof of this property in the semi-succinct setting is the same as the proof in the non-succinct setting. The issue is with the second part: Given sk , the distributions generated by $\text{Ext}^{\text{Open}^*}$ and Open^* are no longer computationally indistinguishable. Diving deeper into our scheme and its analysis, we note that the standard basis outputs produced by $\text{Ext}^{\text{Open}^*}$ and Open^* are actually statistically close, and it is the Hadamard basis outputs that are only computationally indistinguishable.

We next examine the leakage that the decoded messages m_1, \dots, m_ℓ may contain about the secret key, and argue that even given this leakage, the Hadamard basis outputs produced by $\text{Ext}^{\text{Open}^*}$ and Open^* remain computationally indistinguishable. To this end, we will need to use additional properties about Mahadev’s measurement protocol, and thus recall it in Section 2.1 below. Jumping ahead, we mention that one property that we rely on is the fact that in Mahadev’s protocol, Out does not use the secret key when generating standard basis outputs (and the secret key is only used to generate Hadamard basis outputs).

Recall that in our commitment scheme, the secret key consists of two parts, (sk, sk') , since we apply Mahadev’s protocol twice (where sk is for a single qubit state and sk' is for an $(n + 1)$ -qubit state). We mention that when opening in the standard basis, the output m can only leak information about sk' . This is the case since to open in the standard basis, we first use sk' to generate a standard basis opening z for Mahadev’s protocol, and then use Mahadev’s Out algorithm to decode z , which as mentioned above, can be done publicly without the secret key sk (since it is a standard basis opening). Importantly, we show that the computational indistinguishability of the Hadamard basis opening only relies on the fact that sk is secret, and does not rely on the secrecy of sk' . Thus, the remaining problem, which is at the heart of the technical complication, is the leakage of the Hadamard basis openings on sk . We note that in Mahadev’s protocol, the Hadamard basis openings may leak the entire sk . What saves us in our setting is the fact that we tie the hands of the adversary when opening in the Hadamard

basis. To explain this in more detail we need to recall Mahadev’s measurement protocol.

2.1 Mahadev’s Measurement Protocol

As mentioned, Mahadev’s measurement protocol [16] uses a noisy TCF family.¹³ In this overview, for the sake of simplicity, we describe her scheme assuming we have a noiseless TCF family, which is a function family associated with algorithms

$$(\text{Gen}_{\text{TCF}}, \text{Eval}_{\text{TCF}}, \text{Invert}_{\text{TCF}})$$

where Gen_{TCF} is a PPT algorithm that takes as input the security parameter 1^λ and outputs a key pair (pk, sk) ; Eval is a poly-time deterministic algorithm that takes as input the public key pk , and a pair (b, x) where $b \in \{0, 1\}$ is a bit and $x \in \{0, 1\}^n$ (where $n = \text{poly}(\lambda)$), and outputs a value y , and $\text{Eval}(pk, \cdot)$ is a two-to-one function where every y in the image has exactly two preimages of the form $(0, x_0)$ and $(1, x_1)$; $\text{Invert}_{\text{TCF}}$ takes as input the secret key sk and an element y in the image and it outputs the two preimages $((0, x_0), (1, x_1))$.

In what follows we show how Mahadev uses a TCF family to construct a measurement protocol. The following protocol slightly differs from Mahadev’s scheme, and in particular the basis choice is not determined during the key generation algorithm. The measurement protocol consists of algorithms $(\text{Gen}, \text{Commit}, \text{Open}, \text{Out})$ defined as follows:

- Gen is identical to Gen_{TCF} ; it takes as input the security parameter 1^λ and outputs a key pair (pk, sk) .
- Commit takes as input pk and a single-qubit pure state $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ and generates

$$|\psi'\rangle = \alpha_0 |0, x_0\rangle + \alpha_1 |1, x_1\rangle$$

such that $\text{Eval}(pk, (0, x_0)) = \text{Eval}(pk, (1, x_1)) = y$, and outputs y as the commitment string.

- Open takes as input the post-committed state $|\psi'\rangle$ and a basis $b \in \{0, 1\}$; if $b = 0$ it returns the outcome z of measuring $|\psi'\rangle$ in the standard basis, which is of the form (b, x_b) , and if $b = 1$ it returns the outcome z of measuring $|\psi'\rangle$ in the Hadamard basis.
- Out takes as input (sk, y, b, z) , and if $b = 0$ it checks that $\text{Eval}(pk, z) = y$ and if this is the case it outputs the first bit of z , and otherwise it outputs \perp . If $b = 1$ it outputs $z \cdot (1, x_0 \oplus x_1)$ where $((0, x_0), (1, x_1)) = \text{Invert}_{\text{TCF}}(y)$.

Recall that, as explained in the introduction, Mahadev’s measurement protocol is not fully binding. The issue is that a cheating prover can produce any opening in the Hadamard basis, and will never be rejected. For instance, a cheating prover could commit to $|+\rangle$ honestly, apply a Z to the first qubit of the post-commitment state, and then open to $|-\rangle$.

2.2 Our Single-Qubit Commitment Scheme

We convert Mahadev’s protocol into a binding commitment scheme by adding another step to the commitment algorithm, as described in the beginning of Section 2. More specifically, our commitment

¹³As mentioned above, her work, as well as followup works, use a dual-mode TCF family; we avoid this technicality.

scheme consists of algorithms (Gen, Commit, Open, Out) defined as follows:

- Gen(1^λ) generates $n+2$ TCF keys $(pk_i, sk_i)_{i \in \{0,1,\dots,n+1\}}$, where each $(pk_i, sk_i) \leftarrow \text{Gen}_{\text{TCF}}(1^\lambda)$, and outputs $pk = (pk_0, pk_1, \dots, pk_{n+1})$ and $sk = (sk_0, pk_1, \dots, sk_{n+1})$.
- Commit($pk, |\psi\rangle$) operates as follows:
 - (1) Parse $pk = (pk_0, pk_1, \dots, pk_{n+1})$.
 - (2) Apply Mahadev's measurement protocol to commit to $|\psi\rangle = \alpha_0 |0\rangle + \alpha_1 |1\rangle$ w.r.t. pk_0 ; i.e., generate

$$|\psi'\rangle = \alpha_0 |0, x_0\rangle + \alpha_1 |1, x_1\rangle$$

such that $\text{Eval}(pk_0, (0, x_0)) = \text{Eval}(pk_0, (1, x_1)) = y_0$.

- (3) Compute

$$H^{\otimes(n+1)} |\psi'\rangle = \sum_{\mathbf{d} \in \{0,1\}^{n+1}} \beta_{\mathbf{d}} |\mathbf{d}\rangle,$$

- (4) Use Mahadev's measurement protocol to commit qubit-by-qubit to the above $(n+1)$ -qubit state, w.r.t. public keys pk_1, \dots, pk_{n+1} to obtain the state

$$\sum_{\mathbf{d} \in \{0,1\}^{n+1}} \beta_{\mathbf{d}} |\mathbf{d}\rangle |x'_{1,d_1}\rangle \dots |x'_{n+1,d_{n+1}}\rangle$$

and strings y_1, \dots, y_{n+1} such that for every $i \in [n+1]$,

$$\text{Eval}(pk_i, (0, x'_{i,0})) = \text{Eval}(pk_i, (1, x'_{i,1})) = y_i.$$

- (5) Output $(y_0, y_1, \dots, y_{n+1})$, and (for simplicity) rearrange the post-commitment state to be

$$\sum_{\mathbf{d} \in \{0,1\}^{n+1}} \beta_{\mathbf{d}} |\mathbf{d}_1, x'_{1,d_1}\rangle \dots |\mathbf{d}_{n+1}, x'_{n+1,d_{n+1}}\rangle$$

- Open takes as input the post-commitment state ρ and a basis $b \in \{0,1\}$. If $b = 1$ (corresponding to opening in the Hadarmard basis) then it outputs the measurement of the state ρ in the standard. If $b = 0$ (corresponding to opening in the stanadard basis) then it outputs the measurement of the state ρ in the Hadamard basis.
- Out takes as input the secret key $sk = (sk_0, sk_1, \dots, sk_{n+1})$, a commitment string $y = (y_0, y_1, \dots, y_{n+1})$, a basis $b \in \{0,1\}$ and an opening string $z \in \{0,1\}^{(n+1)^2}$ and does the following:
 - (1) If $b = 1$ then parse

$$\mathbf{z} = (\mathbf{d}_1, x'_{1,d_1}, \dots, \mathbf{d}_{n+1}, x'_{1,d_1})$$

and check that for every $i \in [n+1]$ it holds that

$$y_i = \text{Eval}(pk_i, (\mathbf{d}_i, x'_{i,d_i})).$$

If all these checks pass then output $\mathbf{d} \cdot (1, x_0 \oplus x_1)$, where $((0, x), (1, x_1)) = \text{Invert}(sk_0 y_0)$. Otherwise, output \perp .

- (2) If $b = 0$ then parse $\mathbf{z} = (z_1, \dots, z_{n+1})$, and for every $i \in [n+1]$ compute

$$((0, x'_{i,0}), (1, x'_{i,1})) = \text{Invert}(sk_i, y_i) \text{ and } m_i = z_i \cdot (1, x'_{i,0} \oplus x'_{i,1})$$

If $\text{Eval}(pk_0, (m_1, \dots, m_{i+1})) = y_0$ then output m_1 , and otherwise output \perp .

Analyzing the leakage. We next analyze the leakage that a cheating QPT algorithm Commit* and a cheating QPT algorithm Open* obtain by, given $pk = (pk_0, pk_1, \dots, pk_{n+1})$, generating a commitment string $y = (y_1, \dots, y_\ell)$, where each $y_i = (y_{i,0}, y_{i,1}, \dots, y_{i,n+1})$, a basis (b_1, \dots, b_ℓ) and an opening $\mathbf{z} = (z_1, \dots, z_\ell)$, and obtaining outputs $m_i = \text{Out}(sk, y_i, b_i, z_i)$ for every $i \in [\ell]$. Denote by

$$I = \{i : b_i = 0\} \text{ and } J = \{i : b_i = 1\}.$$

We distinguish between the leakage obtained from $\{m_i\}_{i \in I}$ and that obtained from $\{m_i\}_{i \in J}$. As mentioned above, $\{m_i\}_{i \in I}$ only leaks information about sk_1, \dots, sk_{n+1} , since sk_0 is not used when computing $\{m_i\}_{i: b_i=0}$. For $i \in J$, it holds that

$$m_i = \mathbf{d}_i \cdot (1, \mathbf{x}_{i,0} \oplus \mathbf{x}_{i,1}),$$

where

$$((0, \mathbf{x}_{i,0}), (1, \mathbf{x}_{i,1})) = \text{Invert}_{\text{TCF}}(sk_0, y_{i,0})$$

and

$$\mathbf{z}_i = (\mathbf{d}_{i,1}, x'_{i,1,d_1}, \dots, \mathbf{d}_{i,n+1}, x'_{i,n+1,d_{n+1}}).$$

This may leak information about sk_0 . In particular, if we use the underlying (noisy) TCF family from [4], along with an adversarially chosen $\mathbf{d} = (\mathbf{d}_1, \dots, \mathbf{d}_\ell)$ then $\{m_i\}_{i \in J}$ may leak part of the secret key which breaks the indistinguishability between the output produced by Open* and Ext^{Open*}.

We get around this problem by arguing that in our scheme if \mathbf{z} is accepted then it must be the case that the ‘‘important’’ bits of \mathbf{d} have min-entropy $\omega(\log \lambda)$.¹⁴ For this we rely on the fact that the underlying TCF family has the adaptive hardcore bit property, which the (noisy) TCF family from [4] was proven to have under the LWE assumption. We actually need the stronger condition that the ‘‘important’’ bits of \mathbf{d} have min-entropy $\omega(\log \lambda)$ even given some auxiliary input (which comes into play due to the fact that we are opening many qubits). We prove this for the specific NTCF family from [4]. Specifically, we prove that under the LWE assumption, the NTCF family from [4] has a property which we refer to as the *distributional strong adaptive hardcore bit property*. We argue that this property, together with the min-entropy property of \mathbf{d} , implies that the leakage obtained from $\mathbf{d}_i \cdot (\mathbf{x}_{i,0} \oplus \mathbf{x}_{i,1})$ is benign and does not break the indistinguishability between the output produced by Open* and Ext^{Open*}.

In more detail, for Mahadev's measurement protocol, the proof that the Hadamard outputs of Ext^{Open*} and Open* are computationally indistinguishable relies on the adaptive hardcore bit property, which states that for every QPT adversary A ,

$$\Pr[A(pk) = (b, \mathbf{x}_b, \mathbf{d}, \mathbf{d} \cdot (1, \mathbf{x}_0 \oplus \mathbf{x}_1))] \leq \frac{1}{2} + \text{negl}(\lambda)$$

where $((0, \mathbf{x}_0), (1, \mathbf{x}_1)) = \text{Invert}(sk, \text{Eval}(pk, b, \mathbf{x}_b))$. We need to argue that this holds even if A gets as auxiliary input a bunch of elements of the form

$$(b_i, \mathbf{x}_{i,b_i}, \mathbf{d}_i \cdot (1, \mathbf{x}_{i,0} \oplus \mathbf{x}_{i,1})).$$

While in general this is not true, we prove that it is true for the (noisy) TCF family from [3], if each \mathbf{d}_i has $\omega(\log \lambda)$ min-entropy (even conditioned on $(\mathbf{d}_1, \dots, \mathbf{d}_{i-1})$), under the LWE assumption.

¹⁴We emphasize that this is not the case for Mahadev's scheme, since in her scheme every Hadamard opening \mathbf{d} is accepted.

2.3 Succinct Commitments

As mentioned in the introduction, our main result is a *succinct* commitment scheme, where Commit commits to an ℓ -qubit state by generating a *succinct* classical commitment string that consists of only $\text{poly}(\lambda)$ many bits, and Open generates an opening to any qubit $i \in [\ell]$ in any basis $b \in \{0, 1\}$, where the opening consists of only $\text{poly}(\lambda)$ many bits. Importantly, the guarantee we provide is that even if Open* only opens to a few qubits, we should still be able to extract the entire ℓ -qubit quantum state from Open*.¹⁵ This seems impossible to do, since how can we extract information about qubits that were never opened? Indeed, to achieve this we need to change the syntax.

We add to the syntax an *interactive test phase*. Similar to the test round in Mahadev’s protocol, our test phase is executed with probability $1/2$, and if it is executed then after the test phase the protocol is terminated and the opening phase is never run. This is the case since the test phase destroys the quantum state. Importantly, we allow the test phase to be *interactive*. It is this interaction that allows us to extract a long ℓ -qubit state from Open*. Loosely speaking, in this test phase, we choose at random $b \leftarrow \{0, 1\}$ and ask the prover to provide an opening to all the ℓ -qubits in basis b^ℓ . To ensure that the protocol remains succinct, we ask for the openings to be sent in a succinct manner, using a Merkle hash. Then the prover and verifier engage in a succinct interactive argument where the prover proves knowledge of the committed openings. For this we use Kilian’s protocol and the fact that it is a proof-of-knowledge even in the post-quantum setting [8, 21]. Then the verifier sends the prover the secret key sk and the prover and verifier engage in a succinct interactive argument where the prover proves that the committed openings are accepted (w.r.t. sk). This is also done using the Kilian protocol.

In addition, we allow the commit phase to be interactive. This allows Commit to first generate a non-succinct commitment y , and send its Merkle hash, denoted by rt . Then the committer can run a succinct proof-of-knowledge interactive argument, to prove knowledge of a preimage of y . Importantly, the proof-of-knowledge must be *state-preserving*, which means that we can extract y without destroying the state. Such a state-preserving proof-of-knowledge protocol was recently constructed in [15]. This interactive commitment phase allows us to reduce the binding of the succinct commitment scheme to that of the semi-succinct one. This part of the analysis is similar to [1].

2.4 Applications

We construct a succinct interactive argument for **QMA** and a compiler that converts any X/Z PCP into a succinct interactive argument, both under the LWE assumption. For simplicity we do not use our succinct commitment scheme to construct these succinct interactive arguments. Rather we use our *semi-succinct* commitment scheme to construct a *semi-succinct* interactive argument. We then rely on a black-box transformation from [1] which shows a generic transformation for converting any semi-succinct interactive argument for **QMA** into a fully succinct one.¹⁶

¹⁵This guarantee is important for our applications, as we will see in Section 2.4.

¹⁶We mention that this transformation was used (in a non-black-box way to convert our semi-succinct commitment scheme into a succinct one.

An important point to note is that the argument systems we construct have negligible soundness, even though the binding guarantee of our commitment only holds for provers with a probability close to 1 of being accepted. We do this by applying sequential repetition to our semi-succinct protocol to drive down the soundness error, before applying the black-box transformation of [1]. A drawback of this approach (shared with all known succinct argument systems for **QMA**) is that the honest prover must hold polynomially many copies of the **QMA** witness state, or the X/Z PCP state.

The formal statements of these results are contained in the full version of the paper.

Compiling an X/Z PCP into a semi-succinct interactive argument. Our compiler uses a succinct commitment in a straightforward way. The succinct interactive argument proceeds as follows:

- (1) The verifier generates a key pair $(pk, sk) \leftarrow \text{Gen}(1^\lambda)$ corresponding to the underlying semi-succinct commitment scheme.
- (2) The prover commits to the X/Z PCP $|\pi\rangle$ by generating a classical commitment string $y \leftarrow \text{Commit}(pk, |\pi\rangle)$. It sends y to the verifier.
- (3) With probability $1/2$ the verifier behaves as the PCP verifier and chooses small set of indices (i_1, \dots, i_k) along with basis choices (b_1, \dots, b_k) ; with probability $1/2$ the verifier chooses a random $b \leftarrow \{0, 1\}$ and sends b to the prover.
- (4) If the prover receives a bit b then it opens the entire PCP in the standard basis if $b = 0$ and in the Hadamard basis if $b = 1$. Otherwise, if the prover receives a set of indices (i_1, \dots, i_k) along with basis choices (b_1, \dots, b_k) then the prover opens these locations in the desired basis.

Completeness follows immediately from the completeness of the underlying semi-succinct commitment scheme. To argue soundness, fix a cheating prover P^* that is accepted with high probability. We rely on the soundness property of the underlying commitment scheme to argue that there exists a QPT extractor that extracts a state $|\pi^*\rangle$ from P^* , such that on a random challenge produced by the PCP verifier (for which P^* succeeds in opening with high probability), the output of P^* is close to the the outcome obtained by measuring $|\pi^*\rangle$ directly, which implies that $|\pi^*\rangle$ is an X/Z PCP that is accepted with high probability, implying that the soundness property holds.

*Semi-succinct interactive argument for **QMA**.* To obtain a semi-succinct argument, we follow the blueprint of Mahadev [16]. Namely, we first convert the **QMA** witness into one that can be verified by measuring only in the X/Z basis. For this we rely on a result due to Fitzsimons, Hajdušek, and Morimae [10] which shows how to convert multiple copies of the **QMA** witness into an ℓ -qubit state $|\pi\rangle$ that can be verified by measuring it only in the X/Z basis. Importantly, this state can be verified by measuring it in a random basis $(b_1, \dots, b_\ell) \leftarrow \{0, 1\}^\ell$. Armed with this tool, the semi-succinct interactive argument proceeds as follows:

- (1) The verifier generates a key pair (pk, sk) corresponding to the underlying semi-succinct commitment scheme.
- (2) The prover converts its (multiple copies) of the **QMA** witness into a state $|\pi\rangle$ by relying on the [10] result, and computes $y \leftarrow \text{Commit}(pk, |\pi\rangle)$.

- (3) With probability $1/2$ the verifier chooses at random a seed $s \in \{0, 1\}^\lambda$ and sends s to the prover, and with probability $1/2$ the verifier chooses a random $b \leftarrow \{0, 1\}$ and sends b to the prover.
- (4) If the prover receives a bit b then it sends the opening of the commitment in the basis b^ℓ . If it receives a seed s then it uses a pseudorandom generator to deterministically expand s to a pseudorandom string (b_1, \dots, b_ℓ) and sends an opening of the commitment in basis (b_1, \dots, b_ℓ) .
- (5) The verifier uses its secret key to compute the output corresponding to this opening. If any of the openings are rejected it rejects. Otherwise, in the case that it sent a seed, it accepts if the verifier from [10] would have accepted.

To argue soundness, fix a cheating prover P^* that is accepted with high probability. We first rely on the soundness property of the underlying commitment scheme to argue that the QPT extractor extracts a state $|\pi^*\rangle$ from P^* , such that for any choice of basis $\mathbf{b} = (b_1, \dots, b_\ell)$ for which P^* succeeds in opening with high probability, the output corresponding to these openings are computationally indistinguishable from measuring $|\pi^*\rangle$ in basis \mathbf{b} . By the soundness of the underlying scheme [10] we note that for a random basis (b_1, \dots, b_ℓ) , the state would be rejected with high probability. Hence it must also be the case if the basis is pseudorandom, as otherwise one can distinguish a pseudorandom string from a truly random one.

Acknowledgements

This work was done in part while SG, AN, and AV were participants at the Simons Institute 2023 Summer Cluster on Quantum Computing, and we thank the Simons Institute and the organizers for the opportunity. Yael Kalai is supported by DARPA under Agreement No. HR00112020023. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Government or DARPA. Anand Natarajan is supported by NSF CAREER Grant CCF-2339948. Agi Villanyi acknowledges support by the Doc Bedard fellowship from the Laboratory for Physical Sciences through the Center for Quantum Engineering and the National Science Foundation Graduate Research Fellowship under Grant No. 2141064. Sam Gunn is supported by a Google PhD Fellowship and the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Quantum Systems Accelerator.

References

- [1] James Bartusek, Yael Tauman Kalai, Alex Lombardi, Fermi Ma, Giulio Malavolta, Vinod Vaikuntanathan, Thomas Vidick, and Lisa Yang. 2022. Succinct Classical Verification of Quantum Computation. doi:10.48550/ARXIV.2206.14929 1, 3, 5, 6, 7, 10
- [2] James Bartusek, Fuyuki Kitagawa, Ryo Nishimaki, and Takashi Yamakawa. 2023. Obfuscation of Pseudo-Deterministic Quantum Circuits. Cryptology ePrint Archive, Paper 2023/252. <https://eprint.iacr.org/2023/252> <https://eprint.iacr.org/2023/252>. 6
- [3] Bob Blakley, G. R. Blakley, Agnes Hui Chan, and James L. Massey. 1993. Threshold Schemes with Disenrollment. 540–548. doi:10.1007/3-540-48071-4_38 9
- [4] Zvika Brakerski, Paul F. Christiano, Urmila Mahadev, Umesh V. Vazirani, and Thomas Vidick. 2018. A Cryptographic Test of Quantumness and Certifiable Randomness from a Single Quantum Device. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7–9, 2018*, Mikkel Thorup (Ed.). IEEE Computer Society, 320–331. doi:10.1109/FOCS.2018.00038 4, 5, 9
- [5] Gilles Brassard, David Chaum, and Claude Crépeau. 1988. Minimum Disclosure Proofs of Knowledge. *J. Comput. Syst. Sci.* 37, 2 (1988), 156–189. doi:10.1016/0022-0000(88)90005-0 1
- [6] David Chaum, Ivan Damgård, and Jeroen van de Graaf. 1988. Multiparty Computations Ensuring Privacy of Each Party’s Input and Correctness of the Result. 87–119. doi:10.1007/3-540-48184-2_7 1
- [7] Nai-Hui Chia, Kai-Min Chung, and Takashi Yamakawa. 2020. Classical Verification of Quantum Computations with Efficient Verifier. 181–206. doi:10.1007/978-3-030-64381-2_7 6
- [8] Alessandro Chiesa, Fermi Ma, Nicholas Spooner, and Mark Zhandry. 2021. Post-Quantum Succinct Arguments: Breaking the Quantum Rewinding Barrier. Cryptology ePrint Archive, Paper 2021/334. <https://eprint.iacr.org/2021/334>. 3, 10
- [9] Amos Fiat and Adi Shamir. 1987. How to Prove Yourself: Practical Solutions to Identification and Signature Problems. 186–194. doi:10.1007/3-540-47721-7_12 1
- [10] Joseph F. Fitzsimons, Michal Hajdušek, and Tomoyuki Morimae. 2018. Post hoc Verification of Quantum Computation. *Phys. Rev. Lett.* 120 (Jan 2018), 040501. Issue 4. doi:10.1103/PhysRevLett.120.040501 10, 11
- [11] Oded Goldreich, Silvio Micali, and Avi Wigderson. 1986. Proofs that Yield Nothing But their Validity and a Methodology of Cryptographic Protocol Design (Extended Abstract). 174–187. doi:10.1109/SFCS.1986.47 1
- [12] Oded Goldreich, Silvio Micali, and Avi Wigderson. 1987. How to Play any Mental Game or A Completeness Theorem for Protocols with Honest Majority. 218–229. doi:10.1145/28395.28420 1
- [13] Sam Gunn, Nathan Ju, Fermi Ma, and Mark Zhandry. 2022. Commitments to Quantum States. Cryptology ePrint Archive, Paper 2022/1358. <https://eprint.iacr.org/2022/1358> <https://eprint.iacr.org/2022/1358>. 1, 3, 6
- [14] Yael Kalai, Alex Lombardi, Vinod Vaikuntanathan, and Lisa Yang. 2022. Quantum Advantage from Any Non-Local Game. arXiv:2203.15877 [quant-ph] 6
- [15] Alex Lombardi, Fermi Ma, and Nicholas Spooner. 2022. Post-Quantum Zero Knowledge, Revisited or: How to Do Quantum Rewinding Undetectably. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*. IEEE, 851–859. doi:10.1109/FOCS54457.2022.00086 3, 10
- [16] Urmila Mahadev. 2018. Classical Homomorphic Encryption for Quantum Circuits. In *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7–9, 2018*, Mikkel Thorup (Ed.). IEEE Computer Society, 332–338. doi:10.1109/FOCS.2018.00039 1, 2, 3, 4, 5, 6, 7, 8, 10
- [17] Ralph C Merkle. 1987. A digital signature based on a conventional encryption function. In *Conference on the theory and application of cryptographic techniques*. Springer, 369–378. 1
- [18] Tony Metger, Anand Natarajan, and Tina Zhang. 2024. Succinct arguments for QMA from standard assumptions via compiled nonlocal games. (2024). To appear. 6
- [19] Silvio Micali. 1994. CS Proofs (Extended Abstracts). 436–453. doi:10.1109/SFCS.1994.365746 1
- [20] Thomas Vidick. 2020. Interactions with Quantum Devices (Course). <http://users.cms.caltech.edu/~vidick/teaching/fsmp/fsmp.pdf>. 5, 6, 7
- [21] Thomas Vidick and Tina Zhang. 2021. Classical Proofs of Quantum Knowledge. In *Advances in Cryptology - EUROCRYPT 2021 - 40th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Zagreb, Croatia, October 17–21, 2021, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 12697)*, Anne Canteaut and François-Xavier Standaert (Eds.). Springer, 630–660. doi:10.1007/978-3-030-77886-6_22 10

Received 2024-11-04; accepted 2025-02-01