

# Quantum Advantage from Any Non-Local Game

Yael Kalai\*    Alex Lombardi†    Vinod Vaikuntanathan‡    Lisa Yang§

September 30, 2022

## Abstract

We show a general method of compiling any  $k$ -prover non-local game into a single-prover interactive game maintaining the same (quantum) completeness and (classical) soundness guarantees (up to negligible additive factors in a security parameter). Our compiler uses any quantum homomorphic encryption scheme (Mahadev, FOCS 2018; Brakerski, CRYPTO 2018) satisfying a natural form of correctness with respect to auxiliary (quantum) input. The homomorphic encryption scheme is used as a cryptographic mechanism to simulate the effect of spatial separation, and is required to evaluate  $k - 1$  prover strategies (out of  $k$ ) on encrypted queries.

In conjunction with the rich literature on (entangled) multi-prover non-local games starting from the celebrated CHSH game (Clauser, Horne, Shimony and Holt, Physical Review Letters 1969), our compiler gives a broad framework for constructing mechanisms to classically verify quantum advantage.

**Keywords:** Non-local games, CHSH, Quantum advantage, Fully Homomorphic Encryption.

---

\*MIT and Microsoft Research. E-mail: yael@microsoft.com.

†MIT. E-mail: alexjl@mit.edu. Research supported by a Charles M. Vest Grand Challenges Fellowship, an NDSEG Fellowship, and grants of the third author.

‡MIT. E-mail: vinodv@mit.edu. Research supported in part by DARPA under Agreement No. HR00112020023, a grant from the MIT-IBM Watson AI, a grant from Analog Devices and a Microsoft Trustworthy AI grant. 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.

§MIT. E-mail: lisayang@mit.edu. Research supported by an NSF Graduate Research fellowship and grants of the third author.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Our Results . . . . .	4
1.2	Technical Overview . . . . .	5
<b>2</b>	<b>Preliminaries</b>	<b>9</b>
2.1	Non-Local Games . . . . .	10
2.2	Quantum Homomorphic Encryption . . . . .	11
<b>3</b>	<b>Our Compiler: From Non-Local Games to Interactive Protocols</b>	<b>12</b>
3.1	Our Transformation, $k = 2$ case. . . . .	13
3.2	Extension to $k$ -Player Games . . . . .	17
3.3	Achieving $1 - \text{negl}(\lambda)$ Quantum-Classical Gap . . . . .	20
3.4	Non-Interactive Protocols in the ROM . . . . .	22
<b>4</b>	<b>Protocols for Verifying Quantum Advantage</b>	<b>23</b>
4.1	Compiling the CHSH Game . . . . .	23
4.2	Future Directions . . . . .	24
<b>A</b>	<b>QHE Correctness with Respect to Auxiliary Input</b>	<b>29</b>
A.1	The [Mah18a] QFHE Scheme . . . . .	31

# 1 Introduction

Quantum computing promises to usher in a new era in computer science. There are several ongoing large-scale efforts by Google, Microsoft, IBM and a number of startups to demonstrate *quantum advantage*, that is, to construct a quantum machine capable of computations that no classical machine can replicate with comparable speed. Central to the demonstration of quantum advantage is the question of verification. That is, can a classical machine<sup>1</sup> *check* that a given device is performing computations that no classical device is capable of?

Such verification schemes for quantum advantage are central to the ongoing race to construct non-trivial quantum computers. Several mechanisms have been proposed to classically verify quantum advantage:

1. **The Algorithmic Method:** Ask the quantum machine to do a computation, such as factor a large number [Sho97], that is believed to be (sub-exponentially) hard for classical machines [LJMP90]. This is the most straightforward way to check quantum advantage, yet it is concretely very expensive and requires millions of (high-fidelity) qubits which is likely infeasible for near-term quantum machines (see [GE21] for a recent, optimized estimate).
2. **The Sampling Method:** Ask the quantum machine to (approximately) sample from a distribution, which is presumably hard to accomplish classically [BJS10, AA10]. This results in tasks that are feasible for near-term quantum machines [AAB<sup>+</sup>19], yet suffers from two problems. First, checking that the output is correct (at least in the current proposals) is an exponential-time computation that quickly becomes infeasible for a classical machine. Secondly, the question of whether approximate sampling is classically hard is less clear and has come under question [LLL<sup>+</sup>21], despite hardness results for exact sampling [BFNV18].
3. **The Interactive Method:** The most recent example of a way to demonstrate quantum advantage relies on using interactive proofs and the fact that quantum algorithms cannot necessarily be “rewound” which, in turn, is a consequence of the *quantum no-cloning principle*. Starting with the work of Brakerski, Christiano, Mahadev, Vazirani and Vidick [BCM<sup>+</sup>18], there have been a handful of proposals reducing the number of rounds [BKVV20], improving computational efficiency and near-term feasibility [KCVY21], and extending to more powerful guarantees such as self-testing [MV21] and quantum delegation [Mah18b]. The practicality of the quantum machine in these protocols (eg. [ZKML<sup>+</sup>21]) seems to fall somewhere between (1) and (2), i.e., the prover is more practical than with the algorithmic method but less than the sampling method. Compared to (2), verification by the classical machine is efficient.

Meanwhile, from more than half a century ago, starting with the celebrated works of John Bell [Bel64] and Clauser, Horne, Shimony and Holt [CHSH69] to explain the Einstein-Podolsky-Rosen (EPR) paradox, there is a rich body of work that came up with several methods to test that two or more spatially separated, non-communicating quantum devices can generate (testable) correlations that no pair of classical devices can produce. This immediately gives us a test of quantum advantage

---

<sup>1</sup>One could relax this requirement to allow a verifier with *limited* quantum capabilities, for example, preparing and transmitting single-qubit states [ABOE08, FK17]. In this paper, we deal exclusively with classical verifiers.

of a very different kind, one whose validity relies on the (non-falsifiable) physical assumption that devices do not communicate.

4. **The Non-Local Method:** Here, quantum advantage comes from *entanglement*. Ask for a pair (or more) of spatially separated, but entangled, quantum devices to perform a computation that no classical machine can replicate without communication. A classical example is the famed CHSH game [CHSH69], or other violations of the Bell inequality. A rich class of non-local games have since been constructed and this is an active area of current research with many breakthroughs in the last few years (e.g. [Mer90, Ara01, Har93, GHZ07, RUV13] up until [JNV<sup>+</sup>20]). While many of these games are attractive and super-efficient (e.g. using only a small constant number of qubits), this suffers from possible “loopholes”. That is, *how does one ensure that the provers do not subliminally communicate in some way?* Indeed, constructing so-called loophole-free demonstrations of Bell inequality violations is an active area of research [HBD<sup>+</sup>15].

## 1.1 Our Results

A tantalizing question that comes out of the discussion above is the relation between techniques used to construct *multi-prover* non-local games and the ones used to construct *single-prover* proofs of quantumness. On the one hand, the former relies on entanglement whereas the latter relies on quantum non-rewinding. On the other hand, the intellectual foundations of single-prover protocols for quantum advantage sometimes come from non-local games. Indeed, Kahanamoku-Meyer, Choi, Vazirani and Yao [KCVY21], in a recent work, call their interactive single-prover protocol for quantum advantage a “computational Bell test”. Yet, there have been no formal connections between multi-prover non-local games and single-prover interactive protocols for quantum advantage.

The central result of this work is the construction of a *simple* and *general* compiler that gives us an interactive protocol demonstrating quantum advantage starting from two building blocks.

**Theorem 1.1.** *Given*

- Any  $k$ -prover non-local game quantum value  $c$  and classical value  $s$ ; and
- Any quantum homomorphic encryption scheme satisfying correctness with respect to auxiliary input (Definition 2.3), that can implement homomorphic evaluation of at least  $k - 1$  provers in the non-local game.

*there is a  $2k$ -round<sup>2</sup> single-prover interactive game with completeness  $c$ , realized by a quantum polynomial-time algorithm, and soundness  $s + \text{negl}(\lambda)$ , against any classical polynomial-time algorithm. Here,  $\lambda$  is a security parameter that governs the security of the quantum homomorphic encryption scheme.*

This gives us a rich class of protocols by mixing and matching any non-local game (CHSH, Magic Square, Odd-Cycle Test, GHZ, and many more) and any powerful enough quantum homomorphic encryption scheme ([Mah18a, Bra18] both satisfy the appropriate correctness property; see Appendix A).

---

<sup>2</sup>We refer to a message from the prover to the verifier or the verifier to the prover as a round.

For example, an instantiation of our method using the CHSH game and the quantum FHE scheme of Mahadev [Mah18a] gives us a protocol that already improves on the state of the art in several respects:

- Compared to [BCM<sup>+</sup>18], the protocols that come out of our framework do not require the adaptive hardcore bit property, and therefore can be instantiated from a variety of assumptions including Ring Learning with Errors (Ring-LWE), resulting in better efficiency.
- Compared to [BKVV20], our protocols do not require random oracles. [BKVV20] uses random oracles in a crucial way to obtain an instantiation from Ring-LWE, bypassing the adaptive hardcore bit property. They simultaneously achieve a two-round protocol in contrast to the 4-round protocol of [BCM<sup>+</sup>18].
- Compared to [KCVY21], our protocols require 4 rounds as opposed to their 6.

Our framework elucidates interactive protocols for quantum advantage by separating them into an information-theoretic component (namely, the non-local game) and a cryptographic component (namely, the QFHE). We believe that the richness and extreme simplicity of our transformation could eventually lead to a protocol that concretely improves on the family of protocols in (3) above in a significant way. Philosophically, a remarkable aspect of our framework is that it uses cryptography as a bridge to connect two apparently different sources of quantumness, namely *entanglement* (used to construct non-local games) and *non-rewinding* (which underlies the reason why single-prover proofs of quantumness as in [BCM<sup>+</sup>18] work).

We proceed to describe a self-contained outline of our transformation below.

## 1.2 Technical Overview

Our framework is remarkable in its simplicity, drawing inspiration from similar compilers in the classical world [BMW98, BC12, KRR13, KRR14]. For the rest of the introduction, we focus on compiling *two-prover* non-local games and describe how to extend our construction to  $k$  provers in Section 3.2. The basic idea is to convert any non-local game into a single prover protocol by sending the prover the queries sent to *both* provers in the non-local game. It is easy to see that if we do this “in the clear,” we lose soundness, since by giving both queries to a single prover, we lose the spatial separation condition. This is precisely where the encryption scheme comes into play.

**Using Cryptography to Enforce Spatial Separation.** We enforce spatial separation using an encryption scheme. Specifically, in our protocol the verifier samples two queries  $q_1, q_2$  by emulating the verifier in the non-local game, and the interactive protocol proceeds in rounds, as follows:

1. The verifier sends  $\text{Enc}(q_1)$  to the prover.
2. The prover sends an encrypted answer  $\text{Enc}(a_1)$ .
3. The verifier then sends the second query  $q_2$  (in the clear).
4. The prover replies with an answer  $a_2$ .

The verifier accepts if and only if the decrypted transcript would be accepted in the non-local game.

**Completeness.** We use the properties of the underlying encryption scheme to argue that any strategy of the (quantum) provers in the non-local game can be emulated in our protocol (partially “under the hood” of the encryption scheme). For this, we need our homomorphic encryption scheme to have the following two properties:

- It needs to support the quantum operations done by the first prover in the non-local game.
- It needs to be *entanglement preserving* in the following sense: if a quantum circuit  $C$  is defined using a register  $A$  and is homomorphically evaluated using a state  $|\Psi\rangle_{AB}$  that has entanglement between  $A$  and another register  $B$  (which is not input to  $C$ ), the decrypted output  $y$  should remain correlated with  $B$  (as it would be if  $C$  were evaluated using the  $A$  register of  $|\Psi\rangle$ ). This property can be enforced via a natural “aux-input correctness” property (Definition 2.3).

We view correctness with respect to auxiliary input as a natural requirement that one should expect to hold for a “typical” QHE scheme. In Appendix A, we show that it follows from “plain” correctness under mild structural assumptions; namely, if decryption is “bit-by-bit” and homomorphic evaluation is “local.” In particular, this implies that seminal homomorphic encryption schemes due to Mahadev and Brakerski [Mah18a, Bra18] satisfy these two properties.

**Remark 1.2.** *We note that a 2-message version of this protocol was used by Kalai, Raz and Rothblum [KRR13, KRR14] (and in follow-up works) to obtain a 2-message delegation scheme for classical computations. In this 2-message variant, the verifier simply sends all of the queries in the first message each encrypted under independently chosen keys, and the prover computes each answer homomorphically. This 2-message variant has the desired completeness property described above.*

*However, it cannot be used as a protocol that has a gap between classical and quantum winning probabilities. This is the case since the soundness of the resulting 2-message delegation scheme is proven only if the original non-local game has non-signaling soundness, namely, assuming the provers cannot cheat via a non-signaling (and therefore, a quantum) strategy. In other words, classical provers may be able to implement non-signaling strategies, and these strategies can have value which is at least as high as the quantum value. Thus, we cannot argue that in the [KRR14]-like 2-message protocol, a classical cheating prover cannot emulate this non-signaling strategy. Instead, we consider the 4-message variant described above.*

*This 4-message protocol bears some resemblance to (although is significantly simpler than) the classical succinct argument system of [BC12], which is also concerned with compiling MIPs into single-prover protocols.*

**Soundness.** We argue that any computationally bounded *classical* cheating prover in the 4-message protocol can be converted into a *local* prover strategy in the 2-player game, with roughly

the same acceptance probability (see Theorem 3.2). This allows us to argue quantum advantage: Namely, an honest quantum prover can obtain the (quantum) value of the 2-player game, whereas any classical prover can obtain only the classical value of the 2-player game.

We prove our desired soundness condition with a security reduction that *rewinds* the (classical) adversary, something that cannot be done with a general quantum prover.

**Special Case: the CHSH Game.** To illustrate how our compiler preserves soundness, we first analyze a simple special case of our transformation where the 2-prover game is the CHSH game [CHSH69]. In the CHSH game, the verifier sends i.i.d. random bits  $x, y$  to  $P_1, P_2$  (respectively); the provers are then supposed to return bits  $a, b$  such that  $a \oplus b = xy$ . The best local strategy for this game has success probability  $3/4$  (e.g., if  $P_1, P_2$  assume that  $xy = 0$ ). This can be proved with a straightforward direct analysis:

- Any (say, deterministic) local strategy can be described by bits  $a_0, a_1, b_0, b_1$  (corresponding to the two provers' responses when the verifier sends 0 or 1).
- The strategy is correct on query pair  $(x, y)$  if  $a_x \oplus b_y = xy$ .
- If the strategy is correct on *both*  $(x, 0)$  and  $(x, 1)$ , then

$$a_x \oplus b_1 = x = x \oplus a_x \oplus b_0,$$

which can only occur if  $x = b_0 \oplus b_1$ . Thus, at least one of the four pairs  $(x, y)$  must be answered incorrectly.

By making use of *rewinding*, it is possible to give a direct computational analogue of this soundness analysis for our 4-message compiled protocol. In the 4-message protocol, an adversarial prover  $P^*$  is given as input a ciphertext  $\hat{x}$  encrypting a bit  $x$ , and returns a ciphertext  $\hat{a}$  (which contains a message  $a = a(x)$  that the verifier decrypts). Then, the verifier sends a bit  $y$  and the prover replies with a bit  $b$ . Unlike the 2-prover setting, this bit  $b$  *may* depend on  $x$ ; however, since  $x$  is encrypted,  $b$  itself should not reveal *information* about  $x$ . However, one must be careful when arguing security because the verifier's *decision* (in the interactive game) makes use of the decryption key  $sk$ , which *can* reveal information about  $x$ .

Nevertheless, there is a simple argument showing that a classical prover cannot convince  $V$  with probability higher than  $3/4 + \text{negl}$ :

- Let  $P^*$  be an adversarial prover (deterministic without loss of generality), and fix:
  - A ciphertext  $\hat{x}$  sent by the verifier in the first round (sampled as a random encryption of a random bit), and
  - A ciphertext  $\hat{a}$  returned by the prover in the second round.
- Moreover, let  $b_0$  and  $b_1$  denote the prover's responses on challenge  $y = 0$  and  $y = 1$  respectively (given fixed  $\hat{x}, \hat{a}$ ).

- Then, if the verifier accepts *both*  $b_0$  and  $b_1$  (on challenges 0 and 1), we have

$$b_1 \oplus a = x = b_0 \oplus a \oplus x,$$

where  $a = \text{Dec}(\hat{a})$ . As before, this implies that  $b_0 \oplus b_1 = x$ !

- Thus, if  $P^*$  is more than  $3/4$ -convincing, we will break the security of the encryption scheme by guessing  $x$  with non-negligible advantage over  $1/2$ , given only  $\hat{x}$  (and the encryption key):
  - Given a ciphertext  $\hat{x}$ , run  $P^*$  on  $\hat{x}$  to obtain  $\hat{a}$ .
  - Run  $P^*$  on  $y = 0$ , obtaining  $b_0$ , and then *rewind*  $P^*$  and run  $P^*$  on  $y = 1$ , obtaining  $b_1$ .
  - Output  $b_0 \oplus b_1$ .

A straightforward calculation tells us that our advantage of guessing  $x$  in this reduction (compared to random guess) is non-negligible if  $P^*$  wins the interactive game with probability non-negligibly higher than  $3/4$ .

This proves the security of our transformation for the special case of the CHSH game.

**The General Case.** Having illustrated the utility of our transformation in a special case, we now analyze its soundness for *any* 2-prover game.

Fix a classical cheating prover  $P^*$  for the 4-message protocol, and assume without loss of generality that it is deterministic (otherwise, fix its random coins to ones that maximize the success probability). The corresponding local provers  $(P_1^*, P_2^*)$  for the 2-player game are defined as follows:

- Fix the first 2 messages in the 4-message protocol, by choosing any  $q_1'$  (e.g., it can be the all-zero string), computing  $\text{ct}_1 = \text{Enc}(q_1')$  and emulating  $P^*$  to compute  $\text{ct}_2 = P^*(\text{ct}_1)$ .
- $P_2^*$  has  $(\text{ct}_1, \text{ct}_2)$  hardwired.<sup>3</sup> On input a query  $q_2$ , it emulates the response of  $P^*$ , conditioned on the first two messages being  $(\text{ct}_1, \text{ct}_2)$ , to obtain  $a_2 = P^*(\text{ct}_1, \text{ct}_2, q_2)$ . It outputs  $a_2$ .
- $P_1^*$  also has  $(\text{ct}_1, \text{ct}_2)$  hardwired (and thus can evaluate  $P_2^*$  on an arbitrary input). On input  $q_1$  it computes for every possible  $a_1$  the probability (over  $q_2$  sampled conditioned on  $q_1$ ) that the verifier accepts  $(q_1, a_1, q_2, P_2^*(q_2))$  and sends  $a_1$  with the maximal acceptance probability.

To argue that  $(P_1^*, P_2^*)$  convinces the verifier to accept with essentially the same probability that  $P^*$  does, we rely on the security of the encryption scheme. Namely, we argue (by contradiction) that if the verifier accepts with probability significantly smaller than the acceptance probability of  $P^*$  in the 4-message protocol, then there exists an adversary  $\mathcal{A}$  that breaks semantic security of the underlying encryption scheme.

Specifically,  $\mathcal{A}$  is given a challenge ciphertext  $\text{ct}$ , which it will use to define  $P_1^*$  and  $P_2^*$  (as above), and will use  $P_1^*$  and  $P_2^*$  in his attack. The first barrier is that  $P_1^*$  may not be efficient, which results with  $\mathcal{A}$  being inefficient. Thus, we first argue that (a good enough approximation of)  $P_1^*$  can be emulated in time  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda)$ , which results in  $\mathcal{A}$  running in that time as

---

<sup>3</sup>We note that it suffices to hardwire  $\text{ct}_1$  since  $\text{ct}_2 = P^*(\text{ct}_1)$  can be computed from  $\text{ct}_1$ .

well. Loosely speaking, this is done by estimating for every  $a_1$ , the probability of acceptance, and taking  $a_1$  with the maximal acceptance probability. This explains the exponential blowup in  $a_1$ . To estimate the probability that  $(q_1, a_1)$  is accepted (w.r.t.  $P_2^*$ ) we do an empirical estimation by choosing many  $q_2$ 's from the residual query distribution (conditioned on  $q_1$ ) and compute the fraction of  $q_2$ 's for which  $V(q_1, q_2, a_1, P^*(q_2)) = 1$ . We use the Chernoff bound to argue that this empirical estimation is close to the real probability of acceptance. Note that this estimation requires sampling  $q_2$  from the residual query distribution (conditioned on  $q_1$ ), and it is not clear that this can be done efficiently. Thus, we hardwire into  $\mathcal{A}$  many such samples for each and every possible  $q_1$ , which results with the exponential blowup in  $q_1$ .<sup>4</sup>

Given this  $2^{|q_1|+|a_1|}\text{poly}(\lambda)$ -time implementation of  $P_1^*$ , the hardness reduction proceeds as follows. Loosely speaking, given  $q_1, q_1'$ , ct where ct is either  $\text{Enc}(q_1)$  or  $\text{Enc}(q_1')$  with equal probability, we run the 2-prover game with  $(P_1^*, P_2^*)$  corresponding to ct and with the first query being  $q_1$ , and then run it again with the first query being  $q_1'$ . If the verifier accepts exactly one of these executions, then we guess  $b$  to be the one corresponding to the winning execution; otherwise, we output a random guess for  $b$ .

Note that in our attack we broke the security in time  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda)$  and thus need to assume that the encryption scheme has that level of security. For constant-size games, this is equivalent to polynomial security; in general, this requires assuming the sub-exponential security of the underlying FHE. However, as discussed in [Section 3.3.1](#), it is possible to rely on polynomially-secure FHE for certain superconstant-size non-local games such as parallel repetitions of a constant-size game. We refer the reader to [Section 3](#) for the details, and to the proof of [Theorem 3.2](#) for the formal analysis.

## 2 Preliminaries

We let PT denote deterministic polynomial time, PPT denote probabilistic polynomial time and QPT denote quantum polynomial time. For any random variables  $A$  and  $B$  (possibly parametrized by a security parameter  $\lambda$ ), we use the notation  $A \equiv B$  to denote that  $A$  and  $B$  are identically distributed, and  $A \approx_s B$  to denote that  $A$  and  $B$  have negligible ( $\lambda^{-\omega(1)}$ ) statistical distance. We use  $A \approx_c B$  to denote that  $A$  and  $B$  are computationally indistinguishable, namely for every PPT distinguisher  $D$ ,  $|\Pr[D(A) = 1] - \Pr[D(B) = 1]| = \lambda^{-\omega(1)}$ .

**Quantum States.** Let  $\mathcal{H}$  be a Hilbert space of finite dimension  $2^n$  (thus,  $\mathcal{H} \simeq \mathbb{C}^{2^n}$ ). A (pure,  $n$ -qubit) quantum state  $|\Psi\rangle \in \mathcal{H}$  is an element of the form

$$|\Psi\rangle = \sum_{b_1, \dots, b_n \in \{0,1\}} \alpha_{b_1, \dots, b_n} |b_1, \dots, b_n\rangle$$

where  $\{|b_1, \dots, b_n\rangle\}_{b_1, \dots, b_n \in \{0,1\}}$  forms an orthonormal basis of  $\mathcal{H}$ ,  $\alpha_{b_1, \dots, b_n} \in \mathbb{C}$  and

$$\sum_{b_1, \dots, b_n \in \{0,1\}} |\alpha_{b_1, \dots, b_n}|^2 = 1.$$

---

<sup>4</sup>However, this blowup can be avoided for many non-local games; for example, if  $q_1$  and  $q_2$  are independent, or if the game is a parallel repetition of a constant-size game. For the latter case, see [Section 3.3.1](#) for more details.

We refer to  $n$  as the number of registers of  $|\Psi\rangle$ .

We denote by  $M(|\Psi\rangle)$  the outcome of measuring  $|\Psi\rangle$  in the standard basis (throughout this work we are only concerned with standard basis measurements). For any set of registers  $\mathcal{I} \subseteq [n]$ , we denote by  $M_{\mathcal{I}}(|\Psi\rangle)$  the outcome of measuring only the  $\mathcal{I}$  registers of  $|\Psi\rangle$  in the standard basis.

A *mixed state*  $\rho$  over  $\mathcal{H}$  is a density operator (we use  $\mathbf{S}(\mathcal{H})$  to denote the space of density operators on  $\mathcal{H}$ ) normalized so that  $\text{Tr}(\rho) = 1$ . Every pure state  $|\Psi\rangle$  has a corresponding rank 1 mixed state  $|\Psi\rangle\langle\Psi|$  such that for any PSD projection  $\Pi$ , we have  $\|\Pi|\Psi\rangle\|^2 = \text{Tr}(\Pi|\Psi\rangle\langle\Psi|)$ .

We sometimes divide the registers of  $\mathcal{H}$  into named registers, denoted by calligraphic upper-case letters, such as  $\mathcal{A}$  and  $\mathcal{B}$ , in which case we also decompose the Hilbert space into  $\mathcal{H} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ , so that each pure quantum state  $|\Psi\rangle$  is a linear combination of quantum states  $|\Psi_{\mathcal{A}}\rangle \otimes |\Psi_{\mathcal{B}}\rangle \in \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ .

For any mixed state  $\rho$ , we denote by  $\rho_{\mathcal{A}}$  the reduced density operator

$$\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}}(\rho) \in \mathbf{S}(\mathcal{H}_{\mathcal{A}}),$$

where  $\text{Tr}_{\mathcal{B}}$  is the ‘‘partial trace’’ linear operator defined by

$$\text{Tr}_{\mathcal{B}}(|\Psi_{\mathcal{A}}\rangle\langle\Psi_{\mathcal{A}}| \otimes |\Psi_{\mathcal{B}}\rangle\langle\Psi_{\mathcal{B}}|) = |\Psi_{\mathcal{A}}\rangle\langle\Psi_{\mathcal{A}}| \cdot \text{Tr}(|\Psi_{\mathcal{B}}\rangle\langle\Psi_{\mathcal{B}}|).$$

**Quantum Circuits and Locality.** A quantum circuit is a sequence of elementary quantum gates (taken from some complete basis) and measurement operations. These operations are applied to an initial state  $|\psi\rangle$  and result in some final state  $|\psi'\rangle$ . For further definitions related to quantum circuits and gate types, we refer the reader to [NC00].

One direct consequence of this model is a form of **locality**: if  $|\psi\rangle = |\psi\rangle_{\mathcal{A}\mathcal{B}}$  is shared between two registers ( $\mathcal{A}$  and  $\mathcal{B}$ ), operations of the form  $C \otimes \text{Id}_{\mathcal{B}}$  (where  $C$  is a quantum circuit acting only on  $\mathcal{A}$ ) can be performed *without possession of the register  $\mathcal{B}$* .

## 2.1 Non-Local Games

**Definition 2.1** (*k*-Player Non-local Game). A *k*-player non-local game  $\mathcal{G}$  consists of a PPT sampleable query distribution  $\mathcal{Q}$  over  $(q_1, \dots, q_k) \in (\{0, 1\}^n)^k$  and a polynomial time verification predicate  $\mathcal{V}(q_1, q_2, \dots, q_k, a_1, a_2, \dots, a_k) \in \{0, 1\}$ , where each  $a_i \in \{0, 1\}^m$ . The classical value and the quantum value of  $\mathcal{G}$  are defined below.

- **Classical (local) value:** The classical (or local) value  $v$  of  $\mathcal{G}$  is defined as:

$$\max_{P_1, \dots, P_k: \{0, 1\}^n \rightarrow \{0, 1\}^m} \Pr_{(q_1, \dots, q_k) \leftarrow \mathcal{Q}} [\mathcal{V}(q_1, \dots, q_k, P_1(q_1), \dots, P_k(q_k)) = 1]$$

- **Quantum (entangled) value:** The quantum (or entangled) value  $v^*$  of  $\mathcal{G}$  is defined as:

$$\sup_{\substack{|\Psi\rangle \in \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_k \\ U_i \in U(\mathcal{H}_i \otimes \mathbb{C}^n), 1 \leq i \leq k}} \Pr_{(q_1, \dots, q_k) \leftarrow \mathcal{Q}} [\mathcal{V}(q_1, \dots, q_k, a_1, \dots, a_k) = 1]$$

where the maximum is over all *k*-partite states  $|\Psi\rangle \in \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_k$  (each  $\mathcal{H}_i$  is an arbitrary finite-dimensional Hilbert space), unitaries  $U_i$  acting on  $\log \dim_{\mathbb{C}} \mathcal{H}_i + n$  qubits

respectively, and answers  $a_i$  computed by applying  $U_1 \otimes \dots \otimes U_k$  to  $|\Psi\rangle \otimes |q_1 \dots q_k\rangle$ , and measuring the first  $m$  qubits in each  $\mathcal{H}_i$ ; that is,

$$(a_1, a_2, \dots, a_k) \leftarrow M_{\mathcal{I}}(U_1 \otimes \dots \otimes U_k (|\Psi\rangle \otimes |q_1, \dots, q_k\rangle))$$

for  $\mathcal{I} = \{1, \dots, m\} \times [k]$ . We remark that without loss of generality,  $|\Psi\rangle$  can be (and is above) taken to be a pure state.

We are interested in non-local games where quantum strategies can win with probability strictly more than any classical strategy, that is ones for which  $v^* > v$ .

**Remark 2.2.** In this work, it is crucial to consider the complexity of the honest provers in a non-local game, which is often not a parameter of interest in the literature but instead hidden in the description of the provers as a tuple of unitaries  $(U_1, \dots, U_k)$  and state  $|\Psi\rangle$ .

We explicitly consider non-local games where the each prover's unitary  $U_i$  can be implemented as a quantum circuit  $C_i$  of size polynomial in a security parameter  $\lambda$ . When  $\mathcal{G}$  has constant size (such as the CHSH or Magic square games) this holds automatically, but this is a non-trivial requirement when the size of  $\mathcal{G}$  grows with  $\lambda$ .

## 2.2 Quantum Homomorphic Encryption

We define the notion of quantum homomorphic encryption [Mah18a, Bra18] that is central to our framework. Unlike in [Bra18], our definition requires a form of *correctness with respect to auxiliary input*. However, we show that this definition holds for QHE schemes satisfying mild additional requirements, and in particular holds for the [Mah18a, Bra18] schemes.

**Definition 2.3** (Quantum Homomorphic Encryption (QHE)). *A quantum homomorphic encryption scheme  $\text{QHE} = (\text{Gen}, \text{Enc}, \text{Eval}, \text{Dec})$  for a class of quantum circuits  $\mathcal{C}$  is a tuple of algorithms with the following syntax:*

- *Gen is a PPT algorithm that takes as input the security parameter  $1^\lambda$  and outputs a (classical) secret key  $\text{sk}$  of  $\text{poly}(\lambda)$  bits;*
- *Enc is a PPT algorithm that takes as input a secret key  $\text{sk}$  and a classical input  $x$ , and outputs a ciphertext  $\text{ct}$ ;*
- *Eval is a QPT algorithm that takes as input a tuple  $(C, |\Psi\rangle, \text{ct}_{\text{in}})$ , where  $C : \mathcal{H} \times (\mathbb{C}^2)^{\otimes n} \rightarrow (\mathbb{C}^2)^{\otimes m}$  is a quantum circuit,  $|\Psi\rangle \in \mathcal{H}$  is a quantum state, and  $\text{ct}_{\text{in}}$  is a ciphertext corresponding to an  $n$ -bit plaintext. Eval computes a quantum circuit  $\text{Eval}_C(|\Psi\rangle \otimes |0\rangle^{\text{poly}(\lambda, n)}, \text{ct}_{\text{in}})$  which outputs a ciphertext  $\text{ct}_{\text{out}}$ . If  $C$  has classical output, we require that  $\text{Eval}_C$  also has classical output.*
- *Dec is a PT algorithm that takes as input a secret key  $\text{sk}$  and ciphertext  $\text{ct}$ , and outputs a state  $|\phi\rangle$ . Additionally, if  $\text{ct}$  is a classical ciphertext, the decryption algorithm outputs a classical string  $y$ .*

The above syntax is more general than the form required in [Bra18]; we elaborate on this difference below. We require the following two properties from  $(\text{Gen}, \text{Enc}, \text{Eval}, \text{Dec})$ :

- **Correctness with Auxiliary Input:** For every security parameter  $\lambda \in \mathbb{N}$ , any quantum circuit  $C : \mathcal{H}_A \times (\mathbb{C}^2)^{\otimes n} \rightarrow \{0, 1\}^*$  (with classical output), any quantum state  $|\Psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ , any message  $x \in \{0, 1\}^n$ , any secret key  $\text{sk} \leftarrow \text{Gen}(1^\lambda)$  and any ciphertext  $\text{ct} \leftarrow \text{Enc}(\text{sk}, x)$ , the following states have negligible trace distance:

*Game 1.* Start with  $(x, |\Psi\rangle_{AB})$ . Evaluate  $C$  on  $x$  and register  $A$ , obtaining classical string  $y$ . Output  $y$  and the contents of register  $B$ .

*Game 2.* Start with  $\text{ct} \leftarrow \text{Enc}(\text{sk}, x)$  and  $|\Psi\rangle_{AB}$ . Compute  $\text{ct}' \leftarrow \text{Eval}_C(\cdot \otimes |0\rangle^{\text{poly}(\lambda, n)}, \text{ct})$  on register  $A$ . Compute  $y' = \text{Dec}(\text{sk}, \text{ct}')$ . Output  $y'$  and the contents of register  $B$ .

- **$T$ -Classical Security:** For any two messages  $x_0, x_1$  and any  $\text{poly}(T(\lambda))$ -size classical circuit ensemble  $\mathcal{A}$ :

$$\left| \Pr \left[ \mathcal{A}(\text{ct}_0) = 1 \mid \begin{array}{l} \text{sk} \leftarrow \text{Gen}(1^\lambda) \\ \text{ct}_0 \leftarrow \text{Enc}(\text{sk}, x_0) \end{array} \right] - \Pr \left[ \mathcal{A}(\text{ct}_1) = 1 \mid \begin{array}{l} \text{sk} \leftarrow \text{Gen}(1^\lambda) \\ \text{ct}_1 \leftarrow \text{Enc}(\text{sk}, x_1) \end{array} \right] \right| \leq \text{negl}(T(\lambda)) .$$

In words, “correctness with auxiliary input” requires that if QHE evaluation is applied to a register  $\mathcal{A}$  that is a part of a joint (entangled) state in  $\mathcal{H}_A \otimes \mathcal{H}_B$ , the entanglement between the QHE evaluated output and  $\mathcal{B}$  is preserved.

**Remark 2.4.** A *quantum fully homomorphic encryption* (QFHE) is a QHE for the class of all poly-size quantum circuits. While [Mah18a, Bra18] construct QFHE (with security against quantum distinguishers), weaker forms of QHE may yield more efficient quantum advantage protocols (see Section 4.2 for discussion).

**Remark 2.5.** In our definition of security, we only consider classical attacks. Classical security is sufficient for the purposes of this work as protocols for quantum advantage are required to have quantum completeness and classical soundness.

For the purposes of this paper, it suffices to know the following claim about the instantiability of Definition 2.3.

**Claim 2.6.** The [Mah18a, Bra18] QFHE schemes satisfy Definition 2.3 with correctness holding for the class of all poly-size quantum circuits.

Claim 2.6 can be verified by inspecting the constructions given in [Mah18a, Bra18]. In Appendix A, we show mild generic conditions under which a QFHE scheme satisfies correctness with respect to auxiliary input, and sketch a proof of Claim 2.6 .

### 3 Our Compiler: From Non-Local Games to Interactive Protocols

In this section, we show how to use a quantum homomorphic encryption scheme satisfying aux-input correctness (Definition 2.3) to convert a 2-prover non-local game into a *single-prover* interactive protocol with *computational soundness*.

**Definition 3.1** (Single-Prover Computationally Sound Interactive Game). *A single-prover computationally sound (CS) interactive game  $\mathcal{G}$  consists of an interactive PPT verifier  $\mathcal{V}$  that takes as input a security parameter  $1^\lambda$  and interacts with an interactive prover. The classical (computationally sound) value and the quantum (computationally sound) value of  $\mathcal{G}$  are defined below.*

- *Classical CS value:  $\mathcal{G}$  has classical CS value  $\geq v$  if and only if there exists an interactive polynomial-size Turing machine  $\mathcal{P}$  such that for every  $\lambda \in \mathbb{N}$ ,*

$$\Pr\left[(\mathcal{P}, \mathcal{V})(1^\lambda) = 1\right] \geq v$$

*where the probability is over the random coin tosses of  $\mathcal{V}$ , and where  $(\mathcal{P}, \mathcal{V})(1^\lambda) \in \{0, 1\}$  denotes the output bit of  $\mathcal{V}(1^\lambda)$  after interacting with  $\mathcal{P}$ .*

- *Quantum CS value:  $\mathcal{G}$  has quantum CS value  $\geq v^*$  if and only if there exists a Hilbert space  $\mathcal{H}$  and a quantum state  $|\Psi\rangle \in \mathcal{H}$  and an interactive QPT prover  $\mathcal{P}$  such that for every  $\lambda \in \mathbb{N}$ ,*

$$\Pr\left[(\mathcal{P}(|\Psi\rangle), \mathcal{V})(1^\lambda) = 1\right] \geq v^*$$

*where the probability is over the randomness of  $\mathcal{P}$  and  $\mathcal{V}$ .*

### 3.1 Our Transformation, $k = 2$ case.

Fix a quantum homomorphic encryption scheme  $\text{QHE} = (\text{Gen}, \text{Enc}, \text{Eval}, \text{Dec})$  for a class of quantum circuits  $\mathcal{C}$  (e.g. [Mah18a, Bra18], see Claim 2.6). We present a PPT transformation  $\mathcal{T}$  that converts any 2-prover non-local game  $\mathcal{G} = (\mathcal{Q}, \mathcal{V})$  into a single-prover computationally sound interactive game  $\mathcal{T}^{\mathcal{G}}$  (associated with security parameter  $\lambda$ ), defined as follows.

1. The verifier samples  $(q_1, q_2) \leftarrow \mathcal{Q}$ ,  $\text{sk} \leftarrow \text{Gen}(1^\lambda)$ , and  $\hat{q}_1 \leftarrow \text{Enc}(\text{sk}, q_1)$ . In the first round the verifier sends  $\hat{q}_1$  and in the third round he sends  $q_2$ .
2. The verifier, upon receiving  $\hat{a}_1$  from the prover in the first round, and  $a_2$  in the second round, accepts if and only if  $\mathcal{V}(q_1, q_2, \text{Dec}(\text{sk}, \hat{a}_1), a_2) = 1$ .

**Theorem 3.2.** *Fix any QHE scheme for a circuit class  $\mathcal{C}$ , and any 2-player non-local game  $\mathcal{G} = (\mathcal{Q}, \mathcal{V})$  with classical value  $v$  and quantum value  $v^*$ , such that the value  $v^*$  is obtained by a prover strategy  $(C_1^*, C_2^*)$  with a quantum state  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$  where  $C_1^*(|\Psi\rangle_A, \cdot) \in \mathcal{C}$ . Denote by  $|q_1|$  and  $|a_1|$  the lengths of the query and answer of the first prover, respectively. If the underlying QHE encryption scheme is  $T$ -secure, for  $T(\lambda) = 2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda)$ , then the following holds:*

1. *The quantum CS value of  $\mathcal{T}^{\mathcal{G}}$  is at least  $v^*$ .*
2. *The classical CS value of  $\mathcal{T}^{\mathcal{G}}$  is at most  $v + \text{negl}(\lambda)$ .*

*Proof.* Fix any 2-player non-local game  $\mathcal{G} = (\mathcal{Q}, \mathcal{V})$  with classical value  $v$  and quantum value  $v^*$ , such that the value  $v^*$  is obtained by a prover strategy  $(C_1^*, C_2^*)$  with the (joint) quantum state  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ , where  $C_1^* \in \mathcal{C}$ .

1. **The quantum CS value of  $\mathcal{T}^{\mathcal{G}}$  is at least  $v^*$ .** Consider the following BQP prover  $\mathcal{P}^*$ :
  - (a) In the first round, upon receiving  $\hat{q}_1$ ,  $\mathcal{P}^*$  computes  $\text{ct}' \leftarrow \text{Eval}(\cdot_{\mathcal{A}}, C_1^*, \hat{q}_1)$  on the register  $\mathcal{A}$  of  $|\Psi\rangle_{\mathcal{AB}}$ , and sends  $\text{ct}$ . As internal state,  $\mathcal{P}^*$  retains the contents of register  $\mathcal{B}$ .
  - (b) In the second round, upon receiving  $q_2$ ,  $\mathcal{P}^*$  uses its internal state  $\rho_{\mathcal{B}}$  to compute and send  $a'_2 \leftarrow C_2^*(\cdot_{\mathcal{B}}, q_2)$ .

We argue that

$$\Pr[\mathcal{V}(q_1, q_2, a'_1, a'_2) = 1] = v^* - \text{negl}(\lambda),$$

in the probability space where:

- $(q_1, q_2) \leftarrow \mathcal{Q}$
- $\text{sk} \leftarrow \text{Gen}(1^\lambda)$  and  $\hat{q}_1 \leftarrow \text{Enc}(q_1, \text{sk})$
- $\text{ct}' \leftarrow \text{Eval}(\cdot_{\mathcal{A}}, C_1^*, \hat{q}_1)$  applied to  $|\Psi\rangle_{\mathcal{AB}}$  and  $\rho$  set to the contents of  $\mathcal{B}$ .
- $a'_1 = \text{Dec}(\text{sk}, \text{ct}')$ , and
- $a'_2 \leftarrow C_2^*(\rho, q_2)$

By the aux-input correctness of QHE ([Definition 2.3](#)) and the fact that  $C_1^* \in \mathcal{C}$ , we see that for every  $q_1$ , the mixed state consisting of the distribution over  $(a'_1, \rho)$  above is the same (up to negligible trace distance) as what would have been obtained by applying  $C_1 \otimes \text{Id}_{\mathcal{B}}$  to  $(|\Psi\rangle_{\mathcal{AB}}, q_1)$ . By the contractivity of trace distance (with respect to the map defined by  $C_2^*(\cdot_{\mathcal{B}}, q_2)$  and  $\mathcal{V}$ ), we conclude that

$$\Pr[\mathcal{V}(q_1, q_2, a'_1, a'_2) = 1] = \Pr[\mathcal{V}(q_1, q_2, a_1, a_2) = 1] \pm \text{negl}(\lambda) = v^* \pm \text{negl}(\lambda),$$

as desired.

2. **The classical CS value of  $\mathcal{T}^{\mathcal{G}}$  is at most  $v + \text{negl}(\lambda)$ .** Suppose for the sake of contradiction that the classical value of  $\mathcal{T}^{\mathcal{G}}$  is  $v' = v + \delta$  for a non-negligible  $\delta = \delta(\lambda)$ . This implies that there exists a (deterministic) poly-size classical prover  $\tilde{\mathcal{P}}$  such that for every  $\lambda \in \mathbb{N}$ ,

$$\Pr\left[(\tilde{\mathcal{P}}, \mathcal{V}(\mathcal{T}^{\mathcal{G}}))(1^\lambda) = 1\right] = v',$$

where  $\mathcal{V}(\mathcal{T}^{\mathcal{G}})$  denotes the verifier in the protocol  $\mathcal{T}^{\mathcal{G}}$ . Next, for every  $\lambda \in \mathbb{N}$  we convert  $\tilde{\mathcal{P}}$  into (local) classical provers  $(\mathcal{P}_1, \mathcal{P}_2) = (\mathcal{P}_1(\lambda), \mathcal{P}_2(\lambda))$  such that there exists a negligible function  $\mu$  such that for every  $\lambda \in \mathbb{N}$ ,

$$\Pr[(\mathcal{P}_1, \mathcal{P}_2, \mathcal{V}) = 1] \geq v' - \mu(\lambda).$$

Since  $v' - \mu(\lambda) > v$  for sufficiently large  $\lambda$ , this contradicts the fact that the classical value of  $\mathcal{G}$  is at most  $v$ .

To that end, for every  $i \in \{1, 2\}$  we denote by  $\mathcal{Q}_i$  the residual distribution of  $\mathcal{Q}$  corresponding to player  $\mathcal{P}_i$ . Namely,  $\mathcal{Q}_i$  samples  $(q_1, q_2) \leftarrow \mathcal{Q}$  and outputs  $q_i$ . Similarly, we denote by  $\mathcal{Q}|_{q_1}$  to be the distribution that samples  $(q'_1, q'_2) \leftarrow \mathcal{Q}$  conditioned on  $q'_1 = q_1$ , and outputs  $q'_2$ .

We next define  $(\mathcal{P}_1, \mathcal{P}_2)$ :

- (a) Choose  $q'_1 \leftarrow \mathcal{Q}_1$  and generate  $\text{sk} \leftarrow \text{Gen}(1^\lambda)$ . Let  $\text{ct}_1 \leftarrow \text{Enc}(\text{sk}, q'_1)$  and  $\text{ct}_2 = \tilde{\mathcal{P}}(\text{ct}_1)$  (i.e.,  $\text{ct}_2$  is the first message sent by  $\tilde{\mathcal{P}}$  upon receiving  $\text{ct}_1$  from the verifier).
- (b)  $\mathcal{P}_2$  has the ciphertexts  $(\text{ct}_1, \text{ct}_2)$  hardwired into it. On input  $q_2$ , it simply emulates the response of  $\tilde{\mathcal{P}}$  given the first three messages  $(\text{ct}_1, \text{ct}_2, q_2)$  to obtain  $a_2$ . It outputs  $a_2$ .
- (c)  $\mathcal{P}_1$  also has  $(\text{ct}_1, \text{ct}_2)$  hardwired into it. On input  $q_1$ , it computes and outputs  $a_1$  that maximizes the probability of the verifier accepting (w.r.t.  $\mathcal{P}_2$  defined above). Namely, it outputs

$$a_1 = \text{argmax}_{a_1} \Pr_{q_2 \leftarrow \mathcal{Q}|q_1} [\mathcal{V}(q_1, q_2, a_1, \mathcal{P}_2(q_2)) = 1].$$

We next argue that there exists a negligible function  $\mu = \mu(\lambda)$  such that for every  $\lambda \in \mathbb{N}$ ,

$$\Pr[(\mathcal{P}_1(\lambda), \mathcal{P}_2(\lambda), \mathcal{V}) = 1] \geq v' - \mu(\lambda),$$

as desired. To this end, suppose for the sake of contradiction that there exists a non-negligible  $\epsilon = \epsilon(\lambda)$  such that for every  $\lambda \in \mathbb{N}$ ,

$$\Pr[(\mathcal{P}_1(\lambda), \mathcal{P}_2(\lambda), \mathcal{V}) = 1] \leq v' - \epsilon(\lambda). \quad (1)$$

We construct an adversary  $\mathcal{A}$  of size  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda/\epsilon)$  that breaks the semantic security of the underlying encryption scheme with advantage  $\frac{\epsilon}{4}$ .

The adversary  $\mathcal{A}$  will use his challenge ciphertext  $\text{ct}$  to define  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , and will use  $\mathcal{P}_1$  and  $\mathcal{P}_2$  in his attack. Note that  $\mathcal{P}_2$  can be efficiently emulated in time  $\text{poly}(\lambda)$  (assuming  $\lambda$  is larger than the communication complexity of  $\mathcal{G}$ ). However,  $\mathcal{P}_1$  may not be efficient. In what follows we show that the maximization problem implicit in  $\mathcal{P}_1$  can be *approximated* in (non-uniform) time  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda/\epsilon)$ . More specifically, we show that there exists a function  $F$ , that takes as input a ciphertext  $\text{ct}$  and a query  $q_1 \in \text{Support}(\mathcal{Q}_1)$ , it runs in time  $2^{|a_1|+|q_1|} \cdot \text{poly}(\frac{\lambda}{\epsilon})$ , and for every  $q_1 \in \text{Support}(\mathcal{Q}_1)$ ,

$$\Pr_{q_2 \leftarrow \mathcal{Q}|q_1} [\mathcal{V}(q_1, q_2, \mathcal{P}_1(q_1), \mathcal{P}_2(q_2)) = 1] - \Pr_{q_2 \leftarrow \mathcal{Q}|q_1} [\mathcal{V}(q_1, q_2, F(\text{ct}, q_1), \mathcal{P}_2(q_2)) = 1] \leq \frac{\epsilon}{2}, \quad (2)$$

where  $\mathcal{P}_1, \mathcal{P}_2$  are defined w.r.t. the ciphertext  $\text{ct}$ .

In what follows, we describe  $F$  as having randomized advice, but we will later set its advice to be "the best possible", and thus obtain a deterministic function. For every possible  $q_1 \in \text{Support}(\mathcal{Q}_1)$ , we hardwire  $N = \frac{9(\lambda+|a_1|)}{\epsilon^2}$  queries  $q_{2,1}, \dots, q_{2,N}$  sampled independently from the distribution  $\mathcal{Q}|q_1$ .

$F(\text{ct}, q_1)$  is computed by approximating for every  $a_1$  the probability

$$p_{q_1, a_1} = \Pr_{q_2 \leftarrow \mathcal{Q}|q_1} [\mathcal{V}(q_1, q_2, a_1, \mathcal{P}_2(q_2)) = 1]$$

by its empirical value

$$p'_{q_1, a_1} = \frac{1}{N} |\{i : \mathcal{V}(q_1, q_{2,i}, a_1, \mathcal{P}_2(q_{2,i})) = 1\}|.$$

It outputs  $a_1$  with the maximal value of  $p'_{q_1, a_1}$ .

Note that (as a circuit)  $F$  is of size  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda)$ , since it has hardwired into it  $N \cdot 2^{|q_1|}$  queries hardwired ( $N$  for each possible  $q_1$ ), and on input  $(\text{ct}, q_1)$  it runs in time  $2^{|a_1|} \cdot N \cdot \text{poly}(\lambda)$ . Thus, its total size is as desired.

By a Chernoff bound,<sup>5</sup>

$$\Pr \left[ |p'_{q_1, a_1} - p_{q_1, a_1}| > \frac{\epsilon}{3} \right] \leq 2^{-2N(\epsilon/3)^2} = 2^{-\lambda} \cdot 2^{-|a_1|}$$

From the equation above (and applying a union bound over all  $a_1$ ) indeed the difference between the two probabilities in Equation (2) is at most  $\frac{\epsilon}{3} + 2^{-\lambda} \leq \frac{\epsilon}{2}$ , as desired.

We are now ready to define our adversary  $\mathcal{A}$  that will use  $(F, P_2)$  to break semantic security. Specifically,  $\mathcal{A}$  takes as input a tuple  $(q_{1,0}, q_{2,0}, q_{1,1}, q_{2,1}, \text{ct})$ , where  $(q_{1,0}, q_{2,0}), (q_{1,1}, q_{2,1}) \leftarrow \mathcal{Q}$ , and  $\text{ct}$  is distributed by choosing  $\text{sk} \leftarrow \text{Gen}(1^\lambda)$  and  $b^* \leftarrow \{0, 1\}$ , and sampling  $\text{ct} \leftarrow \text{Enc}(\text{sk}, q_{1, b^*})$ . It guesses  $b^*$  as follows:

- (a) For every  $b \in \{0, 1\}$ , run  $F(\text{ct}, q_{1, b})$  in time  $2^{|q_1|+|a_1|} \cdot \text{poly}(\lambda/\epsilon)$  and compute  $a_{1, b} = F(\text{ct}, q_{1, b})$ .
- (b) If there exists  $b \in \{0, 1\}$  such that

$$\mathcal{V}(q_{1, b}, q_{2, b}, a_{1, b}, \mathcal{P}_2(q_{2, b})) = 1 \quad \wedge \quad \mathcal{V}(q_{1, 1-b}, q_{2, 1-b}, a_{1, 1-b}, \mathcal{P}_2(q_{2, 1-b})) = 0$$

then output  $b$ . Else output a random  $b \leftarrow \{0, 1\}$ .

Note that by definition of  $\mathcal{P}^*$  and  $(\mathcal{P}_1, \mathcal{P}_2)$ , it holds that for  $b = b^*$ ,

$$\Pr_{(q_{1, b}, q_{2, b}) \leftarrow \mathcal{Q}} [\mathcal{V}(q_{1, b}, q_{2, b}, \mathcal{P}_1(q_b), \mathcal{P}_2(q_{2, b})) = 1] \geq v'.$$

Thus, by Equation (2), it holds that for  $b = b^*$ ,

$$\Pr_{(q_{1, b}, q_{2, b}) \leftarrow \mathcal{Q}} [\mathcal{V}(q_{1, b}, q_{2, b}, F(\text{ct}, q_{1, b}), \mathcal{P}_2(q_{2, b})) = 1] \geq v' - \frac{\epsilon}{2}. \quad (3)$$

On the other hand, by our contradiction assumption (Equation (1)) it holds that for  $b = 1 - b^*$ ,

$$\Pr_{(q_{1, b}, q_{2, b}) \leftarrow \mathcal{Q}} [\mathcal{V}(q_{1, b}, q_{2, b}, F(\text{ct}, q_{1, b}), \mathcal{P}_2(q_{2, b})) = 1] \leq v' - \epsilon. \quad (4)$$

Denote by  $E_{\text{Good}}$  the event that

$$\mathcal{V}(q_{1, b}, q_{2, b}, F(\text{ct}, q_{1, b}), \mathcal{P}_2(q_{2, b})) = 1 \quad \text{for} \quad b = b^*$$

and

$$\mathcal{V}(q_{1, b}, q_{2, b}, F(\text{ct}, q_{1, b}), \mathcal{P}_2(q_{2, b})) = 0 \quad \text{for} \quad b = 1 - b^*$$

<sup>5</sup>The form of Chernoff bound that we use here is that for  $X_1, \dots, X_N$  identically and independently distributed in  $\{0, 1\}$  with expectation  $\mu$ , it holds that  $\Pr \left[ \left| \frac{1}{N} \sum_{i=1}^N X_i - \mu \right| > \delta \right] \leq 2^{-2N\delta^2}$ .

Similarly, denote by  $E_{\text{Bad}}$  the event that

$$\mathcal{V}(q_{1,b}, q_{2,b}, F(\text{ct}, q_{1,b}), \mathcal{P}_2(q_{2,b})) = 0 \quad \text{for } b = b^*$$

and

$$\mathcal{V}(q_{1,b}, q_{2,b}, F(\text{ct}, q_{1,b}), \mathcal{P}_2(q_{2,b})) = 1 \quad \text{for } b = 1 - b^*$$

Denote by  $E$  the event that

$$\mathcal{V}(q_{1,b}, q_{2,b}, F(\text{ct}, q_{1,b}), \mathcal{P}_2(q_{2,b})) = 1 \quad \forall b \in \{0, 1\}$$

Then by Equation (3),

$$\Pr[E_{\text{Good}}] + \Pr[E] \geq v' - \frac{\epsilon}{2},$$

and by Equation (4),

$$\Pr[E_{\text{Bad}}] + \Pr[E] \leq v' - \epsilon,$$

which together imply that

$$\Pr[E_{\text{Good}}] - \Pr[E_{\text{Bad}}] \geq \frac{\epsilon}{2}. \tag{5}$$

Thus,

$$\begin{aligned} \Pr[b = b^*] &\geq \\ &\frac{1}{2} \cdot (1 - \Pr[E_{\text{Good}} \cup E_{\text{Bad}}]) + \Pr[E_{\text{Good}}] \geq \\ &\frac{1}{2} + \frac{1}{2}(\Pr[E_{\text{Good}}] - \Pr[E_{\text{Bad}}]) = \\ &\frac{1}{2} + \frac{\epsilon}{4}, \end{aligned}$$

as desired, where the first equation follows from the definition of  $E_{\text{Good}}$  and  $E_{\text{Bad}}$  and the definition of  $\mathcal{A}$ , the second equation follows from the union bound, the third equation follows from Equation (5). This contradicts the security of QHE; thus, we conclude the desired bound on the classical CS value of  $\mathcal{T}^{\mathcal{G}}$ .  $\square$

### 3.2 Extension to $k$ -Player Games

In this section, we generalize [Theorem 3.2](#) to  $k$ -player games for  $k > 2$ . We begin with a construction that is a  $2k$ -round analogue of the transformation  $\mathcal{T}$  from [Theorem 3.2](#): given any  $k$ -player non-local game  $\mathcal{G}$ , we define the following interactive game  $\mathcal{T}^{\mathcal{G}}$ :

1. The verifier samples  $(q_1, \dots, q_k) \leftarrow \mathcal{Q}$ ,  $\text{sk}_1, \dots, \text{sk}_{k-1} \leftarrow \text{Gen}(1^\lambda)$ , and  $\hat{q}_i \leftarrow \text{Enc}(\text{sk}_i, q_i)$  for each  $1 \leq i \leq k-1$ .
2. For each  $1 \leq i \leq k-1$ , in round  $2i-1$  the verifier sends  $\hat{q}_i$ . In round  $2i$  the prover responds with a ciphertext  $\hat{a}_i$ .
3. In round  $2k-1$  the verifier sends  $q_k$ ; in round  $2k$  the prover responds with some string  $a_k$ .

4. The verifier decrypts each  $\hat{a}_i$  with  $\text{sk}_i$  and accepts if and only if the transcript  $(q_1, a_1, \dots, q_k, a_k)$  is accepting according to  $\mathcal{G}$ .

We prove the following theorem.

**Theorem 3.3.** *Fix any QHE scheme (satisfying correctness with respect to auxiliary inputs) for a circuit class  $\mathcal{C}$ , and any  $k$ -player non-local game  $\mathcal{G} = (\mathcal{Q}, \mathcal{V})$  with classical value  $v$  and quantum value  $v^*$ , such that the value  $v^*$  is obtained by a prover strategy  $(C_1^*, \dots, C_k^*)$  with a quantum state  $|\Psi\rangle \in \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_k$  with each  $C_i^* \in \mathcal{C}$  (except possibly  $C_k^*$ ). Denote by  $|q_i|$  and  $|a_i|$  the lengths of the query and answer of  $P_i$ , respectively. If the underlying QHE encryption scheme is  $T$ -secure, for  $T(\lambda) = 2^{\sum_{i=1}^{k-1} (|q_i| + |a_i|)} \cdot \text{poly}(\lambda)$ , then the following holds:*

1. The quantum CS value of  $\mathcal{T}^{\mathcal{G}}$  is at least  $v^*$ .
2. The classical CS value of  $\mathcal{T}^{\mathcal{G}}$  is at most  $v + \text{negl}(\lambda)$ .

*Proof.* We briefly sketch the quantum CS value of  $\mathcal{T}^{\mathcal{G}}$ . Given a  $k$ -tuple of entangled provers  $\mathcal{P}_1, \dots, \mathcal{P}_k$  (with shared state  $|\Psi\rangle_{\mathcal{A}_1, \dots, \mathcal{A}_k}$ ), we define the following prover  $\mathcal{P}$  for the interactive game:

- $\mathcal{P}$  initially has internal state  $|\Psi\rangle_{\mathcal{A}_1, \dots, \mathcal{A}_k}$ .
- Given  $\hat{q}_i$  (for each  $1 \leq i \leq k-1$ ),  $\mathcal{P}$  homomorphically evaluates the circuit defining  $\mathcal{P}_i$  on register  $\mathcal{A}_i$  and  $\hat{q}_i$  (tracing out any ancilla registers).  $\mathcal{P}$  sends the encrypted answer  $\hat{a}_i$  to the verifier.
- Given  $q_k$ ,  $\mathcal{P}$  evaluates the circuit defining  $\mathcal{P}_k$  on  $\mathcal{A}_k$  (and  $q_k$ ), and sends the answer  $a_k$  to the verifier.

Analogously to [Theorem 3.2](#), the aux-input correctness of QHE implies that the verifier will accept with probability  $v^*(\mathcal{P}_1, \dots, \mathcal{P}_k) \pm \text{negl}(\lambda)$ , where  $v^*(\mathcal{P}_1, \dots, \mathcal{P}_k, |\Psi\rangle)$  denotes the value of strategy  $(\mathcal{P}_1, \dots, \mathcal{P}_k, |\Psi\rangle)$ . In more detail, we invoke aux-input correctness and the contractivity of trace distance  $k-1$  times sequentially (starting with auxiliary registers  $(\mathcal{A}_2, \dots, \mathcal{A}_k)$  and removing one  $\mathcal{A}_i$  each time).

We now bound the classical value of  $\mathcal{T}^{\mathcal{G}}$  via the following argument. Suppose that a (computationally bounded) classical interactive prover  $\tilde{\mathcal{P}}$  (deterministic without loss of generality) has value  $v'$  in  $\mathcal{T}^{\mathcal{G}}$ . We will construct local provers  $(P_1^*, \dots, P_k^*)$  winning  $\mathcal{G}$  with probability at least  $v' - \text{negl}(\lambda)$ .

To this end, we sample  $(q'_1, \dots, q'_k) \leftarrow \mathcal{Q}$ , secret keys  $\text{sk}'_1, \dots, \text{sk}'_{k-1} \leftarrow \text{Gen}(1^\lambda)$  and  $\text{ct}'_i \leftarrow \text{Enc}(\text{sk}'_i, q'_i)$  for  $1 \leq i \leq k-1$ . Each prover  $P_i^*$  has  $\text{ct}'_1, \dots, \text{ct}'_i$  hardwired into its description. The prover  $P_k^*$  simply emulates the last message function of  $\tilde{\mathcal{P}}$ ; namely, upon receiving  $q_k$  it emulates  $\tilde{\mathcal{P}}$  assuming that the first  $k-1$  messages from the verifier were  $\text{ct}'_1, \dots, \text{ct}'_{k-1}$ . We next define  $P_1^*, \dots, P_{k-1}^*$  recursively starting with  $P_1^*$ .

Assuming we have already defined  $P_1^*, \dots, P_{\ell-1}^*$  (this includes the base case  $\ell = 1$ ), we define  $P_\ell^*$  and an *interactive* prover  $\tilde{P}_{\ell+1, \dots, k}$  that has  $\text{ct}'_1, \dots, \text{ct}'_\ell$  hardwired to its description, and is sequentially given  $\hat{q}_{\ell+1}, \dots, \hat{q}_{k-1}, q_k$  as inputs and returns  $\hat{a}_{\ell+1}, \dots, \hat{a}_{k-1}, a_k$  as outputs.

- $\tilde{P}_{\ell+1,\dots,k}$  simply emulates  $\tilde{P}$  using hard-coded  $ct'_1, \dots, ct'_\ell$ .
- $P_\ell^*$  is given as input  $q_\ell$  and outputs an optimum of the following maximization problem:

$$a_\ell^* = \arg \max_{a_\ell} \Pr_{\{q_j\}_{j \neq \ell} \leftarrow \mathcal{Q}|q_\ell} [\mathcal{V}(q_1, a_1, \dots, q_k, a_k) = 1],$$

where  $a_j = P_j^*(q_j)$  for all  $j < \ell$ , and for all  $j > \ell$ ,  $a_j$  is obtained by running  $\tilde{P}_{\ell+1,\dots,k}$  on encryptions (under fresh secret keys) of  $q_{\ell+1}, \dots, q_k$  and then (unless  $j = k$ ) decrypting the resulting answers.

Note that by construction,  $P_k^* = \tilde{P}_k$ , and  $P_1^*, \dots, P_k^*$  are indeed local. Moreover, just as in the proof of [Theorem 3.2](#), we can *approximately* solve the maximization problems defined in  $P_1^*, \dots, P_{k-1}^*$  with functions  $F_1, \dots, F_{k-1}$  that can be implemented in time  $2^{\sum_{i=1}^{k-1} |q_i| + |a_i|} \cdot \text{poly}(\lambda)$ . This is done, given an inverse polynomial error  $\epsilon$ , by hard-coding for each  $q_i$ ,  $N = \frac{18k^2(\lambda + |a_i|)}{\epsilon^2}$  samples  $\{q_j^{(\ell)}\}_{j \neq k}$  (for  $1 \leq \ell \leq N$ ) from  $\mathcal{Q}|q_i$ , and will result in provers  $F_1, \dots, F_{k-1}, P_k^*$  that attain value matching  $P_1^*, \dots, P_k^*$  up to error  $\epsilon/4$ .

Thus, to complete the proof of [Theorem 3.3](#) it remains to prove the following claim.

**Claim 3.4.** *The tuple  $(P_1^*, \dots, P_k^*)$  has success probability at least  $v' - \text{negl}(\lambda)$ .*

*Proof.* Assume that  $(P_1^*, \dots, P_k^*)$  has success probability at most  $v' - \epsilon$  for some non-negligible  $\epsilon$ . We first replace  $P_i^*$  by  $F_i$  defined above, and obtain that  $(F_1, \dots, F_k)$  has success probability at most  $v' - 3\epsilon/4$ . We now derive a contradiction by a hybrid argument. Specifically, for every  $j$ , we define the quantity

$$\text{Hyb}_j = \Pr_{\substack{q_1, \dots, q_k \leftarrow \mathcal{Q} \\ \text{for } i \leq j: a_i = F_i(q_i) \\ \text{for } i > j: a_i = \text{Dec}(\hat{a}_i), \hat{a}_i \text{ output by } \tilde{P}_{j+1, \dots, k}}} [\mathcal{V}(q_1, a_1, \dots, q_k, a_k) = 1]$$

Note that  $\text{Hyb}_0$  is equal to the success probability of  $\tilde{P}$ , which is equal to  $v'$  by assumption, while  $\text{Hyb}_{k-1}$  is equal to the value of  $(F_1, \dots, F_k)$ .

We now claim that  $\text{Hyb}_j > \text{Hyb}_{j-1} - \frac{\epsilon}{4k} - \text{negl}(\lambda)$  for every  $j$ . To prove this, we will reduce from the security of QHE with respect to ciphertext  $ct'_j$ ; note that  $F_1, \dots, F_{j-1}$  do not depend on  $ct'_j$ . Ciphertexts  $ct'_1, \dots, ct'_{j-1}$  will remain fixed for this entire argument, while  $ct'_{j+1}, \dots, ct'_{k-1}$  are not used by any algorithms in  $\text{Hyb}_{j-1}$  or  $\text{Hyb}_j$ .

Define the auxiliary quantity  $\text{Hyb}'_j$  to be the same as  $\text{Hyb}_j$ , except that  $ct'_j$  is sampled as  $\text{Enc}(\text{sk}'_j, q_j)$ , where  $q_j$  is the input sent to  $F_j$  in the experiment. Note that  $\text{Hyb}'_j > \text{Hyb}_{j-1} - \frac{\epsilon}{4k}$ , because the particular choice of  $a_j^* = \text{Dec}(\tilde{P}_{j,\dots,k}(ct'_j))$  in the maximization problem defining  $P_j^*$  would have value  $\text{Hyb}_{j-1}$  (as this strategy matches the value of  $(F_1, \dots, F_{j-1}, \tilde{P}_{j,\dots,k})$ ), and  $F_j$  approximates the  $P_j^*$  maximization problem up to error  $\frac{\epsilon}{4k}$ .

Moreover, it holds that  $\text{Hyb}'_j - \text{Hyb}_{j-1} = \text{negl}(\lambda)$ , or this would result in an efficient test distinguishing encryptions of  $q_j$  vs. encryptions of  $q'_j$  (by an analogous reduction as in the proof of [Theorem 3.2](#)).

Thus, by a hybrid argument,<sup>6</sup> we conclude that  $\text{Hyb}_{k-1} > \text{Hyb}_0 - \frac{\epsilon}{4} - \text{negl}(\lambda) = v' - \frac{\epsilon}{4} - \text{negl}(\lambda)$ , contradicting our initial assumption. This completes the proof of the claim.  $\square$

This completes the proof of [Theorem 3.3](#).  $\square$

### 3.3 Achieving $1 - \text{negl}(\lambda)$ Quantum-Classical Gap

[Theorems 3.2](#) and [3.3](#) show how to convert a  $k$ -prover non-local game  $\mathcal{G}$  into a  $2k$ -round interactive argument for quantum advantage. The simplest instantiation of this paradigm is to compile a constant-size non-local game  $\mathcal{G}_0$  (such as CHSH, Magic Square, Odd-Cycle Test, GHZ), which will result in a constant-round protocol with constant gap between the quantum and classical values of the game.

We now discuss various methods of obtaining *optimal* quantum-classical gap: namely, quantum value  $1 - \text{negl}(\lambda)$  and classical value  $\text{negl}(\lambda)$ .

1. **Sequential Repetition.** The simplest method is repeating  $\mathcal{T}(\mathcal{G}_0)$   $\lambda$  times in sequence. If the quantum value of  $\mathcal{T}(\mathcal{G}_0)$  (equivalently, the quantum value of  $\mathcal{G}_0$ ) is  $1 - \text{negl}(\lambda)$ , then this will also hold for a sequential repetition, while if the classical value of  $\mathcal{T}(\mathcal{G}_0)$  is bounded away from 1, then the sequentially repeated game’s classical value will be  $\text{negl}(\lambda)$  by a standard argument.

In the case where the quantum value of  $\mathcal{G}_0$  is less than 1, one can apply **threshold sequential repetition**, in which the verifier for the repeated game accepts if at least a  $\theta$ -fraction of the copies of  $\mathcal{T}(\mathcal{G}_0)$  would be accepted. Choosing  $\theta$  strictly between the classical and quantum values of  $\mathcal{G}$  will “polarize” the classical and quantum values to  $(\text{negl}(\lambda), 1 - \text{negl}(\lambda))$ .

2. **Random-Terminating Parallel Repetition.** Option 1 increases the round complexity of the interactive game to  $O(\lambda)$ , which is somewhat undesirable. Ideally, it would be possible to execute  $\lambda$  *parallel* copies of  $\mathcal{T}(\mathcal{G}_0)$  (preserving the round complexity), but it is known that parallel repetition fails to amplify the (classical) computational values of certain interactive games [\[BIN97\]](#).

However, by appealing to the (slightly more complex operation of) *random-terminating* parallel repetition [\[Hai09\]](#), we can again “polarize” the classical and quantum values of  $\mathcal{T}(\mathcal{G}_0)$  (while preserving the round complexity). Again, we consider a threshold variant of random-terminating parallel repetition in which the verifier accepts based on a threshold  $\theta$  strictly between the quantum and classical values of  $\mathcal{G}_0$ . [\[Hai09\]](#) implies that the classical value of this repeated game will be  $\text{negl}(\lambda)$  (using  $\lambda$  repetitions), while the quantum value of this game will be  $1 - \text{negl}(\lambda)$ .

3. **Plain Parallel Repetition?** Finally, we revisit the question of whether plain (threshold) parallel repetition suffices – this would result in a simpler protocol as compared to Option 2. A lower bound on the quantum value is immediate; what is unclear is a negligible upper

---

<sup>6</sup>As discussed in [\[FM21\]](#) (although context differs slightly here), a hybrid argument can be applied because the collection of indistinguishability claims  $\text{Hyb}'_j \approx \text{Hyb}_{j-1}$  are proved via a universal reduction  $R$  from the security of QHE.

bound on the classical value. However, we observe that (threshold) parallel repetitions of  $\mathcal{T}(\mathcal{G}_0)$  can be analyzed by appealing to [Theorems 3.2](#) and [3.3](#) with respect to (threshold) parallel repetitions of the *non-local game*  $\mathcal{G}_0$ .

In particular, under a sufficiently strong assumption on the QHE, the classical value of this repeated game is at most the classical value of the (threshold) parallel repetition of  $\mathcal{G}_0$  as a non-local game! Thus, if sufficient (threshold) parallel repetition of  $\mathcal{G}_0$  results in a game with negligible classical value, then under a sub-exponential hardness assumption, the parallel repeated interactive game has  $\text{negl}(\lambda)$  classical value.

Since, for example, such parallel repetition theorems are known for 2-player games [[Raz95](#), [Rao08](#)], this resolves the question for 2-player games under a subexponential assumption. We elaborate below on how to avoid this subexponential loss for games  $\mathcal{G}_0$  that exhibit *exponential* hardness amplification under parallel repetition.

### 3.3.1 Avoiding Sub-exponential Loss for Games with Strong Parallel Repetition

We briefly recall the formal definition of (*threshold*) *parallel repeated games*  $\mathcal{G} = \mathcal{G}_0^{t,\theta}$ :

**Definition 3.5** (Threshold Parallel Repetition). *Let  $\mathcal{G}_0$  denote a constant-size  $k$ -prover non-local game. We define the  $t$ -fold  $\theta$ -threshold repeated game  $\mathcal{G} = \mathcal{G}_0^{t,\theta}$  as follows:*

- *Queries are of the form  $(\tilde{q}_1 = (q_{1,1}, \dots, q_{1,t}), \dots, \tilde{q}_k = (q_{k,1}, \dots, q_{k,t}))$ . The  $k$ -tuples  $(q_{1,i}, \dots, q_{k,i})$  are sampled i.i.d.*
- *Answers have the form  $(\tilde{a}_1 = (a_{1,1}, \dots, a_{1,t}), \dots, \tilde{a}_k = (a_{k,1}, \dots, a_{k,t}))$ .*
- *The repeated verifier  $\mathcal{V}$  accepts if at least  $\theta$  fraction of the transcripts  $(q_{1,i}, a_{1,i}, \dots, q_{k,i}, a_{k,i})$  are accepted by the  $\mathcal{G}_0$ -verifier  $\mathcal{V}_0$ .*

**Remark 3.6.** *We remark that if the quantum value of  $\mathcal{G}_0$  is at least  $v^* > \theta + \epsilon$ , then by a Chernoff bound (see [Footnote 5](#)), the quantum value of the repeated game  $\mathcal{G}$  is at least*

$$1 - 2^{2t\epsilon^2}.$$

*On the other hand, if  $k = 2$  and the classical value of  $\mathcal{G}_0$  is at most  $v < \theta - \epsilon$ , then by [[Rao08](#)], the classical value of  $\mathcal{G}$  is at most*

$$2^{-\gamma\epsilon^3 t / |a_1|},$$

*where  $\gamma$  is a constant that can depend on  $\theta$ . Thus, for large enough  $t = O(\lambda)$ ,  $\mathcal{G}$  will have quantum value  $\geq 1 - 2^{-\lambda}$  and classical value  $\leq 2^{-\lambda}$  in the case  $k = 2$ .*

*For general  $k$ , the status of parallel repeated games is considerably less well understood [[Ver96](#), [DHVY16](#), [HR20](#)].*

In this section, we prove a strengthening of [Theorem 3.2](#) for all games  $\mathcal{G}_0$  with strong enough parallel repetition properties:

**Theorem 3.7.** *Let  $\mathcal{G}_0$  be a  $k$ -player non-local game of constant size; i.e., the lengths of queries and answers in  $\mathcal{G}_0$  is  $O(1)$ . Let  $v$  be the classical value of  $\mathcal{G}_0$  and let  $v^*$  be its quantum value, and suppose that  $v^* > v$ . Let  $\mathcal{G}_0^{t,\theta}$  denote the threshold parallel repetition of  $\mathcal{G}_0$  (Definition 3.5), and let  $\theta = \frac{v+v^*}{2}$ .*

Assume that for all  $t$ ,  $v(\mathcal{G}_0^{t,\theta}) = 2^{-ct}$  for a fixed constant  $c$  (that can depend on  $\mathcal{G}_0$ ).

Then, given a polynomially secure QHE scheme QHE, the game  $\mathcal{G} = \mathcal{G}_0^{\lambda,\theta}$  can be converted into a single-prover game that has (computational) quantum value  $1 - \text{negl}(\lambda)$  and (computational) classical value  $\text{negl}(\lambda)$ .

For example, using Theorem 3.7, we can compile arbitrary parallel repeated (constant-size) 2-prover games under polynomial hardness assumptions.

*Proof.* We consider the single prover protocol defined in Theorem 3.3 for the game  $\mathcal{G}$  with the following modification:<sup>7</sup> we sample a fresh QHE key for each parallel instance of  $\mathcal{G}_0$  (this simply corresponds to a parallel repetition of the  $\mathcal{G}_0$ -compiled protocol as defined in Theorem 3.3).

The fact that the quantum value is  $1 - \text{negl}(\lambda)$  follows from the fact that the quantum value of  $\mathcal{G}$  is  $1 - \text{negl}(\lambda)$  (as observed in Remark 3.6) together with Theorem 3.3. We focus on bounding the classical value.

For the sake of contradiction, assume that  $\tilde{\mathcal{P}}$  is a poly-size prover for the compiled game attaining value at least  $\delta$  for non-negligible  $\delta = \delta(\lambda)$ . This means that with non-negligible probability,  $\tilde{\mathcal{P}}$  (implicitly) produces transcripts  $(\tilde{q}_1, \tilde{a}_1, \dots, \tilde{q}_k, \tilde{a}_k)$  such that at least  $\theta$  fraction of the  $\mathcal{G}_0$ -transcripts  $(q_{1,i}, a_{1,i}, \dots, q_{k,i}, a_{k,i})$  are accepting.

Let  $t' = O(\log \lambda)$  be defined so that the classical value of  $\mathcal{G}_0^{t',\theta} < \delta/2$  (such  $t' = O(\log \lambda)$  exists by our assumption on  $\mathcal{G}_0$ ). By an averaging argument, we know that there exists a subset  $S \subset [t']$  with  $|S| = t'$  and queries  $\{q_{1,i}, \dots, q_{k,i}\}_{i \notin S}$  such that  $\tilde{\mathcal{P}}$  wins with probability at least  $\delta$  conditioned on  $\{q_{j,i}\}_{i \notin S}$ .

This allows us to construct a prover  $\tilde{\mathcal{P}}'$  for the QHE-compiled variant of  $\mathcal{G}^{t',\theta}$  that wins with probability  $\delta$ :  $\tilde{\mathcal{P}}'$  simply has  $S, \{q_{j,i}\}_{i \notin S}$  hardcoded and emulates  $\tilde{\mathcal{P}}$  by sampling the  $([t'] - S)$ -slot messages itself (this requires either QHE to be public-key or for the protocol to use independent secret keys for the different slots).

Finally, we see that since the classical value of  $\mathcal{G}^{t',\theta}$  is at most  $\delta/2$  but  $\tilde{\mathcal{P}}'$  wins the interactive game with probability at least  $\delta$ ,  $\tilde{\mathcal{P}}'$  contradicts Theorem 3.3 assuming the polynomial security of QHE.  $\square$

### 3.4 Non-Interactive Protocols in the ROM

We briefly remark on the ability to convert protocols arising from Theorem 3.2 into non-interactive protocols for verifying quantum advantage.

Specifically, suppose that  $\mathcal{G}$  is a 2-player game such that the distribution of  $q_2$  is *uniform and independent of  $q_1$* . Then, we know (by Theorems 3.2 and 3.7) that given an QHE scheme, we can convert  $\mathcal{G}$  into a 4-message interactive game with quantum value  $1 - \text{negl}(\lambda)$  and classical value  $\text{negl}(\lambda)$ . Moreover, the 3rd message of this interactive game is *public-coin*. Therefore, we can apply

<sup>7</sup>If the QHE scheme is *public-key*, this modification is unnecessary.

the Fiat-Shamir heuristic [FS87] to this interactive game to obtain a *non-interactive game* with quantum value  $1 - \text{negl}(\lambda)$  and classical value  $\text{negl}(\lambda)$  in the random oracle model.<sup>8</sup> This template can be instantiated using (for example) the CHSH game or the magic square game to obtain new 2-message quantum advantage protocols in the random oracle model.

We observe that even the “honest” quantum prover (attaining  $1 - \text{negl}(\lambda)$  value) in this non-interactive game only requires classical access to the random oracle (to hash its own classical message).

## 4 Protocols for Verifying Quantum Advantage

In this section, we give a concrete instantiation of our framework and outline directions for future work, focusing on obtaining protocols with  $\text{simple}(r)$  quantum provers. We proceed to describe a concrete instantiation of our blueprint, using the CHSH game and the Mahadev QFHE scheme.

### 4.1 Compiling the CHSH Game

We first recall the CHSH game.

**Definition 4.1** (The CHSH Game). *The CHSH game  $\mathcal{G}_{\text{CHSH}}$  consists of the uniform query distribution  $\mathcal{Q}_{\text{CHSH}}$  over  $(q_1, q_2) \in (\{0, 1\})^2$  and verification predicate  $\mathcal{V}_{\text{CHSH}}(q_1, q_2, a_1, a_2)$  which is 1 if and only if*

$$a_1 \oplus a_2 = q_1 q_2 \pmod{2}.$$

*The classical value of this game is  $v_{\text{CHSH}} = 0.75$  and the quantum value is  $v_{\text{CHSH}}^* = \cos^2(\pi/8) \approx 0.85$ . The optimal quantum strategy is as follows: The two players  $\mathcal{A}, \mathcal{B}$  share an EPR pair*

$$\frac{1}{\sqrt{2}}(|0\rangle_{\mathcal{A}}|0\rangle_{\mathcal{B}} + |1\rangle_{\mathcal{A}}|1\rangle_{\mathcal{B}})$$

*where players  $\mathcal{A}, \mathcal{B}$  have the  $\mathcal{A}, \mathcal{B}$  registers respectively. Upon receiving  $q_1$ , player  $\mathcal{A}$  measures her register  $\mathcal{A}$  in the Hadamard ( $\pi/4$ ) basis if  $q_1 = 0$ , and in the standard basis if  $q_1 = 1$ , and reports the outcome  $a_1 \in \{0, 1\}$ . Player  $\mathcal{B}$  measures her register  $\mathcal{B}$  in the  $\pi/8$ -basis if  $q_2 = 0$  and in the  $3\pi/8$ -basis if  $q_2 = 1$ , and reports the outcome  $a_2 \in \{0, 1\}$ .*

By Theorem 3.2 and Claim 2.6, we have the following corollary:

**Corollary 4.2.** *Consider the [Mah18a] QHE scheme for poly-size circuits in the Toffoli and Clifford basis. Consider the CHSH game  $\mathcal{G}_{\text{CHSH}} = (\mathcal{Q}_{\text{CHSH}}, \mathcal{V}_{\text{CHSH}})$  and quantum strategy in Definition 4.1 where*

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_{\mathcal{A}}|0\rangle_{\mathcal{B}} + |1\rangle_{\mathcal{A}}|1\rangle_{\mathcal{B}})$$

*and  $C_1^*(|\Psi\rangle_{\mathcal{A}}, \cdot) \in \mathcal{C}$ .*

*The single-player computationally sound interactive game  $\mathcal{T}^{\mathcal{G}}$  has:*

---

<sup>8</sup>Our interactive game is privately verifiable, but the classical soundness reduction for Fiat-Shamir extends immediately to this case.

- *quantum CS value* =  $v_{\text{CHSH}}^* \geq 0.85$
- *classical CS value*  $\geq v_{\text{CHSH}} + \text{negl}(\lambda) = 0.75 + \text{negl}(\lambda)$ .

Amplifying this gap can be done by sequential repetition. Alternatively, one can compile a parallel-repeated version of the CHSH game to get a protocol with a large gap between the quantum and classical CS values.

**Prover Efficiency.** The compilation of the CHSH game with Mahadev’s QFHE (Corollary 4.2) results in a conceptually simple 4-round protocol with a relatively simple quantum prover. Here, we analyze the quantum prover’s algorithm.

Returning to the quantum strategy for the CHSH game (Definition 4.1), player  $\mathcal{A}$  applies a controlled Hadamard gate to  $|\bar{q}_1\rangle|\Psi\rangle_{\mathcal{A}}$ , and then measures the  $\mathcal{A}$  register. This can be implemented by a circuit  $C_1^*$  containing Clifford gates and a single Toffoli gate [BFGH08]. Recall that in Mahadev’s scheme ([Mah18a]), evaluating Clifford gates only requires applying the intended Clifford gate to a (Pauli one-time-padded) encryption of the underlying state. To evaluate a Toffoli gate, the Toffoli gate is applied to the encrypted qubit, followed by 3 “encrypted CNOT” operations and 2 Hadamard gates. The bulk of the computational cost of the prover is in the encrypted CNOT operations. Using the trapdoor claw-free functions (TCF) for the classical ciphertexts in the Mahadev QHE scheme, an encrypted CNOT operation consists of creating a uniform superposition over the TCF domain, evaluating the function in superposition, measurements and Clifford gates. That is, this requires  $p(\lambda) := \log|\mathcal{D}| + \log|\mathcal{R}|$  ancilla qubits corresponding to the TCF domain  $\mathcal{D}$  and range  $\mathcal{R}$ . Concretely, the prover’s quantum operations in our compilation of CHSH are as follows.

- The prover creates an EPR pair which involves applying a Hadamard and a CNOT gate.
- The prover receives a classical ciphertext  $\hat{q}_1$  from the verifier in round 1.
- The prover homomorphically evaluates  $C_1^*(|\Psi\rangle_{\mathcal{A}}, q_1)$  (where  $C_1^*$  implements player  $\mathcal{A}$ ’s strategy in CHSH). This uses the constant number of qubits in  $C_1^*$  and  $3p(\lambda)$  ancilla qubits for TCF evaluations. All of its operations are Clifford gates except a single Toffoli gate and 3 invocations of the TCF evaluation algorithm run in superposition.
- The prover sends back a classical ciphertext, and receives a bit  $q_2 \in \{0, 1\}$  in round 2. It measures  $|\Psi\rangle_{\mathcal{B}}$  in the  $\pi/8$  or  $3\pi/8$  basis (depending on  $q_2$ , as per player  $\mathcal{B}$ ’s strategy in CHSH). It sends back the result as  $a_2 \in \{0, 1\}$ . In particular, the prover can discard the remaining qubits right after it computes and sends its message in round 1.

Overall, the prover uses  $3p(\lambda) + O(1)$  qubits, and the complexity of its operations is dominated by the 3 TCF evaluations. We note that designing a more efficient QHE scheme supporting the controlled-Hadamard gate, potentially based on simpler assumptions, is an attractive route to improving the prover efficiency.

## 4.2 Future Directions

Our work suggests several directions for future work.

**Protocols for quantum advantage with very simple quantum provers?** First, could there be a non-local game where, once an appropriate bipartite state is set up, one of the prover strategies can be implemented using only Clifford gates? If this were possible, the complexity of the QHE evaluation reduces drastically. In particular, quantum homomorphic encryption schemes handling only Clifford gates are much simpler than general quantum homomorphic encryption schemes (e.g. they only require applying the intended Clifford gates and additional classical computations). This could lead us to a truly efficient protocol that does not require maintaining superpositions with security parameter number of qubits. On the other hand, if it were to be the case that any non-local game requires *both* provers to be non-Clifford, that would be an interesting outcome as well. To the best of our knowledge, this statement is not known, and does not seem to follow from Gottesman-Knill-type classical simulation of Clifford circuits.

**Simpler homomorphic encryption schemes.** Can we design better somewhat homomorphic encryption schemes for the compilation in Theorem 3.2? We note that the scheme only needs to support the evaluation of *one* of the two provers' strategies in the non-local game, which can often be implemented by a simple circuit. For example, designing a scheme that simply supports the controlled-Hadamard gate would suffice for compiling the CHSH game. This may give a quantum advantage protocol with a simpler prover strategy.

**Protocols based on different assumptions.** We note that for quantum advantage, we only need soundness against classical adversaries. Namely, the homomorphic encryption should have completeness for some quantum gates, but *only needs soundness against classical polynomial-time adversaries*. This opens the door to designing QHEs based on e.g. discrete log style assumptions, or the hardness of factoring, rather than learning with errors. (Indeed, quantum advantage protocols have been constructed using factoring-based TCFs; see [KCVY21]).

**Understanding existing protocols.** The existing interactive protocols for quantum advantage (to our knowledge, [BCM<sup>+</sup>18, BKVV20, KCVY21]) are presented as an all-in-one package. Intuitively, a protocol testing quantumness should have a component testing for quantum resources, e.g. a test of entanglement, and a component that tests computational power, i.e. the cryptography. Can existing protocols be disentangled to two such components? A starting point is [KCVY21] which has some resemblance to our CHSH compilation in Section 4.1, although it does incur two more rounds.

More ambitiously, could we understand any single-prover quantum advantage protocol as compiling a (perhaps contrived)  $k$ -player non-local game via a somewhat homomorphic encryption scheme? We leave it open to understand the reach and the universality of our framework for constructing protocols for quantum advantage.

## Acknowledgements

We thank Fermi Ma for comments regarding QHE correctness on an earlier draft of this paper.

## References

- [AA10] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics, 2010. [3](#)
- [AAB<sup>+</sup>19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando Brandao, David Buell, Brian Burkett, Yu Chen, Jimmy Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Michael Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew Harrigan, Michael Hartmann, Alan Ho, Markus Rudolf Hoffmann, Trent Huang, Travis Humble, Sergei Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, Dave Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod Ryan McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin Jeffery Sung, Matt Trevithick, Amit Vainsencher, Benjamin Villalonga, Ted White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574:505–510, 2019. [3](#)
- [ABOE08] Dorit Aharonov, Michael Ben-Or, and Elad Eban. Interactive proofs for quantum computations, 2008. [3](#)
- [Ara01] P.K. Aravind. Bell’s theorem without inequalities and only two distant observers. In *Optical Fiber Communication Conference and International Conference on Quantum Information*, page EAPA7. Optica Publishing Group, 2001. [4](#)
- [BC12] Nir Bitansky and Alessandro Chiesa. Succinct arguments from multi-prover interactive proofs and their efficiency benefits. pages 255–272, 2012. [5](#), [6](#)
- [BCM<sup>+</sup>18] Zvika Brakerski, Paul F. Christiano, Urmila Mahadev, Umesh V. Vazirani, and Thomas Vidick. A cryptographic test of quantumness and certifiable randomness from a single quantum device. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 320–331. IEEE Computer Society, 2018. [3](#), [5](#), [25](#)
- [Bel64] J. S. Bell. On the einstein podolsky rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964. [3](#)
- [BFGH08] Debajyoti Bera, Stephen Fenner, Frederic Green, and Steve Homer. Universal quantum circuits, 2008. [24](#)

- [BFNV18] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. On the complexity and verification of quantum random circuit sampling. *Nature Physics*, 15(2):159–163, Oct 2018. 3
- [BIN97] Mihir Bellare, Russell Impagliazzo, and Moni Naor. Does parallel repetition lower the error in computationally sound protocols? pages 374–383, 1997. 20
- [BJS10] Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, Aug 2010. 3
- [BKVV20] Zvika Brakerski, Venkata Koppula, Umesh V. Vazirani, and Thomas Vidick. Simpler proofs of quantumness. In Steven T. Flammia, editor, *15th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2020, June 9-12, 2020, Riga, Latvia*, volume 158 of *LIPICs*, pages 8:1–8:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. 3, 5, 25
- [BMW98] Ingrid Biehl, Bernd Meyer, and Susanne Wetzal. Ensuring the integrity of agent - based computations by short proofs. In *Mobile Agents - Second International Workshop, MA '98, Proceedings*, Lecture Notes in Computer Science, pages 183–194. Springer, 1998. 2nd International Workshop on Mobile Agents, MA'98 ; Conference date: 09-09-1998 Through 11-09-1998. 5
- [Bra18] Zvika Brakerski. Quantum FHE (almost) as secure as classical. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology - CRYPTO 2018 - 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part III*, volume 10993 of *Lecture Notes in Computer Science*, pages 67–95. Springer, 2018. 4, 6, 11, 12, 13, 29, 30, 31
- [CHSH69] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. 3, 4, 7
- [DHVY16] Irit Dinur, Prahladh Harsha, Rakesh Venkat, and Henry Yuen. Multiplayer parallel repetition for expander games. *arXiv preprint arXiv:1610.08349*, 2016. 21
- [FK17] Joseph F. Fitzsimons and Elham Kashefi. Unconditionally verifiable blind quantum computation. *Physical Review A*, 96(1), Jul 2017. 3
- [FM21] Marc Fischlin and Arno Mittelbach. An overview of the hybrid argument. Cryptology ePrint Archive, Report 2021/088, 2021. <https://eprint.iacr.org/2021/088>. 20
- [FS87] Amos Fiat and Adi Shamir. How to prove yourself: Practical solutions to identification and signature problems. pages 186–194, 1987. 23
- [GE21] Craig Gidney and Martin Ekerå. How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits. *Quantum*, 5:433, Apr 2021. 3

- [GHZ07] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. Going beyond bell’s theorem, 2007. 4
- [Hai09] Iftach Haitner. A parallel repetition theorem for any interactive argument. pages 241–250, 2009. 20
- [Har93] Lucien Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Phys. Rev. Lett.*, 71:1665–1668, Sep 1993. 4
- [HBD<sup>+</sup>15] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, and R. Hanson. Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres. *Nature*, 526(7575):682–686, 2015. 4
- [HR20] Justin Holmgren and Ran Raz. A parallel repetition theorem for the ghz game. *arXiv preprint arXiv:2008.05059*, 2020. 21
- [JNV<sup>+</sup>20] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Mip<sup>\*</sup>=re, 2020. 4
- [KCVY21] Gregory D. Kahanamoku-Meyer, Soonwon Choi, Umesh V. Vazirani, and Norman Y. Yao. Classically-verifiable quantum advantage from a computational bell test. *CoRR*, abs/2104.00687, 2021. 3, 4, 5, 25
- [KRR13] Yael Tauman Kalai, Ran Raz, and Ron D. Rothblum. Delegation for bounded space. In Dan Boneh, Tim Roughgarden, and Joan Feigenbaum, editors, *Symposium on Theory of Computing Conference, STOC’13, Palo Alto, CA, USA, June 1-4, 2013*, pages 565–574. ACM, 2013. 5, 6
- [KRR14] Yael Tauman Kalai, Ran Raz, and Ron D. Rothblum. How to delegate computations: the power of no-signaling proofs. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 485–494. ACM, 2014. 5, 6
- [LJMP90] Arjen K. Lenstra, Hendrik W. Lenstra Jr., Mark S. Manasse, and John M. Pollard. The number field sieve. In Harriet Ortiz, editor, *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*, pages 564–572. ACM, 1990. 3
- [LLL<sup>+</sup>21] Yong (Alexander) Liu, Xin (Lucy) Liu, Fang (Nancy) Li, Haohuan Fu, Yuling Yang, Jiawei Song, Pengpeng Zhao, Zhen Wang, Dajia Peng, Huarong Chen, Chu Guo, Heliang Huang, Wenzhao Wu, and Dexun Chen. Closing the “quantum supremacy” gap. *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov 2021. 3

- [Mah18a] Urmila Mahadev. Classical homomorphic encryption for quantum circuits. In Mikkel Thorup, editor, *59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018*, pages 332–338. IEEE Computer Society, 2018. [2](#), [4](#), [5](#), [6](#), [11](#), [12](#), [13](#), [23](#), [24](#), [29](#), [31](#), [32](#), [33](#)
- [Mah18b] Urmila Mahadev. Classical verification of quantum computations. pages 259–267, 2018. [3](#)
- [Mer90] N. David Mermin. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.*, 65:3373–3376, Dec 1990. [4](#)
- [MV21] Tony Metger and Thomas Vidick. Self-testing of a single quantum device under computational assumptions. *Quantum*, 5:544, Sep 2021. [3](#)
- [NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. [10](#)
- [Rao08] Anup Rao. Parallel repetition in projection games and a concentration bound. pages 1–10, 2008. [21](#)
- [Raz95] Ran Raz. A parallel repetition theorem. pages 447–456, 1995. [21](#)
- [RUV13] Ben W. Reichardt, Falk Unger, and Umesh V. Vazirani. Classical command of quantum systems. *Nat.*, 496(7446):456–460, 2013. [4](#)
- [Sho97] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, 1997. [3](#)
- [Ver96] Oleg Verbitsky. Towards the parallel repetition conjecture. *Theoretical Computer Science*, 157(2):277–282, 1996. [21](#)
- [ZKML<sup>+</sup>21] Daiwei Zhu, Gregory D. Kahanamoku-Meyer, Laura Lewis, Crystal Noel, Or Katz, Bahaa Harraz, Qingfeng Wang, Andrew Risinger, Lei Feng, Debopriyo Biswas, Laird Egan, Alexandru Gheorghiu, Yunseong Nam, Thomas Vidick, Umesh Vazirani, Norman Y. Yao, Marko Cetina, and Christopher Monroe. Interactive protocols for classically-verifiable quantum advantage, 2021. [3](#)

## A QHE Correctness with Respect to Auxiliary Input

In this section, we discuss the relationship between aux-input QHE correctness ([Definition 2.3](#)) and prior definitions/schemes [[Mah18a](#), [Bra18](#)]. In particular, we sketch a proof of [Claim 2.6](#).

We first recall the formal QFHE definitions from [[Bra18](#)]. Specifically, we now generalize QFHE to allow for encryption/decryption of *states* and homomorphic evaluation of circuit with quantum (rather than classical) output. However, we impose (as done in [[Bra18](#)]) the constraint that encryption and decryption are (qu)bit-by-(qu)bit.

**Definition A.1** (Quantum qubit-by-qubit Homomorphic Encryption (bQHE)). *A quantum (qu)bit-by-(qu)bit homomorphic encryption scheme  $\text{bQHE} = (\text{Gen}, \text{Enc}, \text{Eval}, \text{Dec})$  for a class of quantum circuits  $\mathcal{C}$  is a tuple of algorithms with the following syntax:*

- *Gen is a PPT algorithm that takes as input the security parameter  $1^\lambda$  and outputs a (classical) secret key  $\text{sk}$  of  $\text{poly}(\lambda)$  bits;*
- *Enc is a QPT algorithm that takes as input a secret key  $\text{sk}$  and a qubit  $b$ , and outputs a ciphertext  $\text{ct}$ . Additionally, if  $b$  is a classical bit, the encryption algorithm is PPT and the ciphertext  $\text{ct}$  is classical;*
- *Eval is a QPT algorithm that takes as input a tuple  $(C, |\Psi\rangle, \vec{\text{ct}}_{\text{in}})$ , where  $C : \mathcal{H} \times (\mathbb{C}^2)^{\otimes n} \rightarrow (\mathbb{C}^2)^{\otimes m}$  is a quantum circuit,  $|\Psi\rangle \in \mathcal{H}$  is an auxiliary quantum state, and  $\vec{\text{ct}}_{\text{in}} = (\text{ct}_1, \dots, \text{ct}_n)$  is a tuple of  $n$  ciphertexts. Eval computes a quantum circuit  $\text{Eval}_C(|\Psi\rangle \otimes |0\rangle^{\text{poly}(\lambda, n)}, \vec{\text{ct}}_{\text{in}})$  which outputs a tuple of ciphertexts  $\vec{\text{ct}}_{\text{out}}$ .*
- *Dec is a PT algorithm that takes as input a secret key  $\text{sk}$  and ciphertext  $\text{ct}$ , and outputs a qubit  $b$ . Additionally, if  $\text{ct}$  is a classical ciphertext, the decryption algorithm outputs a bit  $b$ .*

*In addition to security (which again we can restrict to hold against classical adversaries), we require (following [Bra18]) the following correctness property:*

- *For any quantum circuit  $C$  and any  $n$ -qubit state  $|\Phi\rangle = \sum \alpha_{i_1, \dots, i_n} |x_1 \dots x_n\rangle$ , the following two states have negligible trace distance. The first state  $\rho_1$  is defined to be the output of  $C(|\Phi\rangle)$ . The second state  $\rho_2$  is defined by sampling  $\text{sk} \leftarrow \text{Gen}(1^\lambda)$ ,  $\text{ct}_i \leftarrow \text{Enc}(\text{sk}, x_i)$  and computing  $\text{Dec}(\text{sk}, \text{Eval}(C, \text{ct}_1, \dots, \text{ct}_n))$ .*

The QFHE scheme of [Bra18] is shown (in [Bra18]) to satisfy the definition above. We now discuss mild structural hypotheses under which an encryption scheme satisfying [Definition A.1](#) also satisfies [Definition 2.3](#).

- **Allowing un-encrypted inputs:** this first property is merely syntax that can be added without loss of generality. [Definition A.1](#) does not have *any* un-encrypted state (i.e.  $|\Psi\rangle$  in [Definition 2.3](#)). This can be easily rectified by encrypting any additional state  $|\Psi\rangle$  during QFHE evaluation.
- **Locality:** this is the most important structural property. We say that a QHE scheme is *local* if homomorphic evaluation of a circuit of the form  $C_{\mathcal{A}} \otimes \text{Id}_{\mathcal{B}}$ , acting on  $\mathcal{H} \simeq \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ , is identical to homomorphic evaluation of  $C_{\mathcal{A}}$  on (encrypted) register  $\mathcal{A}$  (and acts as identity on the encrypted  $\mathcal{B}$ ).
- **Honest Decryption Correctness:** we say that a QHE scheme satisfies honest decryption correctness if  $\text{Dec}(\text{sk}, \text{Enc}(\text{sk}, b))$  is the identity map on  $b$ . In particular, this implies that the equation  $\text{Dec}(\text{sk}, \text{Enc}(\text{sk}, b))$  holds in the presence of any auxiliary input.

**Claim A.2.** *Let bQHE be a (qu)bit-by-(qu)bit QHE scheme satisfying [Definition A.1](#). Moreover, extend bQHE to allow un-encrypted inputs, and assume that bQHE satisfies [locality](#) and [honest decryption correctness](#). Then, bQHE satisfies [Definition 2.3](#).*

*Proof.* We want to show that bQHE (as described in [Claim A.2](#)) satisfies correctness with respect to auxiliary input. Let  $|\Psi\rangle_{AB}$  be a bipartite state and  $x$  be a (classical) input as in [Definition 2.3](#), and let  $C$  denote a circuit taking as input  $x$  and a state on register  $\mathcal{A}$ . By the *basic* correctness property of bQHE with respect to the circuit  $C \otimes \text{Id}_{\mathcal{B}}$ , we know that the state

$$\text{Dec}(\text{sk}, \text{Eval}(C \otimes \text{Id}_{\mathcal{B}}, \text{Enc}(\text{sk}, x), \text{Enc}(\text{sk}, |\Psi\rangle_{AB})))$$

is negligibly close in trace distance from the state obtained by evaluating  $C \otimes \text{Id}_{\mathcal{B}}$  on  $x, |\Psi\rangle_{AB}$ .

Moreover, by the *locality* of bQHE, we know that  $\text{Eval}(C \otimes \text{Id}_{\mathcal{B}}, \text{Enc}(\text{sk}, x), \text{Enc}(\text{sk}, |\Psi\rangle_{AB}))$  is equivalent to applying  $\text{Eval}(C, \cdot)$  to  $\text{Enc}(\text{sk}, x)$  and the encrypted  $\mathcal{A}$  register of  $\text{Enc}(|\Psi\rangle_{AB})$ . Combining this with the *honest decryption correctness* property on register  $\mathcal{B}$ , we conclude that the state obtained by evaluating  $C \otimes \text{Id}_{\mathcal{B}}$  on  $x, |\Psi\rangle_{AB}$  is negligibly close in trace distance to the state

$$(\text{Dec}(\text{sk}, \text{Eval}(C, \text{Enc}(\text{sk}, x), \text{Enc}(\text{sk}, [|\Psi\rangle]_{\mathcal{A}})), [|\Psi\rangle]_{\mathcal{B}}))$$

where  $[|\Psi\rangle]_{\mathcal{A}}$  and  $[|\Psi\rangle]_{\mathcal{B}}$  denote the  $\mathcal{A}$  and  $\mathcal{B}$  registers of  $|\Psi\rangle$ , respectively. Finally, the syntax for unencrypted inputs tells us that the above indistinguishability implies the *aux-input correctness* property of [Definition 2.3](#).  $\square$

Thus, to prove [Claim 2.6](#), it suffices to observe that the [\[Mah18a, Bra18\]](#) schemes satisfy *locality* and *honest decryption correctness*. This can be checked by inspecting [\[Mah18a, Bra18\]](#); for intuition, we note that any scheme whose homomorphic evaluation is executed in a “gate-by-gate fashion” should be local.

## A.1 The [\[Mah18a\]](#) QFHE Scheme

For the rest of this section, we give a more explicit discussion of the [\[Mah18a\]](#) scheme to sketch why it satisfies our hypotheses. The encryption of a qubit  $|\psi\rangle$  in the scheme consists of a quantum one-time padded qubit  $X^x Z^z |\psi\rangle$  together with two classical ciphertexts per qubit that encrypt  $x$  and  $z$ . To homomorphically evaluate a gate on some qubits, it applies that gate (plus additional operations if necessary) to those qubits in the padded state, and updates the classical ciphertexts for those qubits. The high-order bit is that when evaluating a circuit  $C$ , the qubits which are not acted upon will not be changed, which in turn implies the *locality* property.

We recall Mahadev’s scheme (Scheme 6.1) in more detail (extended slightly to formally allow for encryption of quantum states) and show that it is local. For this discussion, we assume that in a quantum circuit  $C$ , all measurement gates are deferred to the end.

Recall that  $\text{Gen}$  is a key generation algorithm corresponding to a classical homomorphic encryption scheme (that satisfies some desired properties that we elaborate on below). The encryption of a (classical) message  $M \in \{0, 1\}^n$  is the one-time pad applied to  $M$  together with the classical encryption of the pad. Using quantum notation,

$$\text{Enc}(M) = (X^x |M\rangle, \{\hat{x}_i\}_{i \in [n]})$$

where  $x = (x_1, \dots, x_n)$  is randomly chosen in  $\{0, 1\}^n$ , and  $\{\hat{x}_i\}_{i \in [n]}$  are classical ciphertexts where  $\hat{x}_i$  encrypts the pad  $x_i$  on the  $i$ 'th qubit (using the underlying classical homomorphic encryption).

The encryption algorithm,  $\text{Enc}$ , can be extended to encrypt a quantum state  $\Psi$  consisting of  $\ell$  qubits, in which case, the encryption of  $\Psi$  is the quantum one-time pad applied to  $\Psi$ , together with the classical encryptions of the pad. Namely,

$$\text{Enc}(|\Psi\rangle) = (Z^z X^x |\Psi\rangle, \{\hat{x}_i, \hat{z}_i\}_{i \in [\ell]}) \quad (6)$$

where  $z = (z_1, \dots, z_\ell)$  and  $x = (x_1, \dots, x_\ell)$  are randomly chosen in  $\{0, 1\}^\ell$ , and  $\{\hat{z}_i, \hat{x}_i\}_{i \in [\ell]}$  are classical ciphertexts where  $\hat{z}_i, \hat{x}_i$  encrypt the Pauli pad on the  $i$ 'th qubit (using the underlying classical homomorphic encryption).

We next focus on the homomorphic operations. Suppose we wish to homomorphically compute a circuit  $C : \mathcal{H} \times \{0, 1\}^n \rightarrow \mathcal{H}' \times \{0, 1\}^m$  applied to the quantum state  $|\Psi\rangle = |\Psi_0\rangle |M\rangle$ .  $\text{Eval}_C$  homomorphically evaluates each gate in  $C$ , as described in [Mah18a] (Scheme 6.1), and then measures qubits corresponding to the qubits that  $C$  measures at the end. Recall that the input to  $\text{Eval}_C$  is  $(|\Psi_0\rangle | \text{Enc}(M)\rangle, |0^t\rangle)$ .

In what follows, we assume for simplicity that the input to  $\text{Eval}_C$  is  $(\text{Enc} |\Psi\rangle, |0^t\rangle)$ . This is without loss of generality since we can think of  $\Psi_0$  as being encrypted with the trivial pad  $Z^0 X^0$ . Namely,  $\text{Eval}_C$  starts with the quantum state  $Z^z X^x |\Psi\rangle |0^t\rangle$  and  $\text{ct} = \{\hat{x}_i, \hat{z}_i\}_{i \in [\ell]}$  (as defined in Equation (6)). It proceeds gate by gate, while satisfying the following desired locality property: For each gate in  $C$ , suppose the gate implements the unitary  $U \otimes I$ , where  $U$  acts only on a constant number of registers, denoted by  $\mathcal{J}$ , and  $I$  is the identity acting on all the other registers, then  $\text{Eval}_C$  homomorphically evaluates the gate by applying a unitary  $U' \otimes I$  to its padded state, where  $U'$  acts on the same set of registers  $\mathcal{J}$  and on the ancilla registers (initialized to be  $|0^t\rangle$ ), and in addition it updates the classical ciphertexts,  $\text{ct}$ , to encrypt the "correct pads". Note that this locality property implies that scheme is indeed local, as desired.

In what follows we argue that this locality property indeed holds. It suffices to prove that it holds for Clifford gates and for Toffoli gates (since they form a universal gate set).

In the following paragraph we use  $C$  to denote a Clifford gate (instead of a quantum circuit). A Clifford gate  $C$  has the desired property that it preserves the Pauli group under conjugation, i.e., for every  $P_1, P_2$  in the Pauli group there exist  $P_3, P_4$  in the Pauli group such that

$$C(P_1 \otimes P_2)C^\dagger = P_3 \otimes P_4.$$

This implies that for every  $x, z$  there exists  $x', z'$  such that

$$CZ^z X^x |\Psi\rangle = Z^{z'} X^{x'} C |\Psi\rangle.$$

Thus,  $\text{Eval}_C$  homomorphically computes a Clifford gate, by simply applying the exact same Clifford gate to its padded quantum state, and then obtains a quantum one-time pad of  $C |\Psi\rangle$  and homomorphically updates the classical ciphertexts to encrypt  $x', z'$ . Thus, the locality property is indeed satisfied.

The Toffoli gate, denoted by  $T$ , is more complicated, and this is where the additional ancilla registers are used. It cannot be applied to the padded state while only updating the encrypted pads, as was done for Clifford gates, since (as opposed to Clifford gates) it does not preserve Pauli

operators by conjugation. Applying a Toffoli directly to a 3-qubit one-time padded state results in:

$$\begin{aligned} & T(Z^{z_1} X^{x_1} \otimes Z^{z_2} X^{x_2} \otimes Z^{z_3} X^{x_3} |\Psi\rangle) = \\ & T(Z^{z_1} X^{x_1} \otimes Z^{z_2} X^{x_2} \otimes Z^{z_3} X^{x_3}) T^\dagger T |\Psi\rangle = \\ & CNOT_{1,3}^{x_2} CNOT_{2,3}^{x_1} (I \otimes H) CNOT_{1,2}^{z_3} (I \otimes H) (Z^{z_1+x_2z_3} X^{x_1} \otimes Z^{z_2+x_1z_3} X^{x_2} \otimes Z^{z_3} X^{x_1x_2+x_3}) T |\Psi\rangle \end{aligned}$$

where  $CNOT_{i,j}^s$  is the “encrypted CNOT operation” applied to the  $i, j$ ’th qubits, where  $s$  indicates whether to apply the CNOT operation or not, and  $H$  is the Hadamard gate (see [Mah18a] for details).

To homomorphically evaluate a Toffoli gate,  $\text{Eval}_C$  applies the Toffoli gate  $T$  to the corresponding qubits in its padded state and updates the classical ciphertexts,  $\text{ct}$ , to be consistent with the pad

$$Z^{z_1+x_2z_3} X^{x_1} \otimes Z^{z_2+x_1z_3} X^{x_2} \otimes Z^{z_3} X^{x_1x_2+x_3}$$

(exactly as was done for Clifford gates). The remaining piece is to undo the (undesired) operations:

$$CNOT_{1,3}^{x_2} CNOT_{2,3}^{x_1} (I \otimes H) CNOT_{1,2}^{z_3} (I \otimes H).$$

Note that undoing the unitary  $I \otimes H$  is trivial, since all we need to do is multiply by the conjugate transpose, and this unitary (as well as the unitary corresponding to the Toffoli gate  $T$ ) satisfies the desired locality property.

Hence, in what follows we focus on how  $\text{Eval}$  undoes the encrypted CNOT operations (which is the crux of the difficulty). Note that  $CNOT_{i,j} CNOT_{i,j}$  is the identity, and hence to undo the encrypted CNOT operation  $\text{Eval}$  needs to apply  $CNOT_{i,j}^s$  given only a ciphertext encrypting the bit  $s$ . Indeed, the key idea in [Mah18a] is to show how this can be done (locally). For this, she needs the underlying classical encryption scheme to satisfy certain properties.

Specifically, [Mah18a] uses a special classical encryption scheme that is associated with a trapdoor claw-free function family.<sup>9</sup> The encryption scheme has the property that given an encryption of a bit  $s$ , denoted by  $\hat{s}$ , one can efficiently compute a description of a pair of claw free functions  $f_0, f_1 : \{0, 1\} \times \mathcal{R} \rightarrow \mathcal{Y}$  such that for every  $(\mu_0, r_0), (\mu_1, r_1) \in \{0, 1\} \times \mathcal{R}$  such that  $f_0(\mu_0, r_0) = f_1(\mu_1, r_1)$  (a “claw”) it holds that  $\mu_0 \oplus \mu_1 = s$ . Moreover, given  $\hat{s}$  one can efficiently compute the encryption of the trapdoor corresponding to the pair  $(f_0, f_1)$ .

Armed with this encryption scheme, one can compute the encrypted CNOT operation  $CNOT^s$  on a 2-qubit state

$$|\Psi\rangle = \sum_{a,b \in \{0,1\}} \alpha_{a,b} |a, b\rangle,$$

given  $\hat{s}$ , as follows:

1. Classically compute a description of the claw-free pair  $(f_0, f_1)$  corresponding to  $\hat{s}$ .

---

<sup>9</sup>A trapdoor claw-free function family  $\mathcal{F}$  is a family of injective functions, with a PPT algorithm that generates a pair of functions  $f_0, f_1 \in \mathcal{F}$  together with a trapdoor  $\text{td}$ , such that given (a description of)  $(f_0, f_1)$  it is hard to find a claw (i.e., a pair  $x_0, x_1$  such that  $f_0(x_0) = f_1(x_1)$ ), whereas given  $(f_0, f_1)$  together with the trapdoor  $\text{td}$  one can efficiently invert  $f_0$  and  $f_1$ , and in particular, find a claw.

2. Use the ancilla qubits to entangle  $|\Psi\rangle$  with a random claw for  $f_0, f_1$ , by computing

$$\sum_{a,b,\mu \in \{0,1\}, r \in \mathcal{R}} \alpha_{a,b} |a, b\rangle |\mu, r\rangle |f_a(\mu, r)\rangle$$

and measuring the last register to obtain  $y \in \mathcal{Y}$ . Let  $(\mu_0, r_0), (\mu_1, r_1)$  be the two preimages of  $y$ , namely

$$f_0(\mu_0, r_0) = f_1(\mu_1, r_1) = y.$$

Then the remaining state is

$$\sum_{a,b \in \{0,1\}} \alpha_{a,b} |a, b\rangle |\mu_a, r_a\rangle.$$

3. Then XOR  $\mu_a$  into the second register, which results in

$$\begin{aligned} \sum_{a,b \in \{0,1\}} \alpha_{a,b} |a, b \oplus \mu_a\rangle |\mu_a, r_a\rangle &= \\ \sum_{a,b \in \{0,1\}} \alpha_{a,b} (I \otimes X^{\mu_0}) |a, b + a \cdot s\rangle |\mu_a, r_a\rangle &= \\ \sum_{a,b \in \{0,1\}} \alpha_{a,b} (I \otimes X^{\mu_0}) CNOT_{1,2}^s |a, b\rangle |\mu_a, r_a\rangle \end{aligned}$$

where the first equation follows from the fact that  $\mu_0 \oplus \mu_1 = s$  and the second equation follows from the definition of  $CNOT_{1,2}^s$ .

4. Remove the entangled registers  $|\mu_a, r_a\rangle$  by applying the Hadamard transform to these registers to obtain

$$\sum_{a,b,d_0 \in \{0,1\}, d \in \{0,1\}^\ell} \alpha_{a,b} (I \otimes X^{\mu_0}) CNOT_{1,2}^s |a, b\rangle (-1)^{(d_0, d) \cdot (\mu_a, r_a)} |d_0, d\rangle$$

5. Measure the registers  $|d_0, d\rangle$  to obtain the state

$$\begin{aligned} (I \otimes X^{\mu_0}) CNOT_{1,2}^s \sum_{a,b \in \{0,1\}} (-1)^{(d_0, d) \cdot (\mu_a, r_a)} \alpha_{a,b} |a, b\rangle &= \\ (-1)^{(d_0, d) \cdot (\mu_0, r_0)} (I \otimes X^{\mu_0}) CNOT_{1,2}^s \left( \sum_{b \in \{0,1\}} \alpha_{0,b} |0, b\rangle + (-1)^{(d_0, d) \cdot ((\mu_0, r_0) \oplus (\mu_1, r_1))} \sum_{b \in \{0,1\}} \alpha_{1,b} |1, b\rangle \right) &= \\ (-1)^{(d_0, d) \cdot (\mu_0, r_0)} (I \otimes X^{\mu_0}) CNOT_{1,2}^s (Z^{(d_0, d) \cdot ((\mu_0, r_0) \oplus (\mu_1, r_1))} \otimes I) \left( \sum_{a,b \in \{0,1\}} \alpha_{a,b} |a, b\rangle \right) &= \\ (-1)^{(d_0, d) \cdot (\mu_0, r_0)} (Z^{(d_0, d) \cdot ((\mu_0, r_0) \oplus (\mu_1, r_1))} \otimes X^{\mu_0}) CNOT_{1,2}^s \left( \sum_{a,b \in \{0,1\}} \alpha_{a,b} |a, b\rangle \right). \end{aligned}$$

6. Finally, use the (classical) encryption of the trapdoor for  $(f_0, f_1)$  to homomorphically evaluate  $(d_0, d) \cdot ((\mu_0, r_0) \oplus (\mu_1, r_1))$  and  $\mu_0$ , and update the classical encryptions of the Pauli pads.

Note that all these operations (in Steps (1)-(6)) are either classical operations relating to updating the (classically) encrypted pad, or quantum operations applied to the 2 qubits of  $\Psi$  and to additional ancilla qubits. By the deferred measurement principle, we can think of the quantum operations as applying a unitary and only then applying the measurement.

In our context,  $\text{Eval}_C$  does these quantum operations on its large quantum state consisting of  $\ell + t$  qubits. However, all these operations (in Steps (1)-(6)) change only the 2-qubits on which the CNOT is applied and some of the ancilla registers. Overall, the unitary that  $\text{Eval}_C$  applies, when evaluating  $U \otimes I$ , is of the form  $U' \otimes I$  where the unitary  $U'$  is applied only to the qubits that  $U$  would act on and to the ancilla qubits, and  $I$  is the identity unitary applied to all the other registers, as desired.