A computational test of contextuality and, even simpler proofs of quantumness

Atul Singh Arora
QuICS, University of Maryland
College Park, USA
IQIM, Caltech
Pasadena, USA
atul.singh.arora@gmail.com

Kishor Bharti
*A*STAR Quantum Innovation Centre,*
Institute of High Performance Computing
Agency for Science,
*Technology and Research (A*STAR)*
Singapore
kishor.bharti1@gmail.com

Alexandru Cojocaru
School of Informatics,
University of Edinburgh
Edinburgh, UK
QuICS, University of Maryland
College Park, USA
a.cojocaru@ed.ac.uk

Andrea Coladangelo
Paul G. Allen School of
Computer Science and Engineering
University of Washington
Seattle, USA
andrea.coladangelo@gmail.com

***Abstract*—Bell non-locality is a fundamental feature of quantum mechanics whereby measurements performed on "spatially separated" quantum systems can exhibit correlations that cannot be understood as revealing predetermined values. This is a special case of the more general phenomenon of "quantum contextuality", which says that such correlations can occur even when the measurements are not necessarily on separate quantum systems, but are merely "compatible" (i.e. commuting). Crucially, while any non-local game yields an experiment that demonstrates quantum advantage by leveraging the "spatial separation" of two or more devices (and in fact several such demonstrations have been conducted successfully in recent years), the same is not true for quantum contextuality: finding the contextuality analogue of such an experiment is arguably one of the central open questions in the foundations of quantum mechanics.**

In this work, we show that an arbitrary contextuality game can be compiled into an "operational test of contextuality" involving a single quantum device, by only making the assumption that the device is computationally bounded. Our work is inspired by the recent work of Kalai et al. (STOC '23) that converts any non-local game into a classical test of quantum advantage with a single device. The central idea in their work is to use cryptography to enforce spatial separation within subsystems of a single quantum device. Our work can be seen as using cryptography to enforce "temporal separation", i.e. to restrict communication between sequential measurements.

Beyond contextuality, we employ our ideas to design a "proof of quantumness" that, to the best of our knowledge, is arguably even simpler than the ones proposed in the literature so far.

***Index Terms*—contextuality, foundations, computational assumptions, proof of quantumness, homomorphic encryption**

I. Introduction

One of the most intriguing features of quantum mechanics is that, in general, observable properties of a quantum system, usually referred to as "observables", do not seem to hold a precise value until they are measured. In technical jargon, quantum mechanics is not a "local hidden variable theory". While this feature is now well-understood, Einstein, Podolski, and Rosen, in their paper [1], originally conjectured that a local hidden variable explanation of quantum mechanics should exist. It was only years later that Bell [2], and subsequently Clauser, Horne, Shimony, and Holt (CHSH) [3], in their seminal works, proposed an operational test, i.e. an experiment, capable of ruling out a local hidden variable explanation of quantum mechanics. More precisely, they showed that there exists an experiment involving measurements on "spatially separated" quantum systems such that the outcomes exhibit correlations that cannot be explained by a local hidden variable theory. Such an experiment, usually referred to as a Bell test or a *non-local game*, has been performed convincingly multiple times [4]–[9]. Crucially, a Bell test rules out a local hidden variable theory under the assumption that the devices involved in the experiment are non-communicating (which is usually enforced through "spatial separation").

Bell non-locality may be viewed as a special case of the more general phenomenon of quantum contextuality [10], [11], which says that such correlations can occur even when the measurements are not necessarily on separate quantum systems, but are merely "compatible" (i.e. commuting). Contextuality has a long tradition in the foundations of quantum mechanics [12]. However, unlike a non-local game, a more general *contextuality game*[1] does not in general have a corre-

[1]We choose to use the term contextuality "game" here to preserve the analogy with a non-local game. However, in the contextuality literature, the more commonly used term is contextuality "scenario". We refer the reader to the technical overview (Subsection II-B) for details.

2575-8454/24/$31.00 ©2024 IEEE
DOI 10.1109/FOCS61266.2024.00073

sponding operational test. By an operational test, we mean a test that can be carried out on a device by interacting with it classically, and importantly, without having to make bespoke assumptions about its inner workings. Thus, even though some contextuality games are even simpler than non-local games [13], a satisfactory approach to compiling arbitrary contextuality games into operational "tests of contextuality" is missing. This is not for lack of trying—numerous attempts have been made that inevitably have to either resort to strong assumptions about the quantum hardware or assumptions that are hard to enforce in practice, such as the device being essentially *memoryless*.[2] Thus, one of the central open questions in the foundations of quantum mechanics is:

Is there a way to compile an arbitrary contextuality game into an "operational test of contextuality"?

Beyond demonstrating the presence of genuine quantum behaviour, non-local games can be employed to achieve a much more fine-grained control over the behaviour of untrusted quantum devices: for example, they allow a classical user to verify the correctness of full-fledged quantum computations, by interacting with two non-communicating quantum devices [25], [26]. Of course, any guarantee obtained via non-local games hinges on the physical (and non-falsifiable) assumption that the devices involved are non-communicating. To circumvent the need for this assumption, a lot of the attention in recent years has shifted to the computational setting. This exploration was kick-started by Mahadev's seminal work [27] showing, via cryptographic techniques, that the verification of quantum computations can be achieved with a single quantum device, under the assumption that the device is computationally bounded. She and her collaborators [28] later proposed what can be thought of as the analogue of a Bell/CHSH experiment with a single device—they proposed a simple test that an efficient quantum device can pass, but that an efficient classical device cannot. Since then, various works have proposed increasingly efficient "proofs of quantumness" in this setting [29]–[34]. The goal of this line of work is to simplify these tests to the point that they can be implemented on current quantum devices. An experimental realisation of such a proof of quantumness would be a milestone for the field of quantum computation, as it would constitute the first *efficiently verifiable* demonstration of quantum advantage. Towards this goal, the second question that we consider in this work is:

Can contextuality help realise simpler proofs of quantumness?

A. Our results

Our first contribution is a positive answer to the first question. We show that, using cryptographic techniques, an arbitrary contextuality game can be compiled into an "operational test of contextuality" involving a single quantum device, where the only assumption is that the device is computationally bounded.[3] A *contextuality game* involves a single player and a referee (unlike non-local games that always involve more than one player). In an execution of the game, the referee asks the player to measure a *context*—a set of commuting observables—and the player wins if the measurement outcomes satisfy certain constraints. Importantly, a given observable may appear in multiple contexts. If the player uses a strategy where the values of these observables are "predetermined" (which is the analogue of a "local hidden variable" strategy in the non-local game setting), then the highest winning probability achievable is referred to as the *non-contextual value* of the game, denoted by valNC.[4] On the other hand, if the player uses a quantum strategy, then the highest winning probability achievable is the *quantum value*, denoted by valQu (which is strictly greater than valNC for contextuality games of interest).[5] We formalise what we mean by an "operational test of contextuality" by introducing a property called *faithfulness*. Intuitively, this requires that the test has a strong correspondence to some contextuality game. We defer more precise definitions to the technical overview (see Subsection II-E2). We show the following.

Theorem 1 (informal, simplified)**.** *Any contextuality game can be compiled into a single prover, 2-round (i.e. 4-message) protocol such that the protocol is faithful to the contextuality game. Furthermore, the following two conditions hold (assuming standard cryptographic assumptions):*

- *(Completeness) There is a quantum polynomial-time (QPT) prover that wins with probability at least*
$$\frac{1}{2}(1+\mathrm{valQu})-\mathsf{negl}\,.$$
- *(Soundness) Any probabilistic polynomial-time (PPT) prover wins with probability at most*
$$\frac{1}{2}(1+\mathrm{valNC})+\mathsf{negl}\,.$$

Here negl *are (possibly different) negligible functions of a security parameter.*

The bounds in the statement above are for games with contexts of size two (which subsume "2-player" non-local games, see Fig. 1). The general bounds are similar but depend on the size of the contexts and are stated later. Notably, the number of messages, even in the general case (which subsumes non-local games with any number of players), remains constant (i.e. four). The cryptographic assumptions we make are the

[2]These assumptions are typically referred to as "loopholes" in the literature: [14]–[24].

[3]In the contextuality jargon, one might say that all "loopholes" are being replaced by a computational assumption.

[4]We emphasise that a non-contextual strategy is such that, if an "observable" appears in multiple contexts, then the *same* predetermined value should be returned by the player for that observable in *all* contexts in which it appears. This is precisely the difficulty in realising an "operational test of contextuality": how does the referee enforce that the player is consistent across contexts?

[5]A quantum strategy is such that, if an observable appears in multiple contexts, then the *same* observable should be measured to obtain the answer in all such contexts.

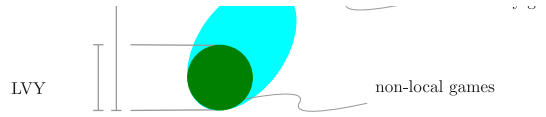


Fig. 1: The compiler in this work compiles a much larger set of games compared to the one in [32].

existence of (1) a quantum homomorphic encryption (QFHE) scheme[6] and (2) a new primitive, the *oblivious pad*, which we introduce below. QFHE can be realised assuming the quantum hardness of the Learning With Errors problem (LWE) [37]. We show how to construct the oblivious pad under the same assumption in the quantum random oracle model—which in turn can be heuristically instantiated using a cryptographic hash function, such as SHA3.[7] The *oblivious pad* can also be constructed from hardness of factoring (in the plain model) but this requires significantly more resources compared to the previously mentioned generic construction.

Our result is motivated by the recent work of Kalai et al. [32] (KLVY from here on) that converts any non-local game into a test of quantumness with a single device. Consider a non-local game with two players, Alice and Bob. The central idea behind the KLVY compiler is to use cryptography to enforce spatial separation within a *single* quantum device, i.e. to enforce that Alice and Bob's measurements occur on separate subsystems. The KLVY compiler relies on the following mechanism to cryptographically enforce spatial separation: Alice's question and answer are encrypted (using a quantum fully homomorphic encryption (QFHE) scheme), while Bob's question and answer are in the clear. The referee, who holds the decryption key, can then test the correlation across the two. Unfortunately, this approach does not extend to contextuality, at least not in any direct way, since in a contextuality game there is no notion of Alice and Bob.

In contrast, our work can be seen as using cryptography to enforce "temporal separation", i.e. to restrict communication between sequential measurements. In a nutshell, our idea is to ask the first question under a homomorphic encryption and the second question in the clear, as in KLVY, but with the following important difference. In KLVY, the quantum device prepares an *entangled* state, whose first half is encrypted, and used to homomorphically answer the first question, while the second half is not, and is used to answer the second question in the clear. In our protocol, there are no separate subsystems between the two rounds: instead, the encrypted post-measurement state from the first round (which results from the measurement performed to obtain the encrypted answer) is *re-used* in the second round. The technical barrier is, of course, the following: how can the post-measurement state be re-used if it is still encrypted? The two most natural approaches do not work:

(i) Providing the decryption key to the quantum device is clearly insecure as it allows the device to learn the first question in the clear.
(ii) Homomorphically encrypting the second question does not work either because the quantum device can correlate its answers to the two questions "under the hood of the homomorphic encryption".

We circumvent this barrier by introducing a procedure that allows the prover to *obliviously* "re-encrypt" the post-measurement state non-interactively. This re-encryption procedure is such that it allows the verifier to achieve the following: the verifier can now reveal some information that allows a quantum prover to access the post-measurement state in the clear, while a PPT prover *does not learn the first question*. More precisely, the verifier does not directly reveal the original decryption key, which would clearly be insecure, as pointed out earlier. Instead, the verifier expects the prover to "re-encrypt" its state using the new procedure, and then the verifier is able to safely reveal the resulting "overall" decryption key. Crucially, while a classical prover is unable to make use of the additional information to beat the classical value valNC, we show that there is an efficient quantum prover that can access the post-measurement state in the clear, and proceed to achieve the quantum value valQu.

The main technical tool that we introduce to formalise this idea, which may find applications elsewhere, is a primitive that we call *oblivious (Pauli) pad*. The oblivious pad takes as input a state $|\psi\rangle$ and a public key pk and produces a Pauli-padded state $X^x Z^z |\psi\rangle$ together with a string s (which can be thought of as encrypting x and z using pk). The string s can be used to recover x, z given the corresponding secret key sk. The security requirement is modelled as the following distinguishing game between a PPT prover and a challenger:

- The challenger generates public and secret keys $(\mathsf{pk}, \mathsf{sk})$ and sends pk to the prover.
- The prover produces a string s and sends it to the challenger.
- The challenger either returns (x, z) (as decoded using s and sk), or a fresh pad $(\tilde{x}, \tilde{z})$ sampled uniformly at random.

We require that no PPT prover can distinguish between the two cases with non-negligible advantage.[8] We show that an oblivious pad can be realised based on ideas from [28] in the

[6]With an assumption on the form of encrypted ciphertexts, which is satisfied by both Mahadev's and Brakerski's QFHE schemes, [35], [36]; see Subsection II-C.

[7]In the random oracle model [38] (ROM), a hash function f is modelled as a uniformly random black-box function: parties can evaluate it by sending a query x and receiving $f(x)$ in return. In the *quantum* random oracle model (QROM), such queries can also be made in superposition. A proof of security in this model is taken to be evidence for security of the protocol when the black-box is replaced by, for example, a suitable hash function f. This is because, informally, any attack on the resulting protocol must necessarily exploit the structure of f.

[8]Note that there is a QPT prover that *can* distinguish these two cases with non-negligible advantage.

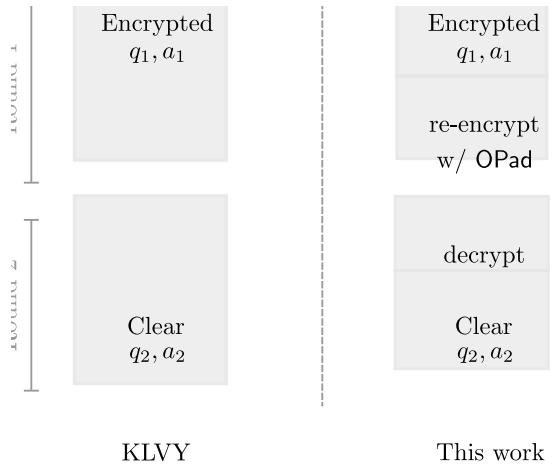


Fig. 2: A schematic comparing the non-local game compiler [32] with our contextuality game compiler. The key idea in KLVY is to ask the first question of a non-local game under a homomorphic encryption and the second one in the clear, with the prover using two entangled subsystems (one that is encrypted, and one in the clear). In our compiler, the oblivious pad (OPad) allows the prover to "re-encrypt" its post-measurement state, just before Round 2. Upon obtaining information about the "re-encryption" that took place, the verifier can then safely reveal the "overall" decryption key in Round 2, allowing the prover to proceed with the next measurement in the clear.

quantum random oracle model.[9] Crucially, while in KLVY the second phase of the protocol happens on a *different* subsystem, the oblivious pad is what allows us to carry out the second phase on the *same* subsystem, as depicted in Figure 2.

Our second contribution streamlines the ideas introduced to prove Theorem 1 in order to obtain a 2-round proof of quantumness relying on the classical hardness of the Learning-with-Errors (LWE) problem [37]. Our construction has the main advantage of being simpler than existing ones in the literature, in the sense explained below. Our proof of quantumness makes use of Noisy Trapdoor Claw-Free functions (NTCFs), introduced in [28] (but it does not require the NTCFs to have an "adaptive hardcore bit" property). It relies on the particular structure of the "encrypted CNOT operation" introduced in Mahadev's QFHE scheme [36]. We informally state our result, but we defer the construction to the technical overview.

Theorem 2 (Informal)**.** *Assuming the classical hardness of LWE, there exists a 2-round (i.e. 4-message) proof of quantumness with the following properties:*

- *It requires only one coherent evaluation of an NTCF, and one layer of single-qubit Hadamard gates.*
- *The quantum device only needs to maintain a single qubit coherent in-between the two rounds.*

Our proof of quantumness can be seen as combining ideas from [32] and [31], [33]. It is simpler than existing proofs of quantumness in the following ways:

- *Single encrypted CNOT operation.* The 2-round proof of quantumness of KLVY is based on a QFHE scheme. Concretely, an implementation of their proof of quantumness based on Mahadev's QFHE scheme requires performing a homomorphic controlled-Hadamard gate, which requires three sequential applications of the "encrypted CNOT operation". Crucially, these three operations require computing the NTCF three times in superposition while maintaining coherence all along. Our 2-round proof of quantumness only requires a single application of the "encrypted CNOT" operation. Moreover, in KLVY, the encrypted subsystem, on which Alice's operations are applied homomorphically needs to remain entangled with a second subsystem, which is used to perform Bob's operations in the clear. In contrast, in our proof of quantumness, the encrypted CNOT operation and the subsequent operations in the clear happen on a single system (of the same size as Alice's in KLVY).
- *Single qubit coherent across rounds and 2 rounds of interaction.* Compared to the 3-round proof of quantumness of [31], ours requires one less round of interaction. However, more importantly than the number of rounds, the protocol from [31] requires the quantum device to keep a superposition over preimages of the NTCF coherent in-between rounds, while waiting for the next message. In contrast, both our proof of quantumness and that of KLVY only require the quantum device to keep a *single* qubit coherent in-between rounds.
- *Simple quantum operations.* The 2-round proof of quantumness of [33] matches ours in that it requires a single coherent application of an NTCF based on LWE, and the quantum device only needs to keep a single qubit coherent in between rounds. However, their protocol requires the prover to perform single-qubit measurements in "rotated" bases coming from a set of a size that scales linearly with the LWE modulus (for which typical parameters are $\approx 10^2$ to 10^3). Their completeness-soundness gap also suffers a loss compared to ours that comes from the use of the "rotated" basis measurements.

To the best of our knowledge, the only aspect in which our proof of quantumness compares unfavourably with existing ones, e.g. [31], is that, based on current knowledge of constructions of NTCFs, our proof of quantumness requires a construction from LWE (in order to implement the "encrypted CNOT" procedure), whereas [31] has the flexibility that it can be instantiated using any TCF, e.g. based on Diffie-Hellman or Rabin's function. While the latter are more efficient to implement, they also generally require a larger

[9] One can also obtain a construction in the plain model, where the honest prover must factor a large integer (e.g. using Shor's algorithm). In fact, one can generically convert any non-interactive proof of quantumness into a construction of a close variant of the oblivious pad. Both of these constructions appear in the full version [39]. These constructions are based on a suggestion by an anonymous QCrypt reviewer.

security parameter (inverting Rabin's function is as hard as factoring, whereas breaking LWE is as hard as worst-case lattice problems). Hence, it is currently still unclear which route is closer to a realisation at scale. In particular, we note that instantiating the construction of [31] with the $x^2 \bmod N$ function by relying on a novel multiplication algorithm leads to a very efficient quantum circuit for the prover as shown in [40]. It is plausible that future improvements to our proof of quantumness might yield constructions from a broader class of hardness assumptions than LWE (e.g. Ring-LWE), which would likely yield a further improvement in concrete efficiency.

Some existing proofs of quantumness are non-interactive (i.e. 1-round), with security in the random oracle model [29], [30], [41]. Likewise, our proof of quantumness can also be made non-interactive by using the Fiat-Shamir transformation [42] (where the computation of the hash function is classical, and does not increase the complexity of the actual quantum computation). The proof of quantumness in [41] has the additional desirable property of being publicly-verifiable, although it currently seems to be more demanding than the others in terms of quantum resources.

B. Future Directions

- *Better contextuality compilers.* Focusing on compilers for contextuality, the following important aspect remains to be strengthened. The current compiler does not in general achieve *quantum soundness* (in the sense that there are contextuality games in which a QPT prover can do much better than the "compiled" quantum value from Theorem 1). Interestingly, recent works show that KLVY does satisfy quantum soundness for certain families of non-local games [43]–[45].
- *Oblivious Pauli pad.* We think that the new functionality of the oblivious Pauli pad (or an inspired generalisation) has the potential to be useful elsewhere, and we leave this exploration to future work. A related question is to realise the oblivious Pauli pad in the plain model without requiring the prover to factor a large integer, perhaps using ideas from [46].
- *More efficient "encrypted CNOT".* Turning to proofs of quantumness, the broad goal is of course to simplify these even further towards experimental implementations. One concrete direction in which our proof of quantumness could be simplified is the following. Currently, we use an NTCF that supports the "encrypted CNOT" operation from [36]. However, our only requirement is that the NTCF satisfies the potentially weaker property that it hides a bit in the xor of the first bit of a pair of preimages. This is because we only need to perform the "encrypted CNOT" operation once, and not repeatedly as part of a full-fledged homomorphic computation. It would be interesting to see if the weaker requirement could be achieved by simpler claw-free functions (from LWE or other assumptions, like DDH or factoring).
- *Testing other sources of quantumness.* Many other sources of quantumness have been identified in the literature, such as generalised contextuality (which allows arbitrary experimental procedures, not just projective measurements, and a broad class of ontological models, not just deterministic ones) [47] and the Leggett-Garg experiment (which is the time analogue of Bell's experiment) [48], [49]. These all suffer from the same limitation as contextuality, and our results raise the following question: can one construct analogous operational single-device tests for these sources of quantumness as well?
- *Testing indefinite causal order.* More ambitiously, one could try to separate quantum mechanics from more general theories such as those with indefinite causal order [50], [51] (i.e. theories that obey causality only locally and that may not admit a definite causal order globally). For instance, a recent result gives evidence (in the black-box model) that indefinite causal order does not yield any relevant advantage over quantum mechanics [52]. Perhaps one can obtain a clear separation under cryptographic assumptions?

Acknowledgements

We are thankful to an anonymous QCrypt reviewer for their observations about the relation between the oblivious pad and non-interactive proofs of quantumness, and for pointing out the limitations of not formalising what it means for a test to be an operational test of contextuality.

We thank Ulysse Chabaud, Mauro Morales and Thomas Vidick for their feedback on early drafts of this work. We thank Yusuf Alnawakhtha, Alexandru Gheorghiu, Manasi Mangesh Shingane, and Urmila Mahadev for various helpful discussions. ASA acknowledges support from the U.S. Department of Defense through a QuICS Hartree Fellowship, IQIM, an NSF Physics Frontier Center (GBMF-1250002) and MURI grant FA9550-18-1-0161. Part of the work was carried out while ASA was visiting the Simons Institute for the Theory of Computing. AC acknowledges support from the National Science Foundation grant CCF-1813814, from the AFOSR under Award Number FA9550-20-1-0108 and from the Engineering and Physical Sciences Research Council through the Hub in Quantum Computing and Simulation grant (EP/T001062/1) and the Quantum Advantage Pathfinder (QAP) research programme (EP/X026167/1). KB is supported by A*STAR C230917003.

II. Technical Overview

In this overview, we start by briefly recalling non-local games, and introducing the notion of contextuality with some examples. We then build up towards our compiler for contextuality games, by first introducing the ideas behind the KLVY compiler, and then describing why new ideas are needed to achieve a compiler in the contextuality setting. We then describe our main novel technical tool, the oblivious pad, and how we use it to realise a compiler for contextuality games. We also discuss the main ideas in the proof. Finally, we describe

how some of the new ideas can be streamlined to obtain a potentially simpler proof of quantumness. The latter can be understood without any reference to contextuality, and the interested reader may wish to skip directly to it (Section II-F).

A. Non-local Games

Let A, B, X, Y be finite sets. A 2-player non-local game is specified by a predicate $\mathsf{pred} : A \times B \times X \times Y \to \{0,1\}$ which indicates whether the players win or not, and a probability distribution $\mathcal{D}_{\text{questions}}$ over the questions, which specifies $\Pr(x,y)$ for $(x,y) \in X \times Y$. The game consists of a referee and two players, Alice and Bob, who can agree on a strategy before the game starts, but cannote communicate once the game starts. The game proceeds as follows: the referee samples questions $(x,y) \leftarrow \mathcal{D}_{\text{questions}}$, sends x to Alice and y to Bob, and receives their answers $a \in A$ and $b \in B$ respectively. Their success probability can be written as

$$\Pr(\text{win}) = \sum_{x,y,a,b} \mathsf{pred}(a,b,x,y) \Pr(a,b|x,y) \Pr(x,y)\,, \quad (1)$$

where the strategy used by Alice and Bob specifies $\Pr(a,b|x,y)$.

Local Hidden Variable Strategy: The "local hidden variable" model allows Alice and Bob to share a classical (random) variable r, and then have their answers be arbitrary, but fixed, functions of their respective questions, i.e.

$$\Pr(a,b|x,y) = \sum_r P_A(a|x,r) P_B(b|y,r) P_R(r)$$

where P_A, P_B and P_R are probability distributions that specify Alice and Bob's strategy. We refer to the optimal winning probability achievable by local hidden variable strategies as the *classical value* of the game.

Quantum Strategy: A quantum strategy allows Alice and Bob to share a state $|\psi\rangle_{AB}$, and use local measurements $M_x^A = \{M_{a|x}^A\}_a$ and $M_y^B = \{M_{b|y}^B\}_b$ to produce their answers (where the measurements can be taken to be projective without loss of generality), so that

$$\Pr(a,b|x,y) = \langle\psi| M_{a|x}^A \otimes M_{b|y}^B |\psi\rangle\,.$$

We refer to the optimal winning probability achievable by quantum strategies as the *quantum value* of the game.

It is well-known that there exist non-local games (like the CHSH game) where the quantum value exceeds the classical value. However, is it necessary to consider spatial separation (i.e. a tensor product structure) to observe such a "quantum advantage"? A partial answer is no: one can consider contextuality.

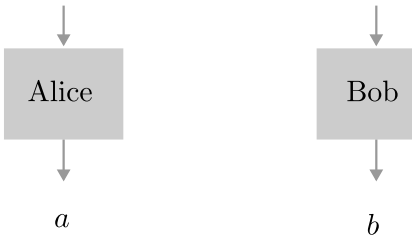


Fig. 3: A two-player non-local game. Alice and Bob get (x,y) from the referee and respond with (a,b). They cannot communicate once the game starts.

B. Contextuality

We start with a slightly informal definition of a contextuality game (see [39] for a formal treatment). Let Q and A be finite sets. Let C^{all} be a set of subsets of Q. We refer to the elements of C^{all} as *contexts*. Suppose for simplicity that all subsets $C \in C^{\text{all}}$ have the same size k. A *contextuality game* is specified by a predicate $\mathsf{pred} : A^k \times C^{\text{all}} \to \{0,1\}$, and a probability distribution $\mathcal{D}_{\text{contexts}}$ over contexts, which specifies $\Pr(C)$ for $C \in C^{\text{all}}$. The game involves a referee and a *single* player, and proceeds as follows:

- The referee samples $C = \{q_1, \ldots, q_k\} \in C^{\text{all}}$ according to $\mathcal{D}_{\text{contexts}}$, and sends C to the player.
- The player responds with answers $\{a_1, \ldots, a_k\} \in A^k$.

The success probability is

$$\Pr(\text{win}) = \sum_{a_1,\ldots,a_k,C} \mathsf{pred}(a_1,\ldots,a_k,C) \Pr(a_1,\ldots,a_k|C) \Pr(C)\,, \quad (2)$$

where the strategy used by the player specifies $\Pr(a_1,\ldots,a_k|C)$.

Non-contextual strategy: This is the analogue of a "local hidden variable" strategy. A *deterministic assignment* maps each question $q \in Q$ to a *fixed* answer a_q. This represents the following strategy: upon receiving the context $C = \{q_1,\ldots,q_k\} \in C^{\text{all}}$, return $(a_{q_1},\ldots,a_{q_k})$ as the answer. A strategy that can be expressed as a convex combination of deterministic assignments is referred to as *non-contextual*, i.e. the answer to a question q is independent of the context in which q is being asked.

Quantum strategy: A quantum strategy $\mathsf{qstrat} = (|\psi\rangle, \mathbf{O})$ is specified by a quantum state $|\psi\rangle$, and a collection of observables $\mathbf{O} = \{O_q\}_{q \in Q}$, such that, for any context $C = \{q_1,\ldots,q_k\} \in C^{\text{all}}$, the observables $O_{q_1},\ldots,O_{q_k}$ are compatible (i.e commuting). The strategy is the following: upon receiving context $C = \{q_1,\ldots,q_k\} \in C^{\text{all}}$, measure observables $O_{q_1},\ldots,O_{q_k}$ on state $|\psi\rangle$, and return the respective outcomes $a_1,\ldots,a_k$.

Quantum mechanics is *contextual* in the sense there are examples of games for which a quantum strategy can achieve a higher winning probability than the best non-contextual strategy. However, crucially, unlike a non-local game, a contextuality game does not directly yield an "operational test"

of contextuality. The issue is that there is no clear way for the referee to enforce that the player's answer to question q is consistent across the different contexts in which q appears!

We informally describe three simple examples: the magic square game (Peres-Mermin) [53]–[56], non-local games, and the KCBS experiment (the contextuality analogue of the Bell/CHSH experiment) [13]. We do this by directly specifying a quantum strategy first and then "deriving" the corresponding game (where the observables are just labels for the questions). In doing so, we abuse the notation slightly and identify questions with observables.

Example 1 (Peres-Mermin (Magic Square) [53], [54])**.** Consider the following set of observables

$$\mathbf{O} := \begin{array}{cccc} \{X\otimes\mathbb{I}, & \mathbb{I}\otimes Z, & X\otimes Z, & \mathbb{I} \\ \mathbb{I}\otimes X, & Z\otimes\mathbb{I}, & Z\otimes X, & \mathbb{I} \\ X\otimes X, & Z\otimes Z, & Y\otimes Y\} & \mathbb{I} \\ \\ \mathbb{I} & \mathbb{I} & -\mathbb{I} & \end{array}$$

(where X, Y, Z are Pauli matrices). They satisfy the following properties: (a) they take ± 1 values (i.e. they have ± 1 eigenvalues), (b) operators along any row or column commute, and (c) the product of observables along any row or column equals $\mathbb{I}$, except along col_3, where it equals $-\mathbb{I}$. If we define the set of contexts by $C^{\mathrm{all}} := \{\mathrm{row}_1, \mathrm{row}_2, \mathrm{row}_3, \mathrm{col}_1, \mathrm{col}_2, \mathrm{col}_3\}$, it is not difficult to see that no deterministic assignment can be such that the condition on the products is satisfied along each row and column. For instance, the assignment

$$\begin{array}{ccc} 1 & -1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & ? \end{array}$$

satisfies all constraints except one: if the question mark is 1, then the condition along col_3 fails, and if it is -1 then the condition along row_3 fails. To satisfy both, somehow the value assigned to the last "observable" has to depend on the context, row_3 or col_3, in which it appears—the assignment has to be "contextual". In quantum mechanics, all of the constraints can be satisfied using the observables described. One can in fact measure these observables on an arbitrary state $|\psi\rangle$ to win with probability 1. One thus concludes that quantum mechanics is "contextual" in this sense.

2-player non-local games can be viewed as a special case of contextuality games as follows (and similarly for non-local games with more players).

Example 2 (2-player non-local games)**.** Given a quantum strategy for a non-local game (as in Subsection II-A), we identify the measurements M_x^A and M_y^B with corresponding observables. Then, define $\mathbf{O} := \{M_x^A \otimes I\}_x \cup \{I \otimes M_y^B\}_y$ and $C^{\mathrm{all}} := \{\{M_x^A \otimes I, I \otimes M_y^B\}\}_{x,y}$. It is not hard to see that the set of non-contextual strategies is the same as the set of local hidden variable strategies.

Some contextuality games can yield a separation between non-contextual and quantum strategies with even smaller quantum systems than what is possible for non-local games. The following example yields a separation using just a qutrit—a single 3-dimensional system (in contrast, non-local games require at least two qubits, i.e. dimension 4, as in the CHSH game). For contextuality 3 dimensions are necessary and sufficient [11].[10]

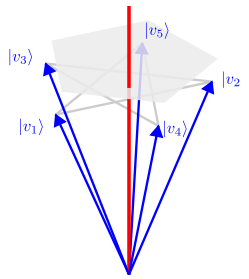


Fig. 4: An illustration of the optimal quantum strategy corresponding to the KCBS Game defined in Example 3. Here the red vector denotes the quantum state $|\psi\rangle$ and the blue ones $|v_q\rangle$ correspond to projective measurements, $\Pi_q = |v_q\rangle\langle v_q|$. Consecutively indexed blue vectors, i.e. $|v_q\rangle, |v_{q+1}\rangle$, are orthogonal (indexing is periodic).

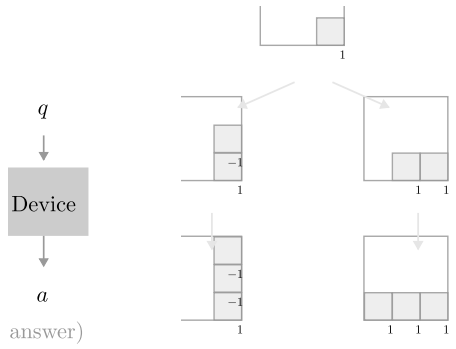
Example 3 (KCBS [13])**.** Consider a 3-dimensional vector space spanned by $\{|0\rangle, |1\rangle, |2\rangle\}$. Let $|\psi\rangle = |0\rangle$ and define five vectors

$$|v_q\rangle := \cos\theta\,|0\rangle + \sin\theta\sin\phi_q\,|1\rangle + \sin\theta\cos\phi_q\,|2\rangle\,,$$

indexed by $q \in \{1,\ldots,5\}$ where $\phi_q = 4\pi q/5$ and $\cos^2\theta = \cos(\pi/5)/(1+\cos(\pi/5))$. The heads of these vectors form a pentagon with $|\psi\rangle$ at the centre, and vectors indexed consecutively are orthogonal, i.e. $\langle v_q|v_{q+1}\rangle = 0$ (where we take the indices to be periodic) as illustrated in Fig. 4. Define $\mathbf{O} := \{\Pi_q\}_{q\in 1\ldots 5}$ where $\Pi_q := |v_q\rangle\langle v_q|$ and $C^{\mathrm{all}} := \{\{\Pi_1,\Pi_2\},\{\Pi_2,\Pi_3\}\ldots\{\Pi_5,\Pi_1\}\}$. The referee asks a context $C \leftarrow C^{\mathrm{all}}$ uniformly at random from the set of all contexts and the player wins if the answer is either $(0,1)$ or $(1,0)$ (i.e. neighbouring assignments should be distinct). It is not hard to check that a non-contextual strategy wins at most with probability $4/5 = 0.8$, while the quantum strategy described above wins with probability $\frac{2}{\sqrt{5}} \approx 0.8944$.

One route towards obtaining an operational test of contextuality is to find a way to enforce that measurements on a single system happen "sequentially", i.e. they are separated in "time" (as opposed to being separated in "space", which is the case for non-local games). We describe one folklore attempt at constructing an operational test, which assumes that the device is "memoryless". This example is not essential to understanding our results, and may be skipped at first read.

[10]There are generalisations of contextuality [47] that can give a separation with dimension 2, but we do not consider these here.



(a) (b)

Fig. 5: The folklore memoryless interpretation of contextuality: a memoryless device (left) and a memoryless operational test, illustrated using the Peres-Mermin magic square (right).

The "memoryless" attempt at an operational test: As denoted in Fig. 5a, consider a device that takes as input a question q, produces an answer a, and then forgets the question. The referee in this case proceeds as follows:

1) Samples a context C from C^{all} with probability $\Pr(C)$.
2) Sequentially asks all the questions $q_1 \dots q_k$ in the context C.
3) Accepts the answers $a_1 \dots a_k$ if the constraint corresponding to C holds, i.e. if $\mathsf{pred}(a_1 \dots a_k, C)$ is true.

Note that the most general deterministic model for the device is one that encodes a "truth table" $\tau : Q \to A$ that maps questions to answers. Since there is no memory, no previous question can affect the way the device answers the current question. The most general quantum device, on the other hand, starts with an initial $|\psi\rangle$ and to each question, assigns an observable O_q, which is measured to obtain an answer a to the question q.

Let us work out an example to illustrate the key point. Consider again the Magic Square game from Example 1. Suppose the first question the referee asks is the value of the bottom right box of the magic square (see Fig. 5b). It can then either ask questions completing the corresponding row or column. For a deterministic *memoryless* device, since a single truth table τ is being used, it is impossible to satisfy all the constraints of the magic square. However, a deterministic device *with memory* can satisfy all the constraints. This is because before answering the last question, it can learn whether the third column is being asked or the third row and thus, it can answer the last question to satisfy the constraint. Thus "classical memory", allows for classical simulation of contextuality in this test, without using any quantum effects. In fact, [57], [58] even quantifies the amount of classical memory needed to simulate contextuality.

Evidently, the glaring limitation of this attempt is that there is no operational way of ensuring that a device is "memoryless". Thus, this has remained a barrier despite the numerous attempts [12], [14]–[24], [59]–[62] and the fact that contextuality has, in general, been a bustling area of investigation [12], [22]–[24], [61], [63].

Our construction follows the same general approach of enforcing separation in "time" instead of in "space", but is a radical departure from the attempt described above. In the subsequent sections, we only assume that the device is *computationally bounded*, and use cryptographic techniques to construct an "operational test of contextuality".

Remark 3 (Criteria for being an Operational Test of Contextuality). We have not yet formally defined what we mean by an "operational test of contextuality". Consider any test that serves as a proof of quantumness. Intuitively, we require that this test, additionally, is *faithtful to some contextuality game* G, i.e. it satisfies the following two properties:

- The test involves asking the prover questions corresponding to G, possibly under some "encoding". The test is, however, allowed to involve other messages unrelated to G.
- Whether the test passes or fails is determined solely by evaluating the predicate corresponding to G, using the "decoded" questions and answers.

Formalising these requirements is a bit more involved and is deferred to Subsection II-E2. However, this intuitive notion suffices for now, as will be clear from our construction in the subsequent sections.

C. Quantum Fully Homomorphic Encryption (QFHE)

We informally introduce fully homomorphic encryption. We start with the classical notion.

Fully Homomorphic Encryption (FHE): A homomorphic encryption scheme is specified by four algorithms, $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec}, \mathsf{Eval})$, as follows:

- Gen takes as input a security parameter 1^λ, and outputs a secret key sk.
- Enc takes as input a secret key sk and a message s, and outputs a ciphertext c. We use the notation $\mathsf{Enc}_{\mathsf{sk}} = \mathsf{Enc}(\mathsf{sk}, \cdot)$.
- Dec takes as input a secret key sk, a ciphertext c and outputs the corresponding plaintext s. We use the notation $\mathsf{Dec}_{\mathsf{sk}} = \mathsf{Dec}(\mathsf{sk}, \cdot)$.

The "homomorphic" property says that a circuit $\mathsf{Circuit}$ can be applied on an encrypted input to obtain an encrypted output, i.e.

- Eval takes as input a circuit $\mathsf{Circuit}$, a ciphertext c, and an auxiliary input aux, and outputs a ciphertext c'. The following is satisfied. Let $c \leftarrow \mathsf{Enc}_{\mathsf{sk}}(s)$ and $c' \leftarrow \mathsf{Eval}(\mathsf{Circuit}, c, \mathsf{aux})$, then $\mathsf{Dec}_{\mathsf{sk}}(c') = \mathsf{Circuit}(s, \mathsf{aux})$.

Crucially, note that the secret key sk is not needed to apply the Eval algorithm. The security condition is the usual one: that an encryption of s should be indistinguishable from an encryption of $s' \neq s$. If Eval supports evaluation for the class

of all polynomial-size circuits (in the security parameter), then the scheme is said to be *fully* homomorphic.[11]

Quantum Fully Homomorphic Encryption (QFHE): For this overview, it suffices to take a QFHE scheme to be the same as an FHE scheme, except that it allows messages and auxiliary inputs to be quantum states, and circuits to be quantum circuits. The QFHE schemes that are relevant to our work [35], [36] satisfy the following additional properties.

1) *Classical encryption/decryption.*
   Gen is a classical algorithm, while $\mathsf{Enc}, \mathsf{Dec}$ become classical algorithms when their inputs are classical. In particular, this means that classical inputs are encrypted into classical ciphertexts. This property is essential when the scheme is deployed in protocols involving classical parties, as will be the case here.[12]
2) *Locality is preserved.*
   Consider an arbitrary bipartite state $|\psi\rangle_{AB}$ and let M^A and M^B be circuits acting on registers A and B respectively with a measurement at the end. The property requires that correlations between the measurement outcomes from M^A and M^B should be the same in the following two cases:
   (i) Register A is encrypted, M_A is applied using Eval, and the result is decrypted. Let a be the decrypted outcome. M^B is applied to register B. Let b be the outcome.
   (ii) Apply M^A on register A and M^B on register B. Let a and b respectively be the outcomes.
3) *Form of Encryption.*
   Encryption of a state $|\psi\rangle$ takes the form $(X^x Z^z |\psi\rangle, \widehat{xz}) \leftarrow \mathsf{Enc}_{\mathsf{sk}}(|\psi\rangle)$, where X^x applies a Pauli X to the i-th qubit based on the value of x_i, and Z^z is defined similarly. $\widehat{xz}$ is an encryption of the pad xz.

All known constructions of QFHE schemes satisfying property 1 also satisfy properties 2 and 3. The KLVY compiler, as one might guess, relies on property 2. Our compiler relies on property 3.

D. The KLVY Compiler

Consider a two-player non-local game specified by question and answer sets X, Y, A, B, a predicate $\mathsf{pred} : A \times B \times X \times Y \to \{0,1\}$ and a distribution over the questions $\mathcal{D}_{\text{questions}}$ that specifies $\Pr(x,y)$. The KLVY compiler takes this non-local game as input and produces the following single-prover game, where the verifier proceeds as follows:

- **Round 1.** Sample $(x,y) \leftarrow \mathcal{D}_{\text{questions}}$, and a secret key sk for a QFHE scheme. Send an encryption c_x of "Alice's question", and get an encrypted answer c_a.
- **Round 2.** Send an encryption c_y of "Bob's question", and get an encrypted answer c_b in the clear. Decrypt the answers using sk, and accept if $\mathsf{pred}(a,b,x,y) = 1$.

The honest prover prepares the entangled state $|\psi\rangle_{AB}$ corresponding to the optimal quantum strategy for the non-local game. It uses QFHE's Eval algorithm on subsystem A to answer question x according to Alice's optimal strategy, and answers question y in the clear using Bob's optimal strategy on subsystem B. More formally, they proceed as in Fig. 6.

Theorem ([32], informal)**.** *Consider a two-player non-local game with classical and quantum values ω_c and ω_q respectively. The corresponding KLVY-compiled single-prover game satisfies the following:*

- *(Completeness) The QPT prover described above makes the verifier accept with probability $\omega_q - \mathsf{negl}(\lambda)$.*
- *(Soundness) Every PPT prover makes the verifier above accept with probability at most $\omega_c + \mathsf{negl}(\lambda)$.*

Here negl *are (possibly different) negligible functions.*

For completeness, we briefly outline the general compiler for non-local games with $k > 2$ players. To compile games with k players, KLVY generalises the procedure above as follows: the compiled game consists of k rounds–each round consisting of a question and an answer (so $2k$ messages in total). The first $k-1$ questions are encrypted (using $k-1$ independent random QFHE keys) and the last question is asked in the clear.

Let us take a moment and build some intuition about the KLVY compiler. At first, one might think that it would be even more secure to also encrypt y. However, as we remarked in the introduction, this is not the case: since the prover can compute any desired answer b using x, a and y under the hood of the QFHE, it can ensure that b is such that $\mathsf{pred}(a,b,x,y) = 1$ (for any non-trivial choice of pred). Instead, the key observation of KLVY is that the verifier is testing the correlation between encrypted answers and answers in the clear, and it turns out that a quantum prover can produce stronger correlations than any PPT prover can. How is this intuition formalised? The key idea is that, since a PPT prover is classical, one can rewind the PPT prover to obtain answers to *all* possible second round questions. This is equivalent to obtaining Bob's entire assignment of answers to questions (corresponding to a fixed encrypted question and answer for Alice from the first round). Suppose for a contradiction that such a PPT prover wins with probability non-negligibly greater than the classical value. Then, it must be that Bob's entire assignment is non-trivially correlated to "Alice's encrypted question". This information can thus be used to obtain a non-negligible advantage in guessing the encrypted question, breaking security of the QFHE scheme.

Now that we understand the KLVY construction and the key insight behind their proof, we will discuss barriers to extending these ideas to contextuality, and our approach to circumvent these.

[11] While homomorphic schemes for restricted families of circuits have been known for some time, a "fully homomorphic scheme" was only discovered somewhat recently in [64].

[12] The first schemes to satisfy this property appeared recently in the breakthrough works [35], [36].

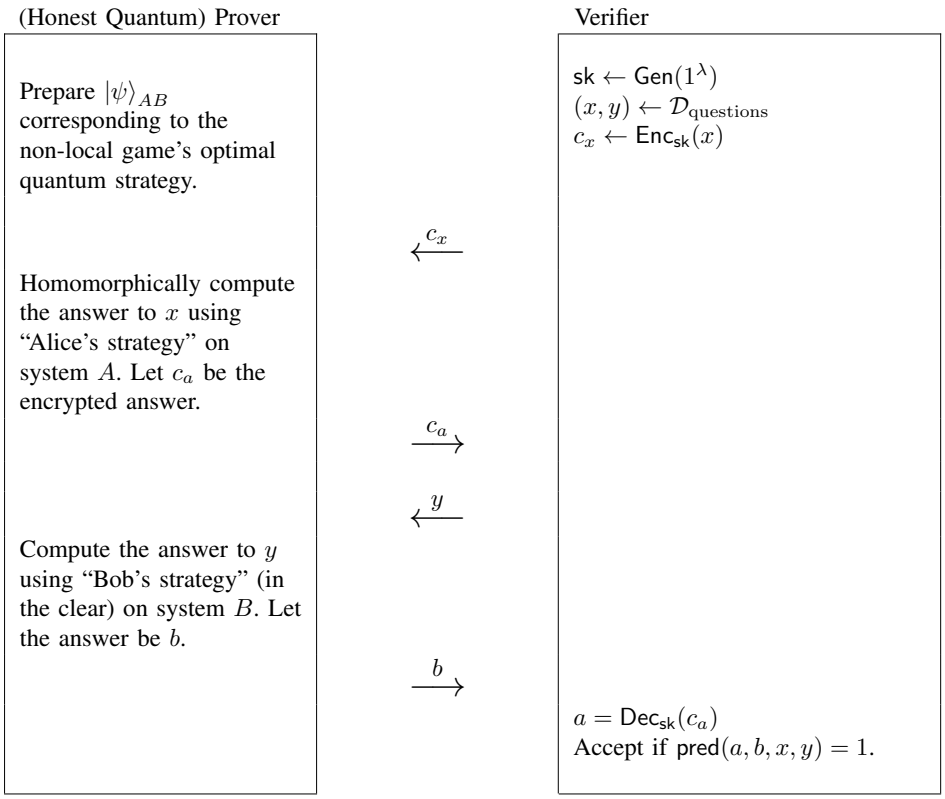


Fig. 6: The KLVY compiler [32] is illustrated above: it takes any two-player non-local game and converts it into a proof of quantumness.

E. Contribution 1 — A Computational Test of Contextuality

For simplicity, let us first restrict to contextuality games with contexts of size 2.

Attempts at extending KLVY to contextuality.: Let us consider compilers for contextuality games where the verifier proceeds in an analogous way as the verifier in the KLVY compiler:

- **Round 1.** Sample a context $C = (q_1, q_2) \leftarrow \mathcal{D}_{\text{contexts}}$. Send a QFHE encryption c_{q_1} of q_1 and receive as a response an encryption c_{a_1} of a_1.
- **Round 2.** *There is no clear way to proceed. Three natural approaches are listed below.*

The honest prover's state after Round 1 is encrypted and has the form $(X^x Z^z |\psi_{(q_1,a_1)}\rangle, \widehat{xz})$ where $\widehat{xz}$ denotes a classical encryption of the strings xz (using Property 3 of the QFHE scheme), as detailed below in Figure 7.

Here are three natural approaches for how to proceed, and how they fail.

1) *Proceed just as KLVY: Ask question q_2 in the clear.* This does not work because, unlike in the original KLVY setup, there is no analogue of system B which is left in the clear. Here the prover holds an encrypted state so it is unclear how q_2 can be answered with any non-trivial dependence on the state under the encryption.
2) *Ask the second question also under encryption.* If the same key is used for encryption, then, as we argued for KLVY, the predicate of the game can be satisfied by computing everything homomorphically. If the keys are independent, then we essentially return to the problem in item 1.
3) *Reveal the value of the (classically) encrypted pad $\widehat{xz}$ and ask question q_2 in the clear.* This has a serious issue: the prover can simply ask for the encrypted pad corresponding to c_{q_1} and thereby learn q_1 (or at least some bits of q_1). Once q_1 is learned, again, the predicate of the game can be trivially satisfied.

The Oblivious Pauli Pad: As mentioned in Section I, our compiler relies on a new cryptographic primitive, that we introduce to circumvent the barriers described above. Here, it helps to be a bit more formal. Let $\mathbf{U} := \{U_k\}_{k \in K}$ be a group of unitaries acting on the Hilbert space $\mathcal{H}$. We take this group to be the set of Paulis $\{X^x Z^z\}_{xz}$, as this makes the primitive compatible with the the form of the QFHE scheme that we will employ later in our compiler. Nonetheless, we use the general notation $\{U_k\}_{k \in K}$, as it simplifies the presentation. We define the *oblivious* $\mathbf{U}$ *pad* as follows.

The *oblivious* $\mathbf{U}$ *pad* is a tuple of algorithms $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ where Gen and Dec are PPT. Let

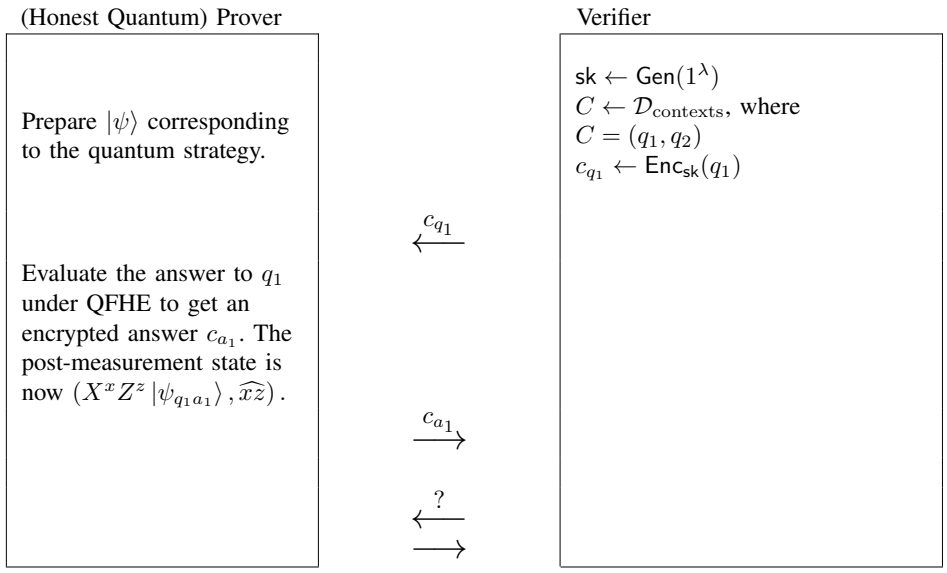


Fig. 7: The figure illustrates that, a priori, it is unclear how to generalise KLVY to compile contextuality games.

$(\mathsf{pk}, \mathsf{sk}) \leftarrow \mathsf{Gen}(1^\lambda)$ be the public and the secret keys generated by Gen. Encryption takes the form $(U_k |\psi\rangle, s) \leftarrow \mathsf{Enc}_{\mathsf{pk}}(|\psi\rangle)$, where $k = \mathsf{Dec}_{\mathsf{sk}}(s)$. Notice the similarity with the post-measurement state in the discussion above (we will return to this in a moment). The security requirement is that no PPT algorithm can win the following security game with probability non-negligibly greater than $1/2$. We depict the oblivious pad primitive in Figure 8.

The security game formalises the intuition that no PPT prover can distinguish between the correct "key" k_1 and a uniformly random "key" k_0. We emphasize, in words, the two distinctive features of this primitive:

- By running Enc, a QPT prover can obtain, given a state $|\psi\rangle$, an encryption of the form $(U_k |\psi\rangle, s)$, where $k = \mathsf{Dec}_{\mathsf{sk}}(s)$.
- There is no way for a PPT prover, given pk, to produce an "encryption" s, for which it has non-negligible advantage at guessing $\mathsf{Dec}_{\mathsf{sk}}(s)$.

We describe informally how to instantiate the primitive in the random oracle model assuming noisy trapdoor claw-free functions (the detailed description is in [39]). The construction builds on ideas from [29].

The key idea is the following. Let f_0, f_1 be a Trapdoor Claw-Free function pair. We take $\mathsf{pk} = (f_0, f_1)$, and sk to be the corresponding trapdoor. Then, $\mathsf{Enc}_{\mathsf{pk}}$ is as follows:

(i) On input a qubit state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$, evaluate f_0 and f_1 in superposition, controlled on the first qubit, and measure the output register. This results in some outcome y, and the leftover state $(\alpha |0\rangle |x_0\rangle + \beta |1\rangle |x_1\rangle)$, where $f(x_0) = f(x_1) = y$.

(ii) Compute the random oracle "in the phase", to obtain $((-1)^{H(x_0)}\alpha |0\rangle |x_0\rangle + (-1)^{H(x_1)}\beta |1\rangle |x_1\rangle)$. Measure the second register in the Hadamard basis. This results in a string d, and the leftover qubit state

$$|\psi_Z\rangle = Z^{d\cdot(x_0 \oplus x_1)+H(x_0)+H(x_1)} |\psi\rangle .$$

(iii) Repeat steps (i) and (ii) on $|\psi_Z\rangle$, but *in the Hadamard basis*! This results in strings y' and d', as well as a leftover qubit state $|\psi_{XZ}\rangle =$

$$X^{d'\cdot(x'_0 \oplus x'_1)+H(x'_0)+H(x'_1)} Z^{d\cdot(x_0 \oplus x_1)+H(x_0)+H(x_1)} |\psi\rangle ,$$

where x'_0 and x'_1 are the pre-images of y'.

Notice that the leftover qubit state $|\psi_{XZ}\rangle$ is of the form $X^x Z^z |\psi\rangle$ where x, z have the following two properties: (a) a verifier in possession of the TCF trapdoor can learn z and x given respectively y, d and y', d', and (b) no PPT prover can produce strings y, d as well as predict the corresponding bit z with non-negligible advantage (and similarly for x). Intuitively, this holds because a PPT prover that can predict z with non-negligible advantage must be querying the random oracle at *both* x_0 and x_1 with non-negligible probability. By simulating the random oracle (by lazy sampling, for instance), one can thus extract a claw x_0, x_1 with non-negligible probability, breaking the claw-free property.

1) OUR COMPILER*:* We finally describe our contextuality game compiler. As mentioned in the introduction, our strategy is still to ask the first question under QFHE encryption and the second question in the clear, with the following crucial difference: the prover is first asked to "re-encrypt" the post-measurement state using the *oblivious pad* functionality (from here one referred to as OPad), and only *after that* the verifier reveals to the prover how to "decrypt" the state, in order to proceed to round 2.

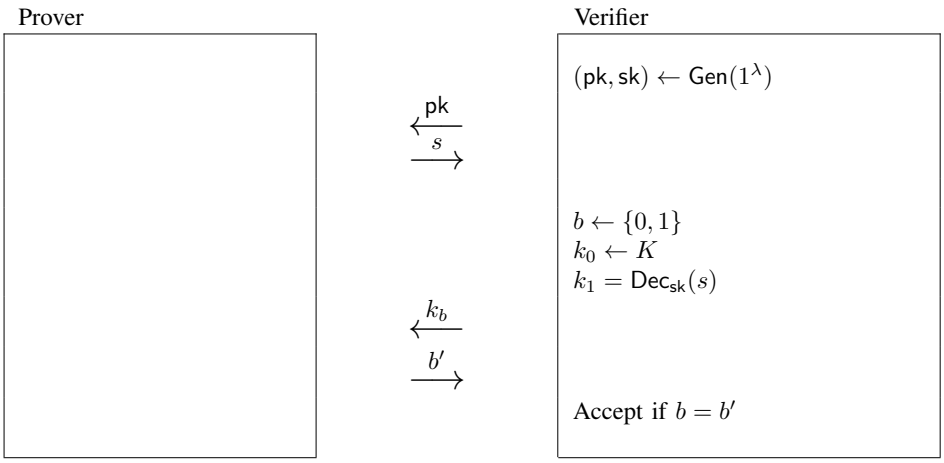


Fig. 8: The security game for the oblivious pad.

The key idea is easy to state, once the notation is clear. To this end, recall that Property 3 of a QFHE scheme ensures that the encryption of a quantum state $|\psi''\rangle$ has the form

$$(U_{k''} |\psi''\rangle, \hat{k}'') \tag{3}$$

where $\hat{k}''$ denotes a classical encryption of k'' (the reason why we use double primes will become clear shortly), and $U_{k''}$ is an element of the Pauli group. Note that using the secret key of the QFHE scheme, one can recover k'' from $\hat{k}''$. Further, let the optimal quantum strategy for the underlying contextuality game consist of state $|\psi\rangle$ and observables $\{O_q\}$. Finally, denote by $|\psi_{a,q}\rangle$ the post-measurement state arising from measuring $|\psi\rangle$ using O_q and obtaining outcome a.

We are now ready to describe our compiler. We first explain it in words and subsequently give a more formal description for clarity. In both cases, we highlight the conceptually new parts in blue.

- **Round 1** *Ask the encrypted question, and have the prover re-encrypt its post-measurement state using* OPad.
  The verifier samples keys for the QFHE scheme and for the OPad. It samples a context $C \leftarrow \mathcal{D}_{\text{contexts}}$ and then uniformly samples questions q_1, q_2 from the context C (note that $q_1 = q_2$ with probability $1/2$ since, for simplicity, we are considering contexts of size 2.)
  - **Message 1** The verifier sends the QFHE encryption c_{q_1} of the first question q_1, together with the public key of the OPad.

  The honest quantum prover obtains the encrypted answer c_{a_1} by measuring, under the QFHE encryption, the state $|\psi\rangle$ using the observable O_{q_1}. It now holds a state of the form in (3), with $|\psi''\rangle = |\psi_{a_1,q_1}\rangle$. Using the public key of the OPad, the prover applies OPad.Enc to the encrypted post-measurement state, $U_{k''} |\psi_{a_1,q_1}\rangle$, to obtain the "re-encrypted" quantum state, $U_{k'} U_{k''} |\psi''\rangle$, where $U_{k'}$ was applied by the OPad, together with a classical string s' that encodes k'. This step is critical to the security of the protocol and is discussed in more detail shortly. Crucially, note that both $U_{k'}$ (coming from the OPad) and $U_{k''}$ (coming from the QFHE) are Paulis.
  - **Message 2** The prover sends the QFHE encrypted answer c_{a_1}, together with the two strings $\hat{k}''$ and s'.
- **Round 2** *Remove the overall encryption, and proceed in the clear.* The verifier recovers k' from s' (using the secret key of the OPad) and k'' from $\hat{k}''$ (using the secret key of the QFHE scheme). It then computes k satisfying $U_k = U_{k'} U_{k''}$.
  - **Message 3** The verifier sends the second question q_2 together with k as computed above.

  The prover measures its quantum state $U_{k'} U_{k''} |\psi''\rangle = U_k |\psi_{a_1,q_1}\rangle$, using observable $U_k O_{q_2} U_k^\dagger$ to obtain an outcome a_2.
  - **Message 4** The prover sends a_2.

  The verifier decrypts c_{a_1} to recover a_1 using the secret key of the QFHE scheme. If $q_1 = q_2$, it accepts if the answers match, i.e. $a_1 = a_2$. If $q_1 \neq q_2$, it accepts if the predicate is true, i.e. $\mathsf{pred}(a_1, a_2, q_1, q_2) = 1$.

We summarise our compiler in Figure 9. Since we now have Gen, Enc, Dec algorithms for both OPad and QFHE, we use prefixes such as OPad.Enc to refer to Enc associated with OPad to avoid confusion.

Our compiler satisfies the following, assuming the underlying QFHE and oblivious pad are secure. We first state a special case of our general result (which is stated later in Theorem 4).

Theorem (restatement of Theorem 1)**.** *Consider a contextuality game* G *with contexts of size* 2*. Let* valNC *and* valQu *be its non-contextual and quantum values respectively. The compiled game (as described above) is faithful to* G*. In particular, it satisfies the following:*

- *(Completeness) The QPT prover described above wins with probability* $\frac{1}{2}(1 + \text{valQu}) - \mathsf{negl}(\lambda)$.

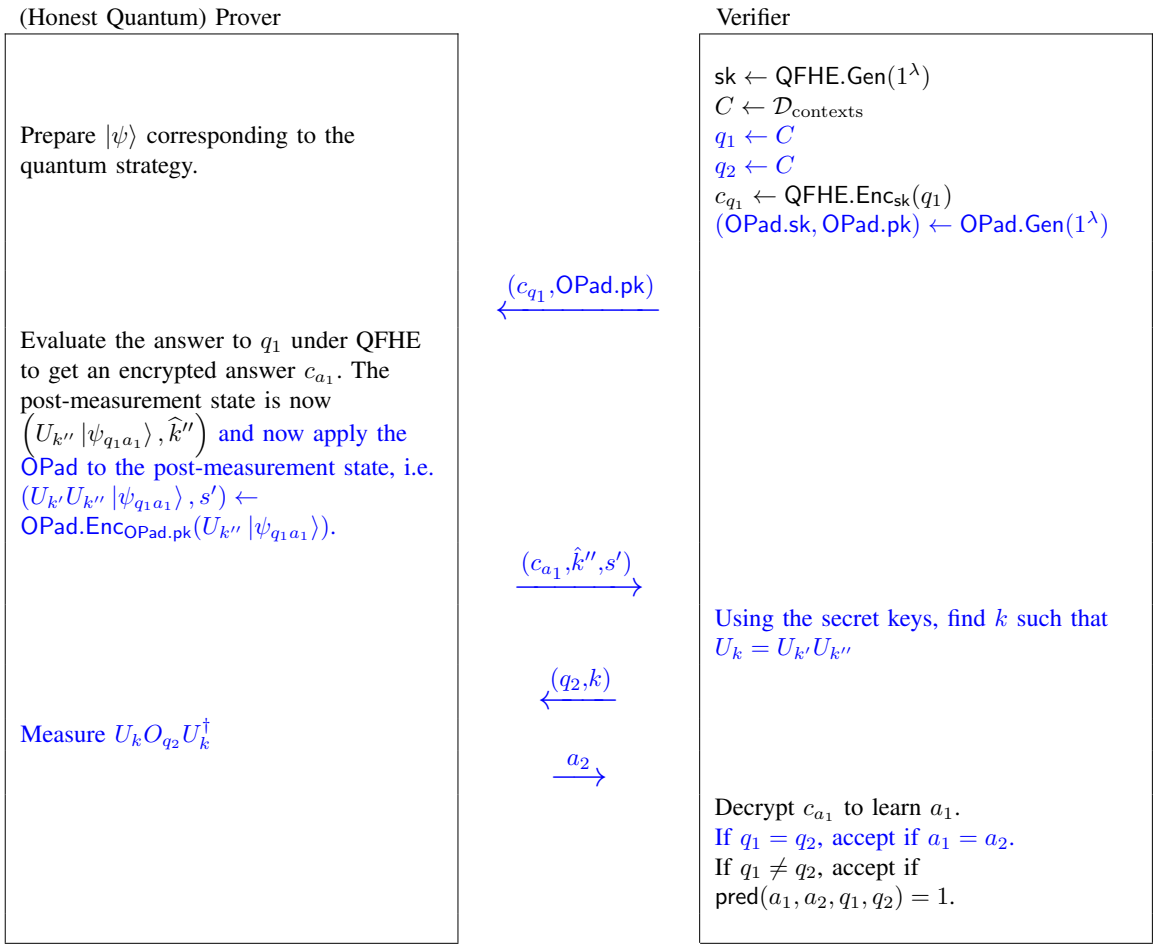


Fig. 9: Our contextuality game compiler. It takes a contextuality game given by $(Q, A, C^{\text{all}}, \mathsf{pred}, \mathcal{D}_{\text{contexts}})$, a QFHE scheme and an OPad and returns an operational test of contextuality.

- *(Soundness) Any PPT prover wins with probability at most* $\frac{1}{2}(1 + \text{valNC}) + \mathsf{negl}(\lambda)$.

Here negl *denote (possibly different) negligible functions.*

Proof sketch: The faithfulness condition as discussed in Remark 3 is evidently satisfied by the compiled game (and it is easy to verify that the formal notion as stated later in Subsection II-E2 is also satisfied). Suppose $\mathcal{A}$ is a PPT algorithm that wins with probability non-negligibly greater than $\frac{1}{2}(1+\text{valNC})$. Observe that one can associate a "deterministic assignment" corresponding to $\mathcal{A}$, conditioned on some fixed first round messages, as follows: simply rewind $\mathcal{A}$ to learn answers to all possible second round questions, obtaining an assignment $\tau : Q \to A$, mapping questions to answers. Let us write τ_{q_1} to make the dependence of the assignment on the first question more explicit (note that the assignment depends on the encrypted question c_{q_1} as well as the encrypted answer c_{a_1}). For the purpose of this overview, suppose also that $\mathcal{A}$ is consistent, i.e. if $q_1 = q_2$, then $a_1 = a_2$ (note that this in particular ensures that, when $q_1 = q_2$, $\mathcal{A}$ wins with probability 1. One can show that an adversary that is not consistent can be turned into an adversary that is consistent and wins with at least the same probability). Now, to win with probability more than $\frac{1}{2}(1 + \text{valNC})$, it must be that the τ_{q_1}'s are different for different q_1's. Otherwise, $\mathcal{A}$'s strategy is just a convex combination of deterministic assignments and this by definition cannot do better than valNC when $q_1 \neq q_2$. But if the distribution over τ_{q_1}s and $\tau_{q_1'}$s is different for at least some $q_1 \neq q_1'$, then one is able to distinguish QFHE encryptions of q_1 from those of q_1'. Thus, as long as the QFHE scheme is secure, no PPT algorithm can win with probability non-negligibly greater than $\frac{1}{2}(1 + \text{valNC})$.

In the above sketch, we glossed over the very important subtlety that, in order to obtain the truth table τ, the reduction needs to provide as input to $\mathcal{A}$ the "correct" decryption key k (as the verifier does in the third message of our compiled game, where k is such that $U_k = U_{k'}U_{k''}$). However, the reduction only sees *encryptions* of k' and k''. So, how does it compute

k without the secret keys? Crucially, this is where the OPad comes into play—it allows the reduction to instead use an independent uniformly random k (not necessarily the "correct" one) when constructing the reduction that breaks the security of the QFHE scheme. The fact that such a k is computationally indistinguishable from the correct one (from the point of view of the prover $\mathcal{A}$) follows precisely from the security of the OPad.

General Compilers: The compiler we described earlier handles contextuality games with contexts of size 2. How does one generalise it to contexts of arbitrary size? Unlike for KLVY, it is not entirely clear what the "correct" way is here.

We design two compilers (which seem incomparable). The first compiler applies universally to all contextuality games. The second applies primarily to contextuality games where the quantum value is 1 (for instance, it works for the magic square but not for KCBS). Notably, both compilers are 4-message protocols.

- $(|C|-1,1)$ compiler:
  - Round 1: Ask $|C|-1$ questions under QFHE.
  - Round 2: Ask 1 uniformly random question in the clear. If the question was already asked in Round 1, check consistency. Otherwise, check the predicate.
- $(|C|,1)$ compiler
  - Round 1: Ask all $|C|$ questions under QFHE.
  - Round 2: Ask 1 uniformly random question in the clear, and check consistency with the questions asked in Round 1.

By now, one would be wary of guessing that asking more questions under QFHE is going to improve the security of the protocol. Indeed, the $(|C|-1,1)$ compiler (which reduces to the one we discussed above for $|C|=2$) is the universal one. We show the following.

Theorem 4 (informal)**.** *Consider an arbitrary contextuality game* G*, and let* valNC *and* valQu *be its non-contextual and quantum values respectively. The compiled game, obtained via the* $(|C|-1,1)$ *compiler, is faithful to* G *and satisfies the following:*

- *(Completeness) There is a QPT prover that wins with probability* $1-\frac{1}{|C|}+\frac{\mathrm{valQu}}{|C|}-\mathsf{negl}$.
- *(Soundness) PPT provers win with probability at most* $1-\frac{1}{|C|}+\frac{\mathrm{valNC}}{|C|}+\mathsf{negl}$.

The compiled game, obtained via the $(|C|,1)$ *compiler, is also faithful to* G*, and satisfies the following:*

- *(Completeness) There is a QPT prover that wins with probability* valQu.
- *(Soundness) PPT provers win with probability at most* $1-\mathsf{const}_1+\mathsf{negl}$*, where* $\mathsf{const}_1=\min_{C\in C^{\mathrm{all}}}\frac{\Pr(C)}{|C|}$ *(this is constant in the sense that it is independent of the security parameter), and* $\Pr(C)$ *denotes the probability of sampling the context* C*.*

Here, negl *are (possibly different) negligible functions.*

We make some brief remarks about the two compilers and defer the details to the main text.

- The $(|C|,1)$ compiler is not universal because, for instance, when applied to KCBS, there is no gap between the PPT and the QPT prover's winning probabilities. In fact, there is a PPT algorithm[13] that does better than the honest quantum strategy. Yet, the compiler does apply to the magic square game, for instance, because $\mathsf{const}_1<1$ and $\mathrm{valQu}=1$. In fact, for the magic square game, this compiler gives a better completeness-soundness gap than the $(|C|-1,1)$ compiler.
- The $(|C|-1,1)$ compiler is universal in the sense that, when applied to any contextuality game with a gap between non-contextual and quantum value, the compiled game will have a constant gap between completeness and soundness. However, the resulting gap is sometimes smaller compared to the previous compiler. It is unclear if one can do better than this, with or without increasing the number of rounds, while preserving universality.

2) CRITERIA FOR BEING AN OPERATIONAL TEST OF CONTEXTUALITY*:* We conclude the overview of our first contribution by formalising what we mean by an operational test of contextuality. Based on the discussion in Subsection II-B, it is reasonable to assert that the most general non-contextual physically relevant model of computation is simply a PPT machine. Thus, any proof of quantumness, i.e. any test that distinguishes a PPT machine from a quantum machine, may be taken to be a proof of contextuality. In other words, quantumness and contextuality become equivalent notions, in accordance with this definition.

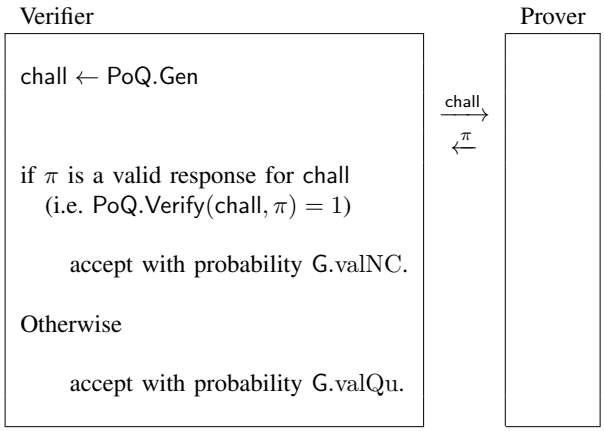
This, however, is unsatisfactory because quantumness could be arising from some other non-classical feature of quantum mechanics which may have no a priori connection to contextuality. For instance, (assuming factoring is hard for a PPT machine) equivalence of quantumness and contextuality would mean that factoring serves as a proof of contextuality. Yet, it is unclear how Shor's algorithm demonstrates contextuality in any direct way.

One proposal could be that a satisfactory test must be "universal" in the sense that corresponding to each contextuality game, there should be a systematic way to construct a corresponding test such that the completeness and soundness values of the test correspond to the quantum and noncontextual value of the underlying contextuality game. A little thought shows that this is not enough, as illustrated by Example 10 below.

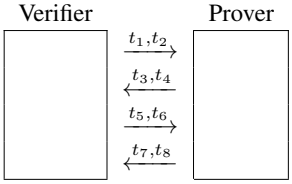
Example 10 is unsatisfactory as a test of contextuality because, even though its soundness and completeness values correspond to the classical and quantum values of the game G, this correspondence is artifical: the prover is not asked a single question that is related to G. We now formalise a stronger and more natural notion of correspondence. To this end, we introduce some notation.

[13]Assuming that Eval is PPT if the circuit and input are classical.

Example 10. This example uses a general non-interactive proof of quantumness as an ingredient: Let $\mathsf{PoQ} = (\mathsf{Gen}, \mathsf{Verify}, \mathsf{Cert})$ denote a proof of quantumness protocol where $\mathsf{Gen}, \mathsf{Verify}$ are PPT algorithms and Cert is a QPT algorithm. Here, Gen generates a challenge, Cert generates a response to the challenge and Verify tests whether the response is valid. For simplicity, we take them to satisfy the property that no PPT machine can make Verify accept while Cert always produces a valid certificate, i.e. PoQ has perfect completeness and soundness. Given a contextuality game G, the protocol is the following: the verifier runs a proof of quantumness protocol PoQ and if the prover passes, the verifier accepts with probability $\mathsf{G}.\mathrm{valQu}$ and if the prover fails, accepts with probability $\mathsf{G}.\mathrm{valNC}$. Evidently, this protocol has completeness $c = \mathsf{G}.\mathrm{valQu}$ and soundness $s = \mathsf{G}.\mathrm{valNC}$. Furthermore, it satisfies Condition 1, but fails to satisfy Condition 2 (see Definition 5).



As a starting point towards a candidate "operational test" $\mathcal{P}$ for G, consider a proof of quantumness protocol $\mathcal{P}$ involving a PPT verifier V and a prover P, with soundness s and completeness c. To be concrete, suppose protocol $\mathcal{P}$ involves four messages. Further suppose that the messages are parsed as follows:



For instance t_1 could be a public key and t_2 could be an encrypted message.

Mapping messages in $\mathcal{P}$ to questions/answers in G.: Let $\mathsf{G} = (Q, A, C^{\mathrm{all}}, \mathrm{pred}, \mathcal{D})$. We require that $\mathcal{P}$ classifies each message into one of three categories: (1) a question in G, (2) an answer in G or (3) other. More precisely, we require that $\mathcal{P}$ specifies a map map_i for each message t_i.
For indices i corresponding to messages *received by the prover* (in our example, t_1, t_2, t_5, t_6), map_i is either

- the constant $\perp$ output map, $\mathsf{map}^{\perp}$ (i.e. $\mathsf{map}^{\perp}$ outputs $\perp$ on all inputs) or
- a map $r_i : C_i \to Q$ from the set of possible messages C_i to a question in the game G.

One can think of r_i as a "decoding map" for the underlying question.
For indices i corresponding to messages *sent by the prover* (in our example t_3, t_4, t_7, t_8), map_i is either

- the constant $\perp$ output map $\mathsf{map}^{\perp}$ or
- a map $s_i : A \to C_i$ from the set of answers A in the game to possible messages C_i.

Similarly, one can think of s_i as an "encoding map" for the underlying answer. Continuing in the same vein as the example above, suppose $t_1 = \mathsf{pk}$ is a public key for some encryption scheme and $t_2 \leftarrow \mathsf{Enc}_{\mathsf{pk}}(q)$ is an encryption of a question $q \in Q$. Then, a natural choice for the corresponding maps is $\mathsf{map}_1 = \mathsf{map}^{\perp}$ and $\mathsf{map}_2 = \mathsf{Dec}_{\mathsf{sk}}$ where sk is the secret key corresponding to pk. Note that we gave one choice for these maps but $\mathcal{P}$ can specify these maps arbitrarily. The non-triviality arises from the requirements we place on $\mathcal{P}$, using these maps.

Faithfulness to G.: In order to capture the fact that protocol $\mathcal{P}$ is faithfully executing an instance of the game G, and not rewarding some other capability of the prover, we consider two families of "simulators". These are computationally unbounded machines and are meant to simulate either a classical or a quantum strategy. The idea is that using $\{\mathsf{map}_i\}_i$ one can isolate the messages that correspond to questions and answers in the game G. One can then define machines that answer these questions non-contextually, or using a quantum strategy, and can behave arbitrarily otherwise. More precisely, we have the following:

- A classical simulator $S_{\tau,\mathsf{aux}}$ is an unbounded machine that is designed to interact with V and responds to "encoded" questions (as specified by $\{\mathsf{map}_i\}_i$) using some non-contextual assignment $\tau : Q \to A$ i.e. the simulator decodes the message t_i using r_i to recover a question $q \in Q$, computes $a = \tau(q)$ and responds with the

encoding $s_j(a)$, for each pair of message indices (i,j) corresponding to a question and its answer. It responds to the remaining messages using an arbitrary strategy specified by[14] aux. We denote by $\mathbf{S}_{\mathrm{NC}} = \{S_{\tau,\mathsf{aux}}\}_{\tau,\mathsf{aux}}$ the set of all classical simulators.

- A quantum simulator $S_{\mathsf{qstrat},\mathsf{aux}}$ is an unbounded machine that is also designed to interact with V and responds to the "encoded" questions using a quantum strategy $\mathsf{qstrat} = (|\psi\rangle, \mathbf{O})$ (see Subsection II-B) but responds to the remaining messages using an arbitrary strategy specified by aux. We denote by $\mathbf{S}_{\mathsf{qu}} = \{S_{\mathsf{qstrat},\mathsf{aux}}\}_{\mathsf{qstrat},\mathsf{aux}}$ the set of all quantum simulators.

We can now state our definition. We ignore negligible additive factors for clarity.

Definition 5 (Operational Test of Contextuality)**.** *We say $\mathcal{P}$ is an operational test of contextuality if there exist s and c (where $s < c$) such that, in addition to being a proof of quantumness with soundness s and completeness c, $\mathcal{P}$ is* faithful to *some* contextuality game G. *We say $\mathcal{P}$ is* faithful to $\mathsf{G} = (Q, A, C^{\mathrm{all}}, \mathsf{pred}, \mathcal{D})$ *if the following hold:*

1) *Well-formedness: $\mathcal{P}$ must specify maps $\{\mathsf{map}_i\}_i$ as described above. Moreover, for any possible set Q' of questions that the verifier V asks in a single execution, i.e.*
$$Q' := \{\mathsf{map}_i(t_i)\}_{i:t_i=r_i}\,, \tag{4}$$
we require that the questions belong to some context, i.e. $Q' \subseteq C$ for some context $C \in C^{\mathrm{all}}$.
2) G*-soundness: For all classical simulators, i.e. $S \in \mathbf{S}_{\mathrm{NC}}$, the probability that the verifier V accepts when interacting with S should be at most the classical value s, i.e.*
$$\Pr\left(1 \leftarrow \langle V, S\rangle\right) \leq s \quad \forall \quad S \in \mathbf{S}_{\mathsf{NC}}.$$
3) G*-completeness: There exists a quantum prover Q_{complete} such that $s < \Pr\left[1 \leftarrow \langle V, Q_{\mathsf{complete}}\rangle\right] \leq c$, and a quantum simulator $S \in \mathbf{S}_{\mathsf{qu}}$, satisfying the following: the transcript produced by the interaction of S with V is distributed identically to the case when Q_{complete} interacts with V. In particular, this implies*
$$\Pr[1 \leftarrow \langle V, S\rangle] = \Pr\left[1 \leftarrow \langle V, Q_{\mathsf{complete}}\rangle\right].$$
4) *Decision Faithfulness: For all $S_{\tau,\mathsf{aux}} \in \mathbf{S}_{\mathsf{NC}}$, consider the questions Q' asked by V (see Eq. (4)) in an execution of $\langle V, S_{\tau,\mathsf{aux}}\rangle$.*
   a) *If Q' is the set of all questions in some context, i.e. $Q' = C$ for some $C \in C^{\mathrm{all}}$, then the verifier outputs $\mathsf{pred}(\tau[C], C)$.*
   b) *Otherwise, the verifier outputs* 1.

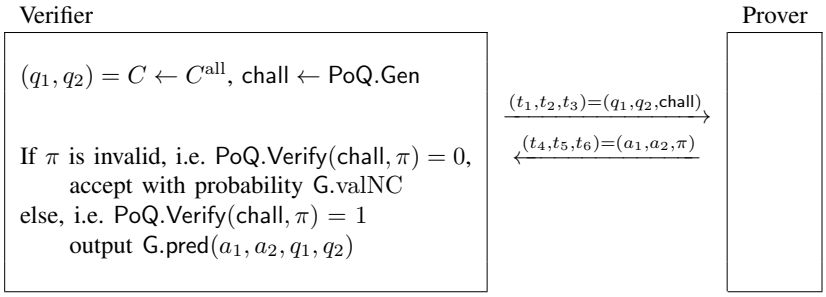
Justification: According to Definition 5, if a proof of quantumness protocol $\mathcal{P}$ is faithful to a contextuality game G, it must specify a mapping $\{\mathsf{map}_i\}_i$ relating it to G. Now, if a prover wins with probability greater than s, from Condition 2, we know that there is no way of interpreting the behaviour of this prover as being consistent with a non-contextual assignment in G (via $\{\mathsf{map}_i\}_i$). By Condition 3, we know that there is a quantum prover whose behaviour is indistinguishable from that of a prover that implements an honest quantum contextual strategy in G (via $\{\mathsf{map}_i\}_i$). And finally, Condition 4, ensures that the criterion for rewarding/penalising a prover is determined solely by the predicate of G (via $\{\mathsf{map}_i\}_i$).

We now introduce these requirements sequentially, illustrating why each one of them plays a crucial role in ruling out unsatisfactory notions of operational tests of contextuality.

- Condition 1 is a very basic requirement. For instance, in Example 10, one could assign map $\mathsf{map}^{\perp}$ to every message in the protocol and this would satisfy 1.
- Condition 2 ensures the following: Suppose that the messages that map to $\perp$ are answered by using unbounded computational power. Even in this case, as long as the messages that correspond to questions/answers in G (as specified by the protocol $\mathcal{P}$) are answered non-contextually, the condition requires that one cannot do better than a PPT machine.
  For instance, Example 10 fails to satisfy this requirement: an unbounded classical simulator $S \in \mathbf{S}_{\mathsf{NC}}$ can easily produce a valid proof of quantumness and make the verifier accept with probability $\mathsf{G}.\mathrm{valQu}$ which is greater than the soundness value $s = \mathsf{G}.\mathrm{valNC}$. Clearly, Example 10 was a very simple construction. Consider the less trivial construction in Example 11 below where the verifier asks questions in the game G, in addition to requesting a proof of quantumness, but only considers the prover's answers when the PoQ π is valid. Using natural maps $\{\mathsf{map}_i\}$ (i.e. $\mathsf{map}_1 = \mathsf{map}_2 = \mathsf{map}_4 = \mathsf{map}_5 = \mathbb{I}, \mathsf{map}_3 = \mathsf{map}_6 = \mathsf{map}^{\perp}$) it is immediate that Condition 2 is satisfied by this construction: even if a simulator $S \in \mathbf{S}_{\mathsf{NC}}$ provides a valid proof of quantumness, the verifier does not accept with probability more than $\mathsf{G}.\mathrm{valNC}$. Yet, this construction is intuitively unsatisfactory as a test of contextuality because it is completely neglecting the answers a_1, a_2 given by a classical prover.
- Condition 3 is also satisfied by Example 11 and this illustrates why the last condition (below) is so crucial.
- Condition 4 is where Example 11 finally fails: Consider two simulators $S_{\tau,\mathsf{aux}}, S_{\tau,\mathsf{aux}'} \in \mathbf{S}_{\mathsf{NC}}$ where the assignment τ corresponds to a game value $v < \mathsf{G}.\mathrm{valNC}$ and aux corresponds to producing the correct proof of quantumness while aux' corresponds to producing an invalid proof of quantumness. Condition 4 (a) requires that $\Pr[1 \leftarrow \langle V, S_{\tau,\mathsf{aux}}\rangle] = \Pr[1 \leftarrow \langle V, S_{\tau,\mathsf{aux}'}\rangle]$ but, $\Pr[1 \leftarrow \langle V, S_{\tau,\mathsf{aux}}\rangle] = v$ and $\Pr[1 \leftarrow \langle V, S_{\tau,\mathsf{aux}'}\rangle] = \mathsf{G}.\mathrm{valNC}$.
- The relevance of Condition 4 (b) is evident from Example 12 below, where the verifier starts by asking for a proof of quantumness and only if this is valid does it ask the questions for the contextuality game G; otherwise it simply rejects. Intuitively, this is unsatisfactory because

[14]Here, aux may be thought of as describing the "program of a Turing Machine" and may involve potentially unbounded computation.

Example 11. Given a contextuality game (with size-two contexts) G, the verifier proceeds as follows (where chall and π are as in Example 10). This protocol satisfies Conditions 1, 2 and 3, but fails to satisfy Condition 4 (a) (see Definition 5).



a classical prover is not even able to see the questions in G, let alone answer them (unlike Example 11). Yet, none of the previous conditions are violated. Such cases are excluded by Condition 4 (b) because it requires that no prover can be penalised unless all questions in some context are asked.

We conclude this discussion by observing that our compilers in Subsection II-E satisfy Definition 5, where the maps $\{\mathsf{map}_i\}_i$ are naturally chosen to be the following: encryption/decryption of the FHE scheme for the encrypted questions and answers, the identity map for the questions and answers in the clear, and $\mathsf{map}^{\perp}$ for everything else.

F. Contribution 2 — An even simpler proof of quantumness

Our proof of quantumness, like many of the existing ones in the literature, is based on the use of Trapdoor Claw-Free Functions (TCF). In our protocol, these are used to realize an "encrypted CNOT" functionality, which is the central building block of Mahadev's QFHE scheme [36]. The "encrypted CNOT" functionality allows a prover to homomorphically apply the gate CNOT^a, while holding a (classical) encryption of the bit a. Formally, our protocol uses Noisy Trapdoor Claw-Free Functions (NTCF, see [39] for a formal treatment), but here we describe our scheme using regular TCFs for simplicity.

The proof of quantumness: Our 2-round proof of quantumness is conceptually very simple. It can be viewed as combining and distilling ideas from the proofs of quantumness in [32], [31] and our contextuality compiler. We provide an informal description here, and we defer a formal description to [39]. At a high level, it can be understood as follows:

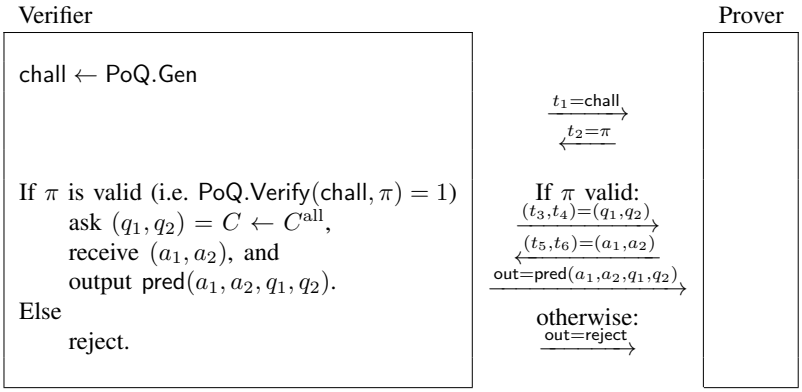
- **Round 1**: *Delegate the preparation of a uniformly random state in $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, unknown to the prover.*
  The verifier samples a bit a uniformly at random.
  - **Message 1**: The verifier sends an appropriate encryption of a to the prover (and holds on to the corresponding secret key).

  The honest prover prepares the two-qubit state $|+\rangle|0\rangle$, along with auxiliary registers required to perform an "encrypted CNOT" operation. It then performs an "encrypted CNOT" operation (from the first qubit to the second), i.e. homomorphically applies CNOT^a, followed by a measurement of the second (logical) qubit.
  - **Message 2**: The prover sends all measurement outcomes to the verifier.

  Since a CNOT gate can be thought of as a deferred measurement in the standard basis, we can equivalently think of the prover's operations as performing an "encrypted measurement" of the first qubit, where the first qubit is being measured or not based on the value of a. Note that, after having performed these operations, the prover holds a *single* qubit. Thanks to the specific structure of the "encrypted CNOT" operation from [36], the resulting "post-measurement" qubit state is encrypted with a Quantum One-Time Pad, and is either:
  - $|+\rangle$ or $|-\rangle$, if $a = 0$ (i.e. no logical CNOT was performed)
  - $|0\rangle$ or $|1\rangle$, if $a = 1$ (i.e. a logical CNOT was performed)

  All in all, at the end of round 1, the honest prover holds a uniformly random state in $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, i.e. a BB84 state. This state is known to the verifier, who possesses a and the secret key. From here on, the protocol no longer uses any encryption, and everything happens "in the clear".
- **Round 2**: *Ask the prover to perform "Bob's CHSH measurement".*
  The astute reader may notice that the qubit held by the prover after Round 1 is distributed identically to "Bob's qubit" in a CHSH game where Alice and Bob perform the optimal CHSH strategy. More precisely, if one imagines that Alice has received her question and performed her corresponding optimal CHSH measurement (which is either in the standard or Hadamard basis), the leftover state of Bob's qubit is a uniformly random state in $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$, where the randomness comes both from the verifier's question (which in our protocol corresponds to the bit a), and Alice's measurement outcome.
  - **Message 3**: The verifier sends a uniformly random bit b to the prover (corresponding to Bob's question in a CHSH game).

Example 12. Given a contextuality game (with size-two contexts) G, the verifier proceeds as above (where chall, π are as in Example 10). It is easy to see that that this protocol satisfies all conditions, except 4 (b), with $s = \mathsf{G.valNC}$ and $q = \mathsf{G.valQu}$.



The prover performs Bob's optimal CHSH measurement corresponding to question b.

- **Message 4**: The prover returns the measurement outcome to the verifier.

The verifier checks that the corresponding CHSH game is won.

In Fig. 13, we show the circuit for the honest quantum prover in our proof of quantumness.

Soundness: An efficient quantum prover can efficiently pass this test with probability $\cos^2(\frac{\pi}{8}) \approx 0.85$, while an efficient classical prover can pass this test with probability at most $3/4$. The proof of classical soundness is fairly straightforward. In essence, a classical prover can be rewound to obtain answers to *both* of the verifier's possible questions in Message 3. If the classical prover passes the test with probability $3/4+\delta$ (which is on average over the two possible questions), then the answers to both questions together must reveal some information about the encrypted bit a (this is a simple consequence of how the CHSH winning conditions is defined). In particular, such a classical prover can be used to guess a with probability $\frac{1}{2} + 2\delta$. This breaks the security of the encryption, as long as δ is non-negligible. We defer the reader to [39] for more details.

Putting the ideas in perspective: We have already discussed in Section I-A how our proof of quantumness compares to existing ones in terms of efficiency. Here, we focus on how our proof of quantumness compares conceptually to [32] and [31]:

- In [32], the prover is asked to create an entangled EPR pair, of which the first half is encrypted, and the second half is in the clear. Then, the prover is asked to perform Alice's ideal CHSH measurement homomorphically on the first half, and Bob's CHSH measurement in the clear on the second half. Our proof of quantumness departs from this thanks to two observations:
  - By leveraging the structure of the "encrypted CNOT operation" from [36], the post-measurement state from Alice's homomorphic measurement can be reused *in the clear* (precisely because the verifier knows what the state is, but the prover does not). So the initial entanglement is not needed. This idea is also the starting point for our contextuality compiler from Section II-E, although for the latter we take this idea much further: we find a way to give the prover the ability to decrypt the leftover state without giving up on soundness. Our proof of quantumness is a baby version of this idea: it leverages the fact that the leftover encrypted state has a special form, namely it is a BB84 state.
  - In order to setup a "CHSH-like correlation" between the verifier and the leftover qubit used by the prover in Round 2, one does not need to compile the CHSH game in its entirety. This compilation, even for the simple CHSH game requires the prover to perform an "encrypted controlled-Hadamard" operation (because Alice's ideal CHSH measurements are in the standard and Hadamard bases). The latter requires three sequential "encrypted CNOT" operations. Instead, our observation is that one can setup this CHSH-like correlation more directly, as we do in Round 1 of our proof of quantumness.
- From a different point of view, our proof of quantumness can also be viewed as a simplified version of [31]. Indeed, observation (ii) is inspired by the proof of quantumness in [31], which introduces the idea of a "computational" CHSH test. One can interpret [31] as setting up an "encrypted classical operation", akin to an "encrypted CNOT", that either entangles two registers or does not, ultimately having the effect of performing an "encrypted

[15]The simplified description of the proof of quantumness before this figure is slightly inaccurate: it states that the prover prepares starts by preparing the two-qubit state $|+\rangle|0\rangle$. Technically, the prover only needs to prepare $|+\rangle$, and the role of the second qubit is performed by the first qubit of the pre-image register (which is initialised as a uniform superposition).

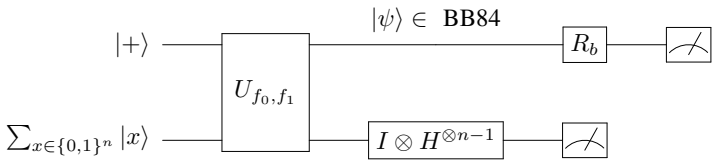


Fig. 13: The prover's circuit. Here, (f_0, f_1) is a pair of trapdoor claw-free functions (with inputs of size n). U_{f_0,f_1} denotes the $(n+1)$-qubit unitary that coherently computes f_0 in the last n qubits if the first qubit is $|0\rangle$, and computes f_1 otherwise. The circuit starts by preparing $|+\rangle$ in the first qubit, and a uniform superposition over all inputs in the next n qubits. The circuit then applies U_{f_0,f_1} (note that we are omitting auxiliary work registers that are required to compute U_{f_0,f_1}), followed by a layer of Hadamard gates on the last $n-1$ qubits. Then, the last n qubits are measured. As a result, the leftover qubit $|\psi\rangle$ in the first register is now a BB84 state (which one it is depends on f_0, f_1, and the measurement outcome). As its second message, the verifier sends a bit b, and the prover applies the rotation R_b defined as follows: $|0\rangle \overset{R_b}{\mapsto} \cos((-1)^b\pi/8)\,|0\rangle + \sin((-1)^b\pi/8)\,|1\rangle$ and $|1\rangle \overset{R_b}{\mapsto} -\sin((-1)^b\pi/8)\,|0\rangle + \cos((-1)^b\pi/8)\,|1\rangle$. Finally, the prover measures the qubit in the standard basis.[15]

measurement". This is achieved via an additional round of interaction. Our proof of quantumness can be thought of as zooming in on this interpretation, and finding a direct way to achieve this without the additional interaction.

References

[1] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" *Phys. Rev.*, vol. 47, pp. 777–780, May 1935. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.47.777

[2] J. S. Bell, "On the einstein podolsky rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, p. 195, 1964.

[3] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," *Physical review letters*, vol. 23, no. 15, p. 880, 1969.

[4] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. Vermeulen, R. N. Schouten, C. Abellán *et al.*, "Loophole-free bell inequality violation using electron spins separated by 1.3 kilometres," *Nature*, vol. 526, no. 7575, pp. 682–686, 2015.

[5] M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán *et al.*, "Significant-loophole-free test of bell's theorem with entangled photons," *Physical review letters*, vol. 115, no. 25, p. 250401, 2015.

[6] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman *et al.*, "Strong loophole-free test of local realism," *Physical review letters*, vol. 115, no. 25, p. 250402, 2015.

[7] M.-H. Li, C. Wu, Y. Zhang, W.-Z. Liu, B. Bai, Y. Liu, W. Zhang, Q. Zhao, H. Li, Z. Wang *et al.*, "Test of local realism into the past without detection and locality loopholes," *Physical review letters*, vol. 121, no. 8, p. 080404, 2018.

[8] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, "Event-ready bell test using entangled atoms simultaneously closing detection and locality loopholes," *Physical review letters*, vol. 119, no. 1, p. 010402, 2017.

[9] S. Storz, J. Schär, A. Kulikov, P. Magnard, P. Kurpiers, J. Lütolf, T. Walter, A. Copetudo, K. Reuer, A. Akin *et al.*, "Loophole-free bell inequality violation with superconducting circuits," *Nature*, vol. 617, no. 7960, pp. 265–270, 2023.

[10] E. Specker, "Die logik nicht gleichzeitig entsc heidbarer aussagen," *Dialectica*, vol. 14, no. 2-3, pp. 239–246, Sep. 1960.

[11] S. Kochen and E. Specker, "The problem of hidden variables in quantum mechanics," *J. Math. Mech.*, vol. 17, pp. 59–87, 1967.

[12] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-Å. Larsson, "Kochen-specker contextuality," *Reviews of Modern Physics*, vol. 94, no. 4, p. 045007, 2022.

[13] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, "Simple test for hidden variables in spin-1 systems," *Physical review letters*, vol. 101, no. 2, p. 020403, 2008.

[14] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wieśniak, and A. Zeilinger, "Experimental non-classicality of an indivisible quantum system," *Nature*, vol. 474, no. 7352, pp. 490–493, 2011.

[15] M. Um, X. Zhang, J. Zhang, Y. Wang, S. Yangchao, D.-L. Deng, L.-M. Duan, and K. Kim, "Experimental certification of random numbers via quantum contextuality," *Scientific reports*, vol. 3, no. 1, p. 1627, 2013.

[16] M. Jerger, Y. Reshitnyk, M. Oppliger, A. Potočnik, M. Mondal, A. Wallraff, K. Goodenough, S. Wehner, K. Juliusson, N. K. Langford *et al.*, "Contextuality without nonlocality in a superconducting quantum system," *Nature communications*, vol. 7, no. 1, p. 12930, 2016.

[17] X. Zhan, P. Kurzyński, D. Kaszlikowski, K. Wang, Z. Bian, Y. Zhang, and P. Xue, "Experimental detection of information deficit in a photonic contextuality scenario," *Physical Review Letters*, vol. 119, no. 22, p. 220403, 2017.

[18] M. Malinowski, C. Zhang, F. M. Leupold, A. Cabello, J. Alonso, and J. Home, "Probing the limits of correlations in an indivisible quantum system," *Physical Review A*, vol. 98, no. 5, p. 050102, 2018.

[19] F. M. Leupold, M. Malinowski, C. Zhang, V. Negnevitsky, A. Cabello, J. Alonso, and J. P. Home, "Sustained state-independent quantum contextual correlations from a single ion," *Physical review letters*, vol. 120, no. 18, p. 180401, 2018.

[20] A. Zhang, H. Xu, J. Xie, H. Zhang, B. J. Smith, M. Kim, and L. Zhang, "Experimental test of contextuality in quantum and classical systems," *Physical review letters*, vol. 122, no. 8, p. 080401, 2019.

[21] M. Um, Q. Zhao, J. Zhang, P. Wang, Y. Wang, M. Qiao, H. Zhou, X. Ma, and K. Kim, "Randomness expansion secured by quantum contextuality," *Physical Review Applied*, vol. 13, no. 3, p. 034077, 2020.

[22] P. Wang, J. Zhang, C.-Y. Luan, M. Um, Y. Wang, M. Qiao, T. Xie, J.-N. Zhang, A. Cabello, and K. Kim, "Significant loophole-free test of kochen-specker contextuality using two species of atomic ions," *Science Advances*, vol. 8, no. 6, p. eabk1660, 2022.

[23] X.-M. Hu, Y. Xie, A. S. Arora, M.-Z. Ai, K. Bharti, J. Zhang, W. Wu, P.-X. Chen, J.-M. Cui, B.-H. Liu *et al.*, "Self-testing of a single quantum system from theory to experiment," *npj Quantum Information*, vol. 9, no. 1, p. 103, 2023.

[24] Z.-H. Liu, H.-X. Meng, Z.-P. Xu, J. Zhou, J.-L. Chen, J.-S. Xu, C.-F. Li, G.-C. Guo, and A. Cabello, "Experimental test of high-dimensional quantum contextuality based on contextuality concentration," *Physical Review Letters*, vol. 130, no. 24, p. 240202, 2023.

[25] B. W. Reichardt, F. Unger, and U. Vazirani, "A classical leash for a quantum system: command of quantum systems via rigidity of chsh games," in *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ser. ITCS '13. New York, NY, USA: Association for Computing Machinery, 2013, pp. 321–322. [Online]. Available: https://doi.org/10.1145/2422436.2422473

[26] A. Coladangelo, A. B. Grilo, S. Jeffery, and T. Vidick, "Verifier-on-a-leash: new schemes for verifiable delegated quantum computation, with quasilinear resources," in *Annual international conference on the theory and applications of cryptographic techniques*. Springer, 2019, pp. 247–277.

[27] U. Mahadev, "Classical verification of quantum computations," in *2018*

IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS). IEEE, 2018, pp. 259–267.

[28] Z. Brakerski, P. Christiano, U. Mahadev, U. Vazirani, and T. Vidick, "A cryptographic test of quantumness and certifiable randomness from a single quantum device," *J. ACM*, vol. 68, no. 5, aug 2021. [Online]. Available: https://doi.org/10.1145/3441309

[29] Z. Brakerski, V. Koppula, U. Vazirani, and T. Vidick, "Simpler Proofs of Quantumness," in *15th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2020)*, ser. Leibniz International Proceedings in Informatics (LIPIcs), S. T. Flammia, Ed., vol. 158. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020, pp. 8:1–8:14. [Online]. Available: https://drops-dev.dagstuhl.de/entities/document/10.4230/LIPIcs.TQC.2020.8

[30] G. Alagic, A. M. Childs, A. B. Grilo, and S.-H. Hung, "Non-interactive classical verification of quantum computation," in *Theory of Cryptography Conference*. Springer, 2020, pp. 153–180.

[31] G. D. Kahanamoku-Meyer, S. Choi, U. V. Vazirani, and N. Y. Yao, "Classically verifiable quantum advantage from a computational bell test," *Nature Physics*, vol. 18, no. 8, pp. 918–924, Aug. 2022. [Online]. Available: http://dx.doi.org/10.1038/s41567-022-01643-7

[32] Y. Kalai, A. Lombardi, V. Vaikuntanathan, and L. Yang, "Quantum advantage from any non-local game," in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, ser. STOC 2023. New York, NY, USA: Association for Computing Machinery, 2023, pp. 1617–1628. [Online]. Available: https://doi.org/10.1145/3564246.3585164

[33] Y. Alnawakhtha, A. Mantri, C. A. Miller, and D. Wang, "Lattice-based quantum advantage from rotated measurements," *arXiv preprint arXiv:2210.10143*, 2022.

[34] Z. Brakerski, A. Gheorghiu, G. D. Kahanamoku-Meyer, E. Porat, and T. Vidick, "Simple tests of quantumness also certify qubits," in *Annual International Cryptology Conference*. Springer, 2023, pp. 162–191.

[35] Z. Brakerski, "Quantum fhe (almost) as secure as classical," in *Advances in Cryptology - CRYPTO 2018: 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2018, Proceedings, Part III*. Berlin, Heidelberg: Springer-Verlag, 2018, pp. 67–95. [Online]. Available: https://doi.org/10.1007/978-3-319-96878-0_3

[36] U. Mahadev, "Classical homomorphic encryption for quantum circuits," *SIAM Journal on Computing*, vol. 52, no. 6, pp. FOCS18–189, 2020.

[37] O. Regev, "On lattices, learning with errors, random linear codes, and cryptography," *J. ACM*, vol. 56, no. 6, sep 2009. [Online]. Available: https://doi.org/10.1145/1568318.1568324

[38] M. Bellare and P. Rogaway, "Random oracles are practical: A paradigm for designing efficient protocols," in *Proceedings of the 1st ACM conference on Computer and Communications Security*, 1993, pp. 62–73. [Online]. Available: https://doi.org/10.1145/168588.168596

[39] A. S. Arora, K. Bharti, A. Cojocaru, and A. Coladangelo, "A computational test of quantum contextuality, and even simpler proofs of quantumness," 2024. [Online]. Available: https://arxiv.org/abs/2405.06787

[40] G. D. Kahanamoku-Meyer and N. Y. Yao, "Fast quantum integer multiplication with zero ancillas," 2024. [Online]. Available: https://arxiv.org/abs/2403.18006

[41] T. Yamakawa and M. Zhandry, "Verifiable quantum advantage without structure," in *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*. IEEE, 2022, pp. 69–74. [Online]. Available: https://doi.org/10.1109/FOCS54457.2022.00014

[42] A. Fiat and A. Shamir, "How to prove yourself: Practical solutions to identification and signature problems," in *Advances in Cryptology — CRYPTO' 86*, A. M. Odlyzko, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 1987, pp. 186–194.

[43] A. Natarajan and T. Zhang, "Bounding the quantum value of compiled nonlocal games: From chsh to bqp verification," in *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. Los Alamitos, CA, USA: IEEE Computer Society, nov 2023, pp. 1342–1348. [Online]. Available: https://doi.ieeecomputersociety.org/10.1109/FOCS57990.2023.00081

[44] Z. Brakerski, A. Gheorghiu, G. D. Kahanamoku-Meyer, E. Porat, and T. Vidick, "Simple tests of quantumness also certify qubits," in *Advances in Cryptology – CRYPTO 2023*, H. Handschuh and A. Lysyanskaya, Eds. Cham: Springer Nature Switzerland, 2023, pp. 162–191.

[45] D. Cui, G. Malavolta, A. Mehta, A. Natarajan, C. Paddock, S. Schmidt, M. Walter, and T. Zhang, "A computational tsirelson's theorem for the value of compiled XOR games," *IACR Cryptol. ePrint Arch.*, p. 348, 2024. [Online]. Available: https://eprint.iacr.org/2024/348

[46] G. D. Kahanamoku-Meyer, S. Choi, U. V. Vazirani, and N. Y. Yao, "Classically-Verifiable Quantum Advantage from a Computational Bell Test." [Online]. Available: http://arxiv.org/abs/2104.00687

[47] R. W. Spekkens, "Contextuality for preparations, transformations, and unsharp measurements," *Physical Review A*, vol. 71, no. 5, p. 052108, 2005.

[48] A. J. Leggett and A. Garg, "Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?" *Physical Review Letters*, vol. 54, no. 9, p. 857, 1985.

[49] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, "Bounding temporal quantum correlations," *Physical review letters*, vol. 111, no. 2, p. 020403, 2013.

[50] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Physical Review A*, vol. 88, no. 2, p. 022318, 2013.

[51] O. Oreshkov, F. Costa, and Č. Brukner, "Quantum correlations with no causal order," *Nature communications*, vol. 3, no. 1, p. 1092, 2012.

[52] A. A. Abbott, M. Mhalla, and P. Pocreau, "Quantum query complexity of boolean functions under indefinite causal order," *arXiv preprint arXiv:2307.10285*, 2023.

[53] A. Peres, "Incompatible results of quantum measurements," *Physics Letters A*, vol. 151, no. 3-4, pp. 107–108, 1990.

[54] N. D. Mermin, "Simple unified form for the major no-hidden-variables theorems," *Physical review letters*, vol. 65, no. 27, p. 3373, 1990.

[55] A. Cabello, "Bell's theorem without inequalities and without probabilities for two observers," *Physical review letters*, vol. 86, no. 10, p. 1911, 2001.

[56] P. K. Aravind, "Quantum mysteries revisited again," *American Journal of Physics*, vol. 72, no. 10, pp. 1303–1307, 2004.

[57] M. Kleinmann, O. Guehne, J. R. Portillo, J.-Å. Larsson, and A. Cabello, "Memory cost of quantum contextuality," *New Journal of Physics*, vol. 13, no. 11, p. 113011, 2011.

[58] A. Cabello, M. Gu, O. Gühne, and Z.-P. Xu, "Optimal classical simulation of state-independent quantum contextuality," *Physical review letters*, vol. 120, no. 13, p. 130401, 2018.

[59] K. Bharti, M. Ray, A. Varvitsiotis, N. A. Warsi, A. Cabello, and L.-C. Kwek, "Robust self-testing of quantum systems via noncontextuality inequalities," *Phys. Rev. Lett.*, vol. 122, p. 250403, Jun 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.122.250403

[60] K. Bharti, M. Ray, A. Varvitsiotis, A. Cabello, and L.-C. Kwek, "Local certification of programmable quantum devices of arbitrary high dimensionality," *arXiv preprint arXiv:1911.09448*, 2019.

[61] Z.-P. Xu, D. Saha, K. Bharti, and A. Cabello, "Certifying sets of quantum observables with any full-rank state," *Phys. Rev. Lett.*, vol. 132, p. 140201, Apr 2024. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.132.140201

[62] D. Saha, R. Santos, and R. Augusiak, "Sum-of-squares decompositions for a family of noncontextuality inequalities and self-testing of quantum devices," *Quantum*, vol. 4, p. 302, 2020.

[63] Z.-H. Liu, *Exploring Quantum Contextuality with Photons*. Springer Nature, 2023.

[64] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, ser. STOC '09. New York, NY, USA: Association for Computing Machinery, 2009, pp. 169–178. [Online]. Available: https://doi.org/10.1145/1536414.1536440