doi:10.1038/nature12035

Classical command of quantum systems

Ben W. Reichardt[1], Falk Unger[2] & Umesh Vazirani[3]

Quantum computation and cryptography both involve scenarios in which a user interacts with an imperfectly modelled or 'untrusted' system. It is therefore of fundamental and practical interest to devise tests that reveal whether the system is behaving as instructed. In 1969, Clauser, Horne, Shimony and Holt proposed an experimental test that can be passed by a quantum-mechanical system but not by a system restricted to classical physics. Here we extend this test to enable the characterization of a large quantum system. We describe a scheme that can be used to determine the initial state and to classically command the system to evolve according to desired dynamics. The bipartite system is treated as two black boxes, with no assumptions about their inner workings except that they obey quantum physics. The scheme works even if the system is explicitly designed to undermine it; any misbehaviour is detected. Among its applications, our scheme makes it possible to test whether a claimed quantum computer is truly quantum. It also advances towards a goal of quantum cryptography: namely, the use of 'untrusted' devices to establish a shared random key, with security based on the validity of quantum physics.

Do the laws of quantum mechanics place any limits on how well a classical experimentalist can characterize the state and dynamics of a large quantum system? As a thought experiment, consider that we are presented with a quantum system, together with instructions on how to control its evolution from a claimed initial state. Our goal is to determine if the system was indeed initialized as claimed, and if its state evolves as instructed.

More formally, we model the quantum system as a black box, with (for example) buttons and light bulbs to allow for classical interactions in binary. Using this limited interface, we wish to characterize the initial state of the system. We also wish to verify that on command—by pressing a suitable sequence of buttons—the system applies a chosen local Hamiltonian, or equivalently a sequence of local quantum gates, and outputs desired measurement results.

A positive answer to this fundamental question would have important consequences. First, as the power of quantum mechanics is harnessed at larger scales—with the advent of quantum computers—it will be useful to evaluate whether a quantum device in fact carries out the claimed dynamics[1,2]. Second, the goal of quantum cryptography is to create cryptographic systems with security premised on basic laws of physics. Although this seemed to have been achieved with quantum key distribution (QKD) and its security proofs[3–5], attackers have repeatedly breached the security of QKD experiments by exploiting imperfect implementations of the quantum devices[6–8]. Rather than relying on *ad hoc* countermeasures, Mayers and Yao's vision[9] of device-independent (DI) QKD, hinted at earlier in ref. 10, relaxes all modelling assumptions about the devices, and even allows for them to have been constructed by an adversary. It instead imagines giving the devices tests that cannot be passed unless they carry out the QKD protocol securely. The challenge at the heart of this vision is for an experimentalist to force untrusted quantum devices to act according to certain specifications. DIQKD has not been shown to be possible; the security proofs, first given in ref. 11, have required the unrealistic assumption that the devices have no memory between trials, or that each party has many, strictly isolated devices[12–20]. A scheme for characterizing and commanding a black-box quantum device would provide a novel approach to achieving DIQKD.

The existence of a general scheme for commanding an adversarial quantum device appears singularly implausible. For example, in an adversarial setting, experiments cannot be repeated exactly to gather statistics, because a system with memory could deliberately deceive the experimentalist. More fundamentally, as macroscopic, classical entities, our access to a quantum system is extremely limited and indirect, and the measurements we apply collapse the quantum state. Furthermore, whereas the dimension of the underlying Hilbert space scales exponentially with the number of particles or can be infinite, the information accessible via measurement grows only linearly[21]. Indeed, as formulated it is impossible to command a single black-box system. Quite simply, it is impossible to distinguish between a quantum system that evolves as desired and a device that merely simulates the desired evolution using a classical computer.

In this Article, we consider a closely related scenario. Suppose we are instead given two devices, each modelled as a black box as above and prevented from communicating with the other. In this setting, with no further assumptions, we show how to command the devices classically. That is, there is a strategy for pushing the buttons such that the answering light bulb flashes will satisfy a prescribed test only if the two devices started in a particular initial quantum state, to which they applied a desired sequence of quantum gates. Moreover, the scheme is theoretically efficient, in the sense that the total effort, measured by the number of button pushes, scales as a polynomial function of the size of the desired quantum circuit. A DIQKD scheme follows, although it is far from practical.

Detailed overview

Rigidity of the CHSH test for quantumness

The starting point for our protocol is the famous Bell experiment[22], and its subsequent 'distillation' by Clauser, Horne, Shimony and Holt[23] (CHSH). Conceptually modelled as a game (Fig. 1), it provides a test for 'quantumness', that is, a way for an experimentalist, whom we shall call Eve, to demonstrate the entanglement of two space-like separated devices, Alice and Bob. According to a Bell inequality, classical devices can win the game with probability at most 3/4. In contrast, quantum devices can win with probability $\omega^* = \cos^2(\pi/8) \approx 85.4\%$, which is optimal by Tsirelson's inequality[24].

[1] Electrical Engineering Department, University of Southern California, Los Angeles, California 90089, USA. [2] Knight Capital Group, Inc., Santa Clara, California 95054, USA. [3] Computer Science Division, University of California, Berkeley, California 94720, USA.

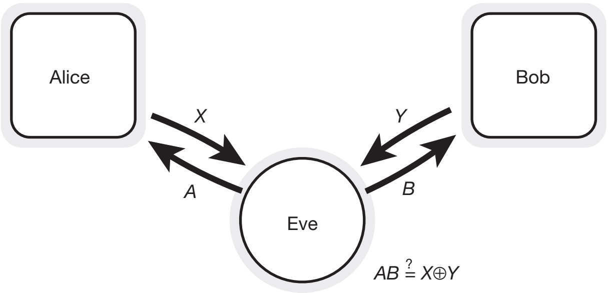

Figure 1 | Test for quantumness. In a CHSH experiment, or 'game', the experimentalist Eve sends random bits A and B to the devices Alice and Bob, respectively, who respond with bits X and Y. The devices 'win' if $AB = X \oplus Y$. Quantum devices can win with probability $\omega^* = \cos^2(\pi/8)$ if they follow an ideal CHSH strategy: on a shared EPR state $|\varphi\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, Bob measures the Pauli operator σ_z if $B = 0$ or σ_x if $B = 1$, and Alice measures $\left(\sigma_z + (-1)^A \sigma_x\right)/\sqrt{2}$.

We prove a robust converse to Tsirelson's inequality, namely a rigidity property of the CHSH game: nearly saturating Tsirelson's bound locks into place the devices' shared state and measurement operators. More precisely, if the devices win with probability $\omega^* - \varepsilon$, then they must share a state that is within a distance $O(\sqrt{\varepsilon})$ of an Einstein–Podolsky–Rosen (EPR) state, possibly in tensor product with an additional ancilla state. Moreover, their joint measurement strategy is necessarily $O(\sqrt{\varepsilon})$-close to the ideal strategy from Fig. 1 (that is, applying Alice's actual measurement operator to the shared state gets within distance $O(\sqrt{\varepsilon})$ of the result of applying her ideal measurement operator to the EPR state tensored with the ancilla; and similarly for Bob). Because each device can locate its qubit (quantum bit) share of the EPR state arbitrarily within its Hilbert space, these statements hold only up to local isometries.

A converse to Tsirelson's inequality for the CHSH game has been shown previously in the exact case[25,26]. Robustness is important for applications, however, because the success probability of a system can never be known exactly. A robust, $\varepsilon > 0$, converse statement has been shown for the game used in the original DIQKD proposal[27]. Recently, robustness has independently been shown for the CHSH game[28,29].

Scalable test for quantumness

We scale up the CHSH test for quantumness to allow us to identify many qubits' worth of entanglement. Consider a protocol in which Eve plays a long sequence of n CHSH games with Alice and Bob, and tests whether they win close to the optimal fraction ω^* of the games. Our main technical result, a multi-game rigidity theorem, establishes that if the devices pass Eve's test with high probability, then at the beginning of a randomly chosen long subsequence of n^χ games, for some constant χ, Alice and Bob must share n^χ EPR states in tensor product, which they measure one at a time using the single-game ideal CHSH operators of Fig. 1. The jth game is played using the jth EPR state, different games being entirely independent. This is a step towards the general vision outlined above, because it characterizes the initial state of many qubits and allows Eve to command the devices to perform certain single-qubit operations. Of course, we cannot hope to characterize the devices' strategies exactly, but only for a suitable notion of approximation.

The difficulty in proving this theorem is that although individual games are typically rigid, the states close to EPR states used in different games could overlap significantly. Furthermore, Alice and Bob's strategy for playing each game—including, for example, the locations of the near EPR states—could depend on the previous games' outcomes. The multi-game rigidity theorem rules out such wayward behaviour.

Verified quantum dynamics

The multi-game rigidity theorem gives strong control over the devices' measurement operators for different games. As described below, combining the CHSH game protocol with protocols for state and process tomography, and for computation by teleportation[30], gives a method for realizing arbitrary dynamics in quantum systems without making assumptions about their internal structure or operations. The dynamics are realized as the joint evolution of two isolated quantum systems, Alice and Bob, mediated by a classical experimentalist, Eve.

The problem of controlling computationally powerful but untrusted resources lies at the foundation of computer science. In the complexity class NP, for example, a polynomial-time routine—the 'verifier'—is allowed one round of interaction with an arbitrarily powerful, but malicious, 'prover'. We show that the same verifier can exploit the power of quantum mechanical provers[31]. In particular, (1) a classical verifier can efficiently simulate a quantum computer by interacting with two untrusted, polynomial-time quantum provers that share entanglement but cannot communicate between themselves. This delegated computation scheme is also 'blind', meaning that each prover learns nothing more about the computation than its length. Furthermore, (2) a classical verifier is as powerful as a quantum verifier in any interactions with multiple quantum provers (formally, the complexity classes QMIP and MIP* are equal).

Previous work introducing this problem has considered a 'semi-quantum' verifier, who manipulates a constant number of qubits while interacting with a prover[1,2,32,33]. Our work is also inspired by a proposal[34] that QMIP should equal MIP*. Although our protocol also uses computation by teleportation, it has a very different form, based on the multi-game rigidity theorem.

Product structure from repeated games

A strategy $\mathcal{S}$ for playing n sequential CHSH games specifies the initial joint state of Alice (A) and Bob (B) as well as their measurement operators for every possible situation. That is, for $D \in \{A, B\}$ and each $j = 1, \ldots, n$, $\mathcal{S}$ specifies the measurement operators used by device D in game (j, h^D_{j-1}), where h^D_{j-1} is a transcript of the device's input and output bits for the first $j - 1$ games. A strategy $\mathcal{S}$ induces a distribution on game transcripts. For two strategies to be 'close', the corresponding distributions on game transcripts should be close in total variation distance and, for almost all transcripts (drawn from either distribution), the resulting quantum states should be close in a suitable norm. We combine these conditions into one by defining for any strategy a block-diagonal density matrix that stores both the classical transcript and the resulting quantum state:

$$\rho_j = \bigoplus_{h_{j-1}} \Pr\left[h_{j-1}\right] \rho_j\left(h_{j-1}\right) \qquad (1)$$

Here $h_{j-1} = \left(h^A_{j-1}, h^B_{j-1}\right)$ is the full transcript for the first $j - 1$ games and $\rho_j(h_{j-1})$ is the state at the beginning of game j conditioned on h_{j-1}. Two strategies $\mathcal{S}$ and $\tilde{\mathcal{S}}$ are close if the associated ρ_j and $\tilde{\rho}_j$ are close in trace distance ($||\ldots||_{\mathrm{tr}}$), for every j.

Assume that for every j and almost all h_{j-1}, the devices' conditional joint strategy at the beginning of game j is 'ε-structured', meaning that the devices win with probability at least $\omega^* - \varepsilon$. Our key theorem establishes that up to local basis changes, the devices' initial state must be close to n EPR states, possibly in tensor product with an irrelevant extra state, and that their total strategy $\mathcal{S}$ must be close to an ideal strategy $\hat{\mathcal{S}}$ that plays game j using the jth EPR state. Because the structure assumption can be established by standard statistical martingale arguments on poly(n) sequential CHSH games, this implies the multi-game rigidity theorem.

The main challenge is to 'locate' the ideal strategy $\hat{\mathcal{S}}$ within Alice and Bob's Hilbert space, that is, to find an isometry on each of their spaces under which their states and measurement operators are close to ideal. However, a priori, we do not know whether $\mathcal{S}$ calls for the devices to measure actual qubits in each step, or, even if so, whether the qubits form EPR states, qubits for different games overlap each other or the locations of the qubits depend on the outcomes of previous games.

A good place to start is the construction in the single-game rigidity theorem that locates the qubits. Consider an ε-structured strategy, consisting of some shared mixed state in $\mathcal{H}_A \otimes \mathcal{H}_B$, and two-outcome projective measurements for each of Eve's possible questions. Truncate the devices' Hilbert spaces to finitely many dimensions, then decompose each space by Jordan's lemma[35] into the direct sum of two-dimensional spaces invariant under the projections. Within each such two-dimensional subspace, adjust the projections so the angle between them matches that of the ideal strategy. This defines a $\{|0\rangle, |1\rangle\}$ basis for each subspace. Aligning the subspaces according to this basis allows each Hilbert space $\mathcal{H}_D$ to be decomposed as the tensor product of a qubit and the remainder. See Supplementary Information and ref. 36 for proof details.

For multiple CHSH games, the given strategy $\mathcal{S}$ can be transformed into a nearby ideal strategy $\hat{\mathcal{S}}$ in a three-step sequence.
Step 1. Replace each device's measurement operators by the ideal operators known to exist from the CHSH rigidity theorem (for a single game). In the resulting strategy $\tilde{\mathcal{S}}$, each device D plays every game (j, h^D_{j-1}) using the ideal CHSH game operators on some qubit, up to a local change in basis. However, the basis change can depend arbitrarily on h^D_{j-1}, and the qubits for different values of j need not be in tensor product.
Step 2. In a 'multi-qubit ideal strategy', $\bar{\mathcal{S}}$, the qubits used in each game can still depend on the local transcripts but must at least lie in tensor product with the qubits from previous games. This imposes a tensor-product subsystem structure that previous DIQKD proofs have assumed. The tensor-product structure is constructed beginning with a trivial transformation on $\tilde{\mathcal{S}}$: to each device, add n ancilla qubits each in state $|0\rangle$. Next, after a qubit has been measured, say as $|\alpha_j\rangle$ in game j, swap it with the jth ancilla qubit, and then rotate this fresh qubit from $|0\rangle$ to $|\alpha_j\rangle$ and continue playing games $j+1, \ldots, n$. This defines a unitary change of basis that places the outcomes for games 1 to j in the first j ancilla qubits, and leaves the state in the original Hilbert space unchanged. At the end of the n games, undo the basis change: swap back the ancilla qubits and undo their rotations. Because qubits are set aside after being measured, the qubits for later games are automatically in tensor product with those for earlier games; the resulting strategy $\bar{\mathcal{S}}$ is multi-qubit ideal.
Step 3. We replace $\bar{\mathcal{S}}$ with an ideal strategy $\hat{\mathcal{S}}$, in which Alice and Bob each play using a fixed set of n qubits. Fix a transcript $\hat{h}_n$, chosen at random. For the first time, change the devices' initial state: replace ρ_1 with $\hat{\rho}_1$, a state having n EPR states in the locations determined by $\hat{h}_n$ in $\bar{\mathcal{S}}$. In $\hat{\mathcal{S}}$, the devices play using these EPR states, regardless of the actual transcript. This $\hat{\mathcal{S}}$ is the desired ideal strategy.

Ideal strategy $\hat{\mathcal{S}}$ is close to $\mathcal{S}$

It remains to be shown that the transformation's three steps incur a small error: $\hat{\mathcal{S}}$ is close to $\mathcal{S}$. A major theme in the analysis is to leverage the known tensor-product structure between $\mathcal{H}_A$ and $\mathcal{H}_B$ to extract a tensor-product structure within each of $\mathcal{H}_A$ and $\mathcal{H}_B$.
Step 1: $\mathcal{S} \approx \tilde{\mathcal{S}}$. Although elementary, explaining this step is useful for establishing some notation. Let ρ_1 be the devices' initial shared state, possibly entangled with the environment. Let $\mathcal{E}^A_j$ and $\mathcal{E}^B_j$ be the super-operators that implement Alice and Bob's respective strategies for game j, let $\mathcal{E}^{AB}_j = \mathcal{E}^A_j \otimes \mathcal{E}^B_j$ and let $\mathcal{E}^{AB}_{j,k} = \mathcal{E}^{AB}_k \cdots \mathcal{E}^{AB}_j$ for $j \leq k$; thus, the state ρ_j of equation (1) equals $\mathcal{E}^{AB}_{1,j-1}(\rho_1)$. For $D \in \{A, B\}$, let $\tilde{\mathcal{E}}^D_j$ be the super-operator in which the actual measurement operators in $\mathcal{E}^D_j$ are replaced with the ideal operators that follow from the CHSH rigidity theorem. Strategy $\tilde{\mathcal{S}}$ is given by ρ_1, $\left\{\tilde{\mathcal{E}}^A_j\right\}$ and $\left\{\tilde{\mathcal{E}}^B_j\right\}$. If Pr[game j is ε-structured] $\geq 1 - \delta$, then $\left\|\mathcal{E}^{AB}_j\left(\rho_j\right) - \tilde{\mathcal{E}}^{AB}_j\left(\rho_j\right)\right\|_{\mathrm{tr}} \leq 2\delta + O\left(\sqrt{\epsilon}\right)$. (This expression combines bounds on the probability of the bad event and the $O(\sqrt{\varepsilon})$ error from the good event.) To achieve the goal, namely showing that $\mathcal{E}^{AB}_{1,n}(\rho_1) \approx \tilde{\mathcal{E}}^{AB}_{1,n}(\rho_1)$ in trace distance, work backwards from game n to game 1 fixing each game's measurement operators one at a time, accumulating an error of $n(2\delta + O(\sqrt{\varepsilon}))$.
Step 2: $\tilde{\mathcal{S}} \approx \bar{\mathcal{S}}$. The key to showing that $\bar{\mathcal{S}}$ is close to $\tilde{\mathcal{S}}$ is the fact that operations on one half of an EPR state can equivalently be performed on the other half, because for any 2×2 matrix M, $(M \otimes I)(|00\rangle + |11\rangle) = (I \otimes M^T)(|00\rangle + |11\rangle)$. This means that the outcome of an ε-structured CHSH game would be nearly unchanged if Bob were hypothetically to perform Alice's measurement before his own. Once Alice's measurement operators for games $j+1$ to n are moved over to Bob's side, they cannot affect the qubit $|\alpha_j\rangle$ from game j on her side. Therefore, undoing the original change of basis restores the ancilla qubits nearly to their initial state $|0^n\rangle$, and $\tilde{\mathcal{S}} \approx \bar{\mathcal{S}}$.

In more detail, define a unitary super-operator $\mathcal{V}_j$ that rotates the jth ancilla qubit to $|\alpha_j\rangle$, depending on Alice's transcript h^A_j. Define a unitary super-operator $\mathcal{T}_j$ to apply $\mathcal{V}_j$ and swap the jth ancilla qubit with the qubit Alice uses in game j (depending on h^A_{j-1}). Alice's multi-qubit ideal strategy is given by

$$\bar{\mathcal{E}}^A_j = \mathcal{T}^{-1}_{1,j-1}\left(1_{\mathbb{C}^{2^n}} \otimes \tilde{\mathcal{E}}^A_j\right)\mathcal{T}_{1,j-1} \qquad (2)$$

We aim to show that the strategy given by ρ_1, $\left\{\bar{\mathcal{E}}^A_j\right\}$ and $\left\{\tilde{\mathcal{E}}^B_j\right\}$ is close to $\tilde{\mathcal{S}}$ up to the fixed isometry that adds $|0^n\rangle\langle 0^n|$ to the state, that is, that $|0^n\rangle\langle 0^n| \otimes \tilde{\mathcal{E}}^{AB}_{1,n}(\rho_1) \approx \bar{\mathcal{E}}^A_{1,n}\left(|0^n\rangle\langle 0^n| \otimes \tilde{\mathcal{E}}^B_{1,n}(\rho_1)\right)$. Define a super-operator $\tilde{\mathcal{F}}^{AB}_j$, in which Alice's measurements are made on Bob's Hilbert space $\mathcal{H}_B$, on the qubit determined by Bob's transcript h^B_{j-1}. Because most games are ε-structured, it follows from the CHSH rigidity theorem that $\tilde{\mathcal{F}}^{AB}_{j+1,k}\left(\tilde{\rho}_{j+1}\right) \approx \tilde{\mathcal{E}}^{AB}_{j+1,k}\left(\tilde{\rho}_{j+1}\right) = \tilde{\rho}_{k+1}$ for $j \leq k$. Because $\tilde{\mathcal{F}}^{AB}_{j+1,k}$ acts on $\mathcal{H}_B$, it does not affect Alice's qubit $|\alpha_j\rangle$ from game j at all, and so this qubit must stay near $|\alpha_j\rangle$ in $\tilde{\rho}_{k+1}$ as well; that is, the trace of the reduced density matrix against the projection $|\alpha_j\rangle\langle\alpha_j|$ stays close to one. Because this holds for every j, $\mathcal{T}^{-1}_{1,n}$ indeed returns the ancillas almost to their initial state $|0^n\rangle$. The $\left\{\tilde{\mathcal{E}}^B_j\right\}$ are symmetrically adjusted to $\left\{\bar{\mathcal{E}}^B_j\right\}$.
Step 3: $\bar{\mathcal{S}} \approx \hat{\mathcal{S}}$. In $\bar{\mathcal{S}}$, Alice and Bob play according to a strategy in which every game uses a qubit in tensor product with the previous games' qubits. However, the qubit's location can depend on previous games' outcomes. We wish to argue that Alice and Bob must play using a single set of n qubits, fixed in advance independent of the transcript.

Intuitively, if the location of Alice's jth qubit depended on h^A_{j-1}, then because the devices cannot communicate with each other, Bob could not know which of his qubits to measure. However, Alice and Bob's transcripts are significantly correlated, and we must show that they cannot use these correlations to coordinate dynamically the locations of their qubits.

For a toy example that illustrates the issue, consider two devices who play the first $n-1$ games honestly and which at the beginning of the last game share two EPR states, $|\varphi\rangle^{\otimes 2}$. Say that for certain functions f and g, Alice uses EPR state $f\left(h^A_{n-1}\right) \in \{0, 1\}$ in game n, and Bob uses EPR state $g\left(h^B_{n-1}\right) \in \{0, 1\}$. For game n to be structured, they need $f\left(h^A_{n-1}\right) = g\left(h^B_{n-1}\right)$ so that they measure the same EPR state. Now Alice and Bob's local transcripts are each uniformly random, separately, but corresponding bits have a constant correlation. To coordinate non-trivially, the best they can do is to set f and g both to the majority function[37]. Even then, though, $\Pr\left[f\left(h^A_{n-1}\right) \neq g\left(h^B_{n-1}\right)\right]$ would be too large. By considering the influences of each input bit on f and g, we can argue that the functions must be nearly constant. Thus, one of the two EPR states is used almost always.

This example gives an essentially classical cheating strategy. The actual devices may be significantly more sophisticated. In particular, small amounts of cheating in earlier games might enable an avalanche of more and more blatant cheating in later games, drastically changing the underlying quantum state. If, for example, Alice knowingly manages to swap her halves of the two last EPR states along some transcripts h^A_{n-1}, then she can use completely different strategies for the last game without having to coordinate with Bob. We control such errors, as in the arguments sketched above, by replacing Alice's superoperator with one acting on Bob's side; locality then isolates the effects of errors. More formal arguments are deferred to the Supplementary Information.

Scheme for verified quantum dynamics

Our scheme for verified quantum dynamics is based on the idea of computation by teleportation, which reduces computation to preparing certain resource states and applying Bell measurements[30] (Fig. 2f). Say that Eve wants to simulate a quantum circuit C, over the gate set $\{H, G, \mathrm{CNOT}\}$, where H is the Hadamard gate, $G=\exp\left(-i\pi\sigma_y/8\right)$ and CNOT is the controlled NOT. Eve asks Bob to prepare for Alice many copies of $|0\rangle\otimes(I\otimes H)|\varphi\rangle\otimes(I\otimes G)|\varphi\rangle\otimes\mathrm{CNOT}_{2,4}(|\varphi\rangle\otimes|\varphi\rangle)$, where $|\varphi\rangle=(|00\rangle+|11\rangle)/\sqrt{2}$. He can do so by applying one-, two- and four-qubit measurements to his halves of the shared EPR states and reporting the results to Eve. If he plays honestly, Alice's shares of the EPR states collapse into the desired resource states, up to simple corrections. Each resource state corresponds to a basic operation in C. Eve wires these up by repeatedly directing Alice to make a Bell measurement connecting the output of one operation to the input of the next operation in C. After each G gate, an H correction might be required.

Of course, Alice and Bob might not follow directions. To enforce honest play, Eve runs this protocol only a small fraction of the time, and otherwise chooses uniformly between three alternative protocols sketched in Fig. 2. Let $m=|C|^{O(1)}$ and $n=m^{O(1)}$.

Protocol 1. In the 'state tomography' protocol, Eve chooses K uniformly from $\{1, \ldots, n/m\}$. She referees $K-1$ blocks of m CHSH games. Then, in the Kth block of m games, Eve asks Bob to prepare the resource states, in a random order, while continuing to play CHSH games with Alice. Eve rejects if the tomography statistics are inconsistent; for each multi-qubit Pauli operator, the number of measurement outcomes reported by Alice should be close to its expected value for honest play. We prove that if Alice plays honestly and Eve accepts with high probability, then on most randomly chosen small subsets of the resource state positions, Alice's reduced state before her measurements is close to the correct tensor product of resource states.

Protocol 2. In the 'process tomography' protocol, Eve again chooses K uniformly from $\{1, \ldots, n/m\}$ and referees $K-1$ blocks of m CHSH games. In the Kth block of m games, Eve asks Alice to make Bell measurements on random pairs of qubits, while continuing to play CHSH games with Bob. If Alice's reported result for any pair of qubits is inconsistent with Bob's outcomes, Eve rejects. Then, if Bob plays honestly and Eve accepts with high probability, Alice must also have applied the Bell measurements honestly.

Protocol 3. In this protocol, Eve simply referees n sequential CHSH games with both devices and rejects if they do not win at least $(1-\varepsilon)\omega^*n$ games.

From Bob's perspective, the process tomography and computation protocols are indistinguishable, as are the state tomography and CHSH game protocols. From Alice's perspective, the state tomography and computation protocols are indistinguishable, as are the process tomography and CHSH game protocols. The devices must behave identically in indistinguishable protocols. The multi-game rigidity theorem therefore provides the base for a chain of implications which implies that if Eve accepts with high probability, then the devices must implement C honestly.

Four main technical problems obstruct these claims. First, in the state tomography protocol, if Bob is dishonest then Alice gets an arbitrary m-qubit state, and there is no reason why it should split into a tensor product of repeated, constant-qubit states. Nonetheless, we argue using martingales that if the counts of Alice's different measurement outcomes roughly match their expectations with high probability, then for most reported measurement outcomes from Bob and for most subsystems j, Alice's conditional state reduced to her jth subsystem is close to what it should be.

Second, saturating Tsirelson's inequality for the CHSH game implies only that Alice is honestly making Pauli σ_x and σ_z measurements on her half of an EPR state. Tomography also requires σ_y measurements. To sidestep this issue, we generalize a theory of ref. 38 and prove that there is a large class of states, including the necessary resource states, that are all robustly determined by only σ_x and σ_z measurements.

A third and bigger problem, though, is that we want to characterize the operations that each device applies to the shared EPR states, and not just the states that these operations create on the other device's

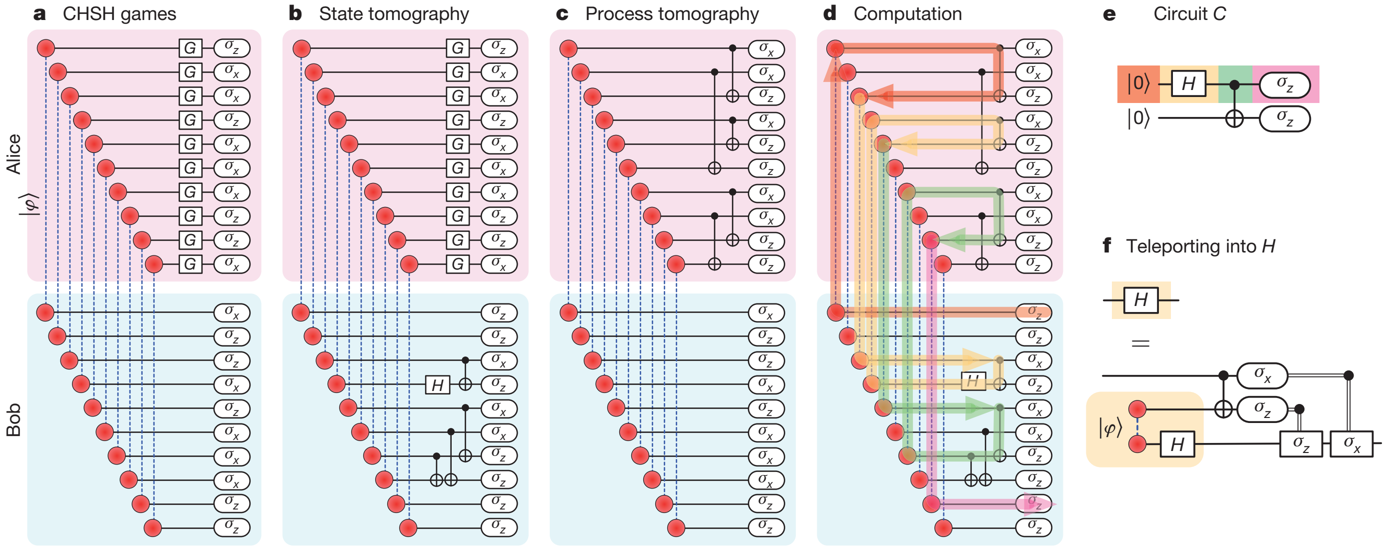

Figure 2 | Sub-protocols for verified quantum dynamics. To delegate a quantum computation, Eve runs a random one of four sub-protocols with Alice (top row, **a**–**d**) and Bob (bottom row, **a**–**d**). **a**, Playing many CHSH games ensures that the devices play honestly, measuring in each game an EPR state $|\varphi\rangle$ on two qubits (red dots). **b**, **c**, This lets Eve apply state (**b**) or process (**c**) tomography to characterize more complicated multi-qubit operations. **d**, **e**, By adaptively combining these operations (**d**), Eve directs a quantum circuit C (**e**). The operations along the zig-zagging logical path of the first qubit of C are in **d** highlighted using the same colours as in **e**. **f**, Each gate of C is implemented through teleportation; in this simpler example, H is applied by a Bell measurement on half of the resource state $(I\otimes H)|\varphi\rangle$.

side. The distinction is the same as that between process and state tomography. Essentially, the problem is that the correct states could be generated by incorrect processes. Moreover, as for sequential CHSH games, Bob's strategy in early tomography rounds might be sufficiently dishonest as to allow him in later rounds to apply completely dishonest operators. A key observation to avoid this problem is that it is enough to certify the states prepared by one device and the processes applied by the other. Then, because a broad class of states can be certified, for applications it suffices to certify a much smaller set of operations. We restrict consideration to Pauli stabilizer measurements[39]. For Pauli operators in the stabilizer of a state, the measurement outcome is deterministic. Therefore, if Alice reports the wrong stabilizer syndrome in even a single round, Eve can reject. Our process certification analysis is similar to the arguments used in step 2 of the proof of the multi-game rigidity theorem. We argue that Alice's earlier measurements cannot usually overly disturb the qubits intended for use in later measurements, by moving Alice's measurement super-operators over onto Bob's halves of the EPR states.

Finally, the verifier's questions in the state and process tomography protocols are non-adaptive, whereas in computation by teleportation the questions must be chosen adaptively on the basis of previous responses. This is an attack vector in some related protocols[36]. However, we argue that the devices can learn nothing from the adaptive questions. This follows because computation by teleportation can be implemented exactly equivalently either by choosing Bob's state preparation questions non-adaptively and Alice's process questions adaptively, or vice versa.

The proof that QMIP = MIP* follows along similar lines. Begin with a k-prover protocol with a quantum verifier. We may assume that there are two rounds of quantum messages from the provers, one before and one after the verifier broadcasts a random bit[40]. To convert to a protocol with a classical verifier, Eve, add two new provers, Alice and Bob. Eve teleports the original k provers' messages to Alice, and directs Alice and Bob together to apply the quantum verifier's acceptance predicate.

Discussion

By characterizing the device strategies that can win many successive CHSH games, we have shown how a fully classical party can direct the actions of two untrusted quantum devices. The simplest case is DIQKD, free of the independence assumptions needed in previous analyses. Following the pattern established in refs 9, 10, the QKD devices begin with shared entanglement and the two experimentalists act together as 'Eve'. They gather statistics as in the verified computation protocol to certify the devices' shared state and measurement operators, and extract secret key material from a random block of games. Two major challenges are to improve the efficiency of the scheme, to get a constant key rate instead of inverse-polynomial in n, and to tolerate a constant noise rate. More generally, the CHSH multi-game rigidity theorem may be viewed as a quantum analogue of classical multilinearity tests, which are central to the theory of probabilistically checkable proofs; by simple local tests, it guarantees the existence of a special type of large quantum state.

Received 6 September 2012; accepted 19 February 2013.

1. Aharonov, D., Ben-Or, M. & Eban, E. in *Proc. Innovations in Computer Science (ICS)* (ed. Yao, A.) 453–469 (Tsinghua Univ. Press, 2010).
2. Broadbent, A., Fitzsimons, J. F. & Kashefi, E. in *Proc. IEEE Foundations of Computer Science (FOCS)* 517–526 (IEEE Computer Society, 2009).
3. Bennett, C. H. & Brassard, G. in *Proc. IEEE Int. Conf. on Computers, Systems and Signal Processing* 175–179 (IEEE Computer Society, 1984).
4. Lo, H.-K. & Chau, H. F. Unconditional security of quantum key distribution over arbitrarily long distances. *Science* **283,** 2050–2056 (1999).
5. Shor, P. W. & Preskill, J. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.* **85,** 441–444 (2000).
6. Zhao, Y., Fung, C.-H. F., Qi, B., Chen, C. & Lo, H.-K. Quantum hacking: experimental demonstration of time-shift attack against practical quantum key distribution systems. *Phys. Rev. A* **78,** 042333 (2008).
7. Lydersen, L. *et al.* Hacking commercial quantum cryptography systems by tailored bright illumination. *Nature Photon.* **4,** 686–689 (2010).
8. Gerhardt, I. *et al.* Full-field implementation of a perfect eavesdropper on a quantum cryptography system. *Nature Commun.* **2,** 349 (2011).
9. Mayers, D. & Yao, A. in *Proc. IEEE Foundations of Computer Science (FOCS)* 503–509 (IEEE Computer Society, 1998).
10. Ekert, A. K. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.* **67,** 661–663 (1991).
11. Barrett, J., Hardy, L. & Kent, A. No signalling and quantum key distribution. *Phys. Rev. Lett.* **95,** 010503 (2005).
12. Masanes, L., Renner, R., Christandl, M., Winter, A. & Barrett, J. Unconditional security of key distribution from causality constraints. Preprint at http://arXiv.org/abs/quant-ph/0606049 (2006).
13. Acín, A., Massar, S. & Pironio, S. Efficient quantum key distribution secure against no-signalling eavesdroppers. *N. J. Phys.* **8,** 126 (2006).
14. Masanes, L. Universally composable privacy amplification from causality constraints. *Phys. Rev. Lett.* **102,** 140501 (2009).
15. Hänggi, E., Renner, R. & Wolf, S. in *Proc. EUROCRYPT* (ed. Gilbert, H.) 216–234 (LNCS 6110, Springer, 2010).
16. Acín, A. *et al.* Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.* **98,** 230501 (2007).
17. Pironio, S. *et al.* Device-independent quantum key distribution secure against collective attacks. *N. J. Phys.* **11,** 045021 (2009).
18. McKague, M. Device independent quantum key distribution secure against coherent attacks with memoryless measurement devices. *N. J. Phys.* **11,** 103037 (2009).
19. Hänggi, E. & Renner, R. Device-independent quantum key distribution with commuting measurements. Preprint at http://arXiv.org/abs/1009.1833 (2010).
20. Masanes, L., Pironio, S. & Acín, A. Secure device-independent quantum key distribution with causally independent measurement devices. *Nature Commun.* **2,** 238 (2011).
21. Holevo, A. S. Bounds for the quantity of information transmitted by a quantum communication channel. *Probl. Inf. Transm.* **9,** 177–183 (1973).
22. Bell, J. S. On the Einstein-Podolsky-Rosen paradox. *Physics* **1,** 195–200 (1964).
23. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23,** 880–884 (1969).
24. Cirel'son, B. S. Quantum generalizations of Bell's inequality. *Lett. Math. Phys.* **4,** 93–100 (1980).
25. Braunstein, S. L., Mann, A. & Revzen, M. Maximal violation of Bell inequalities for mixed states. *Phys. Rev. Lett.* **68,** 3259–3261 (1992).
26. Popescu, S. & Rohrlich, D. Which states violate Bell's inequality maximally? *Phys. Lett. A* **169,** 411–414 (1992).
27. Magniez, F., Mayers, D., Mosca, M. & Ollivier, H. in *Proc. Int. Coll. on Automata, Languages and Programming (ICALP)* (eds Bugliesi, M. *et al.*) 72–83 (LNCS 4051, Springer, 2006).
28. McKague, M., Yang, T. H. & Scarani, V. Robust self-testing of the singlet. *J. Phys. A* **45,** 455304 (2012).
29. Miller, C. & Shi, Y. Robust self-testing quantum states and binary nonlocal XOR games. Preprint at http://arXiv.org/abs/1207.1819 (2012).
30. Gottesman, D. & Chuang, I. L. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* **402,** 390–393 (1999).
31. Watrous, J. PSPACE has constant-round quantum interactive proof systems. *Theor. Comput. Sci.* **292,** 575–588 (2003).
32. Fitzsimons, J. F. & Kashefi, E. Unconditionally verifiable blind computation. Preprint at http://arXiv.org/abs/1203.5217 (2012).
33. Barz, S. *et al.* Demonstration of blind quantum computing. *Science* **335,** 303–308 (2012).
34. Broadbent, A., Fitzsimons, J. F. & Kashefi, E. QMIP = MIP*. Preprint at http://arXiv.org/abs/1004.1130 (2010).
35. Jordan, C. Essai sur la géométrie à *n* dimensions. *Bull. Soc. Math. Fr.* **3,** 103–174 (1875).
36. Reichardt, B. W., Unger, F. & Vazirani, U. A classical leash for a quantum system: command of quantum systems via rigidity of CHSH games. Preprint at http://arXiv.org/abs/1209.0448 (2012).
37. Mossel, E., O'Donnell, R. & Oleszkiewicz, K. Noise stability of functions with low influences: invariance and optimality. *Ann. Math.* **171,** 295–341 (2010).
38. McKague, M. *Quantum Information Processing with Adversarial Devices*. PhD thesis, Univ. Waterloo (2010).
39. Gottesman, D. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Inst. Technol. (1997).
40. Kempe, J., Kobayashi, H., Matsumoto, K. & Vidick, T. Using entanglement in quantum multi-prover interactive proofs. *Comput. Complex.* **18,** 273–307 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Bering, A. Broadbent, A. Chailloux, M. Christandl, R. Colbeck, T. Ito, R. König, M. McKague, V. Madhavan, R. Renner, S. Sondhi and T. Vidick for discussions. Part of this work was conducted while F.U. was at the University of California Berkeley and B.W.R. was at the Institute for Quantum Computing, University of Waterloo. B.W.R. acknowledges support from NSERC, ARO-DTO and Mitacs. U.V. acknowledges support from US NSF grant CCF-0905626 and Templeton grant 21674

Author Contributions All authors made significant contributions to the research presented in this paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.W.R. (ben.reichardt@usc.edu).