Rigidity of quantum steering and one-sided device-independent verifiable quantum computation

Rigidity of quantum steering and one-sided device-independent verifiable quantum computation

2017 New J. Phys. 19 023043

(http://iopscience.iop.org/1367-2630/19/2/023043)

IOP Publishing New J. Phys. 19 (2017) 023043 https://doi.org/10.1088/1367-2630/aa5cff

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
21 October 2016

REVISED
15 December 2016

ACCEPTED FOR PUBLICATION
27 January 2017

PUBLISHED
21 February 2017

PAPER

Rigidity of quantum steering and one-sided device-independent verifiable quantum computation

Alexandru Gheorghiu[1], **Petros Wallden**[1] **and Elham Kashefi**[1,2]

[1] School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, United Kingdom
[2] CNRS LTCI, Departement Informatique et Reseaux, Telecom ParisTech, Paris CEDEX 13, France

E-mail: a.gheorghiu@sms.ed.ac.uk

Keywords: quantum steering, self-testing, rigidity, delegated quantum computation, device independence

Abstract

The relationship between correlations and entanglement has played a major role in understanding quantum theory since the work of Einstein *et al* (1935 *Phys. Rev.* **47** 777–80). Tsirelson proved that Bell states, shared among two parties, when measured suitably, achieve the maximum non-local correlations allowed by quantum mechanics (Cirel'son 1980 *Lett. Math. Phys.* **4** 93–100). Conversely, Reichardt *et al* showed that observing the maximal correlation value over a sequence of repeated measurements, implies that the underlying quantum state is close to a tensor product of maximally entangled states and, moreover, that it is measured according to an ideal strategy (Reichardt *et al* 2013 *Nature* **496** 456–60). However, this strong rigidity result comes at a high price, requiring a large number of entangled pairs to be tested. In this paper, we present a significant improvement in terms of the overhead by instead considering quantum steering where the device of the one side is trusted. We first demonstrate a robust one-sided device-independent version of self-testing, which characterises the shared state and measurement operators of two parties up to a certain bound. We show that this bound is optimal up to constant factors and we generalise the results for the most general attacks. This leads us to a rigidity theorem for maximal steering correlations. As a key application we give a one-sided device-independent protocol for verifiable delegated quantum computation, and compare it to other existing protocols, to highlight the cost of trust assumptions. Finally, we show that under reasonable assumptions, the states shared in order to run a certain type of verification protocol must be unitarily equivalent to perfect Bell states.

1. Introduction

Quantum *steering correlations* first appeared in the seminal paper of Einstein *et al* [1] to support their argument that quantum mechanics is incomplete. It was later formally introduced by Schrödinger [4]. The observation made, was that measurements performed on one half of a bipartite entangled state can steer the state of the other half. This means that the reduced state of one side can be correlated with the classical outcome of the other party in a way that is possible only if the two parties shared entanglement (assuming the correctness of quantum mechanics). A similar effect occurs when examining purely classical correlations, which led Bell to derive his inequalities and reveal the non-local character of quantum mechanics [5]. The study of correlations that characterise quantum systems has developed much since then, and now includes the already mentioned non-local and steering correlations and quantum discord correlations [6–8]. Observing these correlations, given suitable assumptions, is an indication that a particular system behaves quantum mechanically and is used to verify the quantumness of that system. Since these correlations are specific to quantum systems, it is also anticipated that they could be the source of certain new practical applications.

In particular, the existence of *non-local correlations*, apart from revealing a counter intuitive feature of nature, has led to the development of *device-independent* protocols for quantum key distribution (QKD), quantum random number generation (QRNG) and verified delegated quantum computation (VDQC)

[3, 9–16]. In these applications two parties, Alice and Bob, do not trust their devices throughout the run of the protocol. Instead, by obtaining non-local correlations between the classical outputs of their devices, they can test their devices and obtain the correct functionality. This is achieved by confirming that the states they were sharing were entangled and measured in such a way that the resulting correlations could not have come from any classical system. Thus, the two parties either obtain a correct and secure result or detect that the devices are working incorrectly (thus being insecure) and abort. Such protocols are highly desired for practical implementation as they provide a higher level of security, unachievable by classical systems. However, there are certain practical issues that hinder their development such as the need for high detection thresholds, high fidelity transmission channels, space-like separation and a high overhead [13, 17].

The practical limitations of device-independent protocols has motivated the revival of research into quantum steering, which has simpler trust assumptions, where the state of one (trusted) side is steered by operations on the other (untrusted) side. The existing research involves the characterisation of steering correlations both analytically and geometrically [18–20], their relationship to other types of correlations [21, 22] and their application to cryptographic tasks such as QKD and QRNG [23, 24]. Experiments testing quantum steering inequalities [25] (loophole-free) and testing local but steerable states [26] have also been performed. In the case of QKD, Branciard *et al* showed in [23], that there is a natural correspondence between the trust assumptions of the protocol and the types of correlations between the two parties. Using this correspondence, Branciard *et al* introduced *one-sided device-independent* QKD, which uses steering correlations in order to distil a shared secret key. In the cryptographic setting, such correlations allow only one device to be untrusted leading to a reduction in the overall experimental requirements of the protocol. To be precise, they showed that in typical device-independent settings, the detection efficiency of Alice and Bob should be above 91.1%, whereas their one-sided protocol lowers that to 65.9%. A similar relation between trust assumptions and correlations is exploited for QRNG as well [24]. In this case it was shown that a detection efficiency of 50% is sufficient for random number generation in the steering setting, versus 70.7% in the device-independent setting.

For VDQC, we will show that using steering correlations leads to a reduction in communication, compared to the device independent setting. However, the VDQC case presents a complication. Firstly, the setting of VDQC is slightly different than that of QKD and QRNG. In the latter two, Alice and Bob are two parties that are working together towards a common objective (obtaining a shared key or certified randomness), using possibly untrusted quantum devices. In VDQC, Alice is a client who is delegating a difficult computation to Bob, an untrusted quantum server. She does not have the resources to perform the quantum computation herself, and while Bob does, he cannot be trusted to do so. So Alice needs a way to verify that Bob is performing the correct quantum computation. To do this, she utilises a quantum device in her local lab which she may or may not trust. Regardless of this, Bob is always assumed to be unstrusted. This is an important distinction from the collaborative setting of QKD and QRNG. Moreover, approaches to QKD and QRNG rely on using a bound on the correlations of the parties' devices in order to derive a bound for a quantity of interest, such as key rate, mutual information, entropy etc [15, 27–31]. In contrast to this, existing protocols for device-independent VDQC use the bound on correlations to recover the underlying quantum state used in the protocol, as well as the operations being performed on this quantum state [9–12]. This allows for the correctness certification (referred to here as verification) of an arbitrary universal quantum computation strictly from the non-local correlations. To achieve this one needs, at the same time, to obtain close to *maximal* non-local correlations *and* not only recover a characterisation of one Bell pair but of a tensor product of Bell pairs. Such a result is possible due to the rigidity of repeated CHSH games, manifesting non-local correlations, as shown in [9]. In this context, rigidity means that if two non-communicating devices playing CHSH games are achieving the optimal win rate, then they share a state which is close to a tensor product of maximally entangled states and, moreover, their strategy is fixed and uniquely determined. In general, we define rigidity as the ability to derive a robust bound on the distance between a real state (the state shared by Alice and Bob) and many copies of some target state (such as a tensor product of Bell pairs) as well as between target measurements and real measurements, from a bound on correlations. This is similar to the concept of *self-testing*, except that for self-testing one only wants to obtain a single copy of the target state from observed correlations. For instance, it is possible to self-test a Bell state and its associated measurements using the CHSH game. So, one way to arrive at rigidity is to consider multiple CHSH games and combine the self-testing results[3]. However, in general, it is necessary to play a large number of games in order to certify few Bell states. For this reason the existing device-independent VDQC protocols have impractically large communication complexity [9, 10, 12].

There is an additional aspect to be mentioned. As stated, in the VDQC setting we have a *trusted* client, also referred to as *verifier*, and an *untrusted* server, also referred to as *prover*. This is the standard cryptographic scenario when considering verification of computation, whether it is quantum or classical [32, 33]. The asymmetry in trust is precisely the same as in the steering scenario which is why VDQC is the most natural

[3] This needs to be done in a non-trivial way since the games might not be independent from each other, in general.

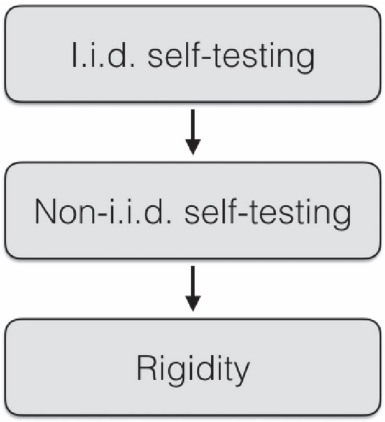

Figure 1. Steps towards rigidity.

application for steering correlations. While it is definitely possible to introduce such an asymmetry in QKD, for example, the typical setting is to have the parties involved be identical in all respects.

In this paper, we relax the trust assumptions and derive a rigidity result for quantum steering correlations, showing that it leads to a reduced overhead compared to the setting of non-local correlations. The result proves that observing steering correlations close to their maximal value, implies a tensor product structure of Bell pairs and fixed measurement operators for the untrusted party, up to local isometry. The only assumptions made in deriving this result are the correctness of quantum mechanics and the fact that one party is completely trusted, having a complete characterisation of her Hilbert space and measurement operators. In particular, the rigidity result makes no independent and identically distributed (i.i.d.) assumption regarding the shared state or strategy of the untrusted party. In analogy to one-sided device-independent QKD and QRNG, this leads us to a one-sided device-independent VDQC protocol, having improved round complexity over the device-independent versions. More generally, the rigidity result is relevant in its own right since it is applicable to any protocol that uses steering correlations.

The structure of this paper is organised as follows. In section 2 we give our main result, which is that using maximal steering correlations we determine, up to a local isometry, a tensor product of Bell pairs and the operations of the untrusted party. To derive this result, we first give a procedure to characterise a single Bell pair from observed correlations (single-shot rigidity), in the i.i.d. setting in section 2.1. This is done by making a protocol for self-testing from steering correlations which gives us a bound on the distance of one shared state from a perfect Bell pair. We also prove the optimality of this bound. We then remove the i.i.d. assumption in section 2.2, thus showing that one can extract a Bell pair from the observed statistics even in the fully adversarial (one-sided) setting. Then, in section 2.3 we use the previous result in order to determine a tensor product structure of Bell pairs and the measurement operators of the untrusted party. These steps are shown in figure 1. Note that rigidity does not follow directly from repeatedly applying the single-shot result as this would require independence. Instead, we use an approach similar to that of [9], by defining a quantum steering game and showing that high win rates in this game determine the states and strategies of the players, up to local isometry.

Lastly, in section 3 we briefly discuss existing approaches to VDQC (section 3.1) and then use the rigidity of quantum steering to construct a one-sided device-independent verification protocol (section 3.2). We also show that for the types of protocols we have considered, the required entangled states should be close to Bell pairs (section 3.3).

2. State and strategy certification via steering

While quantum steering has been studied extensively in the context of verifying entanglement, it is important to elaborate on the subtle difference between verifying entanglement and verifying maximal entanglement and how this relates to the verification of quantum computation. It has already been shown that it is possible to verify, from steering correlations, that a state shared by two parties is entangled. In fact, this type of verification can be done in a fully device-independent way (under certain assumptions), and has been tested experimentally [34, 35]. However, it should be noted that these results use steering correlations as a witness for quantum entanglement. The purpose of a witness is to separate between entangled and non-entangled states and its existence is proven through the violation of a steering inequality, in analogy to a Bell inequality. In our setting, we do not just require correlations that violate a steering inequality, rather we require the correlations to saturate

the maximum achievable value. This enables us to certify the state and strategies used in producing these correlations, which in turn, can be used for verifying quantum computations. To our knowledge, there are no other results which treat the case of saturated steering correlations.

The setting that we will consider is similar to the one used in [21, 23]. This involves two parties, Alice and Bob, where Alice has a trusted measurement device, while Bob has an untrusted measurement device. They share an unknown joint quantum state, $|\psi\rangle$, which without loss of generality can be assumed to be pure. Alice instructs Bob to perform a measurement on their joint state. For example, if the shared state is $|\phi_+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, Alice can instruct Bob to measure the Z observable on his qubit and report the outcome. The measurement steers Alice's qubit to a particular quantum state. She can then measure her state to confirm that her qubit was indeed steered to the expected state. We shall refer to a collection of such measurements as a *steering game* consisting of rounds of single measurements performed by Alice and Bob. They win a round of the game if Bob's reported outcome matches that of Alice. We elaborate more on this in section 2.3.

The correlations between their outcomes are called steering correlations if they cannot be explained by a *local hidden state model*. This happens when the expected value for the correlation of their outcomes obeys certain steering inequalities. However, whenever a steering inequality violation is observed, Alice can conclude that her state was indeed steered by Bob via their shared entanglement. In the rest of this section we prove that violating the steering inequalities maximally (up to order $O(\epsilon)$) leads to recovering a tensor product of Bell pairs with measurement operators close to ideal and we quantify this 'closeness'. Proving this rigidity result is done as follows. First, building on the work regarding self-testing the singlet [36] by McKague *et al*, we derive a robust self-testing protocol of a single Bell state, where one side is trusted, while the other is not. This is achieved by using maximal steering correlations in order to fully characterise, up to a bound of order $O(\sqrt{\epsilon})$, the quantum state shared between the two parties. Importantly, in [36] it is assumed that the quantum states are i.i.d. We also make this assumption in our self-testing result (section 2.1), however we remove it later on (section 2.2). We show that the bound for self-testing is tight up to constant factors. Then, we remove the i.i.d. assumption, arriving at a new bound for characterising the shared state in the completely adversarial setting, from the observed correlations. The way in which we remove the i.i.d. assumption is not specific to steering and can be applied to the non-local setting as well, thus complementing the work from [36]. Using this result, and a game-based argument we show that saturating the steering inequalities enables us to recover the quantum state of a tensor products of Bell pairs and characterise the untrusted measurement operators acting on these states, thus proving the rigidity of quantum steering (section 2.3). Throughout this paper we use $|||\psi\rangle|| = \sqrt{\langle\psi|\psi\rangle}$ as the l^2-norm and $\mathrm{TD}(\rho, \sigma) = \frac{1}{2}\mathrm{Tr}(|\rho - \sigma|)$ as trace distance. Additionally, for the trace distance of pure states we will write $\mathrm{TD}(|\psi\rangle, |\phi\rangle) = \mathrm{TD}(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|)$.

2.1. One-sided device-independent self-testing (i.i.d.)

We start by proving a one-sided device-independent version of robust self-testing, whereby the measurement statistics allow us to determine the existence of a single Bell pair. Specifically, as in [36], successfully self-testing a maximally entangled state between Alice and Bob means:

- There exist local bases (local isometry) in which their shared state can be viewed as a Bell pair, possibly in tensor product with some additional state.
- We can infer the existence of local (physical) observables on Alice and Bob's side, which act non-trivially on the shared state.

This is similar to the works of [11, 36–42] and in fact we adopt a similar notation to that of [36]. The main difference with respect to those works, is that in our case we trust Alice's measurements. The specific observables we consider for her are Pauli X and Y. We will find that Bob must also measure these observables to saturate the correlations of measurement outcomes. The Bell state under consideration is the XY-determined Bell state $|\psi_+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. The result can be generalised for any pair of non-commuting observables and any Bell state. It should be mentioned that a Bell state has a local hidden variable model for Pauli basis measurements by both parties, but it does not have a local hidden state model. This highlights the difference between non-local and steering correlations and emphasises the importance of trusting Alice's system in order characterise the shared state and Bob's measurements. The reason for our choice of observables and state is so that we can easily use this result for a VDQC protocol, presented in the next section, that uses XY-plane states, as explained in section 3. Note that Alice and Bob need to perform multiple measurements in order to approximate the expectation values of their observables. Since we do not trust Bob, we cannot in general assume the independence of his measurements. However, we will make an i.i.d. assumption in the beginning, to prove our main self-testing theorem. We then remove the assumption by modelling the measurement process as a martingale and

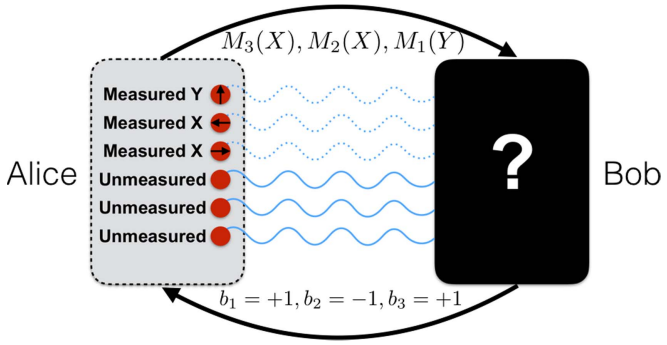

Figure 2. Setting for i.i.d. self-testing.

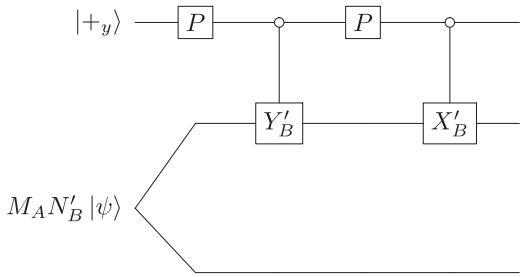

Figure 3. Local isometry Φ.

using the Azuma–Hoeffding inequality [43, 44], as is also done in [11, 15]. A schematic illustration of our setting can be found in figure 2.

We start by proving a theorem analogous to theorem 1 of [36]. Just as in [36], the primed observables denote untrusted operators. The shared state, which is also assumed to be untrusted is denoted as $|\psi\rangle$. The independence (i.i.d.) assumption means that Alice and Bob share the same state $|\psi\rangle$ in each round of measurements. Furthermore, the untrusted measurement observables of Bob can be assumed to be the same each time. This is because his observables can include action on his private ancilla. Since Alice's side is trusted, we assume that she has a single qubit measurement device and so, in each round of measurement, her part of $|\psi\rangle$ is a single qubit state. Given this setting, the theorem is stated as follows:

Theorem 1. *Suppose that from the observed correlations of measurements performed by Alice and Bob and knowing that Alice is measuring the $\{X, Y\}$ observables (denoted X_A, Y_A), one can deduce the existence of local observables $\{X'_B, Y'_B\}$ on Bob's side, with eigenvalues ± 1, which act on a bipartite state $|\psi\rangle$ such that:*

$$||(X_A - X'_B)|\psi\rangle|| \leqslant \gamma_1, \tag{1}$$

$$||(Y_A - Y'_B)|\psi\rangle|| \leqslant \gamma_1, \tag{2}$$

$$||(X'_B Y'_B + Y'_B X'_B)|\psi\rangle|| \leqslant \gamma_2. \tag{3}$$

Then there exists a local isometry $\Phi = I \otimes \Phi_B$ and a state $|\text{junk}\rangle_B$ such that

$$||\Phi(M_A N'_B\ |\psi\rangle) - |\text{junk}\rangle_B M_A N_B\ |\psi_+\rangle_{AB}|| \leqslant \varepsilon \tag{4}$$

with M_A, $N_B \in \{I, X, Y\}$, $N'_B \in \{I, X'_B, Y'_B\}$, $\varepsilon = 3\gamma_1 + \gamma_1^2/4 + 2\gamma_2$, $|\psi_+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$.

Proof sketch. The proof relies on finding an isometry which, given conditions (1)–(3), maps $|\psi\rangle$ to an almost perfect Bell state. Similar to [36], M_A and N'_B are the physical observables of Alice and Bob which act on the shared state. The isometry we considered is illustrated in figure 3, where $P = \frac{1}{\sqrt{2}}(X + Y)$ and the control gates act on the target state when the control qubit is in the $|-_y\rangle$ state instead of the $|1\rangle$ state, and act as identity when the control is in the $|+_y\rangle$ state instead of the $|0\rangle$ state. Here, $|+_y\rangle$ and $|-_y\rangle$ are the two eigenstates for the Pauli Y operator, corresponding to the $+1$ and -1 eigenvalues, respectively. The fact that we are using these states and the P operator is a consequence of shifting everything to the XY-plane of the Bloch sphere instead of the more familiar XZ-plane. It should be noted that M_A is trusted and acts on Alice's part of the shared state, whereas N'_B, acting on Bob's part of $|\psi\rangle$, is untrusted. However, the action of N'_B is equivalent to the honest N_B acting on the

ancilla introduced by Φ. Having the isometry, we write out its action on the state $M_A N'_B \ |\psi\rangle$ and use inequalities (1)–(3) together with the trace preserving properties of the operators and triangle inequalities to prove condition (4). The full proof of theorem 1 can be found in appendix A. □

Our next result is to show that conditions (1)–(3) are satisfied if an almost maximal violation of a particular steering inequality occurs. As mentioned before, the requirement for maximal violation is in contrast to previous work on entanglement detection. In that case, one uses the steering inequality as an entanglement witness to separate the Hilbert space of possible shared states into a subspace of entangled states manifesting steering correlations and its complement. Violating the inequality determines that the shared state lies in the subspace of entangled steerable states. For example, similar to the works of [45, 46], assuming Bob measures local observables X'_B and Y'_B, one could consider the inequality:

$$|\langle\psi| \ X_A X'_B + Y_A Y'_B \ |\psi\rangle| \leqslant \sqrt{2}\,.$$

This inequality holds, whenever there is a local hidden state model for Bob's system. If this is not the case, then the state is steerable. In our case, we do not simply require a violation of this inequality, but we require a (close to) maximal violation. For the above inequality, the maximum achievable violation allowed by quantum mechanics is 2, which also corresponds to its mathematical maximum. Requiring maximal saturation is represented as:

$$|\langle\psi| \ X_A X'_B + Y_A Y'_B \ |\psi\rangle| \geqslant 2 - \epsilon$$

for some small ϵ. Given that this inequality holds we have:

Theorem 2. *Suppose Alice measures the observables X_A, Y_A and that Bob measures the observables X'_B and Y'_B with eigenvalues ± 1, on the state $|\psi\rangle$, such that*

$$|\langle\psi|(X_A X'_B + Y_A Y'_B)|\psi\rangle| \geqslant 2 - \epsilon,$$

where $0 < \epsilon < 1$. Then the conditions of theorem 1 are satisfied with $\gamma_1 = \sqrt{2\epsilon}$ and $\gamma_2 = 4\sqrt{\epsilon}$.

Proof sketch. The proof reduces to expanding inequality 2 and using the properties of the observables, to arrive at the bounds of conditions (1)–(3) from the previous theorem. Concretely, we see that the correlation of local observables that we consider:

$$|\langle\psi|(X_A X'_B + Y_A Y'_B)|\psi\rangle|$$

is simply a sum of two expectation values which are upper bounded by unity (because the observables have ± 1 eigenvalues). Hence, to saturate the absolute value of this quantity, it must be the case that both expectation values are saturated i.e. lower bounded by $1 - \epsilon$ or upper bounded by $-1 + \epsilon$. We will only examine the first case since the second is analogous, so we will drop the absolute value of the expression and simply consider:

$$\langle\psi| \ X_A X'_B + Y_A Y'_B \ |\psi\rangle \geqslant 2 - \epsilon.$$

By expressing each expectation as a trace norm we arrive at conditions (1) and (2). To prove condition (3) we use the Cauchy–Schwarz inequality and the commutators $[X_A, \ Y_A]$ and $[X'_B, \ Y'_B]$, respectively. The full proof can be found in appendix B. □

Using the results of theorems 1 and 2, we find that strong correlations between Alice and Bob's measurement outcomes, given that we trust Alice to be measuring the X and Y observables, determine the shared state between Alice and Bob as a Bell state, under local isometry $\Phi = I_A \otimes \Phi_B$. Additionally, notice that if the steering correlations are $O(\epsilon)$ close to ideal (maximal), we can bound the shared state of Alice and Bob as being $O(\sqrt{\epsilon})$ close to the ideal. The same asymptotic bound is achieved in the case of CHSH games, where both Alice and Bob's outcomes are untrusted. One could expect that the bound we obtained from the steering inequalities is not a tight bound and that it is possible to do better because we trust Alice's measurements. We prove, that in fact this is not the case and the $O(\sqrt{\epsilon})$ bound is actually tight:

Theorem 3. *Suppose that Bob's observables X'_B and Y'_B with eigenvalues ± 1, acting on a state $|\psi\rangle$, are such that:*

$$|\langle\psi|(X_A X'_B + Y_A Y'_B)|\psi\rangle| \geqslant 2 - \epsilon,$$

where $0 < \epsilon < 1$. Then, up to constant factors, the bound of theorem 1 (i.e. inequality (4) with $\varepsilon = O(\sqrt{\epsilon})$) is tight.

Proof sketch. Theorem 3 essentially shows that the $O(\sqrt{\epsilon})$ bound for the closeness of $|\psi\rangle$ to the ideal Bell state $|\psi_+\rangle$ is optimal. The proof relies on finding a state and local observables for Bob such that the steering correlation is saturated up to order $O(\epsilon)$, but the state is $O(\sqrt{\epsilon})$ deviated from the ideal Bell state. We consider such a state which is exactly $\sqrt{\epsilon}$-close to the ideal $|\psi_+\rangle$ Bell state and which saturates the correlation of observables to $2 - \epsilon$. Furthermore, we consider local observables for Bob which are deviated, as a function of ϵ, from the ideal observables. Hence, these deviated observables tend to the ideal X and Y observables in the limit where $\epsilon \to 0$. The specific state and Bob's local observables are given in appendix C. □

The exact bound for closeness can be computed by simply inserting the constants γ_1 and γ_2 from theorem 2 in the calculation for ε of theorem 1. This yields a distance of $\varepsilon = (3\sqrt{2} + 8)\sqrt{\epsilon} + \epsilon/2$.

This result has important consequences for the application we consider in the next section, the verification of quantum computation. It essentially imposes a restriction on how good the fidelity of the entangled state is, as resulting from the observed correlations. In VDQC we require a complete characterisation of the entangled states that are used and in particular, we require them to be close to Bell pairs. This leads to the requirement of a very tight saturation of the steering correlations. Contrast this to the QKD or QRNG settings where less than maximally entangled states suffice [27, 28, 47–49].

It should be noted that we used the same notation for Bob's observables in both theorems 2 and 1 since for this particular case they coincide. In general, however, we would have to construct Bob's observables in theorem 1 from the observed correlations, as is done in [36]. An important corollary to these three theorems is the following:

Corollary 1. *The results of theorems 1–3 hold if instead of X and Y, Alice measures the anti-commuting single-qubit observables A_0 and A_1, having eigenvalues ± 1 and Bob measures observables B_0, B_1 having eigenvalues ± 1.*

Proof. In the proof of theorem 1 we only made use of the anti-commutation properties of the X, Y observables on Alice's side as well as the action of the two operators on the eigenstates of Y. For general observables, A_0 and A_1 this translates to using their anti-commutation properties and the action of the two on the eigenstates of A_1, for example. Essentially the proof of theorem 1 only changes by relabelling X_A as A_0 and Y_A as A_1. On Bob's side, the situation is similar. By relabelling X'_B as B_0 and Y'_B as B_1 we again have conditions (1)–(3) for these observables, which are then used to construct the isometry and prove the result of theorem 1.

For theorem 2, using the same relabelling we have the inequality:

$$|\langle\psi|\ A_0B_0 + A_1B_1\ |\psi\rangle| \geqslant 2 - \epsilon.$$

Since the proof of theorem 2, like theorem 1, only relied on the anti-commutation properties and the action of the observables on the eigenstates of one of them, the relabelling does not change anything and the results go through as before.

Lastly, for theorem 3 we use that fact that the observables A_0 and A_1 are linear combinations of Pauli matrices. Depending on which Bell state is stabilised by the actions of these observables we can consider a state that is $O(\sqrt{\epsilon})$-deviated from that Bell state, similar to the state considered for the proof of theorem 3. Analogously, we will have $O(\epsilon)$-deviated observables for each Pauli matrix. For example, in the proof of theorem 3 we considered the $O(\epsilon)$-deviated versions of the X and Y observables. Bob's observables will therefore be linear combinations of these $O(\epsilon)$-deviated Paulis so as to be $O(\epsilon)$-close to A_0 and A_1, respectively. □

The result of theorem 2 assumes ideal expectation values for the observables of Alice and Bob. Of course, in practice, after performing a finite number of measurements we obtain an approximation of these expectation values. This can be properly taken into account by considering independent random variables (since we are in the i.i.d. setting) associated with the measurement process and using a Chernoff inequality to bound their expectation values. We do not give a full derivation of this here, since we will give the more general derivation for the non-i.i.d. case in the next section (for which the proof can be found in appendix D). Instead, we simply state the result of this finite analysis: for a fixed $\epsilon > 0$, we require at least $(1/\epsilon^2)\log(1/\epsilon)$ measurements in order to certify that the closeness of each shared state is $O(\sqrt{\epsilon})$ to a perfect Bell pair. One can also compute the number of measurements as a function of the desired trace distance for the Bell states. If we denote this distance as $D = c\sqrt{\epsilon}$, then the number of measurements must be at least $(2c^4/D^4)\log(c/D)$. In our case $c \approx 12.3$, so that if we wanted the trace distance to be, for example, $D = 0.1$, we would require at least 2.2×10^9 measurements.

2.2. Removing the independence assumption

We proceed to remove the i.i.d. assumption from the previous statements. The following theorem essentially states that if Alice and Bob are asked to perform a sequence of measurements, and we notice a close to maximal steering inequality violation from their outcomes, we can conclude that the state shared in a typical round of measurement is close to a Bell pair. By 'typical round' we mean a uniformly randomly chosen round. A similar result is obtained in [50], with the essential differences that their non-i.i.d. result shows that at least one state is close to an ideal Bell pair while both parties are untrusted. In our case, Alice is trusted throughout this process and, without loss of generality, we can assume that she chooses the measurement settings for each round. The notation $\mathrm{Tr}_{-i}(\cdot)$ indicates that we are tracing out everything apart from the quantum states that are measured in round i. We also use the notation $\mathrm{Tr}_{-R}(\cdot)$, which generalises the previous notation for a set, R, of rounds (i.e. tracing out all states except those which are used in rounds $i \in R$).

Theorem 4. *Suppose Alice and Bob are required to perform K rounds of measurement and also that:*

- *Prior to the measurements, the shared state of Alice and Bob is assumed to be σ, which can be either pure or mixed. The state σ is thus the global state which will be used for all K rounds*[4].
- *Alice chooses a random set of size $K/2$, consisting of distinct indices from 1 to K and denoted $R_0 = \{i | i \in_R \{1 \ldots K\}\}$, $|R_0| = K/2$. We also denote $R_1 = \{1 \ldots K\} \setminus R_0$, to be the complement of R_0.*
- *We denote $\rho_i = \mathrm{Tr}_{-i}(\mathcal{E}^{AB}_{1,i-1}(\sigma))$ the reduced state of Alice and Bob in round i, and:*

$$\rho_{\mathrm{avg}} = \frac{1}{K}\sum_{i=1}^{K}\rho_i$$

as the averaged state. Here $\mathcal{E}^{AB}_{1,i-1}$ denotes the action (measurements) of Alice and Bob on the state sigma *up to round i.*

- *In round i, let $r_i = 0$ iff $i \in R_0$, otherwise $r_i = 1$. Alice measures the observable A_{r_i} on her half of ρ_i. A_0 and A_1 are anti-commuting single-qubit observables having ± 1 eigenvalues.*
- *In round i, Bob is asked to measure B_{r_i}. B_0 and B_1 have ± 1 eigenvalues.*
- *We denote a_i and b_i, respectively, as the outcomes of their measurements in round i. We also denote $\hat{C}_i = a_i b_i$ as their correlation for round i.*
- *We denote $\hat{C}^0 = \frac{1}{K/2}\sum_{i\in R_0}\hat{C}_i$ and $\hat{C}^1 = \frac{1}{K/2}\sum_{i\in R_1}\hat{C}_i$ as the averaged correlations for the cases where both Alice and Bob are asked to measure the first observable, or both are asked to measure the second, respectively.*

If, for some given $\epsilon > 0$ and suitably chosen $K = \Omega((1/\epsilon^2)\log(1/\epsilon))$, it is the case that $\hat{C}^0 + \hat{C}^1 \geqslant 2 - \epsilon$ (or, alternatively, $\hat{C}^0 + \hat{C}^1 \leqslant -2 + \epsilon$) then there exists a local isometry Φ such that, for a randomly chosen ρ_i, with probability at least $1 - O(\epsilon^{1/6})$:

$$\mathrm{TD}(\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_i))),\ \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)) \leqslant O(\epsilon^{1/6}), \tag{5}$$

where $\mathcal{E}^{AB}$ is some combination of the A_0, A_1, B_0, B_1 operators and $\hat{\mathcal{E}}^{AB}$ is the analogous combination of the ideal operators (i.e. A_0, A_1, I and the ideal operators for Bob, which for the XY-plane case we considered, are I, X, Y), as in theorem 1, and where junk *is Bob's private system apart from the ancilla introduced by Φ. Alternatively, we have that there exists some state $\tilde{\rho}_{\mathrm{junk}}$ such that:*

$$\mathrm{TD}(\Phi(\mathcal{E}^{AB}(\rho_i)),\ \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)\tilde{\rho}_{\mathrm{junk}}) \leqslant O(\epsilon^{1/12}). \tag{6}$$

Proof sketch. The proof is broken down into several parts. First, we show that the average observed correlations $\hat{C}^0$ and $\hat{C}^1$ approximates the ideal quantum correlation for the averaged state. The averaged state can be thought of as the state shared by Alice and Bob in each round of measurements, such that the average correlations of outcomes from this state match those observed in the real experiment (i.e. $\hat{C}^0$ and $\hat{C}^1$). Proving this step is done along similar lines to the approaches of [11, 15]. The measurement process of Alice and Bob can be viewed as a stochastic process with bounded increment, i.e. a martingale. The specific martingale we consider encodes the correlations of their measurement outcomes. While the individual measurements need not be independent, we can still prove that this observed correlation is, with high probability, close to the ideal quantum correlation. To do this, we use the Azuma–Hoeffding inequality for martingales [43, 44]. We then use the result of theorem 2 to show the closeness of the averaged state to an ideal Bell state. Lastly, we prove equation (5) by using an optimisation argument together with properties of trace distance and density matrices. Equation (6) follows from this through an application of the gentle measurement lemma [9]. The full proof is given in appendix D. □

As in the i.i.d case, for a fixed $\epsilon > 0$, we require $\Omega((1/\epsilon^2)\log(1/\epsilon))$ measurements however the closeness of a typical state, to a perfect Bell pair, is of order $O(\epsilon^{1/6})$ in this case. For a better comparison we will also compute the number of measurements as a function of the desired trace distance for a typical state. The proofs in appendix D show that the exact number of measurements required is $(8/\epsilon^2)\log(1/\epsilon)$, yielding a distance $D = c^{1/3}\epsilon^{1/6}$, where c is the constant from the i.i.d. bound. Thus, the number of measurements must be at least $(8c^4/D^{12})\log(c^2/D^6)$. For our case, where $c \approx 12.3$ if we again take $D = 0.1$, we would require at least 3.4×10^{18} measurements. While this is too great for current experimental applications, the current bounds are most likely not tight and can be improved. In fact, better numeric bounds have been obtained for the i.i.d. setting, as we

[4] In the ideal setting where everything is trusted, σ would be a $2K$-qubit state consisting of K Bell pairs.

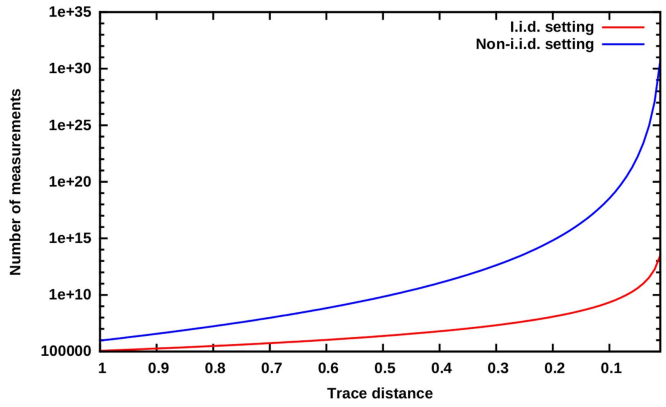

Figure 4. Comparison of i.i.d. and non-i.i.d. settings.

explain in section 2.4. To give an example, the result of [51], obtains a bound where $c = 1.19$ leading to 2.2×10^{14} measurements. We show a comparative plot for the number of measurements required in the i.i.d. setting, versus in the non-i.i.d. setting in figure 4. The graphs are represented as functions of the desired (decreasing) trace distance.

It should be noted that in the proof of theorem 4, we did not use the fact that Alice is trusted except when applying the self-testing result (theorem 2). Thus, a similar theorem can be proven in the case where both Alice and Bob are untrusted. In that case, one could simply use the self-testing results of [12, 36, 41, 42] for the i.i.d. setting, and then obtain a statement about the closeness of a typical state to the ideal one in the non-i.i.d. setting using our techniques. For example, if we were to use theorem 2 from [36] we could once again establish from the measurement statistics that a typical state shared by Alice and Bob is close to a Bell state. This result completes the work of [36] for the non-i.i.d. setting.

It should additionally be noted that throughout this section we not only assumed that Alice's device is trusted but that it also measures the ideal Pauli operators. This could seem unreasonable from an experimental perspective, however note that any (fixed) deviation on Alice's measurement operator can be incorporated into ϵ. In other words, assume Alice's ideal operator is A and the deviated one is δA, such that:

$$\mathrm{TD}(A \otimes B\ |\psi\rangle,\ \delta A \otimes B\ |\psi\rangle) < \delta.$$

It is thus the case that the action of Alice's operators is δ-close to the action of the ideal operators which produce ϵ saturation. Hence, δ can be added to ϵ and viewed as a contribution to the total variation from maximal correlations. However, if such a deviation exists we should consider what happens when $\delta \leqslant \epsilon$ and when $\delta > \epsilon$, respectively. If $\delta \leqslant \epsilon$, then the error on Alice's device is smaller than the precision to which we wish to estimate the saturation of the correlations. Therefore, the saturation can still be considered of order $O(\epsilon)$ and the bounds on the states follow as in the ideal case. However, if $\delta > \epsilon$ then the saturation cannot be estimated within the desired precision. This means that there will be an intrinsic limitation on the determined closeness of the shared states as given by a saturation of order $O(\delta)$.

2.3. Rigidity of quantum steering

We now proceed to prove rigidity of quantum steering games in a manner similar to that of [9]. In this setting, we assume that there is an unknown quantum state shared between Alice and Bob. By asking Alice and Bob to perform repeated measurements we would like to show that this state is close to a tensor product of Bell pairs possibly in tensor product with some additional state, up to local isometry. Additionally, we will show that Bob is essentially performing the correct measurements (recall that we are assuming Alice's measurements are trusted). Self-testing allows us to certify one Bell state and the local observables of Bob in a one-shot manner. Intuitively, it seems that we could perform sequential self-tests in order to certify multiple Bell pairs, thus recovering the tensor product structure. For example, according to theorem 4 we can conclude that after K rounds of measurement the reduced state in one of the rounds is, with high probability, close to a Bell state. We would then simply repeat this procedure N times in order to recover N Bell pairs, each one selected at random from each set of K rounds. This intuitive argument does not hold. As is also noted in [9], Bob's strategy and the untrusted states he prepares for a certain set of measurements can overlap with the states and strategy from another set. Thus, if the reduced state for round i is close to a Bell pair, and the reduced state for round j is close to a Bell pair, we cannot implicitly conclude that the reduced for both rounds i and j is close to two Bell pairs.

We resolve this, in a manner similar to [9], by first defining a steering game akin to the CHSH game. We will then use theorem 4 to prove that the real strategy Alice and Bob use to play the steering game is close to the ideal

strategy. Intuitively, the steering game is one in which we ask Alice and Bob to measure specific observables on their shared state and check to see if their outcomes match. For example, to saturate inequality 2 we would ask Alice and Bob to both measure the X observable or to both measure the Y observable. Since we trust Alice, we know that she is behaving honestly, whereas Bob could deviate from the ideal strategy. Note that we are assuming that there is a referee asking Alice and Bob to perform these measurements. Since Alice is trusted, we could have her act as the referee and send the instructions to Bob. Indeed that will be the case when we consider the verification setting. However, the two situations are equivalent. For this reason, in our formal treatment of the game we shall consider Alice and Bob to be the players and that there exists a referee which instructs them on what to do. The definition is as follows:

Definition 1. We say that a game consisting of players Alice and Bob is a K-round steering game with threshold $T \leqslant K$ iff the following conditions are satisfied:

- Alice and Bob share a joint unknown quantum state $|\psi\rangle$.
- The game has K rounds.
- In round i, Alice is instructed by the referee to measure her half of $|\psi\rangle$ with either the A_0 or A_1 two-outcome, single-qubit, anti-commuting observables, having ± 1 eigenvalues and record her measurement outcome (keeping it secret from Bob).
- Alice's measurement device is fully trusted to perform the correct measurement, moreover she has a complete characterisation of the device's Hilbert space.
- In round i, Bob is instructed by the referee to measure his half of $|\psi\rangle$ with either the B_0 or B_1 two-outcome observables, having ± 1 eigenvalues and reports his outcome to Alice.
- Alice and Bob win the current round iff their outcomes are identical.
- Alice and Bob win the game iff they win at least T rounds.

Note that in the previous definition we match the conditions of theorem 4. Unlike Alice, Bob is untrusted and so his observables are unknown to Alice (and the referee). Moreover, the state $|\psi\rangle$ is unknown and we can assume that it was prepared by Bob. We now define the correlation value of the game.

Definition 2. Let W be the number of rounds that Alice and Bob win in a K-round steering game. The correlation value for the game is defined as the fraction W/K.

It is useful to make the following observation: if we assume that Alice and Bob are measuring the same state, $|\phi\rangle$, in each round, then the correlation value of the game would be:

$$\frac{1}{2}\langle\phi|\ A_0B_0 + A_1B_1\ |\phi\rangle.$$

In general, this might not be the case, since Bob is free to use any state in each round. In accordance with theorem 4, the correlation value of the game is then an estimate for the correlation of the averaged state. We use this fact to conclude something about the correlation of the reduced state in a randomly chosen round. This fact will be used to prove rigidity.

Alice and Bob will be asked to play multiple steering games. Following the notation of [9] we will denote ρ as their shared state for all these games (unlike $|\psi\rangle$ which is the state for one game) and for a specific game, j, we denote the operator associated with Alice's actions (measurements) as $\mathcal{E}_j^A$ and the operator associated with Bob's action as $\mathcal{E}_j^B$. Thus, for N steering games, the triplet $\mathcal{S} = (\rho, \{\mathcal{E}_j^A\}, \{\mathcal{E}_j^B\})$ encodes the *strategy* of Alice and Bob, where $j \leqslant N$. The ideal strategy, which we denote as $\mathcal{S}_{\mathrm{id}}$, is the one in which ρ_{id} is a tensor product of Bell states and Alice and Bob perform the measurements they are instructed to perform in each round. We need to show that $\mathcal{S} \approx \mathcal{S}_{\mathrm{id}}$. This was shown in [9] for CHSH games by extracting a tensor structure in the individual Hilbert spaces of Alice and Bob from the tensor structure of their two spaces. In our case, because we trust Alice and her device, we already have a characterisation of her strategy and Hilbert space. Therefore, we need only use this to characterise Bob's strategy. As in [9], we do this by exploiting the following symmetry property of the Bell state $(M \otimes I)|\psi_+\rangle = I \otimes (XM^TX)|\psi_+\rangle$, for any 2×2 matrix M. This essentially allows us to shift Bob's measurements to Alice's side, thus eliminating any dependence of his outcomes on previous qubits and establishing a tensor product structure. We will use this to create a strategy in which only Alice performs measurements. In analogy to [9], we start by defining an ϵ-structured steering game.

Definition 3. We say that a K-round steering game is ϵ-structured iff the game correlation value is greater than $1 - \epsilon$.

Alternatively, we can set the winning threshold for the game to $T = (1 - \epsilon)K$. Thus, winning a steering game is equivalent to having an ϵ-structured steering game. Similarly, one can define an ϵ-structured strategy.

Definition 4. A general strategy $\mathcal{S}$ for N K-round steering games is ϵ-structured iff for an arbitrary game $j \leqslant N$ we have that $\Pr(\text{game } j \text{ is } \epsilon\text{-structured}) \geqslant 1 - \epsilon$.

Note that this means that the majority of steering games, in a strategy, are ϵ-structured. Lastly, we adapt a definition from [9] to characterise the closeness of two strategies. It should be noted that we will define two strategies as being close if the reduced actions of Alice and Bob on N randomly chosen rounds, in the two strategies, are close to each other. This contrasts the definition from [9], where two strategies are considered close if the actions of Alice and Bob for all rounds are close to each other. The reason for the difference is that we are only interested in establishing a tensor product of N Bell states and that Alice and Bob are performing ideal measurements on those N states. However, since we have N steering games, each consisting of K rounds, we will have NK rounds in total. Showing that all of these are close to ideal would lead to a large overhead which we would like to avoid.

Definition 5. Let $\mathcal{S} = (\rho, \{\mathcal{E}_j^A\}, \{\mathcal{E}_j^B\})$ and $\bar{\mathcal{S}} = (\bar{\rho}, \{\bar{\mathcal{E}}_j^A\}, \{\bar{\mathcal{E}}_j^B\})$ be two strategies for playing N sequential K-round steering games. For $\epsilon \geqslant 0$, we say that strategy $\bar{\mathcal{S}}$ ϵ-simulates strategy $\mathcal{S}$ iff they both use the same Hilbert spaces and for all j:

$$\mathrm{TD}(\mathrm{Tr}_{-R}(\mathcal{E}_{1,j}^{AB}(\rho)), \mathrm{Tr}_{-R}(\bar{\mathcal{E}}_{1,j}^{AB}(\bar{\rho}))) \leqslant \epsilon,$$

where R is a set of N round indices, each chosen at random from the N steering games. Additionally:

$$\mathcal{E}_{1,j}^{AB} = \mathcal{E}_1^{AB} \circ \mathcal{E}_2^{AB} \circ \ldots \mathcal{E}_j^{AB},$$

$$\bar{\mathcal{E}}_{1,j}^{AB} = \bar{\mathcal{E}}_1^{AB} \circ \bar{\mathcal{E}}_2^{AB} \circ \ldots \bar{\mathcal{E}}_j^{AB}$$

And:

$$\mathcal{E}_j^{AB} = \mathcal{E}_j^A \circ \mathcal{E}_j^B \qquad \bar{\mathcal{E}}_j^{AB} = \bar{\mathcal{E}}_j^A \circ \bar{\mathcal{E}}_j^B.$$

Whenever we have that $\mathcal{S}$ ϵ-simulates $\bar{\mathcal{S}}$, or an isometric extension of $\bar{\mathcal{S}}$, we will write $\mathcal{S} \approx \bar{\mathcal{S}}$. This leads us to the main result:

Theorem 5. *Let $\mathcal{S} = (\rho, \{\mathcal{E}_j^A\}, \{\mathcal{E}_j^B\})$ be Alice and Bob's ϵ-structured strategy for playing N sequential K-round steering games. Let $\mathcal{S}_{\mathrm{id}} = (\rho_{\mathrm{id}}, \{\mathcal{E}_{\mathrm{id}\,j}^A\}, \{\mathcal{E}_{\mathrm{id}\,j}^B\})$ be Alice and Bob's ideal strategy for playing N sequential steering games. We have that $\mathcal{S}$ $O(N\epsilon^{1/6})$-simulates an isometric extension of $\mathcal{S}_{\mathrm{id}}$.*

Proof sketch. Firstly, it should be noted that for all j, $\mathcal{E}_j^A = \mathcal{E}_{\mathrm{id}\,j}^A$. This is because Alice is trusted and always playing according to the ideal strategy. The proof then consists of two steps. First we show that if the real strategy $\mathcal{S}$ is ϵ-structured, then $\mathcal{S} \approx \mathcal{S}_g$, where $\mathcal{S}_g$ is a strategy in which Alice plays honestly and also guesses Bob's measurement outcomes. The guesses provided by Alice are taken to be Bob's outcomes for each game while the action on his subsystem is taken to be identity. The proof of this step relies on characterising the evolution of the quantum state ρ in the two strategies and using theorem 4 together with the ϵ-structured nature of the strategies. We then show that $\mathcal{S}_g \approx \mathcal{S}_{\mathrm{id}}$. To do this, note that in the guessing strategy we have effectively removed the problem of adaptivity. Since there is no untrusted Bob in $\mathcal{S}_g$, the original argument, of sequentially repeating self-testing, goes through. This allows us to show that $\mathcal{S}_g$ is close to a similar guessing strategy, denoted $\hat{\mathcal{S}}_g$, which uses ideal Bell pairs. Lastly, this strategy is trivially close to the ideal one. We can then combine these results to show that $\mathcal{S} \approx \mathcal{S}_{\mathrm{id}}$. The full proofs are given in appendix E. □

Since N represents the number of Bell pairs we wish to certify, we see that in order to obtain a decreasing error, we require $\epsilon = O(N^{-6})$. We also know from theorem 4 that given ϵ, the number of games required is of order $K = \Omega((1/\epsilon^2)\log(1/\epsilon))$. Therefore, we must have that $K = \Omega(N^{12}\log(N))$. Since each steering game comprises of K rounds, we have KN rounds in total, or $\Omega(N^{13}\log(N))$ rounds of steering games. This becomes important for use in VDQC.

2.4. Comparison with other approaches

Before presenting the VDQC application, we briefly compare our approach to similar results. As mentioned, this paper builds on the work of self-testing the singlet of McKague *et al* [36]. While we use similar techniques to theirs, we assume that Alice is trusted and thus arrives at an improved bound for the closeness of $|\psi\rangle$ to an ideal Bell pair ($O(\sqrt{\epsilon})$ versus $O(\epsilon^{1/4})$). On the other hand, the results of [9] and [50] do arrive at an $O(\sqrt{\epsilon})$ bound for self-testing. However, we know that their techniques cannot improve the asymptotic closeness, since we have shown in theorem 3 that this bound is optimal up to constant factors.

When comparing the exact bounds, we obtained $(3\sqrt{2}+8)\sqrt{\epsilon}+\epsilon/2$, which is smaller compared to that of [9], approximately $270\sqrt{\epsilon}$. The result of [50] obtained numerically an even smaller factor of $\sqrt{2.2\epsilon}$ by using a semidefinite programme. Their technique could, in principle, be used to improve our approach as well. We also mention the result of Šupić and Hoban [51], which appeared concurrently with our own. They also consider the case of self-testing from steering correlations, obtaining an analytic bound of $13\sqrt{\epsilon}$ and a numeric bound of $1.19\sqrt{\epsilon}$.

Furthermore, [50] also considers removing the i.i.d. assumption. Their approach is based on hypothesis testing, however the end result is to show that *at least* one state, out of all measured states, was close to a perfect Bell pair. In our case, theorem 4 establishes that *a typical* state, out of all measured states, was close to a perfect Bell pair. This is necessary in order to prove the rigidity result and certify a tensor product of Bell pairs.

For the rigidity of steering correlations we employed similar techniques to those of [9] to show that the strategy (consisting of states and measurements) associated with the real scenario is close to that of the ideal scenario. This is done by considering intermediate strategies and showing that they are close to both the real and the ideal one and therefore that the latter strategies are close to each other. The major difference with [9] is that because Alice is trusted, in our case, there was no added overhead in proving the closeness of these strategies to each other. This then lead to a reduced closeness bound.

3. Verified delegated quantum computation

The idea of VDQC is that a computationally weak verifier wants to delegate a computation to a powerful (quantum) prover, and at the same time be able to verify the correctness of the result received. In characterising VDQC protocols, we use the formalism of *interactive proof systems* [32, 52]. For our setting, the prover is restricted to polynomial time quantum computations. Ideally we would like the verifier to be a fully classical computer, however it is still an open problem if this is possible, when there is a single prover [53]. Instead, it is known that if the verifier has some minimal quantum capabilities, he is able to verify the prover's computation [32, 54]. Alternatively, if there are multiple non-communicating provers sharing entanglement it is possible to do verification with a fully classical verifier [9, 10, 12]. It should be noted that in certain cases not all the provers are quantum computers. For example, in the protocols of [10, 11] there are two provers, one is a quantum computer (or server) and the other is an untrusted measurement device. Thus, throughout this paper we will refer explicitly to the provers by their role (i.e. quantum server, measurement device etc).

3.1. Existing approaches to VDQC

The first approaches to VDQC relied on having a classical verifier with a minimal quantum device. This device could be either a constant size quantum computer [32], or a single qubit preparation device [54]. In both cases, the verifier has both a classical and a quantum communication channel with the server. The quantum channel is used only in an initial phase to send quantum states to the server. Verification of the computation is then performed via classical communication only. Importantly, the quantum communication is *offline*, meaning it can be performed before the verifier even decides what computation she would like to perform (she must, however, fix the size of the computation to be performed). We will briefly explain one of these protocols, namely the one by Fitzsimons and Kashefi [54], which we shall refer to as the FK protocol. The reason for choosing this protocol is that many other VDQC protocols that were later developed have been based on FK [10, 11, 55–58, 63]. Furthermore the FK protocol is currently the optimal protocol from the point of view of the client's quantum capability requirements. Hence it is a good starting point in further reducing the trust assumptions as we intend to do in this section. We therefore use our results on quantum steering to modify one of the FK-based protocols.

The FK protocol is expressed in the measurement-based quantum computing (MBQC) model of computation and uses universal blind quantum computation (UBQC) [59] as a basis for verification. We succinctly explain the basic ideas behind these concepts as further details can be found here [59, 60]. In MBQC, computation is achieved through a sequence of adaptive single-qubit measurements performed on a highly entangled state known as a *graph state*. It is possible to make this graph state highly regular, by having all qubits prepared in the $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ state and entangling them using the controlled-Z operation into a

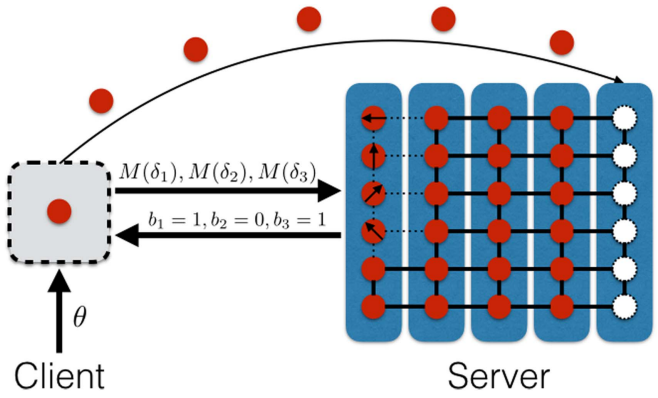

Figure 5. UBQC protocol.

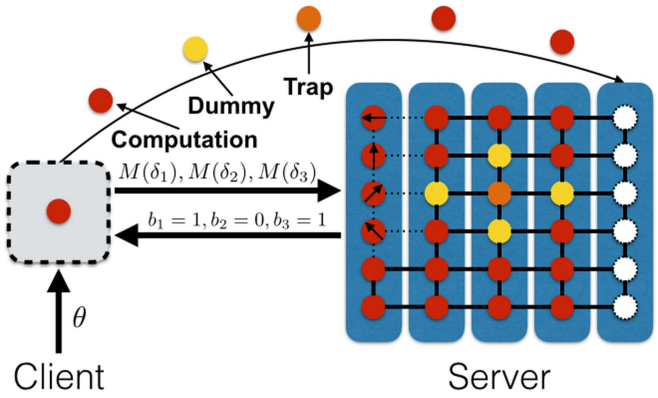

Figure 6. FK verification protocol.

brickwork structure [54, 59]. The qubits are then measured in the basis $\{|+_\theta\rangle, |-_\theta\rangle\}$, where $|\pm_\theta\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm e^{i\theta}|1\rangle)$ and θ is chosen adaptively from the set $D = \{0, \pi/4, \ldots, 7\pi/4\}$.

In the following we give the main idea behind UBQC. A trusted client sends to an untrusted quantum server rotated qubits of the form $|+_\theta\rangle$ with angles θ chosen randomly from the set D. The server is then supposed to entangle these qubits in a generic graph state structure and then measure them in the basis $\{|+_\delta\rangle, |-_\delta\rangle\}$, $\delta \in D$, as instructed by the client. The measurement angles are selected and adapted in order to perform a specific computation. Having no knowledge of the initial rotation angles (the θ's), the measurement angles (the δ's) will appear random to the server and so he will have no information about the computation being performed, apart from an upper bound on its size (given by the number of qubits). We illustrate this in figure 5. This blind computation procedure can be modified in order to perform verification as well, leading to the FK protocol.

In this case, the client, now known as a verifier, will also send computational basis states $\{|0\rangle, |1\rangle\}$ to the server, interleaved randomly with the rotated qubits. The purpose of these states, called dummy qubits, is to isolate certain rotated qubits from the rest of the graph state qubits. Isolation is achieved because the controlled-Z operation when used with a dummy and a rotated qubit will not perform entanglement. The isolated qubits are called *traps* because the verifier will instruct the server to measure these qubits in their preparation basis (i.e. the measurement angle δ will match the rotation angle θ for each trap qubit), thus yielding a deterministic outcome. Because of blindness, the position of the traps is hidden from the server and so he is unaware if he is performing a trap measurement or a computation measurement. This allows the verifier to test, on average, the behaviour of the server and abort the protocol if he behaves maliciously. Schematically, this is shown in figure 6.

As mentioned, the FK protocol uses a verifier with a trusted preparation device, while the untrusted server is performing entangling and measuring operations. We refer to this as a *prepare and measure* type of protocol. There is another class of protocols that have the server prepare the graph state and send the qubits one by one to the verifier who performs single qubit measurements [56, 61, 62]. In this case, the verifier has a trusted measurement device instead of a preparation device. We refer to the latter class of protocols as *measurement-only* type of protocols. The main drawback of this approach is that the quantum communication is online (occurs

during the time that the quantum computation is performed). This means that the verifier must commit to the computation she wishes to perform when the quantum communication commences. In contrast to this, the FK protocol is offline and therefore the verifier does not need to make such a commitment. Instead, after the quantum communication has occurred, she is free to choose any computation up to a fixed size and communicate only classically with the server.

All of the approaches mentioned so far, relied on having a verifier with a trusted quantum device. It is also possible to have a fully classical verifier, with no quantum device, if we allow for multiple entangled quantum servers. Such is the case with the protocols presented in [9, 12]. Alternatively, the verifier can have an untrusted measurement device and share entanglement with a single quantum server. This scenario is a *device-independent* version of the the prepare and measure protocols. The two existing approaches are presented in [10, 11]. We will briefly describe the operation of these device-independent protocols, since a slight modification of them will lead to our *one-sided device-independent* protocol.

The setting of device-independent VDQC is the following:

- The verifier has an untrusted measurement device. The device can measure the observables $\{X', Y', Z', D', E', F'\}$, where $D' = \frac{1}{\sqrt{2}}(X' + Z')$, $E = \frac{1}{\sqrt{2}}(X' + Y')$, $F = \frac{1}{\sqrt{2}}(Y' + Z')$. The observables are primed to indicate they are untrusted.
- The server is instructed to initially prepare Bell states and send half of each state to the verifier.
- The measurement device and the server are assumed to be non-communicating.
- The verifier interacts only classically with both of these devices.

The verification protocol consists of the following two stages:

1. Verified state preparation—In this stage the verifier will use the shared entangled states with the server in order to prepare single qubit states on the server's side. These are the states which, in the FK protocol, were sent to the server having been prepared by the trusted quantum device and which will later be used to perform verified computation. In this case, because the measurement device is also untrusted, the verifier will have to interact with the two devices in order to certify the preparation of correct states on the server's side. In the protocol of [10], this is done using the rigidity property of CHSH games, whereas the protocol of [11] achieves this using a modified version of the Mayers–Yao self-test.
2. *Verified computation*—Having prepared the rotated states and dummy qubits on the server's side, the verifier proceeds to run the FK protocol as if she had sent the qubits to the server. It can be shown that the server is still blind following the preparation stage. Since this stage results in the preparation of imperfect states (ϵ-close to the ideal), it was necessary to show that the FK protocol is robust to deviations in the initial state.

It is worth pointing-out here, that any deviation of the server on the correctly prepared input of the FK, is detected by the verification mechanism of the FK protocol. In other words, the two stages can be separated, and the verified preparation can happen any time earlier than the verified computation. The main disadvantage of the device-independent protocols is the large round complexity in the state preparation stage, leading to an overall large complexity compared to other approaches. The main reason behind this blow-up, is the number of measurements that need to be performed in order to certify correct state preparation and additionally the measurements that need to be performed to test the honesty of the two devices. A major factor for this increase, is the need to certify the tensor product structure of Bell states shared, before the verifier uses them to prepare the states for the FK on the server's side. By allowing the verifier's device to be trusted, we can eliminate this last set of measurements and gain an advantage in terms of round complexity, leading to a one-sided device-independent VDQC protocol. In this protocol, similar to the measurement-only protocols, the verifier has a trusted measurement device. However, in our case, the (trusted) measurements can occur offline, since they are only involved in the verified state preparation stage.

3.2. Verification based on steering correlations

The setting we have, is that the verifier trusts her measuring device, but the server and the shared state, prepared by the server, are not trusted. Since our setting is identical, in terms of trust assumptions, to the one-sided device-independent self-testing scenario, we will sometimes refer to the verifier as Alice and the server as Bob. If Alice knew that the shared state with Bob is a tensor product of perfect Bell states, she could measure her side and collapse the state of the server to the desired input of FK. But since the state is untrusted, we will use the results of

the previous sections to characterise the shared state from steering correlations. If Alice observes a close to maximal saturation of the steering correlations, she can conclude that the shared state is a tensor product of Bell pairs. This follows from the rigidity result of theorem 5. The result is analogous to that of [9, 10] where the rigidity of CHSH games is used in the same fashion. All of this is encapsulated in protocol 1.

Protocol 1. One-sided device-independent verification protocol.

Assumptions

The verifier wants to delegate a quantum computation described by the graph G and specific measurement angles from the set $\{0, \pi/4 \ldots 7\pi/4\}$ chosen to define a desired computation. For the FK protocol, used in stage 2, the graph G is encoded as a different graph, $\mathcal{G}$, having M qubits. The verifier has a trusted single-qubit measurement device.

Stage 1: Verified state preparation

1. Verifier instructs the server to prepare $\Theta(M^{13}\log(M))$ Bell pairs and send half of each pair to her.
2. Verifier measures a random subset of M qubits in the $\{|+_\theta\rangle, |-_\theta\rangle\}$ basis, where $\theta \in \{0, \pi/4 \ldots 7\pi/4\}$, or the computational basis, without instructing the server to measure them.
   – For the remaining qubits (which are of order $\Theta(M^{13}\log(M))$) she again measures in the $\{|+_\theta\rangle, |-_\theta\rangle\}$ or the computational basis, but she also instructs the server to measure the corresponding entangled partner in the same basis.
   – The server reports the measurement outcomes and if the results are not the same as the verifiers, the protocol is aborted.

Stage 2: Verified computation (Robust FK)

3. If the protocol is not aborted in the previous stage, the verifier runs the (robust) FK protocol (given in [10]) with the server, using graph $\mathcal{G}$ (i.e. the qubits from the first step of Stage 1are treated as if they had been sent by the verifier to be entangled into the graph $\mathcal{G}$ and run the FK protocol).

During the verified state preparation stage, the verifier confirms that all her measurement outcomes agree with the measurement outcomes reported by the server. The number of measurement rounds, i.e. $\Theta(M^{13}\log(M))$, is chosen so that the results of theorems 4 and 5 lead to a decreasing error. As explained, the theorems require us to have $\Theta(M^{12}\log(M))$ rounds of measurement per steering game, and since we have M steering games we end up with $\Theta(M^{13}\log(M))$ rounds. Contrast this to the round complexities in the fully device-independent scenarios which are of order $O(M^c)$, where $c > 8192$ for [9] and $c > 2048$ for [10], respectively. Even though these do not follow from tight bounds, it seems clear that the added trust of the verifier's measurement device leads to a significant reduction in complexity.

Next we examine the verified computation stage. As the input obtained from the previous stage is not exact, but close in trace distance to the ideal state, we need to consider the robust version of the FK derived in [10]. Moreover, to obtain the optimal complexity, instead of using a dotted-complete graph as was done in [10, 54], we will use the optimised resource construction introduced in [58], where the number of qubits in the encoded graph, $\mathcal{G}$, is linear in the number of qubits in the computation graph G. This leads us to the following:

Lemma 1. *Protocol 1, for a computation of size M, utilises $\Theta(M^{13}\log(M))$ Bell pairs and has $\Theta(M^{13}\log(M))$ round complexity (rounds of interaction between the server and the verifier). The probability that the verifier accepts the outcome of the protocol, assuming honest behaviour by the server, is unity. If η is the probability that the verifier accepts an incorrect outcome in the FK protocol (verified computation stage), the overall probability of accepting an incorrect outcome for protocol 1 is $\eta + \lambda^{-1}$, for some constant $\lambda > 1$ fixed by the verifier.*

Proof. As mentioned, the FK protocol with the encoding (resource state) defined in [58], or using the resource construction procedure of [63], has linear round complexity. Therefore, the round complexity of the verified computation stage is $\Theta(M)$. The verified preparation stage requires $\Theta(M^{13}\log(M))$ Bell pairs which are used to test the saturation of steering inequalities. Therefore, the overall round complexity is $\Theta(M^{13}\log(M))$. In the honest run, the verified state preparation stage leads to preparing the ideal input, and therefore the probability of accepting an honest run is unity. This is because in the honest setting, the server prepares ideal Bell pairs and all states are measured correctly, leading to the correct rotated qubits on his side, which are then used in the FK protocol. The FK protocol also has probability of acceptance unity, when the server behaves honestly [54].

On the other hand, a dishonest run involves deviations in both state preparation and verified computation stages. In the first stage, a dishonest server would prepare states that are at most $O(\epsilon^{1/6})$-deviated from the ideal by theorem 4. Here, $\epsilon = \Theta(1/M^6)$, in order to have a decreasing error in simulating the ideal strategy. Thus, the deviation per Bell pair is fixed however, the verifier can choose many measurement rounds for the steering games in the state preparation stage, reducing the deviation in the overall (tensor product) state. Concretely, if she chooses to run $\lambda^{12}M^6$ steering games, then $\epsilon = \lambda^{-12}M^{-6}$, for some fixed λ. The overall deviation, which is of order $M\epsilon^{1/6}$, will then be of order λ^{-2}. This constitutes the deviation in the state preparation stage. As explained in [10], the robustness of the FK protocol implies that if the computation stage has probability of acceptance,

given dishonest behaviour, η and there is a deviation of order α in the state preparation phase, then the overall probability of accepting an incorrect outcome will be $\eta + \sqrt{\alpha}$. In our case, $\alpha = \lambda^{-2}$, hence the probability becomes $\eta + \lambda^{-1}$. □

The resulting protocol's round complexity is an improvement, over the device-independent approaches of [9, 10, 12]. The reason for this is that we benefited from the fact that Alice is trusted in the rigidity result. In particular, while there is a cost for removing the i.i.d. assumption, in the semi-trusted setting we have no extra cost to recover the tensor product structure from theorem 5. This leads to a better complexity when we extract the tensor product structure of Bell pairs, than the complexity obtained in the completely untrusted (device-independent) setting of [9]. On the other hand, in the case of FK where the verifier has a trusted preparation device and sends qubits to the server, the round complexity and number of used quantum states are linear in the size of the computation [58]. In our case, the main sources for overhead are:

- Self-testing with i.i.d. states which gives a tight bound for the closeness of the state up order $\sqrt{\epsilon}$.
- Gathering sufficient statistics in the non-i.i.d. setting to get a good estimate of the true quantum correlations for the averaged state. This required $\Omega((1/\epsilon^2)\log(1/\epsilon))$ rounds of measurements.
- Inferring the closeness of a typical state from the closeness of the averaged state, which gives a bound of order $\epsilon^{1/6}$.

3.3. Verification from partially entangled states

It is known that if the source of quantum states is trusted, then classical-quantum correlations are sufficient for verification [11, 54]. Alternatively, in *online* verification protocols (where computation and quantum communication take place at the same time), it is possible to have an untrusted entanglement source and states which are less than maximally entangled. Our setting is that of *offline* verification with an untrusted entanglement source.

We have seen that in both device-independent verification and one-sided device-independent verification we can characterise the tensor structure of Bell pairs between the verifier and the server from correlations. In both cases, saturating an inequality involving correlations leads to a bound on the trace distance between the shared state and perfect Bell pairs, up to an isometry. While this is sufficient, it does not seem necessary to use Bell states. In fact in QKD and random number generation, as mentioned previously, other types of states can also be used [27, 28, 47–49] and so it is interesting to examine if this is also the case for verification. Here, interestingly, we show that under reasonable assumptions about the verification protocol, the entangled states must be unitarily equivalent to Bell states.

Quantum steering derives its name from the idea that with an entangled state one party could steer the state of another party through local measurement. The person who performs the steering is untrusted which in our case corresponds to the quantum server. The verifier, having a trusted measurement device, can check through local measurement, if the server behaved honestly and steered the state correctly. We consider such steerable states, ρ_{AB}, shared between Alice (the verifier) and Bob (the server), which could be useful for verification in the one-sided device-independent scenario. Concretely, the specific type of verification we consider is one which is offline, measurement-only and blind. This is akin to the setting of our steering-based verification protocol, however we no longer assume the use of the FK protocol in the verified computation stage. Instead, we assume independent verification, i.e. a black-box type verification protocol obeying the three conditions we previously stated. Moreover, we assume that the verifier wishes to prepare a specific quantum input for her computation. This input is assumed to be as general as possible. Given these conditions, we wish to know what properties are required of a typical bipartite shared state ρ_{AB}. Since we are interested in blind verification, it must be the case that $\mathrm{Tr}_A(\rho_{AB}) = \rho_B = I/2$. Additionally, because the verifier needs to prepare her quantum input on the server's side, and this can consist of a varied set of states, we will also assume that ρ_{AB} must be *completely steerable* (i.e. can be steered to any state). We use the definition of complete steering from [64].

Definition 6. [64] A bipartite state, ρ_{AB}, shared by Alice and Bob is completely steerable by Bob iff for any positive operators $\{\sigma_a\}$, satisfying $\sum_a \sigma_a = \mathrm{Tr}_A(\rho_{AB})$, there exist a POVM $\{E_a\}$, such that $\sigma_a = \mathrm{Tr}_A((E_a \otimes I)\rho_{AB})$.

Adding the blindness assumption to this, we follow up with the definition:

Definition 7. A bipartite state, ρ_{AB}, shared by Alice and Bob is totally steerable by Bob iff ρ_{AB} is completely steerable by Bob and $\mathrm{Tr}_A(B) = I/2$.

Before proving our result, we state a useful lemma from [64], concerning completely steerable states:

Lemma 2. *[64] A bipartite state, ρ_{AB} is completely steerable by Bob iff exists a purification ρ_{ABC} such that $\rho_{BC} = \rho_B \otimes \rho_C$, where $\rho_{BC} = \mathrm{Tr}_A(\rho_{ABC})$, $\rho_B = \mathrm{Tr}_{AC}(\rho_{ABC})$, $\rho_C = \mathrm{Tr}_{AB}(\rho_{ABC})$.*

This is similar to QKD, where an adversary can have a purification of the state shared by Alice and Bob however he is uncorrelated with Alice or Bob. This is due to the monogamy of entanglement. Our main theorem concerning ρ_{AB} will justify why this monogamy appears in the case of complete steerability as well. We note that the monogamy of quantum steering has been studied in other works [65, 66].

Theorem 6. *A bipartite state, ρ_{AB} is totally steerable by Bob iff ρ_{AB} is maximally entangled.*

Proof sketch. The proof relies on considering ρ_{AB} in matrix form and expressing constraints on its matrix elements using lemma 2 and the fact that $\rho_B = I/2$. Solving the system of constraints leads to a density matrix which can easily be shown to correspond to a pure, maximally entangled state. The complete proof is given in appendix F. □

This theorem reveals that the ρ_{AB} state is unitarily equivalent to a maximally entangled Bell state. Therefore, the monogamy of this state is in fact due to the monogamy of Bell states. Importantly, it shows that two-qubit states which are totally steerable are perfect Bell pairs. This implies that if we are to use total steerability in order to perform verification, we must certify Bell states in particular, rather than any other steerable state. Hence, for the case of offline, measurement-only, blind quantum verification, the quantum resources used for preparing the quantum input, must be unitarily equivalent to Bell states. Intuitively, this can be understood as follows. The verifier wishes to send a certain quantum input to the server. Since this input is as general as possible, the best way to do this is via the general teleportation protocol using Bell states. This also introduces a natural one-time padding to the states, which keeps the protocol blind. It is clear that in order to have perfect fidelity for the teleportation, the shared states must be equivalent to perfect Bell pairs.

4. Conclusion

We have shown that, in analogy to the rigidity of non-local correlations via CHSH games, we can prove a rigidity property of steering correlations. This allows us to establish a tensor product of Bell pairs in a setting of one trusted party and one untrusted party. However, in the case of steering, the extra cost to obtain a tensor product of Bell states is smaller than in the analogous situation for non-local correlations. While in both cases this overhead is polynomial, the additional trust assumptions of the steering setting allowed for a significantly reduced degree for this polynomial. We arrived at this result by first considering self-testing in the one-sided device-independent setting, using quantum steering correlations. The self-testing result, which makes an i.i.d assumption about the shared state, gives an optimal bound for the trace distance between this shared state and a Bell pair. We then removed the i.i.d. assumption in a generic way which can be applied to any type of self-testing result, allowing for the determination of a single Bell state (one-shot) in a fully adversarial setting. Combining this with a game-based approach for characterising the states and strategies of the two parties, has lead to the rigidity result. It should be noted, that the game-based approach is also used to prove the rigidity of non-local correlations. However, in our setting, because we trust one of the parties it is simpler to obtain a characterisation of the untrusted party's strategy and the shared states. Thus, our result has reduced overhead compared the analogous result for CHSH game rigidity.

We considered an application of rigidity to a verifiable quantum computation protocol. The reason for choosing this particular application is that for verifiable quantum computation the quantum states themselves should be recovered, and thus the necessity to obtain the full tensor product structure of Bell states. This is in contrast to QKD and QRNG where one does not need to recover the quantum state explicitly, rather an information theoretic quantity such as entropy, mutual information, key-rate etc. Using the rigidity of steering correlations we constructed a one-sided device-independent protocol for verifiable delegated quantum computation. The protocol we obtained has fewer requirements than the fully device-independent protocol thus being closer to a practical application. Lastly, we have shown that a certain class of states which are useful for verification, totally steerable states, are necessarily maximally entangled. This for example, rules-out the use of local but steerable states [26] for verification. Since establishing the existence of the tensor product structure of maximal entanglement requires the collection of a significant number of statistical samples, this result gives some indication to the general difficulty of quantum verification.

Acknowledgments

We would like to thank Damian Markham for useful discussions. We also acknowledge discussions with Matty J Hoban on his and Ivan Šupić's independent work on self-testing using quantum steering [51], that appeared on the arXiv shortly after our preprint. We would also like to thank Dominique Unruh and Kristiina Rahkema for reading an earlier draft and suggesting improvements. AG and PW thank CNRS Telecom ParisTech for their hospitality during a visit where this work started. EK acknowledges funding through EPSRC grants EP/N003829/1 and EP/M013243/1.

Appendix A. Proof of theorem 1

In this section we give a complete proof of theorem 1 which characterises one-sided device-independent self-testing. Consider the following isometry:

$$\Phi(|\psi\rangle) = \frac{1}{2}(I + Y'_B)|\psi\rangle|+_y\rangle + \frac{\mathrm{i}}{2}X'_B(I - Y'_B)|\psi\rangle|-_y\rangle. \tag{7}$$

An illustration of this isometry is given in figure 3, where the upper part is Bob's system and the lower part is Alice's system. It should be noted that the control gates act on the target when the control qubit is in the $|-_y\rangle$ state, and act as identity when the control qubit is $|+_y\rangle$. This is in contrast to the standard convention in which the control is a computational basis state. Here $|+_y\rangle$ and $|-_y\rangle$ are the eigenstates of the Pauli Y operator and $P = \frac{1}{\sqrt{2}}(X + Y)$. We can clearly see that $\Phi = I_A \otimes \Phi_B$, where Φ_B is determined by the combination of X'_B and Y'_B operators, from expression (7), which only act on Bob's system. We proceed to show that when conditions (1)–(3) are satisfied, we obtain condition (4). First, we show that:

$$||\Phi(M_A N'_B\ |\psi\rangle) - M_A N_B \Phi(|\psi\rangle)|| \leqslant 2\gamma_2.$$

Because M_A only acts on Alice's system, whereas the isometry is local on Bob's system, M_A trivially commutes to the left so that $\Phi(M_A N'_B\ |\psi\rangle) = M_A \Phi(N'_B\ |\psi\rangle)$. Now consider the possible choices for N'_B. If $N'_B = I$, the relation holds trivially. If $N'_B = Y'_B$, since Y'_B is hermitian and unitary we have that:

$$\Phi(Y'_B\ |\psi\rangle) = \frac{1}{2}(I + Y'_B)|\psi\rangle|+_y\rangle - \frac{\mathrm{i}}{2}X'_B(I - Y'_B)|\psi\rangle|-_y\rangle. \tag{8}$$

At the same time, the ideal Pauli operator Y_B, acting on Bob's ancilla, has the following effect:

$$Y_B\Phi(|\psi\rangle) = \frac{1}{2}(I + Y'_B)|\psi\rangle|+_y\rangle - \frac{\mathrm{i}}{2}X'_B(I - Y'_B)|\psi\rangle|-_y\rangle. \tag{9}$$

This is because $Y\ |+_y\rangle = |\ +_y\ \rangle$ and $Y\ |-_y\rangle = -|\ -_y\ \rangle$ and we notice that the two expressions are identical. Lastly, when $N'_B = X'_B$ we have:

$$\begin{aligned}\Phi(X'_B\ |\psi\rangle) = \frac{1}{2}(I + Y'_B)X'_B\ |\psi\rangle|+_y\rangle \\ + \frac{\mathrm{i}}{2}X'_B(I - Y'_B)X'_B\ |\psi\rangle|-_y\rangle.\end{aligned}$$

And the action of the ideal operator yields:

$$X_B\Phi(|\psi\rangle) = \frac{\mathrm{i}}{2}(I + Y'_B)|\psi\rangle|-_y\rangle + \frac{1}{2}X'_B(I - Y'_B)|\psi\rangle|+_y\rangle. \tag{10}$$

This is because $X\ |+_y\rangle = \mathrm{i}\ |\ -_y\ \rangle$ and $X\ |-_y\rangle = (-\mathrm{i})|\ +_y\ \rangle$. Using the approximate anti-commutation of X'_B and Y'_B, as given by condition (3), we notice that commuting X'_B to the left in $\Phi(X'_B\ |\psi\rangle)$ will lead to the same expression as for $X_B\Phi(|\psi\rangle)$ up to $2\gamma_2$ error. Thus:

$$||\Phi(M_A N'_B\ |\psi\rangle) - M_A N_B \Phi(|\psi\rangle)|| \leqslant 2\gamma_2. \tag{11}$$

We therefore, only need to examine the closeness of $\Phi(|\psi\rangle)$ to the ideal Bell state tensored with some junk state. Start by considering the state:

$$\begin{aligned}|\phi\rangle = \frac{1}{2}(|\psi\rangle|+_y\rangle + Y_A\ |\psi\rangle|+_y\rangle \\ + \mathrm{i}X_A\ |\psi\rangle|-_y\rangle - \mathrm{i}Y_AX_A\ |\psi\rangle|-_y\rangle).\end{aligned} \tag{12}$$

We will show that $\Phi(|\psi\rangle)$ and $|\phi\rangle$ are close in trace distance. Firstly, from conditions (1) and (2) using suitable triangle inequalities and the unitarity of operators X_A and Y_A which do not increase trace distance, it can be shown that:

$$||(X'_B Y'_B - Y_A X_A)|\psi\rangle|| \leqslant 2\gamma_1.$$

Expanding the trace distance of $\Phi(|\psi\rangle)$ and $|\phi\rangle$ we have:

$$\begin{aligned}||\Phi(|\psi\rangle) - |\phi\rangle|| = \frac{1}{2}||(Y'_B - Y_A)|\psi\rangle|+_y\rangle \\ + \mathrm{i}(X'_B - X_A)|\psi\rangle|-_y\rangle - \mathrm{i}(X'_B Y'_B - Y_A X_A)|\psi\rangle|-_y\rangle||.\end{aligned}$$

And using the above results it follows that:

$$||\Phi(|\psi\rangle) - |\phi\rangle|| \leqslant 2\gamma_1.$$

Let us now rewrite $|\phi\rangle$. Given that we trust Alice's side of the $|\psi\rangle$ state, we can express it as follows:

$$|\psi\rangle = a\,|\alpha\rangle_B|+_y\rangle_A + b\,|\beta\rangle_B|-_y\rangle_A,$$

where $|a|^2 + |b|^2 = 1$ and the states $|\alpha\rangle$ and $|\beta\rangle$ are normalised. Here, the first part denotes Bob's system, for which we can make no assumptions, and the second part is Alice's qubit. The reason for choosing Pauli-Y eigenstates on Alice's side is to simplify the calculation. We could have expanded her system in any basis since a local unitary on her system does not change the result. Substituting this into the $|\phi\rangle$ and labelling the ancillary qubit introduced by this isometry with the label Φ we get:

$$\begin{aligned}|\phi\rangle = \frac{1}{2}(a\,|+_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A + b\,|+_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A) \\ + \frac{1}{2}Y_A(a\,|+_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A + b\,|+_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A) \\ + \frac{1}{2}X_A(a\,|-_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A + b\,|-_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A) \\ - \frac{1}{2}Y_A X_A(a\,|-_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A + b\,|-_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A).\end{aligned} \tag{13}$$

Using the following identities:

$$\begin{aligned}X\,|+_y\rangle = \mathrm{i}\,|-_y\rangle \quad X\,|-_y\rangle = -\mathrm{i}\,|+_y\rangle \\ Y\,|+_y\rangle = |+_y\rangle \quad Y\,|-_y\rangle = -|-_y\rangle.\end{aligned}$$

We reduce $|\phi\rangle$ to:

$$\begin{aligned}|\phi\rangle = \frac{1}{2}(a\,|+_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A + b\,|+_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A) \\ + \frac{1}{2}(a\,|+_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A - b\,|+_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A) \\ - \frac{1}{2}(a\,|-_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A - b\,|-_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A) \\ - \frac{1}{2}(a\,|-_y\rangle_\Phi\,|\alpha\rangle_B|-_y\rangle_A + b\,|-_y\rangle_\Phi\,|\alpha\rangle_B|+_y\rangle_A).\end{aligned} \tag{14}$$

The terms with b coefficient cancel out and we are left with:

$$|\phi\rangle = a\,|\alpha\rangle_B(|+_y\rangle_\Phi|+_y\rangle_A - |-_y\rangle_\Phi|-_y\rangle_A).$$

This state is equivalent to:

$$|\phi\rangle = a\sqrt{2}\,|\alpha\rangle_B\,|\psi_+\rangle_{AB}.$$

We would like to equate this to $|\mathrm{junk}\rangle_B\,|\psi_+\rangle_{AB}$, however, the state we have is unnormalized unless $a = 1/\sqrt{2}$. We therefore compute a bound on $|a|$ to determine the error introduced by the unnormalized state. Condition (1) can be rewritten as:

$$1 - \gamma_1^2/2 \leqslant \langle\psi|\,X_A X'_B\,|\psi\rangle.$$

By expanding $|\psi\rangle$, applying the operators X_A, X'_B and using the facts that $|a|^2 + |b|^2 = 1$ and that X'_B is hermitian and has ± 1 eigenvalues, we obtain:

$$\sqrt{1 - \gamma_1^2/2} \leqslant a\sqrt{2} \leqslant \sqrt{1 + \gamma_1^2/2}.$$

And since for small γ_1 we know that $\sqrt{1 - \gamma_1^2/2}$ approaches $1 - \gamma_1^2/4$ and $\sqrt{1 + \gamma_1^2/2}$ approaches $1 + \gamma_1^2/4$ we have that the norm of $|\phi\rangle$, can change from unity by an order of $\gamma_1^2/4$. Thus, it follows that:

$$||\Phi(|\psi\rangle) - |\mathrm{junk}\rangle_B\,|\psi_+\rangle_{AB}|| \leqslant 3\gamma_1 + \gamma_1^2/4.$$

Lastly, together with inequality (11) and a triangle inequality, we get:

$$||\Phi(M_A N'_B\ |\psi\rangle) - |\text{junk}\rangle_B M_A N_B\ |\psi_+\rangle_{AB}|| \leqslant 3\gamma_1 + \gamma_1^2/4 + 2\gamma_2. \tag{15}$$

Appendix B. Proof of theorem 2

Theorem 2 shows that saturating the correlation of observables on Alice and Bob's side, with Alice being trusted, leads to the necessary conditions of theorem 1 which imply that the shared state is close, up to local isometry, to a Bell state. Similar to theorem 1 we start the proof by denoting B_0 as X'_B and B_1 as Y'_B. Splitting equation (2), we have:

$$\langle\psi|\ X_A X'_B\ |\psi\rangle + \langle\psi|\ Y_A Y'_B\ |\psi\rangle \geqslant 2 - \epsilon. \tag{16}$$

However, it is clear that:

$$-1 \leqslant \langle\psi|\ X_A X'_B\ |\psi\rangle \leqslant 1$$
$$-1 \leqslant \langle\psi|\ Y_A Y'_B\ |\psi\rangle \leqslant 1$$

So, it follows that:

$$\langle\psi|\ X_A X'_B\ |\psi\rangle \geqslant 1 - \epsilon$$
$$\langle\psi|\ Y_A Y'_B\ |\psi\rangle \geqslant 1 - \epsilon.$$

This allows us to derive conditions (1) and (2), since:

$$||(X_A - X'_B)|\psi\rangle|| = \sqrt{2}\sqrt{1 - \langle\psi|\ X_A X'_B\ |\psi\rangle} \leqslant \sqrt{2\epsilon},$$
$$||(Y_A - Y'_B)|\psi\rangle|| = \sqrt{2}\sqrt{1 - \langle\psi|\ Y_A Y'_B\ |\psi\rangle} \leqslant \sqrt{2\epsilon}.$$

Hence, in theorem 1, $\gamma_1 = \sqrt{2\epsilon}$. Let us now denote:

$$S = X_A X'_B + Y_A Y'_B.$$

Computing S^2 and using the fact that $X_A Y_A = \mathrm{i} Z_A$ we obtain:

$$S^2 = 2 + \mathrm{i} Z_A [X'_B, Y'_B].$$

Since $[X_A, Y_A] = 2\mathrm{i} Z_A$, we can alternatively write this as:

$$S^2 = 2 + \frac{1}{2}[X_A, Y_A][X'_B, Y'_B].$$

The Cauchy–Schwarz inequality together with inequality (16) give us:

$$\langle\psi|\ S^2\ |\psi\rangle \geqslant |\langle\psi|\ S\ |\psi\rangle|^2 \geqslant (2 - \epsilon)^2.$$

Substituting S^2:

$$\langle\psi|\ 2 + \frac{1}{2}[X_A, Y_A][X'_B, Y'_B]|\psi\rangle \geqslant 4 - 4\epsilon + \epsilon^2 \geqslant 4 - 4\epsilon.$$

Hence:

$$\langle\psi|[X_A, Y_A][X'_B, Y'_B]|\psi\rangle \geqslant 4 - 8\epsilon.$$

Expanding the commutators yields:

$$\langle\psi|\ X_A Y_A X'_B Y'_B\ |\psi\rangle - \langle\psi|\ X_A Y_A Y'_B X'_B\ |\psi\rangle$$
$$- \langle\psi|\ Y_A X_A X'_B Y'_B\ |\psi\rangle + \langle\psi|\ Y_A X_A Y'_B X'_B\ |\psi\rangle \geqslant 4 - 8\epsilon.$$

By splitting into terms, as we did with inequality (16), we have that:

$$\langle\psi|\ X_A Y_A X'_B Y'_B\ |\psi\rangle \geqslant 1 - 8\epsilon, \tag{17}$$
$$\langle\psi|\ Y_A X_A Y'_B X'_B\ |\psi\rangle \geqslant 1 - 8\epsilon, \tag{18}$$
$$\langle\psi|\ X_A Y_A Y'_B X'_B\ |\psi\rangle \leqslant 8\epsilon - 1, \tag{19}$$
$$\langle\psi|\ Y_A X_A X'_B Y'_B\ |\psi\rangle \leqslant 8\epsilon - 1. \tag{20}$$

Now using the fact that $X_A Y_A + Y_A X_A = 0$, we have:

$$||(X'_B Y'_B + Y'_B X'_B)|\psi\rangle||$$
$$= ||(X'_B Y'_B + X_A Y_A + Y_A X_A + Y'_B X'_B)|\psi\rangle||.$$

And using a triangle inequality, we have:

$$\begin{aligned}&||(X'_B Y'_B + Y'_B X'_B)|\psi\rangle|| \\ &\leqslant ||(X_A Y_A + X'_B Y'_B)|\psi\rangle|| + ||(Y_A X_A + Y'_B X'_B)|\psi\rangle||.\end{aligned}$$

Additionally:

$$\begin{aligned}&||(X_A Y_A + X'_B Y'_B)|\psi\rangle|| \\ &= \sqrt{2 + \langle\psi|\ X_A Y_A Y'_B X'_B\ |\psi\rangle + \langle\psi|\ Y_A X_A X'_B Y'_B\ |\psi\rangle}\,.\end{aligned}$$

And from inequalities (19) and (20) we get that:

$$||(X_A Y_A + X'_B Y'_B)|\psi\rangle|| \leqslant 4\sqrt{\epsilon}.$$

Similarly, using inequalities (17) and (18), we have:

$$||(Y_A X_A + Y'_B X'_B)|\psi\rangle|| \leqslant 4\sqrt{\epsilon}.$$

Which leads to:

$$||(X'_B Y'_B + Y'_B X'_B)|\psi\rangle|| \leqslant 8\sqrt{\epsilon}.$$

Thus satisfying condition (3), with $\gamma_2 = 8\sqrt{\epsilon}$, and concluding the proof.

Appendix C. Proof of theorem 3

Theorems 1 and 2 show that if the correlation of local observables is saturated up to order $O(\epsilon)$, the shared state is close, up to local isometry, to a Bell state up to order $O(\sqrt{\epsilon})$. Theorem 3 shows that this bound is tight, up to constant factors. We prove this theorem by contradiction. Assume the bound of theorem 1 is not tight and it is possible to derive an asymptotically better bound for the shared state of Alice and Bob. In particular, this means that there is no state $|\psi\rangle$ which is $O(\sqrt{\epsilon})$-close to the $|\psi_+\rangle$ state and there are no observables B_0 and B_1 such that inequality 3 is satisfied. However, letting $\epsilon' = \epsilon/2$, consider the following state:

$$|\psi\rangle = \frac{1}{\sqrt{2}}(\sqrt{1+\sqrt{\epsilon'}}\ |01\rangle + \sqrt{1-\sqrt{\epsilon'}}\ |10\rangle).$$

We have that:

$$|||\psi\rangle - |\psi_+\rangle|| = \sqrt{2 - \langle\psi|\psi_+\rangle - \langle\psi_+|\psi\rangle}\,.$$

Notice that:

$$\langle\psi|\psi_+\rangle = \langle\psi_+|\psi\rangle = \frac{1}{2}(\sqrt{1+\sqrt{\epsilon'}} + \sqrt{1-\sqrt{\epsilon'}}).$$

Substituting this into the previous expression, and taking the first order term we have:

$$|||\psi\rangle - |\psi_+\rangle|| = O(\sqrt{\epsilon'}).$$

Consider also the observables:

$$B_0 = \begin{pmatrix} -\sqrt{\epsilon'} & \sqrt{1-\epsilon'} \\ \sqrt{1-\epsilon'} & \sqrt{\epsilon'} \end{pmatrix}$$

$$B_1 = \begin{pmatrix} 0 & \sqrt{\epsilon'} - \mathrm{i}\sqrt{1-\epsilon'} \\ \sqrt{\epsilon'} + \mathrm{i}\sqrt{1-\epsilon'} & 0 \end{pmatrix}.$$

One can check that $B_0 = B_0^\dagger$, $B_1 = B_1^\dagger$, $B_0 B_0^\dagger = B_1 B_1^\dagger = I$ and that the two matrices have eigenvalues ± 1. Moreover, we can see that as $\epsilon' \to 0$ we have that $B_0 \to X$ and $B_1 \to Y$. Importantly, we have that:

$$\langle\psi|\ X_A B_0\ |\psi\rangle = \langle\psi|\ Y_A B_1\ |\psi\rangle = 1 - \epsilon'.$$

And therefore:

$$\langle\psi|(X_A B_0 + Y_A B_1)|\psi\rangle = 2 - 2\epsilon' = 2 - \epsilon.$$

Thus, inequality 3 is saturated. But this should not be possible under the assumption that the bound on $|\psi\rangle$'s closeness to $|\psi_+\rangle$ is not tight. Therefore, the assumption is false and the $O(\sqrt{\epsilon})$ bound is tight for this type of steering inequality. Note that this result still holds under local isometry since the isometry is, by definition, distance preserving and so under the local isometry the state is still $O(\sqrt{\epsilon})$-close to a Bell pair.

Appendix D. Proof of theorem 4

Theorem 4 shows that from the observed outcomes of Alice and Bob, after K rounds of measurement, we can conclude something about their shared quantum state in a single round, even without assuming independence. Simply stated, if the shared state of Alice and Bob is σ, we need not assume that σ is a tensor product of identical states. Nevertheless, denoting the state of round i as $\rho_i = \mathrm{Tr}_{-i}(\mathcal{E}^{AB}_{1,i-1}(\sigma))$, from their shared correlations we can deduce that ρ_i is close in trace distance to a perfect Bell state (under a suitable isometry). The proof consists of a number of steps:

1. Firstly, we show that the observed correlations of Alice and Bob's, given fixed measurement settings, is a good estimate for the true quantum correlation assuming they shared multiple copies of the averaged state $\rho_{\mathrm{avg}} = \frac{1}{K}\sum_{i=1}^{K}\rho_i$.
2. Secondly, we use the previous result to estimate the correlations for the two measurement settings under consideration. We then use self-testing to show that if the correlations are close to the maximal value, the averaged state is close to a Bell state, under a suitable local isometry.
3. Lastly, we prove that if the averaged state is close to a Bell state (possibly in tensor product with some mixed state), then a typical state ρ_i is also close to that pure state and we compute the exact bound for the trace distance.

We start by proving the first step:

Lemma 3. *Assume Alice and Bob are asked to perform n rounds of measurement of the two-outcome observables with ± 1 eigenvalues, A and B, respectively. We denote the outcomes of their measurements as $\{a_i\}$ and $\{b_i\}$ and $\hat{C}_i = a_i b_i$ as their correlation for round i. Additionally, let $H_i = \{(a_j, b_j)\,|\, j < i\}$ be the history of their measurement outcomes up to, but not including round i. Finally, letting $C_i = E(\hat{C}_i|H_i)$ to be the conditional expectation value of the correlation given the previous history of outcomes, we have that for any $\delta > 0$:*

$$\Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n}C_i - \frac{1}{n}\sum_{i=1}^{n}\hat{C}_i\right| \geqslant \delta\right) \leqslant \exp(-\delta^2 n/8).$$

Proof. The variable C_i represents the true correlation of the outcomes in round i, as determined by the shared state of Alice and Bob. If the shared state in round i is ρ_i then $C_i = \mathrm{Tr}(AB\ \rho_i)$. As mentioned, while we trust Alice and know that she is indeed measuring the observable A, we can still assume that Bob is measuring the observable B in each round. This is because the observable B is unrestricted (apart from being a two-outcome observable) and can in principle act on Bob's ancilla as well. Furthermore, we make no assumption about the state ρ_i, since it is prepared by Bob. Another way in which we can express C_i is using its definition, which leads us to:

$$C_i = \Pr(a_i = b_i|H_i) - \Pr(a_i \neq b_i|H_i).$$

We now define the random variables:

$$X_j = \sum_{i=1}^{j}(C_i - \hat{C}_i).$$

Notice that for any $j \leqslant n$, $|X_{j+1} - X_j| \leqslant 2$ (because $\hat{C}_i = \pm 1, -1 \leqslant C_i \leqslant 1$), $E(X_j) \leqslant \infty$ and:

$$E(X_{j+1} - X_j|H_{j+1}) = C_{j+1} - C_{j+1} = 0.$$

Therefore, $\{X_j\}$ forms a martingale. We can therefore apply the Azuma–Hoeffding inequality [43], in a manner analogous to [11, 15]. Setting $j = n$, we have that for any $t > 0$:

$$\Pr(|X_n| > t) \leqslant \exp(-t^2/8n).$$

Expanding, we have that:

$$\Pr\left(\left|\sum_{i=1}^{n}(C_i - \hat{C}_i)\right| > t\right) \leqslant \exp(-t^2/8n).$$

For some $\delta > 0$, let $t = n\delta$. This yields:

$$\Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n}C_i - \frac{1}{n}\sum_{i=1}^{n}\hat{C}_i\right| > \delta\right) \leqslant \exp(-\delta^2 n/8).$$

Thus concluding the proof of lemma 3. □

We move on to the second step:

Lemma 4. *Suppose Alice and Bob are required to perform K rounds of measurement and also that:*

- *Prior to the measurements, the shared state of Alice and Bob is assumed to be σ.*
- *Alice chooses a random set of size $K/2$, consisting of indices from 1 to K and denoted $R_0 = \{i | i \in_R \{1 \ldots K\}\}$, $|R_0| = K/2$. We also denote $R_1 = \{1 \ldots K\} \setminus R_0$, to be the complement of R_0.*
- *We denote $\rho_i = \mathrm{Tr}_{-i}(\mathcal{E}^{AB}_{1,i-1}(\sigma))$ the reduced state of Alice and Bob in round i, and $\rho_{\mathrm{avg}} = \frac{1}{K}\sum_{i=1}^{K} \rho_i$ as the averaged state. Here $\mathcal{E}^{AB}_{1,i-1}$ denotes the action (measurements) of Alice and Bob on the state σ up to round i.*
- *In round i, let $r_i = 0$ iff $i \in R_0$, otherwise $r_i = 1$. Alice measures the observable A_{r_i} on her half of ρ_i. A_0 and A_1 are anti-commuting single-qubit observables having ± 1 eigenvalues.*
- *In round i, Bob is asked to measure B_{r_i}. B_0 and B_1 have ± 1 eigenvalues.*
- *We denote a_i and b_i, respectively, as the outcomes of their measurements in round i. We also denote $\hat{C}_i = a_i b_i$ as their correlation for round i.*
- *We denote $\hat{C}^0 = \frac{1}{K/2}\sum_{i \in R_0} \hat{C}_i$ and $\hat{C}^1 = \frac{1}{K/2}\sum_{i \in R_1} \hat{C}_i$ as the averaged correlations for the cases where both Alice and Bob are asked to measure the first observable, or both are asked to measure the second, respectively.*

If, for some given $\epsilon > 0$ and suitably chosen $K = \Omega((1/\epsilon^2)\log(1/\epsilon))$, it is the case that $\hat{C}^0 + \hat{C}^1 \geqslant 2 - \epsilon$ (or, alternatively, $\hat{C}^0 + \hat{C}^1 \leqslant -2 + \epsilon$) then there exists an isometry Φ and a mixed state ρ_{junk} such that:

$$||\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}})) - \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)\rho_{\mathrm{junk}}|| \leqslant O(\sqrt{\epsilon}),$$

where $\mathcal{E}^{AB}$ is some combination of the A_0, A_1, B_0, B_1 operators and $\hat{\mathcal{E}}^{AB}$ is the analogous combination of the ideal operators I, X, Y, as in theorem 2.

Proof. We will prove the case $\hat{C}^0 + \hat{C}^1 \geqslant 2 - \epsilon$ since for $\hat{C}^0 + \hat{C}^1 \leqslant -2 + \epsilon$ the derivation is similar. Additionally, we only consider the case $\mathcal{E}^{AB} = I$, since the other cases follow from the linearity of the operators. The previous lemma, essentially shows us that the observed average correlation is a good estimate for the average true correlation. Specifically, it is the case that $\hat{C}^b$, $b \in \{0, 1\}$, is close to the quantum correlation $\mathrm{Tr}(A_b B_b \rho_{\mathrm{avg}})$. Consider now a state $|\zeta\rangle$ which is a purification of ρ_{avg}. We can then write the quantum correlation as $\langle\zeta| A_b B_b |\zeta\rangle$. Using these results, if our estimate of the true correlation is of precision (closeness) $\delta > 0$, then it is the case that:

$$\langle\zeta| A_0 B_0 |\zeta\rangle + \langle\zeta| A_1 B_1 |\zeta\rangle \geqslant 2 - \epsilon - \delta.$$

With probability $1 - \exp(-\delta^2 K/16)$. Let $\delta = \epsilon$ so that we have:

$$\langle\zeta| A_0 B_0 |\zeta\rangle + \langle\zeta| A_1 B_1 |\zeta\rangle \geqslant 2 - O(\epsilon).$$

Using theorem 2, it follows that there exists a local isometry Φ and a state $|\mathrm{junk}\rangle$ such that, with probability $1 - \exp(-\epsilon^2 K/16)$, we have:

$$||\Phi(|\zeta\rangle) - |\psi_+\rangle|\mathrm{junk}\rangle|| \leqslant O(\sqrt{\epsilon}).$$

This also implies:

$$\mathrm{TD}(\Phi(|\zeta\rangle), |\psi_+\rangle|\mathrm{junk}\rangle) \leqslant O(\sqrt{\epsilon}).$$

As mentioned, we are only considering the case of I acting on the state $|\zeta\rangle$. Of course, the argument proceeds identically, when considering $M_A N'_B |\zeta\rangle$, as in theorem 1, leading to the $\mathcal{E}^{AB} \neq I$ cases. It should be noted that from the construction of Φ (in theorem 1), in the case where the shared state is a purification of some mixed state (as is the case with $|\zeta\rangle$ and ρ_{avg}), the isometry does not act on the quantum states used for purification. Therefore, we can trace out those states, and since this operation cannot increase trace distance we have that:

$$\mathrm{TD}(\Phi(\rho_{\mathrm{avg}}), |\psi_+\rangle\langle\psi_+| \rho_{\mathrm{junk}}) \leqslant O(\sqrt{\epsilon}).$$

With probability $1 - \exp(-\epsilon^2 K/16)$. We can incorporate this probability into the trace distance, and we have that:

$$\mathrm{TD}(\Phi(\rho_{\mathrm{avg}}), |\psi_+\rangle\langle\psi_+| \rho_{\mathrm{junk}}) \leqslant O(\sqrt{\epsilon}) + \exp(-\epsilon^2 K/16).$$

Setting $K = -(16/\epsilon^2)\log(\sqrt{\epsilon}) = (8/\epsilon^2)\log(1/\epsilon)$ we are left with:

$$\mathrm{TD}(\Phi(\rho_{\mathrm{avg}}), |\psi_+\rangle\langle\psi_+| \otimes \rho_{\mathrm{junk}}) \leqslant O(\sqrt{\epsilon}).$$

Analogously, we get:

$$\mathrm{TD}(\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}})), \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)\rho_{\mathrm{junk}}) \leqslant O(\sqrt{\epsilon}).$$

Concluding the proof of lemma 4. □

The final step towards proving theorem 4 is to prove the following statement:

Lemma 5. *Let ρ be a general quantum state which we can write as $\rho = \frac{1}{n}\sum_{i=1}^{n}\rho_i$, for some $n > 0$ and density matrices ρ_i. If it is the case that, for some pure state $|\phi\rangle$ and $\gamma > 0$:*

$$\mathrm{TD}(\rho, |\phi\rangle\langle\phi|) \leqslant \gamma.$$

Then for a uniformly at random chosen $i \in_R \{1 \ldots n\}$, with probability at least $1 - \gamma^{1/3}$ we have that:

$$\mathrm{TD}(\rho_i, |\phi\rangle\langle\phi|) \leqslant O(\gamma^{1/3}).$$

Proof. Starting with the bound on the trace distance and the relation between trace distance and fidelity (when one state is pure), we obtain:

$$\begin{aligned} 1 - F^2(\rho, |\phi\rangle) &\leqslant \mathrm{TD}(\rho, |\phi\rangle\langle\phi|) \leqslant \gamma \\ 1 - \gamma &\leqslant |\langle\phi| \rho |\phi\rangle| = F^2(\rho, |\phi\rangle) \\ &\leqslant \frac{1}{n}\sum_i F^2(\rho_i, |\phi\rangle) := \frac{1}{n}\sum_i q_i, \end{aligned} \tag{21}$$

where the second inequality follows from convexity and for convenience we defined $q_i = F^2(\rho_i, |\phi\rangle)$. This gives a lower bound on the average q_i (average fidelity squared). To provide an upper bound on the average q_i, we do the following. We let p be the fraction of i's such that $q_i \leqslant 1 - \gamma - \delta$, where $\delta \in [0, 1 - \gamma]$. Since $q_i \leqslant 1$ it follows:

$$(1 - \gamma - \delta) \cdot p + 1 \cdot (1 - p) \geqslant \frac{1}{n}\sum_i q_i. \tag{22}$$

From equations (21) and (22) it follows that

$$\begin{aligned} (1 - \gamma - \delta) \cdot p + 1 \cdot (1 - p) &\geqslant 1 - \gamma \\ \frac{\gamma}{\gamma + \delta} &\geqslant p. \end{aligned}$$

Now, using the fact that $\mathrm{TD}(\rho_i, |\phi\rangle\langle\phi|) \leqslant \sqrt{1 - q_i}$, we note that with probability $(1 - p)$ we have that $q_i \geqslant 1 - \gamma - \delta$ and thus for these cases $\mathrm{TD}(\rho_i, |\phi\rangle\langle\phi|) \leqslant \sqrt{\gamma + \delta}$. By choosing $\delta = \gamma^{2/3} - \gamma$ we have that with probability at least $1 - \gamma^{1/3}$:

$$\mathrm{TD}(\rho_i, |\phi\rangle\langle\phi|) \leqslant \sqrt{\delta + \gamma} = \gamma^{1/3}.$$

□

We can now use these lemmas to prove theorem 4.

Proof of theorem 4. We have the same assumptions as in lemma 4 and from it we know that after $K = \Omega((1/\epsilon^2)\log(1/\epsilon))$ rounds of measurements, if the observed correlations are saturated up to order ϵ we have:

$$\mathrm{TD}(\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}})), \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)\rho_{\mathrm{junk}}) \leqslant O(\sqrt{\epsilon}).$$

We would now like to apply lemma 5, however because of ρ_{junk} we do not have a pure state in the trace distance expression. Therefore, we trace out the junk subsystem, and since tracing out can only decrease trace distance we have that:

$$\mathrm{TD}(\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}}))), \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)) \leqslant O(\sqrt{\epsilon}).$$

If we denote $\rho = \mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}})))$, we can now apply lemma 5. Note that $\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\cdot)))$ is a linear map, since it is the composition of linear maps. Therefore, it is the case that:

$$\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_{\mathrm{avg}}))) = \sum_{i=1}^{K}\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_i))).$$

Hence, for a randomly chosen ρ_i, with probability at least $1 - O(\epsilon^{1/6})$:

$$\mathrm{TD}(\mathrm{Tr}_{\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_i))), \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)) \leqslant O(\epsilon^{1/6}).$$

Thus we have shown that, under a suitably chosen local isometry and tracing out any additional systems we can 'extract' a single Bell pair from the two device's shared state. If we want to also have a closeness relation for $\Phi(\mathcal{E}^{AB}(\rho_i))$, which includes all of Bob's private subsystem, we can use a corollary of the gentle measurement lemma:

Corollary 2. *[9, 10] Let ρ be a state on $\mathcal{H}_1 \otimes \mathcal{H}_2$, and let π be a pure state on $\mathcal{H}_1$. If for some $\delta \geqslant 0$, $\mathrm{Tr}(\pi\mathrm{Tr}_2(\rho)) \geqslant 1 - \delta$, then*

$$\mathrm{TD}(\rho, \pi \otimes \mathrm{Tr}_1(\rho)) \leqslant 2\sqrt{\delta} + \delta.$$

This leads to:

$$\mathrm{TD}(\Phi(\mathcal{E}^{AB}(\rho_i)), \hat{\mathcal{E}}^{AB}(|\psi_+\rangle\langle\psi_+|)\tilde{\rho}_{\mathrm{junk}}) \leqslant O(\epsilon^{1/12}),$$

where $\tilde{\rho}_{\mathrm{junk}} = \mathrm{Tr}_{-\mathrm{junk}}(\Phi(\mathcal{E}^{AB}(\rho_i)))$. □

Appendix E. Proof of theorem 5

As mentioned, we prove theorem 5 by showing that the actual strategy $\mathcal{S}$ of Alice and Bob is close to a strategy $\mathcal{S}_g$ in which Alice guesses Bob's outcomes, which in turn is close to the ideal strategy $\mathcal{S}_{\mathrm{id}}$.

Lemma 6. *Let $\mathcal{S} = (\rho, \{\mathcal{E}_i^A\}, \{\mathcal{E}_i^B\})$ be Alice and Bob's ϵ-structured strategy for playing N sequential K-round steering games in which Alice plays honestly. Let $\mathcal{S}_g = (\rho, \{\mathcal{E}_i^A\}, \{\mathcal{G}_i^B\})$ be Alice and Bob's ϵ-structured strategy for playing N sequential K-round steering games, in which Alice plays as in $\mathcal{S}$ but also guesses Bob's outcomes. Specifically, Bob's operator will be $\mathcal{G}_i^B$, which yields Alice's guesses for Bob's outcomes. We have that $\mathcal{S}$ $O(N\epsilon^{1/6})$-simulates $\mathcal{S}_g$.*

Proof. Without loss of generality, it can be assume that the initial state ρ is a pure state. Also note that Alice's action is given by $\{\mathcal{E}_i^A\}$ in both strategies, since she is always honest. We have that $\mathcal{S}_g = (\rho, \{\mathcal{E}_i^A\}, \{\mathcal{G}_i^B\})$, where $\mathcal{G}_i^B$ denotes the guessing operator for Bob's outcome in game i. Because Alice is trusted and playing honestly, her guessing strategy for Bob will be to provide the same outcomes as her measurement outcomes. This, of course, ensures that the steering correlations are saturated. Concretely, $\mathcal{G}_i^B$ leaves Bob's system unchanged but gives the same outcome as $\mathcal{E}_i^A$. Denote as $\mathcal{G}_i^{AB} = \mathcal{E}_i^A \circ \mathcal{G}_i^B$ the action of Alice and Bob on the state of the ith game, in strategy $\mathcal{S}_g$ and similarly $\mathcal{E}_i^{AB} = \mathcal{E}_i^A \circ \mathcal{E}_i^B$ the action of Alice and Bob on the state of the ith game, corresponding to the true strategy $\mathcal{S}$. Note that we can assume the same Hilbert space in both strategies. The reason for this is that Alice's side is trusted in the two strategies, so her Hilbert space is determined and fixed. On Bob's side, assume we have Hilbert space $\mathcal{H}_B$ in strategy $\mathcal{S}$. In strategy $\mathcal{S}_g$ we are ignoring Bob's outcomes and replacing them with Alice's guesses. Therefore, we can assume any Hilbert space on Bob's side, so without loss of generality we assume it is $\mathcal{H}_B$. Thus strategies $\mathcal{S}$ and $\mathcal{S}_g$ use the same Hilbert space.

We denote $\mathcal{G}_{1,i}^{AB} = \mathcal{G}_1^{AB} \circ \mathcal{G}_2^{AB} \circ \ldots \mathcal{G}_i^{AB}$ and similarly $\mathcal{E}_{1,i}^{AB} = \mathcal{E}_1^{AB} \circ \mathcal{E}_2^{AB} \circ \ldots \mathcal{E}_i^{AB}$. Additionally, according to definition 5, R is a set of random indices, each index taken from a different steering game. We can write $R = \{\kappa_i | i \in \{1 \ldots N\}, \kappa_i = (i-1)K + r, r \in_R \{1 \ldots K\}\}$. Essentially, κ_i is the selected random round for the ith steering game, out of the total number of rounds in the N games. Thus, $1 \leqslant \kappa_i \leqslant NK$. Lastly, we denote $R_j = \{\kappa_i | 1 \leqslant i \leqslant j\}$ to be the first j random indices (i.e. the randomly selected rounds up to game j). We would like to compute a bound for:

$$\mathrm{TD}(\mathrm{Tr}_{-R_j}(\mathcal{E}_{1,j}^{AB}(\rho)), \mathrm{Tr}_{-R_j}(\mathcal{G}_{1,j}^{AB}(\rho))).$$

We do this inductively, starting with the first steering game. Theorem 4 tells us that for a K-round steering game, if we observe a steering inequality saturation of order $O(\epsilon)$ the reduced state in a randomly chosen round will be $O(\epsilon^{1/6})$ close to a Bell state. Since we know that strategies $\mathcal{S}$ and $\mathcal{S}_g$ are ϵ-structured, it follows that for the first steering game we have that:

$$\mathrm{TD}(\mathrm{Tr}_{-R_1}(\Phi(\mathcal{E}_1^{AB}(\rho))), \hat{\mathcal{E}}_1^{AB}(|\psi_+\rangle\langle\psi_+|)) \leqslant O(\epsilon^{1/6}),$$

$$\mathrm{TD}(\mathrm{Tr}_{-R_1}(\Phi(\mathcal{G}_1^{AB}(\rho))), \hat{\mathcal{G}}_1^{AB}(|\psi_+\rangle\langle\psi_+|)) \leqslant O(\epsilon^{1/6}),$$

where Φ is the isometry of theorem 1, $\hat{\mathcal{E}}_1^{AB}$ and $\hat{\mathcal{G}}_1^{AB}$ are the ideal operators (note that in fact $\hat{\mathcal{G}}_1^{AB} = \mathcal{G}_1^{AB}$ since Alice is honest). However, since $I \otimes M\,|\psi_+\rangle = (XM^TX) \otimes I\,|\psi_+\rangle$, for any operator M, in $\hat{\mathcal{E}}_1^{AB}$ we can 'shift' Bob's action to Alice's side and so we are left with an operator which is ideal on Alice's side (conjugation by X is not a problem, since $XXX = X$ and $XYX = -Y$) and acts as identity on Bob's side. This is precisely the $\hat{\mathcal{G}}_1^{AB}$ operator, hence $\hat{\mathcal{E}}_1^{AB} = \hat{\mathcal{G}}_1^{AB}$. Therefore, both $\mathrm{Tr}_{-R_1}(\Phi(\mathcal{E}_1^{AB}(\rho)))$ and $\mathrm{Tr}_{-R_1}(\Phi(\mathcal{G}_1^{AB}(\rho)))$ are close to the same state, and through a simple application of the triangle inequality it follows:

$$\mathrm{TD}(\mathrm{Tr}_{-R_1}(\Phi(\mathcal{E}_1^{AB}(\rho))),\ \mathrm{Tr}_{-R_1}(\Phi(\mathcal{G}_1^{AB}(\rho)))) \leqslant O(\epsilon^{1/6}).$$

The specific isometry we have considered acts locally in each round (i.e., in each round it only acts on the reduced state of that round and introduces one ancilla qubit). Thus, it is possible in this case to commute the isometry with the tracing out operation, and then remove it completely from the inequality, since it is distance preserving. Of course, the tracing out operation will be acting on a different system (the non-isometrized system), requiring a different notation. In order to avoid complicating the notation, we use the same expression for the tracing out operation, and it is to be understood that we trace out all systems apart from the ones used in round 1. We therefore have:

$$\mathrm{TD}(\mathrm{Tr}_{-R_1}(\mathcal{E}_1^{AB}(\rho)),\ \mathrm{Tr}_{-R_1}(\mathcal{G}_1^{AB}(\rho))) \leqslant O(\epsilon^{1/6}).$$

This is the base case of our induction. Assume now the following holds:

$$\mathrm{TD}(\mathrm{Tr}_{-R_{j-1}}(\mathcal{E}_{1,j-1}^{AB}(\rho)),\ \mathrm{Tr}_{-R_{j-1}}(\mathcal{G}_{1,j-1}^{AB}(\rho))) \leqslant (j-1)O(\epsilon^{1/6}). \tag{23}$$

We would like to show:

$$\mathrm{TD}(\mathrm{Tr}_{-R_j}(\mathcal{E}_{1,j}^{AB}(\rho)),\ \mathrm{Tr}_{-R_j}(\mathcal{G}_{1,j}^{AB}(\rho))) \leqslant jO(\epsilon^{1/6}).$$

For the set of rounds from game j we can again apply theorem 4, and use the closeness bound for the state from round κ_j. Note that for the first game, the shared state of Alice and Bob was $|\psi\rangle$, while the state in game j need not be a pure state. We can, however, still apply theorem 4, regardless of the action of previous games, since the theorem assumes that Alice and Bob share some state σ which can be either pure or mixed. Moreover, as before, we will have that $\hat{\mathcal{E}}_{1,j}^{AB} = \hat{\mathcal{G}}_{1,j}^{AB}$, hence the reduced state in round κ_j from strategy $\mathcal{S}$ is $O(\epsilon^{1/6})$-close to the state in round κ_j from strategy $\mathcal{S}_g$. This together with equation (23) and a triangle inequality lead to:

$$\mathrm{TD}(\mathrm{Tr}_{-R_j}(\mathcal{E}_{1,j}^{AB}(\rho)),\ \mathrm{Tr}_{-R_j}(\mathcal{G}_{1,j}^{AB}(\rho))) \leqslant jO(\epsilon^{1/6}).$$

Since this is true for all $j \leqslant N$, we get:

$$\mathrm{TD}(\mathrm{Tr}_{-R}(\mathcal{E}_{1,N}^{AB}(\rho)),\ \mathrm{Tr}_{-R}(\mathcal{G}_{1,N}^{AB}(\rho))) \leqslant O(N\epsilon^{1/6}).$$

Hence, strategy $\mathcal{S}$ $O(N\epsilon^{1/6})$-simulates strategy $\mathcal{S}_g$. □

Lemma 7. *Let $\mathcal{S}_g = (\rho, \{\mathcal{E}_i^A\}, \{\mathcal{G}_i^B\})$ be Alice and Bob's ϵ-structured strategy for playing N sequential K-round steering games, in which Alice plays as in $\mathcal{S}$ but also guesses Bob's outcomes. Specifically, Bob's operator will be $\mathcal{G}_i^B$, which yields Alice's guesses for Bob's outcomes. Let $\mathcal{S}_{\mathrm{id}} = (\rho_{\mathrm{id}}, \{\mathcal{E}_i^A\}, \{\mathcal{E}_{\mathrm{id}\,i}^B\})$ be the ideal strategy in which ρ_{id} is a tensor product of Bell pairs and Alice and Bob play N sequential K-round steering games ideally (i.e. they measure the same operators on their shared state). We have that $\mathcal{S}_g$ $O(N\epsilon^{1/6})$-simulates an isometric extension of $\mathcal{S}_{\mathrm{id}}$, or $\mathcal{S}_g \approx \mathcal{S}_{\mathrm{id}}$.*

Proof. We shall first consider another guessing strategy $\hat{\mathcal{S}}_g = (\rho_{\mathrm{id}}, \{\mathcal{E}_i^A\}, \{\mathcal{G}_i^B\})$. This strategy is identical to $\mathcal{S}_g$ except for the fact that it uses the ideal state (the tensor product of Bell states) as opposed to the real state, ρ. Note that ρ_{id} must lie in the same Hilbert space as ρ. Without loss of generality, we can use some isometry to take a tensor product of Bell states and map it to the Hilbert space of ρ and denote that state as ρ_{id}. It is easy to see that $\mathcal{S}_g \approx \hat{\mathcal{S}}_g$. On the one hand, both strategies use the same operators for Alice and Bob (and in fact will produce identical statistics). On the other hand, since Alice is effectively guessing for Bob and the action on his subsystem is identity, there is no adaptivity in his strategy or outcomes. Combining this with the fact that the strategies are ϵ-structured, means that we can extract individual Bell pairs from each steering game and hence:

$$\mathrm{TD}(\mathrm{Tr}_{-R}(\mathcal{G}_{1,N}^{AB}(\rho)),\ \mathrm{Tr}_{-R}(\mathcal{G}_{1,N}^{AB}(\rho_{\mathrm{id}}))) \leqslant O(N\epsilon^{1/6}).$$

We now show $\hat{\mathcal{S}}_g \approx \mathcal{S}_{\mathrm{id}}$. First note that, as before, $\{\mathcal{E}_i^A\}$ corresponds to the ideal strategy for Alice and appears in both $\hat{\mathcal{S}}_g$ and $\mathcal{S}_{\mathrm{id}}$ since she is always honest. Additionally, as mentioned in the previous proof, because she is playing honestly according to the ideal strategy, her guesses for Bob's outcomes will be exactly her own measurement outcomes. This mimics the ideal strategy. In the ideal strategy Alice and Bob measure the same operators on their shared state. Moreover, both strategies are using the ideal state ρ_{id}. This directly implies that:

$$\mathrm{TD}(\mathrm{Tr}_{-R}(\mathcal{G}^{AB}_{1,N}(\rho_{\mathrm{id}})),\ \mathrm{Tr}_{-R}(\mathcal{E}^{AB}_{\mathrm{id}\,1,N}(\rho_{\mathrm{id}}))) \leqslant O(N\epsilon^{1/6})$$

Alternatively, we could have used the fact that $\hat{\mathcal{S}}_g$ and $\mathcal{S}_{\mathrm{id}}$ use the same state and both strategies are ϵ-structured, as in the previous lemma, to prove the same thing. Thus, it follows that strategy $\mathcal{S}_g$ $O(N\epsilon^{1/6})$-simulates an isometric extension of strategy $\mathcal{S}_{\mathrm{id}}$, or $\mathcal{S}_g \approx \mathcal{S}_{\mathrm{id}}$. □

It is worth mentioning that in the case of steering, unlike CHSH, the outcomes are deterministic, so Alice can guess perfectly for Bob. This led to a simplified rigidity proof compared to the one for CHSH games from [9]. Finally, we prove theorem 5.

Proof. From lemmas 6 and 7 we have that $\mathcal{S} \approx \mathcal{S}_g$ and that $\mathcal{S}_g \approx \mathcal{S}_{\mathrm{id}}$. Since in lemma 7 we have shown that $\mathcal{S}_g$ $O(N\epsilon^{1/6})$-simulates an isometric extension of $\mathcal{S}_{\mathrm{id}}$ we can consider an isometric extension of $\mathcal{S}_{\mathrm{id}}$, denoted $\mathcal{S}'_{\mathrm{id}}$ that has the same Hilbert space as strategy $\mathcal{S}$. Using a triangle inequality it follows that $\mathcal{S} \approx \mathcal{S}'_{\mathrm{id}}$. Thus $\mathcal{S}$ $O(N\epsilon^{1/6})$-simulates an isometric extension of the ideal strategy $\mathcal{S}_{\mathrm{id}}$. □

Appendix F. Proof of theorem 6

It is clear that a maximally entangled Bell state satisfies the properties of total steerability. We therefore focus on proving that a totally steerable state is maximally entangled. From the two constraints of definition 7 we will express the most general form of ρ_{AB}. We start by considering $|\psi_{ABC}\rangle$ as the 4-qubit purification of ρ_{AB}. All other purifications are equivalent to this one, so this suffices for our purposes. Writing $|\psi_{ABC}\rangle$ in the computational basis, we have:

$$|\psi_{ABC}\rangle = \sum_{i=0}^{15} a_i\ |i\rangle.$$

Of course, we have the additional constraint:

$$\sum_{i=0}^{15} |a_i|^2 = 1.$$

By re-expressing the constraints from definition 7 and lemma 2:

$$\rho_{BC} = \rho_B \otimes \rho_C,$$
$$\rho_B = I/2$$

as constraints on the amplitudes of $|\psi_{ABC}\rangle$ we will have a large bilinear system of equations. From this system we will arrive at the following set of equations:

$$\begin{aligned} a_{2k} &= f \cdot a_{2k+2}, \quad k \in \{0, 1, \ldots 6\}, \\ a_{2k+3} &= -f^* \cdot a_{2k+1}, \quad k \in \{0, 1, \ldots 6\}, \\ a_{4k+1} &= \mathrm{e}^{\mathrm{i}\phi_1} \cdot a_{4k+2}, \quad k \in \{0, 1, 2, 3\}, \\ a_{4k} &= \mathrm{e}^{\mathrm{i}\phi_2} \cdot a_{4k+3}, \quad k \in \{0, 1, 2, 3\}, \end{aligned}$$

where the parameters we introduced are $f \in \mathbb{C}$ and $\phi_1,\ \phi_2 \in [0,\ 2\pi]$. Computing the matrix elements of ρ_{AB}, we arrive at the most general form, given by:

$$\rho_{AB} = \frac{1}{2(|f|^2 + 1)} \begin{pmatrix} |f|^2 & f & f & \mathrm{e}^{\mathrm{i}\phi_1}|f|^2 \\ f^* & 1 & \mathrm{e}^{\mathrm{i}\phi_2} & -f \\ f^* & \mathrm{e}^{-\mathrm{i}\phi_2} & 1 & -f \\ \mathrm{e}^{-\mathrm{i}\phi_1}|f|^2 & -f^* & -f^* & |f|^2 \end{pmatrix}.$$

It can be easily checked that $\mathrm{Tr}(\rho^2_{AB}) = 1$ and therefore ρ_{AB} is a pure state. But since $\rho_B = I/2$, we have that ρ_{AB} is a pure entangled state. We know that all pure entangled states are non-local. Moreover, the condition $\rho_B = I/2$ also implies that the state is maximally non-local and hence a Bell state.

References

[1] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 777–80
[2] Cirel'son B S 1980 Quantum generalizations of Bell's inequality *Lett. Math. Phys.* **4** 93–100
[3] Reichardt B W, Unger R F and Vazirani U 2013 Classical command of quantum systems *Nature* **496** 456–60
[4] Schrödinger E 1936 Probability relations between separated systems *Math. Proc. Camb. Phil. Soc.* **32** 446–52
[5] Bell J S 1964 On the Einstein–Podolsky–Rosen paradox *Physics* **1** 195–200

[6] Zurek W H 2000 Einselection and decoherence from an information theory perspective *Ann. Phys., Lpz.* **9** 855–64
[7] Ollivier H and Zurek W H 2001 Quantum discord: a measure of the quantumness of correlations *Phys. Rev. Lett.* **88** 017901
[8] Henderson L and Vedral V 2001 Classical, quantum and total correlations *J. Phys.* A **34** 6899
[9] Reichardt B W, Unger F and Vazirani U 2012 A classical leash for a quantum system: command of quantum systems via rigidity of CHSH games arXiv:1209.0448
[10] Gheorghiu A, Kashefi E and Wallden P 2015 Robustness and device independence of verifiable blind quantum computing *New J. Phys.* **17** 083040
[11] Hajdušek M, Pérez-Delgado C A and Fitzsimons F J 2015 Device-independent verifiable blind quantum computation arXiv:1502.02563
[12] McKague M 2016 Interactive proofs for BQP via self-tested graph states *Theory of Computing* **12** 1–42
[13] Vazirani U and Vidick T 2014 Fully device-independent quantum key distribution *Phys. Rev. Lett.* **113** 140501
[14] Acín A, Brunner N, Gisin N, Massar S, Pironio S and Scarani V 2007 Device-independent security of quantum cryptography against collective attacks *Phys. Rev. Lett.* **98** 230501
[15] Pironio S *et al* 2010 Random numbers certified by Bell's theorem *Nature* **464** 1021–4
[16] Bouda J, Pawłowski M, Pivoluska M and Plesch M 2014 Device-independent randomness extraction from an arbitrarily weak min-entropy source *Phys. Rev.* A **90** 032313
[17] Pironio S, Acín A, Brunner N, Gisin N, Massar S and Scarani V 2009 Device-independent quantum key distribution secure against collective attacks *New J. Phys.* **11** 045021
[18] Cavalcanti E G, Jones S J, Wiseman H M and Reid M D 2009 Experimental criteria for steering and the Einstein–Podolsky–Rosen paradox *Phys. Rev.* A **80** 032112
[19] Skrzypczyk P, Navascués M and Cavalcanti D 2014 Quantifying Einstein–Podolsky–Rosen steering *Phys. Rev. Lett.* **112** 180404
[20] Jevtic S, Pusey M, Jennings D and Rudolph T 2014 Quantum steering ellipsoids *Phys. Rev. Lett.* **113** 020402
[21] Wiseman H M, Jones S J and Doherty A C 2007 Steering, entanglement, nonlocality, and the Einstein–Podolsky–Rosen paradox *Phys. Rev. Lett.* **98** 140402
[22] Pusey M F 2013 Negativity and steering: a stronger peres conjecture *Phys. Rev.* A **88** 032313
[23] Branciard C, Cavalcanti E G, Walborn S P, Scarani V and Wiseman H M 2012 One-sided device-independent quantum key distribution: security, feasibility, and the connection with steering *Phys. Rev.* A **85** 010301
[24] Passaro E, Cavalcanti D, Skrzypczyk P and Acín A 2015 Optimal randomness certification in the quantum steering and prepare-and-measure scenarios *New J. Phys.* **17** 113010
[25] Wittmann B, Ramelow S, Steinlechner F, Langford N K, Brunner N, Wiseman H M, Ursin R and Zeilinger A 2012 Loophole-free Einstein–Podolsky–Rosen experiment via quantum steering *New J. Phys.* **14** 053030
[26] Saunders D J, Jones S J, Wiseman H M and Pryde G J 2010 Experimental epr-steering using Bell-local states *Nat. Phys.* **6** 845–9
[27] Horodecki M, Horodecki P and Horodecki R 1998 Mixed-state entanglement and distillation: Is there a bound entanglement in nature? *Phys. Rev. Lett.* **80** 5239–42
[28] Horodecki K, Horodecki M, Horodecki P and Oppenheim J 2009 General paradigm for distilling classical key from quantum states *IEEE Trans. Inf. Theory* **55** 1898–929
[29] Kraus B, Gisin N and Renner R 2005 Lower and upper bounds on the secret-key rate for quantum key distribution protocols using one-way classical communication *Phys. Rev. Lett.* **95** 080501
[30] Colbeck R and Kent A 2011 Private randomness expansion with untrusted devices *J. Phys. A: Math. Theor.* **44** 095305
[31] Colbeck R and Renner R 2012 Free randomness can be amplified *Nat. Phys.* **8** 450–3
[32] Aharonov D, Ben-Or M and Eban E 2010 Interactive proofs for quantum computations *Proc. Innovations in Computer Science 2010, ICS2010* pp 453
[33] Gennaro R, Gentry C and Parno B 2010 Non-interactive verifiable computing: outsourcing computation to untrusted workers *Proc. 30th Annual Conf. on Advances in Cryptology, CRYPTO'10* (Berlin: Springer) pp 465–82
[34] Cavalcanti E G, Hall M J W and Wiseman H M 2013 Entanglement verification and steering when Alice and Bob cannot be trusted *Phys. Rev.* A **87** 032306
[35] Kocsis S, Hall M J W, Bennet A J, Saunders D J and Pryde G J 2015 Experimental measurement-device-independent verification of quantum steering *Nat. Commun.* **6** 5886
[36] McKague M, Yang T H and Scarani V 2012 Robust self-testing of the singlet *J. Phys. A: Math. Theor.* **45** 455304
[37] Mayers D and Yao A 2004 Self testing quantum apparatus *Quantum Info. Comput.* **4** 273–86
[38] McKague M 2014 Self-testing graph states *Theory of Quantum Computation, Communication, and Cryptography* (*Lecture Notes in Computer Science* vol 6745) ed D Bacon *et al* (Berlin: Springer) pp 104–20
[39] McKague M and Mosca M 2011 Generalized self-testing and the security of the 6-state protocol *Proc. 5th Conf. on Theory of Quantum Computation, Communication, and Cryptography, TQC'10* (Berlin: Springer) pp 113–30
[40] Miller C A and Shi Y 2013 Optimal robust quantum self-testing by binary nonlocal xor games *Proc. 8th Conf. Theory of Quantum Computation, Communication and Cryptography (TQC 2013)Leibniz International Proceedings in Informatics (LIPIcs)* 254–62
[41] Yang T H and Navascués M 2013 Robust self-testing of unknown quantum systems into any entangled two-qubit states *Phys. Rev.* A **87** 050102
[42] Bamps C and Pironio S 2015 Sum-of-squares decompositions for a family of Clauser–Horne–Shimony–Holt-like inequalities and their application to self-testing *Phys. Rev.* A **91** 052111
[43] Azuma K 1967 Weighted sums of certain dependent random variables *Tohoku Math. J.* 2 **19** 357–67
[44] Hoeffding W 1963 Probability inequalities for sums of bounded random variables *J. Am. Stat. Assoc.* **58** 13–30
[45] Smith D H *et al* 2012 Conclusive quantum steering with superconducting transition-edge sensors *Nat.* **3** 625
[46] Cavalcanti E, Jones S, Wiseman H and Reid M 2009 Experimental criteria for steering and the Einstein–Podolsky–Rosen paradox *Phys. Rev.* A **80** 032112
[47] Allati A E, Baz M E and Hassouni Y 2011 Quantum key distribution via tripartite coherent states *Quantum Inform. Process.* **10** 589–602
[48] Gabriel C, Wittmann C, Sych D, Dong R, Mauerer W, Andersen U L, Marquardt C and Leuchs G 2010 A generator for unique quantum random numbers based on vacuum states *Nat. Photon.* **4** 711–5
[49] Jofre M, Curty M, Steinlechner F, Anzolin G, Torres J P, Mitchell M W and Pruneri V 2011 True random numbers from amplified quantum vacuum *Opt. Express* **19** 20665–72
[50] Bancal J-D, Navascués M, Scarani V, Vértesi T and Yang T H 2015 Physical characterization of quantum devices from nonlocal correlations *Phys. Rev.* A **91** 022115
[51] Šupić I and Hoban J M 2016 Self-testing through epr-steering *New J. Phys.* **18** 075006

[52] Goldwasser S, Micali S and Rackoff C 1989 The knowledge complexity of interactive proof systems *SIAM J. Comput.* **18** 186–208
[53] Aaronson S 2015 The Scott Aaronson 25.00 Prize http:scottaaronson.com/blog/?p=284 (accessed 30 January)
[54] Fitzsimons J F and Kashefi E 2012 Unconditionally verifiable blind computation arXiv:1203.5217
[55] Kapourniotis T, Kashefi E and Datta A 2014 Blindness and verification of quantum computation with one pure qubit *9th Conf. on the Theory of Quantum Computation, Communication and Cryptography (TQC 2014)* vol 27, pp 176–204
[56] Morimae T 2014 Verification for measurement-only blind quantum computing *Phys. Rev.* A **89** 060302
[57] Barz S, Fitzsimons J F, Kashefi E and Walther P 2013 Experimental verification of quantum computation *Nat. Phys.* **9** 727–31
[58] Kashefi E and Wallden P 2015 Optimised resource construction for verifiable quantum computation *J. Phys.* A accepted (https://doi.org/10.1088/1751-8121/aa5dac)
[59] Broadbent A, Fitzsimons J and Kashefi E 2009 universal blind quantum computation *Proc. 50th Annual Symp. on Foundations of Computer Science, FOCS '09* (IEEE Computer Society) pp 517–26
[60] Danos V, Kashefi E and Panangaden P 2007 The measurement calculus *J. ACM* **54** 8
[61] Fujii K and Morimae T 2013 Blind quantum computation protocol in which Alice only makes measurements *Phys. Rev.* A **87** 050301
[62] Hayashi M and Morimae T 2015 Verifiable measurement-only blind quantum computing with stabilizer testing *Phys. Rev. Lett.* **115** 220502
[63] Kapourniotis T, Dunjko V and Kashefi E 2015 On optimising quantum communication in verifiable quantum computing *15th Asian Quantum Information Science Conference* pp 23–5
[64] Pusey M F 2013 Is quantum steering spooky? *PhD Thesis,* Imperial College London, London
[65] Reid M D 2013 Monogamy inequalities for the Einstein–Podolsky–Rosen paradox and quantum steering *Phys. Rev.* A **88** 062108
[66] Milne A, Jevtic S, Jennings D, Wiseman H and Rudolph T 2014 Quantum steering ellipsoids, extremal physical states and monogamy *New J. Phys.* **16** 083017