Robust self-testing of the singlet

M McKague1, T H Yang1 and V Scarani1,2

1 Centre for Quantum Technologies, National University of Singapore, 3 Science drive 2, Singapore 117543
2 Department of Physics, National University of Singapore, 2 Science drive 3, Singapore 117542

E-mail: matthew.mckague@nus.edu.sg, haur@nus.edu.sg and physv@nus.edu.sg

Received 9 July 2012, in final form 21 September 2012
Published 19 October 2012
Online at stacks.iop.org/JPhysA/45/455304

Abstract
In this paper, we introduce a general framework to study the concept of robust self-testing which can be used to self-test maximally entangled pairs of qubits (EPR pairs) and local measurement operators. The result is based only on probabilities obtained from the experiment, with tolerance to experimental errors. In particular, we show that if the results of an experiment approach the Cirel’son bound, or approximate the Mayers–Yao-type correlations, then the experiment must contain an approximate EPR pair. More specifically, there exist local bases in which the physical state is close to an EPR pair, possibly encoded in a larger environment or ancilla. Moreover, in these bases the measurements are close to the qubit operators used to achieve the Cirel’son bound or the Mayers–Yao results.

PACS number: 03.67.Ac

1. Introduction

It is well known by now that the correlations obtained by measuring entangled quantum systems cannot be reproduced with classical resources. In fact, for some of these correlations, a much stronger statement holds: they can be reproduced only by measuring a specific quantum state in a specific way. To date, two such examples are known for the bipartite case. One uses the Clauser–Horne–Shimony–Holt (CHSH) criterion [1] to state the following: if the maximal quantum value \( CHSH = 2\sqrt{2} \) [2] is observed, then the state being measured is necessarily equivalent (in a sense to be made rigorous below) to a maximally entangled state of two qubits, which will be referred to as ‘singlet’ from now on. Moreover, both the measurements on Alice and on Bob must anti-commute [3, 4]. The other criterion is due to Mayers and Yao: it uses a different observation to reach the same conclusion [5]. In particular, Mayers and Yao consider each individual correlation of all the outcomes and from it deduce algebraic relations between the different projectors. From there, one can also self-test the system. Since the Mayers–Yao
correlations cannot reach CHSH = 2√2, the two criteria are inequivalent. Compactly, we shall say that these two criteria realize the self-testing of the singlet and of some measurements.

The possibility of self-testing is all the more remarkable because nothing is assumed \textit{a priori} on the physical system or on the measurements, not even the dimension of the relevant Hilbert space: in principle, these are device-independent assessments, based only on the observed statistical data. Device-independent assessment has been discussed in various scenarios [6–8], including adversarial ones, which may provide the ultimate test of trustfulness. More realistic, and probably more relevant for today’s physics, is a scenario in which neither the experimentalists nor nature are assumed to cheat, but where one wants a simple and direct check that nothing serious is going wrong, that there are no undesired side channels, etc.

In order to be practical, a self-testing procedure must be robust, i.e. tolerate deviations from the theoretically ideal case. A mathematical \textit{tour de force} has recently provided a robustness bound for the Mayers–Yao test [7]. To our knowledge, no robust bound is available for the CHSH test, a situation that plagues the applicability of the corresponding device-independent assessment of entanglement of a source [9] and a measurement [10].

In this paper, we prove a general sufficient criterion for a set of correlations to provide robust self-testing of the singlet. Then we prove that both the CHSH and the Mayers–Yao tests satisfy this criterion and give the explicit bounds. The proofs use rather elementary quantum mathematics, following the simplification of the Mayers–Yao proof by one of us [11–13].

2. Definitions and notation

We are aiming at self-testing the presence of a maximally entangled state of two qubits in unknown devices. This goal calls for a suitable definition. Indeed, there is nothing like an isolated qubit in nature: if one wants to measure the spin of an electron, then the whole electron with its wavefunction should be present; and if the qubit is the polarization of an optical mode, then we ought to allow the whole electromagnetic field to be present. So there will surely be degrees of freedom which do not encode the state of interest, but are nevertheless present. Also, there must be a local frame of reference for each device in order to define the measurements. Because of these two facts, our definition must allow for additional ancillas and local changes of basis. We do so by using an isometry, that is a linear map \( \Phi : \mathcal{H}_1 \to \mathcal{H}_2 \) that preserves inner products. Note that \( \mathcal{H}_1 \) and \( \mathcal{H}_2 \) can have different dimensions. As a concrete example, adding an ancilla and applying a unitary to the total system is an isometry.

Now we are ready to formalize our definition, first proposed using slightly different formalism by Mayers and Yao in [5]. We say that a pair of devices \( A \) and \( B \) hold a pair of maximally entangled qubits, \( |\phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \), if there exists a local isometry \( \Phi = \Phi_A \otimes \Phi_B \) that takes the state \( |\psi\prime\rangle_{AB} \) to

\[
\Phi (|\psi\prime\rangle_{AB}) = |\text{junk}\rangle_{AB}|\phi_+\rangle_{AB} \tag{1}
\]

and physical observables \( M'_A \) and \( N'_B \) operate as

\[
\Phi (M'_A N'_B |\psi\rangle_{AB}) = |\text{junk}\rangle_{AB} M'_A N'_B |\phi_+\rangle_{AB}, \tag{2}
\]

for some \( M_A \) and \( N_B \) to be specified later. In other words, we aim at checking that there exists a choice of local bases such that (i) the state looks like an ancilla tensored with a maximally entangled pair of qubits; and (ii) the measurements act non-trivially only on the pair of qubits.

At this point of time, it is important to clarify our choice of self-testing criteria. In the device-independent scenario, it is impossible to specify any structure on the state space of the devices, even the dimension. Hence, if we wish to identify a qubit state inside the state space of the device, we must somehow introduce the qubit structure. The justification for our
particular definition is twofold. First, in the exact case, as we shall see, a strategy saturates\nCsirleson’s bound if and only if it satisfies our definition (and analogously for the Mayers–\nYao measurements). This implies that any other definition with this property is equivalent\nto ours. Secondly, our robustness bounds can be translated directly into statements about\nprobabilities of outcomes, etc since isometries preserve inner products. Hence, our definition\nis operationally relevant.

A word about notation: we use primed notation ($X'$, $|\psi'$) etc to represent the observables
and states in the actual quantum devices. These will be unknown (even their dimensions)\nexcept for a few properties that we shall specify. Non-primed operators $X$ and $Z$ refer to Pauli
operators.

3. Circuit for self-testing

We start by presenting a set of sufficient conditions to self-test the singlet along with the
associated measurement operators. The state $|\psi'$ can be taken as pure without loss of\ngenerality, since the dimension is not fixed and one can always add the ancillas for purification.
We assume further that the state is always the same in each run of the experiment, which is
reasonable in the non-adversarial scenario. (This assumption could be removed, for instance,
using Azuma’s inequality as in [14].) The measurement settings are denoted by [$A'_0, A'_1, \ldots$] on
Alice’s side and [$B'_0, B'_1, \ldots$] on Bob’s side. For all that follows, it is a crucial assumption that
[$A'_j, B'_j] = 0$: this can ultimately be enforced by the space-like separation of the measurement;
but one may be less demanding and take simple spatial separation as a sufficient guarantee of
commutation.

With these notations and assumptions, the following theorem holds.

**Theorem 1.** Suppose that from the observed correlations, one can deduce the existence of
local observables $\{X'_i, Z'_i\} \text{ (functions of } A'_i\text{)}$ and $\{X'_j, Z'_j\} \text{ (functions of } B'_i\text{)}$ with eigenvalues
$\pm 1$, which act on the bipartite state $|\psi'$ such that

$$
\| (X'_i Z'_i + Z'_i X'_i)|\psi'\rangle \| \leq 2\epsilon_1,
$$
(3)

$$
\| (X'_j Z'_j + Z'_j X'_j)|\psi'\rangle \| \leq 2\epsilon_1,
$$
(4)

$$
\| (X'_i - X'_j)|\psi'\rangle \| \leq \epsilon_2,
$$
(5)

$$
\| (Z'_i - Z'_j)|\psi'\rangle \| \leq \epsilon_2.
$$
(6)

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

$$
\| \Phi (M'_i N'_j |\psi'\rangle - |\text{junk}\rangle_{AB} M_A N_B |\phi_+\rangle_{AB}\| \leq \epsilon,
$$
(7)

for $M, N \in \{I, X, Z\}$ and $\epsilon = (11\epsilon_1 + 5\epsilon_2)/2$.

**Proof.** The details of the derivations are relatively straightforward and hence we include them
in appendix A in order not to interrupt the main ideas within the text. The details are meant to
be read along with the main text and to give enough information for the reader to reproduce
the proofs quickly. Since the proofs follow almost entirely from the definition and elementary
inequalities, a line by line account was not felt to be necessary.

We sketch the proof with the technical details available in appendix A. The isometry is
constructed as shown in figure 1. For the case where $M = N = I$, the isometry gives

$$
\Phi(|\psi'\rangle) = \frac{1}{2} (I + Z'_a)(I + Z'_b)|\psi'\rangle_{00} + \frac{1}{2} X'_a (I + Z'_a)(I - Z'_b)|\psi'\rangle_{01}
+ \frac{1}{2} X'_a (I - Z'_a)(I + Z'_b)|\psi'\rangle_{10} + \frac{1}{2} X'_a (I - Z'_a)(I - Z'_b)|\psi'\rangle_{11}.
$$
(8)
From (6), we know that $Z_A'|\psi'\rangle \approx Z_B'|\psi'\rangle$. The projectors $(I \pm Z_A')$ and $(I \mp Z_B')$ then project the state $|\psi'\rangle$ into approximately orthogonal subspaces. As a result, the terms $(I \pm Z_A')(I \mp Z_B')|\psi'\rangle$ will be small.

Now we are left with the first and last terms in (8), i.e. the correlated parts. However, we need to show that there is no entanglement between the qubit registers and the physical register. To this end, recall (3) and (4), which state that $(X'_A, Z'_A)$ and $(X'_B, Z'_B)$ both anticommute approximately, again in the subspace spanned by $|\psi'\rangle$. Therefore, the last term in (8) is approximately equal to $(I + Z'_A)(I + Z'_B)X'_A X'_B|\psi'\rangle|11\rangle$. Finally, from (5), we have $(I + Z'_A)(I + Z'_B)X'_A X'_B|\psi'\rangle|11\rangle \approx (I + Z'_A)(I + Z'_B)|\psi'\rangle|11\rangle$. This allows us to factor out the Bell state, achieving our definition.

This theorem shows that the premises allow us to conclude that there exist some local bases in which, up to some small error, the state shared by the quantum devices is a singlet together with some ancillas in an unknown state, and the derived operators $X'$ and $Z'$ operate only on the singlet. Additionally, in these local bases, $X'$ and $Z'$ are close to the Pauli $X$ and $Z$.

In the remainder of the paper, we are going to show that (3)–(6) follow from both the CHSH and Mayers–Yao correlation experiments; therefore, both experiments can be used for robust self-testing.

4. Robust self-testing using CHSH

As mentioned above, a robustness bound for the CHSH-based self-testing was missing. We provide it here.

**Theorem 2.** Suppose that the observables $A'_0, A'_1, B'_0$ and $B'_1$ with eigenvalues $\pm 1$, acting on a state $|\psi'\rangle$, are such that

$$
|\langle \psi'|(A'_0B'_0 + A'_1B'_1 + A'_1B'_0 - A'_0B'_1)|\psi'\rangle| \geq 2\sqrt{2} - \epsilon,
$$

where $0 < \epsilon < 1$. Then the conditions of theorem 1 are satisfied with $\epsilon_1 = 2(\epsilon \sqrt{2})^{1/2}$ and $\epsilon_2 = 4(\epsilon \sqrt{2})^{1/4}$.

**Proof.** To establish the theorem, we need to show the existence of four local, Hermitian and unitary operators $X'_A, Z'_A, X'_B$ and $Z'_B$ that satisfy (3)–(6). We are going to show this for

$$
X'_A = A'_0, \quad Z'_A = A'_1,
X'_B = \frac{B'_0 + B'_1}{|B'_0 + B'_1|}, \quad Z'_B = \frac{B'_0 - B'_1}{|B'_0 - B'_1|}.
$$

Figure 1. Local isometry $\Phi$, where $M, N \in \{I, X, Z\}$. 

From (6), we know that $Z'_A|\psi'\rangle \approx Z'_B|\psi'\rangle$. The projectors $(I \pm Z'_A)$ and $(I \mp Z'_B)$ then project the state $|\psi'\rangle$ into approximately orthogonal subspaces. As a result, the terms $(I \pm Z'_A)(I \mp Z'_B)|\psi'\rangle$ will be small.

Now we are left with the first and last terms in (8), i.e. the correlated parts. However, we need to show that there is no entanglement between the qubit registers and the physical register. To this end, recall (3) and (4), which state that $(X'_A, Z'_A)$ and $(X'_B, Z'_B)$ both anticommute approximately, again in the subspace spanned by $|\psi'\rangle$. Therefore, the last term in (8) is approximately equal to $(I + Z'_A)(I + Z'_B)X'_A X'_B|\psi'\rangle|11\rangle$. Finally, from (5), we have $(I + Z'_A)(I + Z'_B)X'_A X'_B|\psi'\rangle|11\rangle \approx (I + Z'_A)(I + Z'_B)|\psi'\rangle|11\rangle$. This allows us to factor out the Bell state, achieving our definition.

This theorem shows that the premises allow us to conclude that there exist some local bases in which, up to some small error, the state shared by the quantum devices is a singlet together with some ancillas in an unknown state, and the derived operators $X'$ and $Z'$ operate only on the singlet. Additionally, in these local bases, $X'$ and $Z'$ are close to the Pauli $X$ and $Z$.

In the remainder of the paper, we are going to show that (3)–(6) follow from both the CHSH and Mayers–Yao correlation experiments; therefore, both experiments can be used for robust self-testing.
where $|M\rangle = \sqrt{M}$ clearly, they are all unitary and Hermitian. Moreover, $|X'_A, Z'_B\rangle = 0$ by construction, thus establishing a tighter version of (4). All the subsequent steps are again somehow pedestrian, so we sketch them here and leave the full details for appendix B.

From (9), a suitable use of the Cauchy–Schwartz and the triangle inequalities leads to

$$\|\langle A'_N, A'_N'\rangle|\psi\rangle\| \leq 2\epsilon_1,$$  
(11)

$$\|\{(B'_N, B'_N')|\psi\rangle\| \leq 2\epsilon_1$$  
(12)

with $\epsilon_1 = 2\sqrt{\epsilon \sqrt{2}}$. Then (3) is established in (11).

The third condition (5) is proved by obtaining first the bound $\|\langle X'_N - (B'_N + B'_N)/\sqrt{2}\rangle|\psi\rangle\| \leq 2(\epsilon \sqrt{2})^{1/4}$, then the same bound for $\|\langle X'_N - (B'_N + B'_N)/\sqrt{2}\rangle|\psi\rangle\|$; both derivations using (9) at one point. The triangle inequality completes the estimate. The proof of (6) follows the same steps.

Note that inequality (7) applies for $M'_N = A_0 = X'_A$ and $M'_A = A_1 = Z'_A$, $N'_B = X'_B$ and $N'_B = Z'_B$. One may want to have a self-testing bound for the operators that are really measured, $B'_0$ and $B'_1$, which are not linear functions of the previous ones. The inequality for $B'_0$ is found by using linearity in the estimations $\|\langle X'_N - (B'_0 + B'_1)/\sqrt{2}\rangle|\psi\rangle\| \leq 2(\epsilon \sqrt{2})^{1/4}$ and $\|\langle Z'_N - (B'_0 + B'_1)/\sqrt{2}\rangle|\psi\rangle\| \leq 2(\epsilon \sqrt{2})^{1/4}$ from the proof, then using the fact that isometries preserve the 2-norm. We obtain

$$\|\Phi (M'_N B'_N)\langle \psi\rangle) - \langle \text{junk}\rangle_{AB} M'_A \frac{X_B + Z_B}{\sqrt{2}} \frac{(\phi_+, AB)\| \leq \sqrt{2}\epsilon + 2\sqrt{2}(\epsilon \sqrt{2})^{1/4}.$$  
(13)

The analogous result holds for $B'_1$.

Thus, we have effectively shown that any experimental settings with CHSH violation close to $2\sqrt{2}$ must contain approximately a Bell pair with respective measurements close to the optimal settings. In the exact case, our result implies that of (3), wherein the authors show that all states which violate maximally CHSH are of even dimensional. Indeed, since a local isometry cannot change the dimension of the support of the state, and since the support of $|\psi\rangle$ has even dimension, the support of $|\psi\rangle$ must also have even dimension. Our result also takes into account the results of [15] showing that there are odd-dimensional states violating CHSH arbitrarily close to the Tsirelson bound. In fact, these states are all close to even-dimensional states, and the difference is captured by our robustness bounds.

5. Robust self-testing using the Mayers–Yao criterion

We now turn to the robustness bound for the Mayers–Yao correlations. The original scenario uses three measurements on Alice’s side and three on Bob’s side; as a matter of fact though, only two measurements are needed by (say) Alice, so we work in this more economic case.

We have the following theorem.

**Theorem 3.** Let $0 < \epsilon < 1$ be given and let a bipartite state $|\psi\rangle$ and observables $X'_A$, $Z'_A$, $X'_B$, $Z'_B$ and $D'_B$ with eigenvalues $\pm 1$ be given such that

$$|\langle \psi| M'_A N'_B |\psi\rangle - \langle \phi_+ | M'_A N'_B |\phi_+\rangle| \leq \epsilon$$  
(14)

holds for all $M \in \{X, Z\}$ and $N \in \{X, Z, D\}$, where $D = (X + Z)/\sqrt{2}$. Then the conditions of theorem 1 are satisfied with $\epsilon_1 = 2(1 + \sqrt{2})(2\epsilon)^{1/4} + 4\sqrt{2\epsilon} + \frac{4 + \sqrt{2}}{2}(2\epsilon)^{3/4}$ and $\epsilon_2 = \sqrt{2\epsilon}.$

3 If $M$ has a subspace with eigenvalue 0, then the eigenvalue of $M^2/M$ in that subspace is taken to be 1.
The proof is included in appendix C. The critical step is to use the strong non-classical correlations between $D_B'$ and both $X_A'$ and $Z_A'$ to establish that the latter approximately anticommute. Interestingly, $D_B'$ is used only in this task, and does not play any role in the other estimations or the construction of the isometry.

6. Discussion

We have presented robust self-testing bounds for the singlet from the two criteria of CHSH and Mayers–Yao. These tools can be used for the device-independent assessment of state preparation and measurement devices, in protocols that are based on these criteria. For instance, in [9] the authors defined the ‘Mayers–Yao’ fidelity, $F_{MY}$ defined as

$$F_{MY}(|\psi\rangle) \equiv \max_{|\psi\rangle} \langle \psi | \psi' | \psi \rangle,$$  \hspace{1cm} (15)

optimized over all states $|\psi\rangle$ giving CHSH violation $2\sqrt{2}$, as a device-independent state estimation parameter based on its CHSH violation. They conjectured a lower bound on this fidelity in terms of the CHSH value, giving a construction which saturates the bound. However, no actual lower bound was given. Our robustness bound for self-testing using CHSH can be straightforwardly converted into such a lower bound, the dominant contribution to which is

$$F_{MY}(|\psi\rangle) \gtrsim 1 - \frac{1}{4}(9\sqrt{2}\epsilon + 2\frac{1}{4} 100 \epsilon + 2\frac{3}{8} 60 \epsilon),$$ \hspace{1cm} (16)

where $\epsilon$ is defined as in (9). This bound is rather loose: the fidelity drops to $F_{MY} = 20\%$ already for $\epsilon \approx 10^{-4}$.

More importantly, the framework introduced here allows one to generalize the concept of self-testing in different ways. For instance, our framework is an interesting contrast to the previous device-independent work with the CHSH inequality [14, 9] as we do not rely on Jordan’s lemma [16] to reduce the high dimension case to the qubit case. Jordan’s lemma only applies in the context of two measurement settings and two outcomes, which limits the applicability of the proof techniques used in these previous papers. The current proof technique thus provides an opening for testing in scenarios with more settings or outcomes, with the Mayers–Yao scenario a concrete example.

Considering the similarities between the CHSH and Mayers–Yao experiments we use here, it is natural to ask whether they can be generalized to a larger class of experiments which can be used to self-test singlets. The framework we use here is a natural starting point for such an enquiry since it is agnostic as to the number of settings or outcomes, requiring only that a pair of anticommuting operators can be found, or constructed as in the CHSH case. It also naturally extends to multipartite scenarios as in [12].

In a way, our result complements the known result that the Bell state with projective measurements is sufficient to achieve the maximal CHSH violation of $2\sqrt{2}$: it is in fact necessary to have a Bell state, in the way we define above. It is natural then to ask whether our result can be extended to complement the known result [17] that pure states of two qubits with projective measurements are sufficient to generate all the extremal correlations in the two input two output settings. We shall answer this interesting question elsewhere [18].

Acknowledgments

We acknowledge discussions with Lana Sheridan, Rafael Rabelo, Melvyn Ho and Cai Yu. This work is funded by the Ministry of Education and the National Research Foundation, Singapore.
Appendix A. Detailed derivation of the bounds used in theorem 1

- Bound for the second term of (8), the one for the third line being analogous:
  \[ \| (I + Z_A')(I - Z_B)|\psi'\rangle \| \leq \| (I - Z_A'Z_B)|\psi'\rangle \| + \| (Z_A' - Z_B)|\psi'\rangle \| \stackrel{(6)}{=} 2\epsilon_2. \]

- Comparison between the first and the fourth lines of (8): we want to bound
  \[ \| X_A'X_B(I + Z_A')(I + Z_B)|\psi'\rangle - (I + Z_A')(I + Z_B)|\psi'\rangle \|. \]

  The trick consists in propagating \( X_A'X_B \) in the first term to the right using (3) and (4). This costs \( 4\epsilon_1 \), and leads to
  \[ \| (I + Z_A')(I + Z_B)(X_A'X_B - I)|\psi'\rangle \|. \]

  Using (5), this can be replaced by zero at the cost of \( 4\epsilon_2 \).

- Bound for \( \| (i'|Z_A'|\psi') \| \), the same holding for \( \| (\psi'|Z_B'|\psi') \| \): this proof uses routinely two arguments: (i) the fact that the operators are unitary and (ii) the fact that if \( \| \psi'\| \leq \epsilon \), then \( \| \chi(\psi) \| \leq \epsilon \) for all normalized \( \chi \). We need to establish two relations. From (i) and (4),
  \[ |Z_A'X_B|\psi') - Z_A'X_B|\psi') \| \leq \epsilon_2. \]

  By inserting \( 0 = X_A'Z_A' - X_A'Z_A' \), the triangle inequality and (3) lead to
  \[ \| Z_A'X_B|\psi') + X_A'Z_A'|\psi') \| \leq 2\epsilon_1 + \epsilon_2. \]

  Using (ii) with \( \chi = Z_A'|\psi' \) and the unitarity of \( X_B' \),
  \[ \| (\psi'|Z_A'|\psi') + \langle \psi'|X_A'X_B'|\psi') \| \leq 2\epsilon_1 + \epsilon_2. \]

  Finally, since the left-hand side is an absolute value, the same holds for the conjugate; hence we find the first relation
  \[ \| (\psi'|Z_A'|\psi') + \langle \psi'|Z_A'X_B'|\psi') \| \leq 2\epsilon_1 + \epsilon_2. \]

  The second relation is
  \[ \| (\psi'|Z_A'|\psi') - \langle \psi'|Z_A'X_B'|\psi') \| \leq \epsilon_2, \]

  obtained simply by combining (i) and (5) in the form \( \| Z_A'|\psi') - Z_A'X_B'|\psi') \| \leq \epsilon_2 \), and then using (ii) with \( \chi = |\psi'\rangle \). The two relations together, by triangle inequality, imply
  \[ \| (\psi'|Z_A'|\psi') \| \leq \epsilon_1 + \epsilon_2. \]

- Bound for the norm of the state: note first that \( (1 + Z_A')^2 = 2(1 + Z_A') \) and similarly with \( Z_B' \). Therefore, we have
  \[ \| (I + Z_A')(I + Z_B')|\psi'\rangle \| = 2\sqrt{1 + \langle \psi'|Z_A'|\psi') + \langle \psi'|Z_B'|\psi') + \langle \psi'|Z_A'Z_B'|\psi')}. \]

  We have derived in the previous bullet
  \[ -(\epsilon_1 + \epsilon_2) \leq \langle \psi'|Z_A'|\psi') \leq \epsilon_1 + \epsilon_2 \]

  and the same for \( Z_B' \). As for the last term, it satisfies
  \[ 1 - \epsilon_2^2/2 \leq \langle \psi'|Z_A'Z_B'|\psi') \leq 1, \]

  where the upper bound is trivial and the lower one is just a rewriting of (6). Neglecting the contribution in \( \epsilon_2^2 \), we find
  \[ \sqrt{1 - \epsilon_1 - \epsilon_2} \leq \frac{\| (I + Z_A')(I + Z_B')|\psi'\rangle \|}{2\sqrt{2}} \leq \sqrt{1 + \epsilon_1 + \epsilon_2}. \]

  With the expansion \( \sqrt{1 + \delta} \leq 1 + \delta/2 \), we find that the error made in normalizing the state is at most \( (\epsilon_1 + \epsilon_2)/2 \) as claimed.
The above bounds for various quantity are used in the following. In the expression for $\Phi(|\psi\rangle\rangle)$ above, the second and third lines are each bounded by $\epsilon_2/2$, while the last line differs from the first by $\epsilon_1+\epsilon_2$. From these, we have

$$
\left\| \Phi(|\psi\rangle\rangle) - \frac{(I+Z_A')(I+Z_B')}{2\sqrt{2}} |\psi\rangle\rangle_{\phi_f} \right\| \leq \epsilon_1 + 2\epsilon_2. \quad (A.1)
$$

This is already the desired form and we would like to identify $\Phi(|\psi\rangle\rangle)$ from the first by $\epsilon_1 + \epsilon_2$, the most tedious estimate being the one that bounds from above both $|\langle\psi|Z_A'|\psi\rangle|$ and $|\langle\psi|Z_B'\psi\rangle|$ with $\epsilon_1 + \epsilon_2$. All in all therefore

$$
\left\| \Phi(|\psi\rangle\rangle) - (\text{junk}) |\phi_+\rangle \right\| \leq \frac{3}{2} \epsilon_1 + \frac{5}{2} \epsilon_2. \quad (A.2)
$$

This is the self-testing bound for the state. In order to derive the bound for the action of the operators, we note that $\Phi(M_0'N_0'|\psi\rangle\rangle) = \frac{1}{2}(I+Z_A')(I+Z_B')M_0'N_0'|00\rangle + (\text{similar terms})$. One starts by propagating $M_0'$ and $N_0'$ to the left using (3) and (4). In the worst case, i.e. when both $M_0'$ and $N_0'$ are not the identity, this preliminary step adds $4\epsilon_1$ to the bound. The resulting expression is analogous to (8); then, one follows the same steps as above.

**Appendix B. Detailed derivation of the bounds used in theorem 2**

- **Exact anti-commutation of $X_A'$ and $Z_B'$**. First note that, $B_0'$ and $B_1'$ being Hermitian and unitary operators, it holds $|B_0' + B_1'| = \sqrt{2 + M}$ and $|B_0' - B_1'| = \sqrt{2 - M}$ with $M = B_0'B_1' + B_0'B_1'$. Thence these two operators commute, being analytic functions of the same operator. Furthermore, both $B_0'$ and $B_1'$ commute with $M$ too, and therefore with both $|B_0' + B_1'|$ and $|B_0' - B_1'|$. Finally, it is easy to show that $B_0' + B_1'$ and $B_0' - B_1'$ anti-commute.

- **Derivation of (11) and (12)**. The square of the CHSH operator is $C^2 = 4 + |A_0|A_1'||B_1'||B_0|$. Therefore, the Cauchy–Schwartz inequality $|\langle\psi|C^2|\psi\rangle| \geq (|\langle\psi|C|\psi\rangle|)^2$ together with (9) gives

$$
|\langle\psi'[A_0', A_1'||B_1'||B_0]|\psi\rangle| \geq 4 - \delta,
$$

with $\delta = 4\sqrt{2}\epsilon - \epsilon^2$. Explicitly, the lhs is the algebraic sum of $|\langle\psi[A_0A_1'||B_1'||B_0]|\psi\rangle|$ and three similar terms, each bounded by 1 in absolute value since each operator has $\epsilon$-norm equal to 1. Therefore, loosely speaking, we have $|\langle\psi'[A_0A_1'||B_1'||B_0]|\psi\rangle| \leq |\langle\psi[A_0A_1'||B_1'||B_0]|\psi\rangle| \approx 1$ and $|\langle\psi'[A_0A_1'||B_1'||B_0]|\psi\rangle| \approx -1$. Now, from the precise relation

$$
|\langle\psi[A_0A_1'||B_1'||B_0]|\psi\rangle| \leq -2 + \delta,
$$

we obtain

$$
\left\| (A_0'A_1' + B_1'B_0)|\psi\rangle \right\| = \sqrt{2 + (\langle\psi[A_0A_1'||B_1'||B_0]|\psi\rangle)^2} \leq \sqrt{3}.
$$

In a similar way, one proves that $\left\| (A_0'A_1' - B_1'B_0)|\psi\rangle \right\| \leq \sqrt{3}$ and $\left\| (A_0'A_1' + B_1'B_0)|\psi\rangle \right\| \leq \sqrt{3}$. Relations (11) and (12) follow from these four, using the triangle inequality, leading to $\epsilon_1 = \sqrt{2} = 2\sqrt{\epsilon^2/2} - O(\epsilon^{3/2})$.

- **Bound for $(X_A' - (B_0' + B_1')/\sqrt{2})|\psi\rangle$**: We open up the norm and use $(B_0' + B_1')^2 = 2 + (B_0, B_1)$ and (12) to obtain

$$
\left\| (X_A' - (B_0' + B_1')/\sqrt{2})|\psi\rangle \right\| \leq \sqrt{2 + \epsilon_1 - \sqrt{2}(\langle\psi[X_A'(B_0' + B_1')]|\psi\rangle)
$$

and we have to find an estimate for the last term.
For this, we start by noting that the definition of the norm and (12) imply $\sqrt{2}\sqrt{1 - \epsilon_1} \leq \| (B_0' \pm B_1') \| \leq \sqrt{2}\sqrt{1 + \epsilon_1}$. In particular, the scalar product with the normalized vector $A_0' \| \psi' \| | \leq \sqrt{2}\sqrt{1 + \epsilon_1}$. From (9), recalling that $X_A' = A_0'$, we find the desired bound

$$
\langle \psi'|X_A'(B_0' + B_1')\| \psi' \rangle \geq \sqrt{2}(1 - \epsilon'),
$$

where $\epsilon' = \epsilon / \sqrt{2 + \sqrt{1 + \epsilon_1} - 1 = \sqrt{\epsilon\sqrt{2} - O(\epsilon^{3/2})}$. All in all,

$$
\| (X_A' - (B_0' + B_1')/\sqrt{2})\| \psi' \| \leq \sqrt{\epsilon_1 + 2\epsilon'} = 2(\epsilon\sqrt{2})^{1/4} - O(\epsilon^{3/2}).
$$

For reference, let us spell out explicitly the hypotheses (14) that are used in the proof

$$
\langle \psi'|X_B'B\| \psi' \rangle \geq 1 - \epsilon
$$

(C.1)

$$
\langle \psi'|Z_B'B\| \psi' \rangle \geq 1 - \epsilon
$$

(C.2)

$$
\langle \psi'|X_A'Z_B'B\| \psi' \rangle \leq \epsilon
$$

(C.3)

$$
\langle \psi'|Z_A'D_B'B\| \psi' \rangle \leq \frac{1}{\sqrt{2}} + \epsilon
$$

(C.4)

$$
\langle \psi'|X_A'D_B'B\| \psi' \rangle \leq \frac{1}{\sqrt{2}} + \epsilon.
$$

(C.5)

The simple opening of the norm in (C.1) and (C.2) leads directly to (5) and (6) in the form

$$
\| X_A'\| \psi' \| - X_B'\| \psi' \| \leq \sqrt{2}\epsilon
$$

(C.6)

$$
\| Z_A'\| \psi' \| - Z_B'\| \psi' \| \leq \sqrt{2}\epsilon.
$$

(C.7)

The two other conditions require a bit more of work. First, we establish

$$
\| X_A' + Z_A'\| \sqrt{2} \| \psi' \| = \sqrt{1 + \langle \psi'|Z_A'X_A'\| \psi' \rangle}
$$

$$
\leq \sqrt{1 + \epsilon + \sqrt{2}\epsilon}:
$$

(C.8)
Indeed, from (C.7) it follows \( \langle \psi' | X' A Z'_A | \psi' \rangle - \langle \psi' | X'_A Z'_B | \psi' \rangle \leq \sqrt{2} \epsilon \) since \( \| \psi' \|_A = 1 \); whence \( \langle \psi' | X'_A Z'_A | \psi' \rangle \leq \epsilon + \sqrt{2} \epsilon \) follows from (C.3).

From (C.8) and the hypotheses (C.4) and (C.5) it follows
\[
\left\| D'_B | \psi' \rangle - \frac{X'_A + Z'_A}{\sqrt{2}} | \psi' \rangle \right\| \leq \sqrt{1 + 2 \sqrt{2} \epsilon + \sqrt{2} \epsilon} = \epsilon'.
\]

Since \( \| D'_B \|_\infty = \| X'_A \|_\infty = \| Z'_A \|_\infty = 1 \), we obtain
\[
\left\| (D'_B)^2 | \psi' \rangle - D'_B \frac{X'_A + Z'_A}{\sqrt{2}} | \psi' \rangle \right\| \leq \epsilon'.
\]
\[
\frac{X'_A + Z'_A}{\sqrt{2}} D'_B | \psi' \rangle - \left( \frac{X'_A + Z'_A}{\sqrt{2}} \right)^2 | \psi' \rangle \leq \sqrt{2} \epsilon'.
\]

Note that the second bound comes from the conservative estimate \( \| (X'_A + Z'_A)/\sqrt{2} \|_\infty \leq \sqrt{2} \), but this is the best one can ensure at this stage: indeed, we know from (C.8) that \( (X'_A + Z'_A)/\sqrt{2} \) is almost unitary when it acts on \( | \psi' \rangle \), but we know nothing about its action on other states.

From the last two estimates, together with the fact that \( (D'_B)^2 \) is the identity, it follows that \( \| (1 - ((X'_A + Z'_A)/\sqrt{2})) | \psi' \rangle \| \leq (1 + \sqrt{2}) \epsilon' \) i.e.
\[
\| X'_A | \psi' \rangle + Z'_A X'_A | \psi' \rangle \| \leq 2(1 + \sqrt{2}) \epsilon',
\]
which establishes (3).

Finally, by evaluating (C.6) on a suitable unit vector we have \( \| Z'_B X'_A | \psi' \rangle - Z'_A X'_A | \psi' \rangle \| \leq \sqrt{2} \epsilon \); analogously, from (C.7) we have \( \| X'_B Z'_A | \psi' \rangle - X'_A Z'_B | \psi' \rangle \| \leq \sqrt{2} \epsilon \). The addition of these two gives
\[
\| Z'_A X'_A | \psi' \rangle - Z'_A X'_B | \psi' \rangle \leq 2 \sqrt{2} \epsilon.
\]

Similarly, we may obtain
\[
\| X'_B Z'_A | \psi' \rangle - X'_A Z'_B | \psi' \rangle \leq 2 \sqrt{2} \epsilon.
\]

From the last two inequalities and (C.9), we reach
\[
\| X'_B Z'_A | \psi' \rangle + X'_B Z'_B | \psi' \rangle \| \leq 2(1 + \sqrt{2}) \epsilon' + 4 \sqrt{2} \epsilon.
\]
which establishes the final condition in (4). The value of \( \epsilon_1 \) given in the main text uses
\[
\epsilon' = (2 \epsilon)^{1/4} \left( 1 + \frac{1 + 2 \sqrt{2}}{2 \sqrt{2}} \sqrt{\epsilon} \right) - O(\epsilon^{5/4}).
\]

References

[1] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Proposed experiment to test local hidden-variable theories Phys. Rev. Lett. 23 880–4

[2] Crellson B S 1980 Quantum generalizations of Bell’s inequality Lett. Math. Phys. 4 93–100

[3] Popescu S and Rohrlich D 1992 Which states violate Bell’s inequality maximally? Phys. Lett. A 169 411–4

[4] Braunstein S L, Mann A and Revzen M 1992 Maximal violation of Bell inequalities for mixed states Phys. Rev. Lett. 68 3259–61

[5] Mayers D and Yao A 2004 Self-testing quantum apparatus Quantum Inf. Comput. 4 273–86

[6] Acin 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

[7] Magniez F, Mayers D, Mosca M and Ollivier H 2006 Self-testing of quantum circuits Proc. 33rd Int. Colloquium on Automata, Languages and Programming (Lecture Notes in Computer Science vol 4051) ed M Bugliesi et al pp 72–83

[8] Pironio S et al 2010 Random numbers certified by Bell’s theorem Nature 464 1021–4
[9] Bardyn C-E, Liew T C H, Massar S, McKague M and Scarani V 2009 Device-independent state estimation based on Bell’s inequalities Phys. Rev. A 80 062327

[10] Rabelo R, Melvyn H, Cavalcanti D, Brunner N and Scarani V 2011 Device-independent certification of entangled measurements Phys. Rev. Lett. 107 050502

[11] McKague M 2010 Quantum information processing with adversarial devices PhD Thesis University of Waterloo

[12] McKague M 2010 Self-testing graph states TQC’11: 6th Conf. on the Theory of Quantum Computation, Communication and Cryptography at press

[13] McKague M and Mosca M 2011 Generalized self-testing and the security of the 6-state protocol Theory of Quantum Computation, Communication, and Cryptography (Lecture Notes in Computer Science vol 6519) ed W van Dam, V Kendon and S Severini (Berlin: Springer) pp 113–30

[14] Pironio S, Acin 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

[15] Liang Y-C and Doherty A 2006 Phys. Rev. A 73 052116

[16] Jordan C 1875 Essai sur la géométrie à n dimensions Bull. Soc. Math. Fr. 3 103

[17] Masanes Ll 2005 Extremal quantum correlations for n parties with two dichotomic observables per site arXiv:quant-ph/0512100

[18] Yang T H et al 2012 in preparation