Opening the black box: how to estimate physical properties from non-local correlations

Tzyh Haur Yang, Tamás Vértesi, Jean-Daniel Bancal, Valerio Scarani, and Miguel Navascués

Abstract

In the device-independent approach to quantum information theory, quantum systems are regarded as black boxes which, given an input (the measurement setting), return an output (the measurement result). These boxes are then treated regardless of their actual internal working. In this paper, we develop SWAP, a theoretical concept which, in combination with known numerical methods for the characterization of quantum correlations, allows us to estimate general physical properties of the black boxes from the observed measurement statistics. As an illustration of the power of our new approach, we provide a robust bound on the state fidelity for the CHSH scenario: a CHSH violation larger than 2.57 gives a fidelity of more than 70% between the black box and the singlet. Also, we prove that maximal violation of CGLMP must be non maximally entangled, and in addition, robustly self test quantum states in the CGLMP scenario for which Jordan’s Lemma cannot be used. Furthermore, this SWAP method can also be extended to estimate in a robust manner the amount of extractable work in CHSH-violating systems, providing a link between thermodynamics and nonlocality. Lastly we also robustly certify entangling measurements in tripartite non-locality scenarios.

1 Introduction

To appreciate and ultimately certify a quantum experiment, a considerable amount of expert information is generally needed. For instance, in typical ion-trap experiments [1] the state of the ion is considered to be (approximately) embedded in a two-dimensional Hilbert space, and the behavior of the ion when subject to sequential laser pulses is well known. Without this knowledge, it would be difficult to make sense of most ion-trap experiments. Remarkably,
there exist situations where expert information is not required in order to take advantage of a (possibly complex) system. Think of the device-independent paradigm, where local systems are regarded as black boxes which admit an input (the measurement setting) and return an output (the measurement outcome), see [2, 3] and references therein. In this description of laboratory experiments, all the accessible information is given by a collection of measurement results, the only assumptions invoked being that individual devices are properly separated from each other (which can be ensured through space-like separation for instance) and that inputs are chosen freely. Several quantum information tasks, such as device-independent quantum cryptography [4, 5, 6, 7, 8] or randomness amplification [9, 10] can be accomplished in this device-independent manner.

In the last years we have seen a proliferation of results where, under the assumption that quantum mechanics holds exactly, some properties of the systems inside the box are estimated. In this respect, Mayers and Yao [11] (see also [12], [13]) showed that there exist lists of statistical data in the device-independent framework which allow us to identify the quantum state and measurement operators involved in the experiment. Moreover, the process can tolerate a small amount of external noise. This notion of determining approximately the state and measurement operators involved via non-locality detection is known as self-testing, and it has inspired a number of works on the subject [14, 15, 16, 17, 18]. In the same spirit, given a Bell inequality violation, Moroder et al. [19] recently described an algorithm to lower bound the negativity of the shared quantum state, see [20] for an extension to steering scenarios.

So far, all current works on self-testing are specific to a given Bell inequality [14, 15, 17], or family of Bell inequalities [18]. In contrast, the method devised by Moroder et al. works in virtually any non-locality scenario. However, it provides access to properties of the measured system only through a specific local operation.

In this paper, we present a unified tool that allows us to guess the general properties of the quantum systems inside the boxes from the experimentally accessible correlations. The process is algorithmic, i.e., it is completely automated, and makes use of the Navascués-Pironio-Acin (NPA) hierarchy for the characterization of quantum correlations [23]. To illustrate the power of this method, we apply it to estimate quantum monotones via convex optimization [27], in the same way that we would calculate such monotones had we known the exact mathematical representation of the state and measurement operators involved in the non-locality experiment. We will illustrate the SWAP concept by using it to estimate the minimum singlet fidelity as a function of the violation of the Clauser-Horne-Shimony-Holt (CHSH) inequality [28, 29]. We will then apply
SWAP to calculate a number of physical properties in a device-independent
way. More specifically, we will: 1) estimate the state and operator fidelity in
Bell scenarios with binary inputs, but with either two or three outputs or more
commonly known as CHSH/tilted-CHSH and CGLMP scenarios; 2) prove the
robustness to noise of the scheme proposed in [37] to certify entangling measure-
ments device-independently; and 3) lower bound the amount of extractable work
in CHSH-violating systems under the assumption of a fully degenerate energy
operator. Also, in connection with the last point, we will show how SWAP can
be used in principle to limit the dimensionality of the quantum states involved
in nonlocality experiments. Finally, we will present our conclusions.

2 Main Concept

2.1 The SWAP concept

Consider the following scenario: two space-like separated parties, call them Al-
ice and Bob, share a bipartite quantum state. Each can perform one out of two
measurements on its corresponding subsystem, with values in \{-1, 1\}. If Alice
and Bob do not have any extra knowledge about the nature of their measure-
ment devices, Alice (Bob) is bound to regard her (his) local experimental setup
as a black box with a binary input \(x \in \{0, 1\}\) (\(y \in \{0, 1\}\)) — the measure-
ment setting — and a binary output \(a \in \{-1, +1\}\) (\(b \in \{-1, +1\}\)) — the measure-
ment outcome. By repeating their experiment many times, Alice and Bob can
estimate \(p(a, b|x, y)\), the probability that they obtain the outcomes \(a, b\) when
they perform the measurements \(x, y\).

In this scenario, a measure of the non-locality of Alice and Bob’s correlations
\(p(a, b|x, y)\) is given by the value of its associated CHSH violation [28]

\[ B_{\text{CHSH}}(p) = \langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle. \] (2.1)

Since we do not put any restriction on the dimensionality of the degrees of
freedom under study, we can work with a pure state and projective measure-
ments without loss of generality. Specifically, there will exist a bipartite state
\(|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B\), and two projectors for each measurement \(\{\Pi_{a=+1}^x, \Pi_{a=-1}^x\}\)
and \(\{\Pi_{b=+1}^y, \Pi_{b=-1}^y\}\), such that

\[ p(a, b|x, y) = \langle \psi | \Pi_a^x \otimes \Pi_b^y | \psi \rangle. \] (2.2)

All our mathematical developments must be based only on these state and
projectors, which describe the physical content of the boxes. Specifically, it is
practical to define from the start the four dichotomic operators \(\{A_x, B_y : x, y =
0, 1\}\) defined as \(A_x = \Pi_{a=+1}^x - \Pi_{a=-1}^x, B_y = \Pi_{b=+1}^y - \Pi_{b=-1}^y\) for which it holds

\[ \langle a_x b_y \rangle = \langle \psi | A_x \otimes B_y | \psi \rangle. \] (2.3)

It is known [15] that, for \(B_{\text{CHSH}}(p)\) close to \(2\sqrt{2}\), Alice’s and Bob’s measures-
ments must be, modulo isometries, close to

\[ \overline{A}_0 = \overline{B}_0 = \sigma_z, \quad \overline{A}_1 = \overline{B}_1 = \sigma_x, \] (2.4)

with \(\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|\) and \(\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|\). Also, their state must be
approximately equal to
\[ |\psi\rangle = \frac{1}{\sqrt{2}} \cos \left( \frac{\pi}{8} \right) (|00\rangle - |11\rangle) + \frac{1}{\sqrt{2}} \sin \left( \frac{\pi}{8} \right) (|01\rangle + |10\rangle). \]  

(2.5)

In the literature it is more customary to write the maximally entangled state as 
\((|00\rangle + |11\rangle)/\sqrt{2}\), whence Bob’s optimal measurements are 
\(B_0 = (\sigma_z^B + \sigma_x^B)/\sqrt{2}\), 
\(B_1 = (\sigma_z^B - \sigma_x^B)/\sqrt{2}\). Actually, both representations are equivalent under local unitaries, and, as we will see, the former is more convenient to our purposes.

Figure 1: Fidelity estimation via partial SWAP. A thought experiment to analyze the black boxes’ physical property.

We would like to verify to what extent the quantum state inside Alice and Bob’s boxes has the form (2.5). This question was first answered in [15], where the authors construct a local isometry for Alice and Bob based on the Bell violation \(B_{CHSH}(\rho) = V\) and use it to certify how close the state \(|\psi\rangle\) is to the reference state, \(|\psi\rangle\).

Following this idea, we imagine the following scheme (Figure 1): 1) Alice and Bob prepare each a trusted qubit in the state \(|0\rangle\); 2) they apply well-chosen local unitaries \(S_A, S_B\) between their boxes, thus bringing their trusted two-qubit system to the state \(\rho_{\text{swap}}\); and 3) they compute the fidelity \(f = \langle \psi|\rho_{\text{swap}}|\psi\rangle\). Clearly, \(f\) will be a lower bound to the actual fidelity of \(|\psi\rangle\) with respect to the reference state \(|\psi\rangle\). Notice again that this scheme is a mathematical possibility and does not need to correspond to operations actually performed in the lab. The only crucial requirement is that \(S_A\) and \(S_B\) be defined from the object that we know to exist in our box, that is ultimately the projectors \(\Pi_x^a\) and \(\Pi_x^b\).

How could we then choose, hypothetically, \(S_A, S_B\) in order to maximize \(f\)? Consider the following operators:

\[ U_A = (I \otimes |0\rangle \langle 0| + A_1 \otimes |1\rangle \langle 1|), V_A = \left( \frac{I + A_0}{2} \otimes I + \frac{I - A_0}{2} \otimes \sigma_x \right), \]

\[ U_B = (I \otimes |0\rangle \langle 0| + B_1 \otimes |1\rangle \langle 1|), V_B = \left( \frac{I + B_0}{2} \otimes I + \frac{I - B_0}{2} \otimes \sigma_x \right). \]  

(2.6)

Here the first Hilbert space corresponds to Alice’s or Bob’s black box, where the operators \(A_x, B_y\) act, while the second represents the state space of Alice’s or Bob’s trusted qubit. Notice that, for whatever choice of dichotomic observables
\( A_0, A_1, B_0, B_1 \), the operators defined in eq. (2.6) are unitary. \( S_A = U_A V_A \), \( S_B = U_B V_B \) are thus valid local unitary transformations. Most crucially, if \( A_x = \overline{A_x}, B_y = \overline{B_y} \), as defined in (2.4), then \( U_A, U_B (V_A, V_B) \) would correspond to Controlled NOT (CNOT) gates \(^{38}\) from the trusted qubits to the boxes (from the boxes to the trusted qubits), and, consequently, \( S_A \otimes S_B |\psi\rangle_{\text{boxes}} |00\rangle_{\text{trust}} = |00\rangle_{\text{boxes}} |\overline{\psi}\rangle_{\text{trust}} \). By \(^{15}\) we thus have that, for \( B_{\text{CHSH}}(p) = 2\sqrt{2} \), \( S_A, S_B \) would copy the state inside the boxes to the trusted qubits, and so the value of \( f \) would be maximum.

Intuitively, one would expect the value of the fidelity after the action of \( S_A \otimes S_B \) to be reasonably high for values of \( B_{\text{CHSH}}(p) \) close to \( 2\sqrt{2} \). Now, it is straightforward that

\[
S_A \otimes S_B |\psi\rangle |00\rangle = \frac{1 + A_0}{2} \frac{1 + B_0}{2} |\psi\rangle |00\rangle + \frac{1 + A_0}{2} B_1 \frac{1 - B_0}{2} |\psi\rangle |01\rangle + \frac{A_1}{2} \frac{1 - A_0}{2} B_0 \frac{1 - B_0}{2} |\psi\rangle |10\rangle + \frac{A_1}{2} \frac{1 - A_0}{2} B_1 \frac{1 - B_0}{2} |\psi\rangle |11\rangle.
\]

For any product \( t \) of Alice and Bob’s measurement operators, define the complex correlation term \( c_t = \langle \psi | t | \psi \rangle \). By tracing out the state of the black box in the system above, the density matrix of the trusted qubits \( \rho_{\text{copied}} \) is then a linear function of the correlation terms \( c = \{ \langle \psi | t | \psi \rangle, \langle \psi | A_0 B_1 A_0 | \psi \rangle, \ldots \} \). For instance,

\[
\begin{align*}
\langle 00 | \rho_{\text{copied}} | 00 \rangle &= \frac{1}{4} (\psi | (1 + A_0) (1 + B_0) | \psi ) = \frac{1}{4} (c_1 + c_{A_0} + c_{B_0} + c_{A_0 B_0}), \\
\langle 01 | \rho_{\text{copied}} | 11 \rangle &= \frac{1}{8} (\psi | (1 - A_0) A_1 (1 + A_0) (1 - B_0) | \psi ) = \\
&= \frac{1}{8} (c_{A_1} + c_{A_1 A_0} - c_{A_1 B_0} - c_{A_1 A_0 B_0} - \ldots )
\end{align*}
\]

Call \( Q \) the set of all complex vectors \( (c_t) \) that admit a quantum representation, i.e., such that \( c_t = \langle \psi | t (A_x, B_y) | \psi \rangle \) for some normalized quantum state \( | \psi \rangle \) and dichotomic operators \( \{ A_x, B_y \} \). The minimum overlap between \( | \psi \rangle \) and \( \rho_{\text{copied}} \) after the action of \( S_A \otimes S_B \) for a CHSH violation \( B_{\text{CHSH}}(p) = V \) is thus equal to

\[
f = \min_{c \in Q} \frac{| \psi \rangle \rho_{\text{copied}}(c) | \psi \rangle}
\]

such that \( c \in Q \)

\[
c_{A_0 B_0} + c_{A_1 B_0} + c_{A_0 B_1} - c_{A_1 B_1} = V.
\]

The reason we are minimizing the fidelity is to consider the worst scenario and thus finding the lower bound of the fidelity. This will serve as a lower bound on the certification of the state \( | \psi \rangle \) to the reference state \( | \overline{\psi} \rangle \).

Optimization problems over the quantum set are difficult in many cases \(^{21}\), and possibly undecidable in others \(^{22}\). To proceed, we require the concept of moment matrix from NPA criterion \(^{23, 25}\): given a vector of complex numbers \( (c_t) \), the moment matrix \( \Gamma^S(c) \) is a matrix whose rows and columns are labeled by sequences of Alice and Bob’s measurement operators belonging to a finite set \( S \) and such that

\[
5
\]
\[\Gamma^n_{s,t} = c_{s,t}, \quad (2.9)\]

for all \(s, t \in S\). It can be verified that, if \((c_t)\) admits a quantum representation, \(\Gamma^S(c) \geq 0\). Conversely, it can be proven that any finite vector \((c)\) that admits extensions \(c' = (c, \tilde{c})\) to bigger sets of sequences \(S'\) with \(\Gamma^{S'}(c') \geq 0\) also admits a quantum representation [25].

Thus the problem of device independent state estimation can be relaxed to a semidefinite program, optimizing the fidelity of the state \(\rho_{\text{swap}}\) to the reference state over all the possible moment matrices of a certain level of hierarchy.

\[
f^S = \min_{\Gamma^n} \langle \psi | \rho_{\text{swap}}(c) | \psi \rangle
\]

such that \(\Gamma^S(c) \geq 0\),

\[
B_{\text{CHSH}}(p) = V. \quad (2.10)
\]

As noted in [25], since the objective function is a linear combination of the moment matrix’ elements with only real coefficients, we can assume that the entries of the moment matrix are real as well. From the above considerations, it follows that \(f^S \leq f\). Moreover, as \(S\) grows, \(f^S \to f\) [25].

How well does this method perform in practice? Figure 2 shows a lower bound on the singlet fidelity as a function of the CHSH violation in two cases: generic boxes and isotropic boxes (i.e., boxes such that \(\langle A_x \rangle = \langle B_y \rangle = 0\) and \(|\langle A_x B_y \rangle| = |\langle A_{x'} B_{y'} \rangle|\) for \(x, y, x', y' = 0, 1\)). For the computation, we built a moment matrix \(\Gamma^S(c)\) with rows and columns labeled by sequences in the set \(S = \{ t : t = (A_x A_{x'}) (B_y B_{y'}) : x, y = -1, 0, 1 \}\). Here, with a slight abuse of notation, we define \(A_{-1} = B_{-1} = 1\). The generic curve grows almost linearly from 0.0732 to 1.0000 as the CHSH value increases from 2 to \(2\sqrt{2}\). In contrast, the isotropic bound is manifestly convex, starting from \(f = 0.3272\) for the local CHSH value. Figure 2 should be compared with previous results in [15, 18] where the fidelity drops to trivial zero as soon as \(|CHSH - 2\sqrt{2}| \gtrsim 10^{-3}\), the improvement is huge.

Although it is strictly speaking a lower bound on the fidelity, the bound for general black box (solid line in Figure 2) is actually tight for the considered swap operation \(S_A, S_B\). Indeed, for every CHSH violation, simple optimization allows one to find two-qubits entangled states and local projective operators such that the fidelity of the swapped state with \(|\psi\rangle\) is indeed given by the curve.

### 2.2 A Better and More General SWAP Concept

In the above scheme for fidelity estimation, the intended action of \(S_A, S_B\) over the state \(|\psi\rangle|00\rangle\) is to copy the state inside the box to the qubit registers. Note, however, that, even in the ideal case, \(S_A \otimes S_B|\psi\rangle|11\rangle \neq |11\rangle|\psi\rangle\). Actually, not even the state of the trusted qudits equals \(|\psi\rangle\). We will hence refer to \(S_A, S_B\) as partial SWAPs, since they swap the contents of the box with the trusted system, but only when the latter is in state \(|00\rangle\). Partial SWAPs will be denoted by an arrow from the box to the trusted system, see Figure 1.

For some applications (Sections 3.3, 3.5) we will need to copy the state in the box into registers in an initial state different from \(|00\rangle\). Some other times we will need to place a non-trivial state inside the box (Section 3.4). This motivates the definition of the full SWAP operators \(S_A \equiv V_A U_A V_A, S_B \equiv V_B U_B V_B\), which, in
Figure 2: Minimal singlet fidelity as a function of CHSH violation obtained with $S_A = U_AV_A$ and $S_B = U_BV_B$. The solid line denotes a lower bound on the fidelity for generic boxes; the dashed one, a lower bound for isotropic boxes.

Figure 3: Full SWAP.
the ideal case where $U, V$ are, respectively, CNOT gates from the trusted system to the box and vice versa, would correspond to the swap operator between the box and the trusted qudits. Full SWAP will be denoted diagrammatically by two arrows: one from the box to the trusted system and another one from the trusted system to the box, as in Figure 3.

This scheme of transferring quantum information from and to the box can be carried further and generalized to other Bell inequalities and other non-locality scenarios with more measurement settings and outcomes, see Appendix A. For that it is required that we have a mathematical guess $(|\psi\rangle, \{E_x^a, F_y^b\} \subset B(C^d \otimes C^d))$ on the physical state and projector operators giving rise to the distribution $P(a, b|x, y) = \langle \psi|E_x^a \otimes F_y^b|\psi\rangle$. We will use this guess to build partial or full SWAPs in order to estimate the properties of the actual state and operators being implemented. In general, the construction of such gates will not be as straightforward, so it will be necessary to introduce new non-commutative unitary variables in our moment matrices which will allow to construct the approximate CNOT unitaries. Such new variables will be related to the measurement operators via extra semidefinite constraints (the so-called localizing matrices [26]).

Finally, note that in principle we can define Bell violation-dependent SWAP operators. For example, rather than taking $A_1$ as Alice’s logical ‘NOT’ operator in the CHSH example above, we could have used the dichotomic operator $A_2$ satisfying

$$A_2 (\cos(\alpha(V))A_0 + \sin(\alpha(V))A_1) \geq 0,$$

where $\alpha(V)$ is an arbitrary function of the CHSH violation $V$. For $V = 2\sqrt{2}$ choosing $\alpha(V) = \pi/2$ allows one to recover $A_2 = A_1$ so that the NOT operator returns the correct SWAP in the case of maximal CHSH violation. Similarly, we can parametrize an auxiliary operator for Bob by a function $\beta(V)$.

Moreover, since we are only interested in certifying the quality of the measured state up to local unitaries, it is sufficient to lowerbound its fidelity with respect to any reference state of the form

$$|\psi_{ME}\rangle = W_A \otimes W_B |\psi\rangle$$

where $W_A$ and $W_B$ are arbitrary single qubit unitaries.

In the case of CHSH, choosing parameters $\alpha$, $\beta$ and unitaries $W_A$, $W_B$ adequately allows one to improve the bounds depicted in figure 2. Moreover, replacing the swaps by identity operators acting on trusted qubits initialized in the state $(\cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle) \otimes |0\rangle$ guarantees that a minimum singlet fidelity of 1/2 can be achieved independently of the CHSH values. The bounds obtained in this way are shown in Figure 4. Both of these curves can be used in experiments to certify how close a produced state is to a maximally entangled two-qubit state.

### 2.3 Dropping Identical and Independently Distributed Assumption

Note that, in order to estimate experimentally the CHSH value of a given box, many identical and independently distributed (i.i.d.) copies are needed. Suppose that, on the contrary, we want to allow for the possibility that the statistics
of each box vary between different rounds and even depend on previous operations conducted by Alice and Bob. Borrowing ideas from [24], in Appendix B it is shown that, under these circumstances, Alice and Bob can use the solid line in Figure 4 in order to disprove the hypothesis that all the distributed boxes contain quantum states with singlet fidelity smaller than or equal to \( f \). Most results proven under the i.i.d. assumption can be similarly translated to fully device-independent claims. For simplicity and without great loss of generality, we will therefore take i.i.d. for granted in the rest of the article.

3 Applications

In the following we show some applications of the SWAP concept to different problems in device-independent quantum information science.

3.1 Device-independent state certification: Partially Entangled Qubits

In the previous section, we used the CHSH inequality to certify how close Alice and Bob’s state \( |\psi\rangle \) is to a maximally entangled qubit pair. Here, we show how our method can be used to certify non-maximally entangled qubit pairs as well. This problem was first considered in [18], where an analytic estimate on the state fidelity was obtained. Here, using the SWAP concept, we derive a much better bound.

The scenario is similar to the case above with the CHSH scenario. The only
difference is that we will use a tilted CHSH inequality of the form

\[ B_\alpha = \alpha \langle A_0 \rangle + \langle A_0 B_0 \rangle + \langle A_1 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle, \]  

(3.1)

where \( 0 \leq \alpha \leq 2 \). The maximum quantum violation of this inequality \([40]\) is given by \( \sqrt{8 + 2\alpha^2} \) and the corresponding optimal qubit strategy is

\[ A_0 = \sigma_z, \]
\[ A_1 = \sigma_x, \]
\[ B_0 = \cos \mu \sigma_z + \sin \mu \sigma_x, \]
\[ B_1 = \cos \mu \sigma_z - \sin \mu \sigma_x, \]
\[ |\psi\rangle = \cos \theta |00\rangle + \sin \theta |11\rangle, \]

(3.2)

where \( \sin 2\theta = \sqrt{(1 - \alpha^2/4)/(1 + \alpha^2/4)} \) and \( \tan \mu = \sin 2\theta \). Thus, we can use the appropriate inequality with the corresponding value \( \alpha \) to device independently certify a reference state of the form \( |\psi\rangle = \cos \theta |00\rangle + \sin \theta |11\rangle \). The details on the construction of partial SWAP can be found in Appendix C.1. Figure 5 shows a plot of the fidelity for different values of \( \alpha \) and different Bell inequality violations. As discussed in Section 2.2 a minimal fidelity of \( \cos^2 \theta \) could also always be obtained in this situation by using identity unitaries instead of the swap operators, however for the remaining of this paper we leave this possibility aside in order to focus our study on the swap operators.

![Figure 5](image-url)

Figure 5: (Color Online) The solid line refers to the standard CHSH case, which interestingly recovers the line in Figure 2. The dashed line refers to the case when \( \alpha = 0.5 \) and the local bound is 2.5, explaining the discontinuity. The dotted-dashed line refers to the case when \( \alpha = 1 \).

The curves are approximately linear, between the local bound and maximum quantum bound. Since the local and quantum bound coincide at \( \alpha = 2 \), the line gets steeper. This is intuitive as \( \alpha \) increases, the range of error tolerable gets smaller. However, it is interesting to note that the robustness is always linear for the three cases.
3.2 Device-independent state certification: Partially Entangled Qutrits

So far, our studied cases concern scenarios where Alice and Bob each have two inputs and two outputs, and they want to certify their state against an entangled pair of qubits. In this section, we illustrate how to extend these ideas to higher dimensional scenarios. It is unclear how one can generalize the methods sketched in [15, 17, 18] for such a purpose. However, using the SWAP method above, it is easy to do it.

The relevant Bell inequality for this case is the CGLMP inequality [30]; it requires two measurement settings on each side, with three measurement outcomes. The inequality reads:

\[
B_{\text{CGLMP}}(p) = p(a < b | x = 1, y = 1) + p(a > b | x = 0, y = 1) + p(a \geq b | x = 1, y = 0) + p(a < b | x = 0, y = 0) \geq 1. \tag{3.3}
\]

The maximum quantum violation is conjectured [31] and verified numerically [23] to be \( B_{\text{CGLMP}}(p) = (12 - \sqrt{33})/9 \approx 0.6950 \). Moreover, it is believed that the maximal quantum violation can only be achieved with the (non-maximally entangled) state and measurement operators \((|\psi\rangle, E_a, F_b)\) described in [31, 30]. Here we will prove this conjecture true.

First, let us write down the reference state

\[
|\psi\rangle = \frac{1}{\sqrt{2 + \gamma^2}}(|00\rangle + \gamma|11\rangle + |22\rangle), \tag{3.4}
\]

where \( \gamma = (\sqrt{11} - \sqrt{3})/2 \). Then the reference measurement operators \( A_a, a = 0, 1, \) measured by Alice and \( B_b, b = 0, 1, \) measured by Bob have the nondegenerate eigenvectors

\[
|k\rangle_{A,a} = \frac{1}{\sqrt{3}} \sum_{j=0}^{2} \exp\left(\frac{2\pi i}{3} j(k + \alpha_a)\right) |j\rangle_A,
\]

\[
|l\rangle_{B,b} = \frac{1}{\sqrt{3}} \sum_{j=0}^{2} \exp\left(\frac{2\pi i}{3} j(-l + \beta_b)\right) |j\rangle_B. \tag{3.5}
\]

We refer the reader to Appendix C.2 for details on the construction of the partial SWAP operators. Figure 6 shows a plot of the minimum fidelity with respect to the reference state \(|\psi\rangle\) as a function of the CGLMP violation.

One can see that as the violation tends to the quantum maximal violation (\( \approx 0.6950 \)), the fidelity of the general black box with respect to the reference state (non maximally entangled state) tends to 1. This shows that any quantum system violating the CGLMP inequality maximally is unitarily equivalent to the non-maximally entangled state described in [31]. Thus, we have proved the conjectured that only non-maximally entangled states can violate CGLMP maximally.

Besides, the graph provides us the first ever robust state certification for black box scenario with number of output more than 2, which is not possible to analyze with the Jordan’s Lemma [39] as used in [16, 17].
Figure 6: Minimum fidelity of the state swapped out the operators defined above. The blue line represents the minimum fidelity obtained from the SDP hierarchy. The hierarchy we used is the smallest hierarchy possible for the problem to be defined.

3.3 Measurement estimation

Suppose that, rather than verifying that $|\psi\rangle$ is close to $|\bar{\psi}\rangle$, we are interested in learning to which degree the actual measurements $\{F_y\}$ that Bob’s box is performing are well described by the matrices $\{\overline{F}_y\}$. One possibility to do so is to put an arbitrary trusted state inside the box and then observe the measurement statistics when we probe it with different measurements $y$. If we obtain the statistics corresponding to $\{\overline{F}_y\}$ we can thus argue that pressing button $y$ induces a dynamics inside the box that allows to perform such quantum measurements. For instance, in the CHSH case we conjecture that Bob’s observables are close to $\overline{B}_0 = \sigma_z$, $\overline{B}_1 = \sigma_x$. To quantify this hypothesis, we define the functional

$$\tau \equiv \frac{1}{2} \{P(0|0,0) + P(1|0,1) + P(0|1,+) + P(1|1,-)\} - 1,$$  \hspace{1cm} (3.6)

where $P(b|y,\varphi)$ denotes the probability to obtain result $b$ when we press button $y$ after applying the full SWAP to Bob’s box and trusted qubit in state $|\varphi\rangle$. $\tau$ is a number ranging from -1 to +1; the latter value $\tau = 1$ being only achievable with the observables $\overline{B}_0 = \sigma_z$, $\overline{B}_1 = \sigma_x$. The scheme is illustrated in Figure 7, where a plot of $\tau$ as a function of the CHSH violation is also shown. This confirms that Bob’s measurement are essentially $\sigma_z$ and $\sigma_x$ when CHSH takes a value close to $2\sqrt{2}$.

3.4 Robust device-independent certification of entangled measurements

In [37] an entanglement swapping-based protocol to certify entangling measurements in a device-independent way is presented. In this protocol, four parties, $A, B, C_A, C_B$ are allowed to conduct one out of two possible dichotomic measurements in their respective subsystems. By showing a maximal violation of the
Figure 7: Estimation of Bob’s measurements. The protocol works in two steps: 1) We implement a full SWAP of Bob’s box and his trusted qubit, that we prepare in state $|\varphi\rangle$. 2) We implement measurement $B_y$ and study the resulting statistics.

CHSH inequality between parties $A,C_A$ and $B,C_B$, the state shared by Alice and Bob is certified to be separable. Then parties $C_A,C_B$ are brought together and a collective dichotomic measurement $C$ is conducted on the joint system $C_A-C_B$. As shown in [37], if, conditioned on the outcome 0, Alice and Bob can violate the CHSH inequality by an amount greater than $\sqrt{2}$, measurement $C$ must be entangling.

Unfortunately, the previous protocol relies on the assumption that parties $A,C_A$ or $B,C_B$ violate CHSH maximally. To make this protocol robust to noise, we propose the scheme depicted in Figure 8. There, a local full SWAP is performed between systems $C_A,C_B$ and two particles maximally entangled to systems 1, 2. A measurement of $C$ is then conducted, and the resulting state $\rho_{12|C=0}$ is considered. Obviously, if $C$ is described by a separable operator, $\rho_{12|C=0}$ cannot be entangled at the end of the procedure.

Figure 9 shows the minimum value of $E \equiv P(C=0)\mathcal{N}(\rho_{12|C=0})$ as a function of the violation of the following three inequalities

$$\begin{align*}
B_1 &= \langle A_0 C_0^A \rangle + \langle A_0 C_1^A \rangle + \langle A_1 C_0^A \rangle - \langle A_1 C_1^A \rangle \leq 2, \\
B_2 &= \langle B_0 C_0^B \rangle + \langle B_0 C_1^B \rangle + \langle B_1 C_0^B \rangle - \langle B_1 C_1^B \rangle \leq 2, \\
B_3 &= \langle (1 + C)(A_0 B_0 + A_0 B_1 + A_1 B_0 - A_1 B_1) \rangle - 2 \langle C \rangle \leq 2.
\end{align*}$$

(3.7)

We fix the Bell violations $B_1 = B_2 = 2\sqrt{2} - \epsilon$ for the first two Bell inequalities above and $B_3 = 1 + \sqrt{2} - \epsilon$ for the last one. We used the SeDuMi package to solve the SDP problem for some small values of $\epsilon$ and obtained the curves labeled by Levels 1, 2, 3 in Fig. 9. The three distinct curves correspond to different relaxation of the SDP problem with respective sizes of the moment matrix $144 \times 144$, $168 \times 168$, $200 \times 200$. 

13
Figure 8: Certification of non-separability of measurement $C$. The scheme works as follows: 1) We perform a full (local) SWAP between the boxes $C_A$, $C_B$ and two particles, each maximally entangled with particles 1, 2, respectively. 2) We perform measurement $C$ and post-select on result $C = 0$. 3) We minimize the negativity of the state $\rho_{12}$.

3.5 SWAP and maximally mixed states: work extraction and dimension estimates

The relation between work and information has been a matter of scientific debate since the dawn of statistical mechanics. In this regard, it has been shown in \([32, 33]\) that the knowledge of the state of a quantum system, as measured by the smooth min and max entropies, can be used to generate work. Since the SWAP tool allows us to acquire knowledge about the state inside the box, it is hence not surprising that it can also lower bound its potential for work extraction.

In Figure 10 we use the full SWAP over two maximally mixed trusted qubits initially inside a Szilard engine in contact with a bath at temperature $T$. Since by definition the state of a maximally mixed qubit is completely unknown, such states can be considered as a free resource. After the interaction with the box, the state of the two qubits gets purified, and they can be rotated to one of the sides of the box with high probability, thus pulling a weight. Under the assumption that the energy operator of the system inside the box is fully degenerate, it can be proven that the amount of work extracted in this way is related to the difference between the maximum and minimum eigenvalues of $\rho_{\text{swap}}$, see Appendix D, and this can be estimated via semidefinite programming. Figure 10 also shows a plot of the CHSH violation vs the minimum amount of work extractable via this scheme. Notice that work extraction is possible as long as the system is non-local. Also, for the maximal CHSH violation, the amount of work extractable per $KT$ is $W/KT = 2\ln(2)$, the work content of two pure qubits \([34, 35]\).

In more complex non-locality scenarios, the use of a full SWAP between Alice’s box and a maximally mixed qudit (or Alice and Bob’s boxes and a pair of maximally mixed qudits) also allows us to infer non-trivial properties
Figure 9: Certification of non-separability of measurement $C$. We minimize the quantity $E = P(C = 0) \mathcal{N}(\rho_{12|C=0})$ in function of the value $\epsilon$, where $\mathcal{N}(\rho_{12|C=0})$ denotes the negativity of the state $\rho_{12|C=0}$. The curves correspond to different levels of the NPA hierarchy. $E > 0$ signals that Charlie’s measurement $C$ is entangling. In the ideal case of $\epsilon = 0$, measurement $C$ is a Bell state measurement providing $E = 1/8$.

Figure 10: Extractable work per $KT$ as a function of the CHSH violation of an isotropic box.
about the dimensionality of the state inside Alice’s box (or Alice and Bob’s boxes). In [36], it is proven that a state $\rho \in B(H)$ can be transformed into a state $\sigma \in B(H)$ by means of unitary transformations over $\rho$ and maximally mixed ancillas iff $\rho$ majorizes $\sigma$, i.e., if $\sum_{i=1}^{k} \lambda_i(\rho) \geq \sum_{i=1}^{k} \lambda_i(\sigma)$, for all $k$ (here $\lambda_i(M)$ denotes the $i^{th}$ greatest eigenvalue of the matrix $M$). Now, let $\rho_{\text{box}} \in B(CD)$ be the state inside the box(es), and consider the transformation $\rho = \rho_{\text{box}} \otimes (I_d/d) \rightarrow \sigma = (I/D) \otimes \rho_{\text{swap}}$ that results after applying a full SWAP between the box and a trusted qudit in the maximally mixed state, followed by the replacement of the state inside the box by a maximally mixed quDit. From [36] it follows that $\lambda_1(\rho)/d \geq \lambda_1(\rho_{\text{swap}})/D$, and so $D$ satisfies

$$D \geq d\lambda_1(\rho_{\text{swap}}).$$

(3.8)

By minimizing $\lambda_1$, we can hence lower bound the dimensionality $D$ of the system inside the box.

4 Conclusion

From the simple but powerful idea of the SWAP tool, we have obtained a robust bound for the singlet fidelity of a quantum black box. This bound, as a function of the CHSH violation of the corresponding black box, provides a quick and practical guide on how close the underlying quantum state is to the singlet.

But the method developed here goes beyond the simple CHSH scenario: for instance, using a tilted CHSH inequality, we have derived robust and practical fidelity bounds on how close the black box is to a partially entangled state for a given inequality violation. Similarly, and going beyond dichotomic observables, we have shown that the CGLMP inequality can be used for robust self testing as well. As a side product, we have proven true the conjecture stated in [31] that only non maximally entangled states can violate CGLMP maximally.

By extending the SWAP unitary into a full SWAP, we showed that the basic tool can be further utilized in much broader contexts. We used the full SWAP methods to certify non-commuting pairs of Pauli measurements. We used it to certify the entangling properties of a joint measurement in a robust way. And last but not least, we used it to bound the amount of work extractable from a quantum black box as a function of its CHSH violation, thus providing an interesting link between thermodynamics and nonlocality.

The problem of building more efficient SWAP operators with minimum computational complexity, i.e. operators that would approximate perfect SWAPs optimally while only involving low degree polynomials of the physical measurement operators, remains open. This topic has been analyzed in Section 2.1, where we showed that violation-dependent linear combinations of Alice’s operators led to better singlet fidelities. Our efforts to find better higher degree approximations were, however, unfruitful. An answer to this problem may be connected to the treatment of reducible (block-diagonal) measurement operators, that, for the sake of brevity, we have skipped in this manuscript. This and other related topics, such as the limitations of self-testing of extremal quantum distributions, will be covered in a future work.
Acknowledgements

This work is funded by the Singapore Ministry of Education (partly through the Academic Research Fund Tier 3 MOE2012-T3-1-009), by the National Research Foundation of Singapore. M.N. acknowledges support from the John Templeton Foundation. T.V. acknowledges financial support from a János Bolyai Grant of the Hungarian Academy of Sciences, the Hungarian National Research Fund OTKA (PD101461), and the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project.

References

[1] D. Leibfried, R. Blatt, C. Monroe and D. Wineland, Rev. Mod. Phys., 75, 281 (2003).
[2] V. Scarani, Acta Physica Slovaca 62, 347 (2012).
[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani and S. Wehner, arXiv:1303.2849 (2013).
[4] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio and V. Scarani, Phys. Rev. Lett. 98, 230501 (2007).
[5] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, New J. Phys. 11, 045021 (2009).
[6] Ll. Masanes, S. Pironio and A. Acín, Nat. Commun. 2, 238 (2011).
[7] E. Hänggi and R. Renner, arXiv:1009.1833 (2010).
[8] S. Pironio, Ll. Masanes, A. Leverrier and A. Acín, Phys. Rev. X 3, 031007 (2013).
[9] R. Colbeck, PhD Dissertation, University of Cambridge (2007), arXiv:0911.3814.
[10] S. Pironio, A. Acín, S. Massar, A.B. de la Giroday, D.N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T.A. Manning and C. Monroe, Nature 464, 1021 (2010).
[11] D. Mayers and A. Yao, Quant. Inf. Comput. 4, 273 (2004).
[12] T. Franz, F. Furrer and R.F. Werner, Phys. Rev. Lett. 106, 250502 (2011).
[13] S. Popescu and D. Rohrlich, Phys. Lett. A 169, 411 (1992).
[14] C.-E. Bardyn, T.C.H. Liew, S. Massar, M. McKague, and V. Scarani, Phys. Rev. A 80, 062327 (2009).
[15] M. McKague, T.H. Yang, and V. Scarani, J. Phys. A: Math. Theor. 45 455304 (2012).
[16] C.A. Miller and Y. Shi, arXiv:1207.1819 (2012).
[17] B.W. Reichardt, F. Unger and U. Vazirani, Nature 496 (2013).
[18] T.H. Yang and M. Navascués, Phys. Rev. A 8, 050102(R) (2013).
[19] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann and O. Gühne, arXiv:1302.1336 (2013).
[20] M.F. Pusey, arXiv:1305.1767 (2013).
[21] T. Ito, H. Kobayashi and K. Matsumoto, arXiv:0810.0693v1 (2008).
[22] T. Fritz, T. Netzer and A. Thom, arXiv:1207.0975 (2012).
[23] M. Navascués, S. Pironio and A. Acín, Phys. Rev. Lett. 98 010401 (2007).
[24] Y. Zhang, S. Glancy and E. Knill, Phys. Rev. A 84, 062118 (2011).
[25] M. Navascués, S. Pironio and A. Acín, New J. Phys. 10 073013 (2008).
[26] S. Pironio, M. Navascués and A. Acín, SIAM J. Optim. Volume 20, Issue 5, pp. 2157-2180 (2010).
[27] S.P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
[28] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).
[29] J. S. Bell, Physics 1 195 (1964).
[30] D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu, Phys. Rev. Lett. 88, 040404 (2002).
[31] A. Acín, T. Durt, N. Gisin, and J. I. Latorre, Phys. Rev. A 65, 052325 (2002).
[32] L. Szilard, Phys. 53 840-56 (1929).
[33] O. C. O. Dahlsten, R. Renner, E. Rieper and V. Vedral, New J. Phys. 13, 053015 (2011).
[34] C.H. Bennett, Int. Jour. Theor. Phys. 21, No 12, (1982).
[35] R. P. Feynman, (Addison-Wesley, Boston, 1998).
[36] M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. A 67, 10 (2003).
[37] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner and V. Scarani, Phys. Rev. Lett. 107, 050502 (2011).
[38] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).
[39] C. Jordan, Bull. Soc. Math. Fr. 3 103 (1875).
[40] A. Acín, S. Massar and S. Pironio, Phys. Rev. Lett. 108, 100402 (2012).
[41] M. Takesaki. *Theory of Operator Algebras I*, Springer (1979).
[42] A. C. Doherty, Y.-C. Liang, B. Toner and S. Wehner, Proceedings of IEEE Conference on Computational Complexity 2008, pages 199–210 (2008).
A The SWAP method

The fact that we can measure experimentally a non-local distribution $P(a, b|x, y)$ implies that there must exist a quantum state $|\psi\rangle$ and projector operators $\{E^x_a, F^y_b\}$, with $\sum_a E^x_a = 1_A$, $\sum_b F^y_b = 1_B$ such that

$$P(a, b|x, y) = \langle \psi| E^x_a \otimes F^y_b |\psi\rangle.$$  \hfill (A.1)

Suppose now that we have a mathematical guess $\left(\overline{\psi}, \{E^x_a, F^y_b\} \subset B(\mathbb{C}^d \otimes \mathbb{C}^d)\right)$ on such quantum state and measurement operators, and we want to test how wrong it can be. Let us assume, for simplicity, that the algebras generated by $\{E^x_a\}$ and $\{F^y_b\}$ are irreducible, i.e., that they cannot be simultaneously block-diagonalized:

$$E^x_a \neq \oplus_k E^x_a(k), F^y_b \neq \oplus_k F^y_b(k).$$  \hfill (A.2)

By the Artin-Wedensburn theorem [41], any matrix in $\mathbb{C}^d \times \mathbb{C}^d$ is hence an element of the algebra generated by Alice’s operators $\{E^x_a\}$ (Bob’s operators $\{F^y_b\}$). Consequently, we can express the unitary operators

$$X \equiv \sum_{n=0}^{d-1} e^{i 2 \pi n/d} |n\rangle\langle n|, \quad P \equiv \sum_{n=0}^{d-1} |n + 1 \pmod{d}\rangle\langle n|$$  \hfill (A.3)

as linear combinations $X_A(E^x_a)$, $P_A(E^x_a)$, $X_B(F^y_b)$, $P_B(F^y_b)$ of products of Alice’s (Bob’s) projector operators.

Now, consider the expressions $X_A(E^x_a)$, $P_A(E^x_a)$ which result when we substitute $\{E^x_a\}$ by general measurement operators $\{E^x_a\}$. In general, $X_A(E^x_a)$ and $P_A(E^x_a)$ will not be unitary. However, by the polar decomposition [41], there always exist unitaries$^1$$\hat{X}_A, \hat{P}_A$ such that

$$\hat{X}^\dagger_A X_A(E^x_a) \geq 0, \quad \hat{P}^\dagger_A P_A(E^x_a) \geq 0.$$  \hfill (A.4)

Such unitaries will satisfy $\hat{X}_A = X_A$, $\hat{P}_A = P_A$ whenever $X_A, P_A$ are unitaries themselves.

$X_A, P_A$ are important because, as in eq. (2.6), the operators

$$U_A = \sum_{k=0}^{d-1} (\hat{P}_A)^k \otimes |k\rangle\langle k|, \quad V_A = \sum_{k=0}^{d-1} (\hat{X}_A)^k \otimes |\psi_k\rangle\langle \psi_k|,$$  \hfill (A.5)

with $|\psi_j\rangle \equiv \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} e^{i 2 \pi kj/d} |k\rangle$, are unitary no matter what unitary operators $\hat{X}_A, \hat{P}_A$ we substitute in the above expression. Moreover, when $\hat{X}_A = X, \hat{P}_A = P, U_A$ and $V_A$ correspond to generalized CNOT gates from Alice’s box to her trusted qudit and from the qudit to Alice’s box, respectively. It can be verified that the gates $U_A V_A, V_A^\dagger U_A V_A$ respectively correspond to the partial and full SWAP in the exact case.

$^1$Technically, there always exist isometries with the said property. However, any isometry $V \in B(\mathcal{H})$ in infinite dimensions can be viewed as a unitary operator in $\mathcal{H} \otimes \mathbb{C}^d$. Indeed, let $V^* V = 1$ and define $U = (\mathbb{I} - VV^*) \otimes |1\rangle\langle 0| + V^\dagger \otimes |1\rangle\langle 1| + V \otimes |0\rangle\langle 0|$. Then $UU^\dagger = U^\dagger U = 1$, and $U|\psi\rangle = (V|\psi\rangle)|0\rangle$. At the level of the moment matrices, we can thus assume that such isometries are unitaries.
Now the only open issue is that, in order to apply the SWAP method, we must extend our optimizations to unitary operators \( \hat{X}_A, \hat{P}_A \) satisfying (A.4). We will take care of that by extending the moment matrix of the system. That is, we will include the non-commutative variables \( \hat{X}_A, \hat{P}_A, \hat{X}_B, \hat{P}_B \) in the sequences of operators that label the rows and columns of \( \Gamma \). Likewise, we will introduce extra semidefinite constraints in our programs, in the form of localizing matrices [26].

Given a polynomial \( f \) of non-commutative variables, or equivalently, a linear combination \( f = \sum_i f_i T \) of sequences of operators \( T \), the localizing matrix \( \Gamma^S(f) \) is a matrix whose rows and columns are numbered by sequences of operators that label the rows and columns of \( \Gamma \). That is, \( \Gamma^S(f) \) must extend our optimizations to unitary operators \( X \) (c.f. Eq. (A.5)) via the relation:

\[
\Gamma^S_{as}(f) = \sum_t f_t c_{atst'}.
\]  

The constraints (A.4) hence translate as:

\[
\begin{align*}
\Gamma^S(\hat{X}_A X_A(E_a^x)) & \geq 0, \Gamma^S(\hat{P}_A P_A(E_a^x)) \geq 0, \\
\Gamma^S(\hat{X}_B X_B(E_b^y)) & \geq 0, \Gamma^S(\hat{P}_B P_B(E_b^y)) \geq 0.
\end{align*}
\]  

(A.7)

Note that requiring \( \Gamma^S(\hat{X}_A X_A(E_a^x)) \) to be positive also implies that it must be hermitian.

Putting all together, we have that the estimation of the fidelity of the state inside the box \( |\psi\rangle \) with respect to \( |\tilde{\psi}\rangle \) can be relaxed to the following SDP:

\[
f^S = \min \langle \psi | \rho_{\text{swap}}(c) | \tilde{\psi} \rangle
\]

such that

\[
\begin{align*}
\Gamma^S(c) & \geq 0, \\
c_{E_tF_t'} = P(a, b|x, y) & \\
\Gamma^S(\hat{X}_A X_A(E_a^x)) & \geq 0, \Gamma^S(\hat{P}_A P_A(E_a^x)) \geq 0, \\
\Gamma^S(\hat{X}_B X_B(E_b^y)) & \geq 0, \Gamma^S(\hat{P}_B P_B(E_b^y)) \geq 0.
\end{align*}
\]  

(A.8)

The sequence of relaxations (A.8) is known to converge to the optimal value compatible with a commutative (i.e., non-tensorial) quantum representation [25, 26, 42]. Consequently, if \( P(a, b|x, y) \) can only be achieved with the quantum representation \( (|\psi\rangle, \{E_a^x, F_b^y\}) \subset B(\mathbb{C}^d \otimes \mathbb{C}^d) \), the above sequence of SDPs should retrieve the exact value of any quantity we are trying to estimate.

A drawback of the previous scheme is that, even if the representation \( (|\psi\rangle, \{E_a^x, F_b^y\}) \subset B(\mathbb{C}^d \otimes \mathbb{C}^d) \) consists of real matrices and vectors we are forced to optimize over complex variables in our SDP, because \( V \) involves complex numbers through \( \xi \) (c.f. Eq. (A.3)) and \( |\psi_j\rangle \) (c.f. Eq. (A.5)), which could appear as complex coefficients in the objective function of Eq. (A.8). This problem can be solved by noticing that \( V_A \) in eq. (A.5) can also be expressed as

\[
V_A = \sum_{k=0}^{d-1} P_k \otimes P_k,
\]  

(A.9)
where \( \{ P_n \} \) is a complete set of orthogonal projectors. \( P_n \), in turn, can be defined via the relations

\[
P_n P_m = \delta_{m,n} P_n, \quad \sum_n P_n = I, \quad (n + 1/2) P_n \geq N(E^+_{x}) P_n \geq (n - 1/2) P_n,
\]

where \( N(E^+_{x}) \) is such that

\[
N(\bar{\Pi}^x_{a}) = d - 1 \sum_{n=0} N|n\rangle \langle n|.
\]

This alternative formulation of the problem makes it necessary to introduce many more variables and localizing matrices, so in some instances it may be preferable to use the complex formulation.

Finally, let us remark that the above schemes for the SWAP method are completely general, and thus unnecessarily inefficient in most interesting cases. As shown in the subsequent Appendices, much simpler schemes involving just real SDPs and one or two localizing matrices are enough to generate full or partial SWAPs in the non-locality scenarios considered in this paper.

## B The non-i.i.d. case

Suppose that Alice and Bob are sequentially distributed pairs of black boxes with two inputs and two outputs. We allow for the possibility that different pairs of boxes exhibit different statistics, which can, in turn, depend on Alice and Bob’s past measurement history. Under these circumstances, we need to disprove the hypothesis \( \Phi \) that all the distributed pairs contain quantum states with singlet fidelity smaller than or equal to \( f \).

Denote by \( F : [2, 2\sqrt{2}] \rightarrow [0, 1] \) the function given by the solid curve in Figure 4 that assigns singlet fidelities to CHSH parameters. Call \( V_0 \) the CHSH violation such that \( F(V_0) = f \), and let \( \Psi \) be a particular box with CHSH violation \( V > V_0 \).

Note that, if Alice and Bob choose their measurement settings \( x,y \) randomly and uniformly, the CHSH Bell inequality can be written as \( \langle B(a, b, x, y) \rangle \leq 2 \), with

\[
B(a, b, x, y) = 4(-1)^{xy}\{\delta_{a,b} - \delta_{-a,b}\}.
\]

Following the lines of [24], define the normalized form of the CHSH inequality \( \tilde{B}(a, b, x, y) \equiv \frac{B(a, b, x, y)}{V_0 + 4} \). Clearly, \( \langle \tilde{B}(a, b, x, y) \rangle_\Psi > 1 \), and \( \tilde{B}(a, b, x, y) \geq 0 \) for all \( x, y, a, b \). Also, \( \langle \tilde{B}(a, b, x, y) \rangle \leq 1 \) for any pair of boxes satisfying hypothesis \( \Phi \).

Next, choose \( 0 < \epsilon < 1 \) such that

\[
R(a, b, x, y) \equiv (1 - \epsilon) + \epsilon \tilde{B}(a, b, x, y)
\]

satisfies

\[
\langle \log[R(a, b, x, y)] \rangle_\Psi > 0.
\]

That such an \( \epsilon \) exists follows from the observation that, for \( \epsilon \ll 1 \),
\[
\langle \log(1 - \epsilon + \epsilon \hat{B}(a, b, x, y)) \rangle_\Psi \approx \epsilon \langle (\hat{B}(a, b, x, y) - 1) \rangle_\Psi > 0. \quad (B.4)
\]

Note that, by construction, \( \langle R(a, b, x, y) \rangle \leq 1 \) under hypothesis \( \Phi \).

Now, suppose that Alice and Bob conduct the CHSH experiment \( n \) times, choosing their inputs \( x, y \) with probability \( 1/4 \) each time, thus obtaining the experimental data \( \{a_k, b_k, x_k, y_k\}_{k=1}^n \). Define the positive random variable \( T = \prod_{k=1}^n R_k \), with \( R_k \equiv R(a_k, b_k, x_k, y_k) \). Under hypothesis \( \Phi \), it can be seen that \( T \leq 1 \) \cite{24}, and so, by Markov’s inequality, \( P(T \geq t) \leq 1/t \). However, in the event that Alice and Bob are actually being distributed \( n \) independent copies of box \( \Psi \), by the central limit theorem, the random variable \( \log(T) = \sum_{k=1}^n \log(R_k) \) is expected to take values in the range \( n \langle \log(R_k) \rangle_\Psi \pm O(\sqrt{n}) \).

From eq. (B.3), we thus have that, with very high probability, \( T \) will grow exponentially with \( n \). In a few experiments, Alice and Bob will hence observe a ridiculously high value of \( T \), and therefore conclude that hypothesis \( \Phi \) must be abandoned.

## C Non-Maximally Entangled State Certification

### C.1 Tilted CHSH

For convenience, we shall work in the local basis such that Bob’s optimal measurements are \( B_0 = \sigma_z \). In this basis, \( B_1 = \cos(2\mu)\sigma_z - \sin(2\mu)\sigma_x \). Note that \( B_1 \) is no longer \( \sigma_x \), as in the CHSH case, thus the partial SWAP \( S_B \) defined in Section 2.1 no longer works. We will have to construct the analog of \( \sigma_x \) by combining the operators \( B_0 \) and \( B_1 \). We thus introduce an auxiliary dichotomic operator \( B_2 \), with \( B_2^2 = I \), and impose relations between \( B_0, B_1, B_2 \) such that \( B_2 \) behaves as if it is \( (\cos(2\mu)B_0 - B_1)/\sin(2\mu) \). Following Appendix A, the appropriate constraint is

\[
B_2 \frac{\cos(2\mu)B_0 - B_1}{\sin(2\mu)} \geq 0. \quad (C.1)
\]

This will force \( B_2 \) to share the same eigenvectors as \( (\cos(2\mu)B_0 - B_1)/\sin(2\mu) \), and will identify both operators in the optimal case.

The swap operators in this case will then be

\[
S_B = (I \otimes |0\rangle\langle 0| + B_2 \otimes |1\rangle\langle 1|) \left( \frac{I + B_0}{2} \otimes I + \frac{I - B_0}{2} \otimes \sigma_x \right), \quad (C.2)
\]

while \( S_A \) is the same as in Section 2.1.

With the introduction of this extra auxiliary operator, our moment matrix is now enlarged with effectively two operators on Alice’s side \( (A_0, A_1) \) and three operators on Bob’s side \( (B_0, B_1, B_2) \). The constraint (C.1) is enforced by imposing that the localizing matrix defined by

\[
\Gamma \left( \frac{B_2 \cos(2\mu)B_0 - B_1}{\sin(2\mu)} \right)_{ss'} = \sum_s \frac{\cos(2\mu)}{\sin(2\mu)} \epsilon_s B_{ss'} - \frac{1}{\sin(2\mu)} \epsilon_s B_{ss'} \quad (C.3)
\]

is positive semidefinite.
C.2 CGLMP

One strategy to achieve the maximal violation of CGLMP is as follows: Alice’s and Bob’s first measurements $x = y = 0$ correspond to the projectors \{\{0\}\{0\}, |1\rangle\langle 1|, |2\rangle\langle 2|\}. The projectors $|\alpha_m\rangle\langle \alpha_m|, |\beta_m\rangle\langle \beta_m|$ corresponding to measurements $x, y = 1$, together with the state to be measured $|\psi\rangle$ are as follows

$$|\alpha_m\rangle = \frac{1}{3} (2|m| + 2|m+1| - |m+2\rangle),$$

$$|\beta_n\rangle = \frac{1}{3} (2|n| + 2|n+1| - |n+2\rangle),$$

$$|\psi\rangle = \frac{1}{3\sqrt{2 + \gamma^2}} \left( (\gamma + \sqrt{3})(|00\rangle + |11\rangle + |22\rangle) + \gamma(|01\rangle + |12\rangle + |20\rangle) + (\gamma - \sqrt{3})(|02\rangle + |10\rangle + |21\rangle) \right),$$

where all addition above performed inside the kets are modulo 3. This strategy is unitarily equivalent to the measurement scheme presented in the main text, which involves complex coefficients.

The above measurements and states shall then be our reference system. To certify the state, as usual, Alice and Bob will each attach a trusted qutrit to the entangled pair. The next step is to construct CNOT operators with which to build a partial SWAP, see Appendix A. The key point is how to build the translation operator $P$ from the measurement projectors defined in Eqn (C.4). There are many choices to do so; we chose the simplest combination:

$$P = \bar{E}_0^0 + 2\bar{E}_2^0 + \frac{1}{2}\bar{E}_1^0 - \frac{3}{2}\bar{E}_0^1 (2\bar{E}_1^1 + \bar{E}_2^1)$$

$$- \frac{3}{2}\bar{E}_1^1 (\bar{E}_1^1 - \bar{E}_2^1) - \frac{3}{2}\bar{E}_2^1 (\bar{E}_1^1 + 2\bar{E}_2^1),$$

which indeed is a translation operator mapping $|0\rangle \rightarrow |1\rangle \rightarrow |2\rangle \rightarrow |0\rangle$. Since Alice and Bob’s optimal operators are identical, the above formula also applies to Bob’s settings if we replace $E$’s by $F$’s.

Now as before, the choice above in Eqn (C.5) is a valid unitary operator only for the optimal strategy in Eqn (C.4). As before, in the device independent scenario, there is no guarantee that it is still a valid unitary operator, and so we must introduce an extra auxiliary operator, $\hat{P}_A$, with the constraint that $\hat{P}_A^1 P(E_a^+) \geq 0$. For Bob’s side, the swap operators are defined exactly the same way as above for Alice. Thus, we require also another auxiliary operator $\hat{P}_B$.

The self-testing of the CGLMP inequality therefore requires two localizing matrices $\bar{\Gamma}(\hat{P}_A^1 P(E_a^+)), \bar{\Gamma}(\hat{P}_B^1 P(F_a^+))$. All three semidefinite matrices can be taken real here, since, for any feasible point $\bar{\Gamma}, \bar{\Gamma}(\hat{P}_A^1 P(E_a^+)), \bar{\Gamma}(\hat{P}_B^1 P(F_a^+))$ of the corresponding complex SDP, the real matrices $\text{Re}\{\bar{\Gamma}\}, \text{Re}\{\bar{\Gamma}(\hat{P}_A^1 P(E_a^+))\}, \text{Re}\{\bar{\Gamma}(\hat{P}_B^1 P(F_a^+))\}$ are also positive semidefinite, satisfy the appropriate linear constraints and return the same state fidelity.

D Work extraction

Let $\rho_{AB}$, with spectral decomposition $\rho_{AB} = \sum_{k=1}^4 \lambda_k |\psi_k\rangle \langle \psi_k|$, describe the state of two particles in a Szilard engine of length $L$ and area $A$, with $\lambda_i \geq \lambda_{i+1}$. 

23
Denoting by $L$ ($R$) the state corresponding to a particle being on the left (right) of the Szilard engine, we can always apply a unitary $U$ over the state $\rho_{AB}$ such that $U|\psi_1\rangle = |L, L\rangle$, $U|\psi_2\rangle = |R, L\rangle$, $U|\psi_3\rangle = |L, R\rangle$, $U|\psi_4\rangle = |R, R\rangle$. It follows that a population of $N$ particle pairs in a Szilard engine, initially in state $\rho_{AB}^\otimes N$, can be brought to a situation where $N_L = N(\lambda_1 - \lambda_4 + 1)$ particles are on the left side of the engine, and $N_R = N(\lambda_4 - \lambda_1 + 1)$ are on the right side. If we place a movable wall connected to a weight on the left side of the engine (in contact with a bath at temperature $T$), the pressure over the wall is equal to

$$P = \frac{N_L K T}{A z} - \frac{N_R K T}{A (L - z)},$$

where $z$ denotes the position of the piston. At constant temperature, the equilibrium position of the piston is $z_{eq} = \frac{N_L L}{N_L + N_R}$. The work extracted in the process of moving the piston from $z = L/2$ to $z_{eq}$ is

$$W = \int_{L/2}^{eq} P Adz = KT \left\{ N_L \ln \left( \frac{2N_L}{N_L + N_R} \right) + N_R \ln \left( \frac{2N_R}{N_L + N_R} \right) \right\} =: W(\lambda_1 - \lambda_4).$$

(D.1)

Finally, the minimum value of $\lambda_4 - \lambda_1$ can be extracted from $\rho_{\text{swap}}$ via the following SDP:

$$\begin{align*}
\min & \quad \mu_1 - \mu_4 \\
\text{s.t.} & \quad \rho_{\text{swap}} - \mu_4 I \geq 0, \\
& \quad \mu_1 I - \rho_{\text{swap}} \geq 0.
\end{align*}$$

(D.2)

Since

$$\begin{align*}
\rho_{\text{swap}} - \mu_4 I & \geq 0 \Rightarrow \Re\{\rho_{\text{swap}}\} - \mu_4 \geq 0, \\
\mu_1 I - \rho_{\text{swap}} & \geq 0 \Rightarrow \mu_1 \geq \Re\{\rho_{\text{swap}}\} \geq 0,
\end{align*}$$

(D.3)

it follows that the free variables in the corresponding SDP can be taken real, as in the previous examples.
CHSH violation vs. Singlet fidelity