Self-testing is a method to infer the underlying physics of a quantum experiment in a black box scenario. As such it represents the strongest form of certification for quantum systems. In recent years a considerable self-testing literature has developed, leading to progress in related device-independent quantum information protocols and deepening our understanding of quantum correlations. In this work we give a thorough and self-contained introduction and review of self-testing and its application to other areas of quantum information.

1 Introduction

In contrast to classical theories, states in quantum physics can be entangled and sets of measurements can be incompatible. As shown by Bell in 1964 [Bel64], these features imply striking observable phenomena. In particular, the outcomes of incompatible measurements made on the local subsystems of an entangled quantum state can exhibit correlations that are provably stronger than any resulting from a classical theory, a phenomenon known as Bell nonlocality. The field of Bell nonlocality has since grown considerably (see [BCP+14] for a recent review article), and the existence of Bell nonlocal correlations in nature is now a well established fact [HBD+15, GVW+15, SMSC+15].

As more was understood about Bell nonlocality, a number of works [SW87, PR92, BMR92, Tsi93] eventually pointed out that there exist Bell nonlocal correlations that—as well as requiring entanglement and incompatibility—can only be produced by making particular sets of incompatible measurements on particular entangled states. These works have since given birth to the field of self-testing, which broadly speaking aims to understand the structure of the set of quantum correlations and identify those correlations that admit a unique physical realisation.

An important milestone in the development of self-testing was the 2004 work of Mayers and Yao [MY04]. This work set the terminology and formalism that was to be adopted by later works, and includes the first usage of the term ‘self-testing’ in this context. A similar idea was already present in [MY98] in a cryptographic context, using the term ‘self-checking’ instead of ‘self-testing’. These early works also introduced the paradigm of device-independence, to which self-testing is intimately related. In particular, a self-testing protocol can be seen as a device-independent—or black box—certification of a quantum system, assuming that the system can be prepared many times in an independent, identically distributed manner. Self-testing is consequently relevant to many device-independent quantum information protocols and has led to related progress in this area. More recently, self-testing has become synonymous with any protocol for certifying any type of quantum system under a small set of assumptions.

In this work we give a up-to-date review of the field of self-testing. We hope that it will be of use to people both unfamiliar with the field, as well as serving as a reference for those within it. The review is organised as follows. In section 2 we give a gentle introduction to device-independence and its connection to self-testing. We then formally introduce self-testing, giving the mathematical definitions in section 3 and a simple example in section 4 that illustrates many important concepts. Sections 5 to 8 are a thorough literature review of state and measurement self-testing, explaining the tools and techniques that are commonly used along the way. In section 9 we review extensions of self-testing to other
scenarios, and in section 10 the application of self-testing to other fields in quantum information theory. In section 11 we cover experimental realisations of self-testing protocols. Finally, in section 12 we discuss some possible future directions for the field and a number of open problems.

We point the reader to the related review articles [MdW16] and [Sca12] where discussions about self-testing can also be found. [MdW16] deals with the classical and quantum certification of both classical and quantum properties of an object, with self-testing being identified as classical certification of quantum properties. [Sca12] provides a pedagogical review of the device-independent approach to quantum physics. Self-testing is discussed as one of device-independent protocols. We also recommend [McK10, Kan17, Kan16] as valuable texts for first time readers.

Contents

1 Introduction 1

2 Self-testing as a device-independent protocol 3

3 Definitions 5
  3.1 Notation 5
  3.2 The self-testing scenario 6
  3.3 Self-testing of states 7
  3.4 Self-testing of measurements 8
  3.5 Self-testing via a Bell inequality and the geometry of the set of quantum correlations 8
  3.6 Robust self-testing 9
  3.7 Generalisations and alternative definitions 10

4 A first example 11
  4.1 Geometrical proof of anticommutativity 13
  4.2 Algebraic proof of anticommutativity 13
  4.3 Swap gate 13
  4.4 Self-testing of measurements 16

5 Self-testing of bipartite states 16
  5.1 Self-testing of two-qubit states 16
  5.2 Self-testing of qudit states 18
  5.3 Self-testing $n$ maximally entangled pairs of qubits 19

6 Self-testing of multipartite states 21
  6.1 Self-testing of graph states from stabilizer operators 21
  6.2 Tailoring Bell inequalities 22
  6.3 Reductions to bipartite methods 23
  6.4 Parallel self-testing of multipartite states 23
  6.5 Self-testing using only marginal information 23

7 Robust self-testing of states 23
  7.1 Robust self-testing methods 24
  7.2 Robust certification of large entanglement 28

8 Self-testing of measurements 29
  8.1 Measurement self-testing results 29
  8.2 Methods in measurement self-testing 30
  8.3 Robust measurement self-testing 32

9 Extensions of self-testing to other scenarios 33
  9.1 Self-testing of quantum gates and circuits 33
  9.2 Semi-device-independent scenarios 34

10 Applications of self-testing 37
  10.1 Device-independent randomness generation 38
  10.2 Device-independent quantum cryptography 39
  10.3 Entanglement detection 40
  10.4 Delegated quantum computing 41
  10.5 Structure of the set of quantum correlations 41

11 Experiments 42

12 Concluding remarks and open questions 43

Acknowledgements 45

Appendix 45

A Self-testing complex measurements 45
A.1 Self-testing complex measurements 45
A.2 Regularisation trick 45
A.3 Swap isometries 46
A.4 Localising matrices in the Swap method 46

B State and measurement assumptions 47
B.1 State 47
B.2 Measurements 47

Accepted in Quantum 2020-23-08, click title to verify
2 Self-testing as a device-independent protocol

The treatment of complex systems as black boxes is a powerful tool in many scientific domains, providing a minimalist level of abstraction that allows one to focus on what a device or system does without the need to model precisely how this is achieved. In quantum information theory, this approach is known as the device-independent (DI) approach.

In order to explain the idea of the device-independent approach we imagine the following scenario. Consider two laboratories, run by two experimenters called Carmela and Deng. In their laboratories (let’s imagine they are quantum optics laboratories) both Carmela and Deng have access to some equipment (e.g. lasers, beamsplitters, waveplates, photon detectors,...) which they can use to perform different experiments. A given experiment consists of a choice of settings (e.g. laser intensity, angle of the waveplates, type of beamsplitter,...) that after a run of the experiment provides a result (e.g. photon detection location, time of detection,...). Furthermore, a source is positioned between the laboratories and emits physical systems (e.g. photons) that are sent to Carmela’s and Deng’s laboratories; see figure 1, left.

Suppose that Carmela and Deng would like to learn if the source is emitting entangled particles (where the entanglement is with respect to the two laboratories). One way to achieve this is to use their equipment to perform tomography of the state of the source, i.e. Carmela and Deng perform a number of experiments each with different settings, collect statistics of the results, and use quantum state tomography to reconstruct the density matrix of the state, which can then be checked to determine if it is entangled (for instance using an entanglement witness). This is indeed what is done in many experiments around the world.

Imagine now however that two computer scientists—called Alice and Bob—are visiting each of the labs. Despite knowing the mathematical definition of entanglement, they will have problems convincing themselves that the source is producing entanglement. Firstly, they do not understand the experimental setup, so they do not know what the different settings do. Moreover, even if they were told what the settings do, they do not have a good understanding of quantum optics. As a result, they will not be able to reconstruct the state of the source in order to check if it is entangled, as was the case for Carmela and Deng.

Alice, however, proposes the following: even though they do not understand what the settings do, they can still change them and observe something. That is, they can simply model their laboratories as black boxes. Each laboratory is treated as a device (a black box) that takes an input (the settings) and returns output (the result), but the physical mechanism behind how this occurs is unknown (see figure 1, centre). Similarly, they do not assume anything about the source; all they know is that it is distributing some physical systems that may or may not be entangled. Alice denotes each of her possible settings as \( x = 0, 1, \ldots \) and Bob denotes each of his possible settings as \( y = 0, 1, \ldots \). Similarly Alice and Bob denote the possible results of their experiments by \( a = 0, 1, \ldots \) and \( b = 0, 1, \ldots \).

After trying the different settings sufficiently many times and collecting statistics, Alice and Bob can estimate the probabilities (also called the correlations)

\[
p(a, b|x, y),
\]

that is, the probabilities to see the results \( a \) and \( b \) given that the settings \( x \) and \( y \) are used. It is important here to stress that although Alice and Bob can estimate these probabilities, they are ignorant about the underlying physics; from their perspective the experiments could have been made on atoms, electrons, neutrinos or any other physical system. This scenario is called the device-independent scenario. Remarkably, even with such little knowledge, Alice and Bob can still conclude that the source emits entangled states.

The trick to achieving this is to use Bell nonlocality, a counter-intuitive phenomenon discovered by John Bell in 1964 [Bel64] (see also box 3.1). At the heart of Bell nonlocality are objects called Bell inequalities. A Bell inequality consists of a function \( I \) of the probabilities \( \{p(a, b|x, y)\} \) such that, for a source producing separable (i.e. non-entangled) states one has

\[
I(\{p(a, b|x, y)\}) \leq \beta.
\]
Figure 1: (left) A source produces systems and distributes them between two laboratories that can perform different
local experiments by varying the settings of their equipment. (centre) In the device-independent approach, each local
laboratory is treated as a black box that takes as input a label that corresponds to a particular choice of settings for the
experiment, and outputs a label that denotes the corresponding result. After repeating the experiments sufficiently
many times, the probabilities \( p(a, b | x, y) \) can be estimated. (right) When self-testing a state and measurements,
one aims to infer the form of the state of the source and measurement operators describing the laboratories from
knowledge of the probabilities alone, i.e. in a black box scenario.

Importantly, the bound \( \beta \) holds for any source producing any kind of physical systems, provided that these systems are not entangled. This is a consequence of the fact that the definition of separability is independent of the physical system in question. Interestingly, Bell inequalities can be violated by entangled systems. That is, for some entangled sources one can achieve \( I(p(a, b | x, y)) > \beta \). This situation can be understood geometrically in the vector space of probabilities, where Bell inequalities are defined by linear hyperplanes; see Fig. 2.

Alice and Bob can therefore do the following. They compare their probabilities against as many Bell inequalities as they know. If they see that one is violated, then it must be that the source is entangled! Moreover, they are able to conclude this despite knowing nothing about how the experiment was actually performed; all the information that was needed was the probabilities \( \{ p(a, b | x, y) \} \). Such a procedure is called a device-independent certification of entanglement.

Suppose now that we change the task: instead of only detecting entanglement, Alice and Bob want to know the particular entangled state of the source. Due to the device-independent scenario, they cannot know the type of physical system that the source is producing. However, they may hope to write down the state vector \( |\psi\rangle \) of the source, without specifying which types of physical degrees of freedom it describes. This often turns out to be possible (up to some local transformations, see the definitions in the following section), as long as one observes the maximum possible violation achievable in quantum theory of a corresponding Bell inequality. Such a procedure is called a device-independent self-test or simply a self-test of the state. Often, this maximal violation also allows one to self-test the measurements, i.e. to determine the form of the measurement operators that describe how the outcomes \( a \) and \( b \) are produced in the local laboratories.

Alice and Bob will nevertheless face a problem: they will never see the maximum violation of a Bell inequality due to experimental noise and finite statistics. At best, they will be able to lower bound the violation up to some statistical confidence. To become a practically relevant protocol, self-testing therefore has to be combined with two other tools: (i) noise-resistant self-testing methods (called robust self-testing; see section 3.6) which give distance upper bounds to the desired state and/or measurements as a function of experimental noise; and (ii) tools of statistical analysis (such as Chernoff bounds) that allow for statistically valid estimations of probabilities and corresponding confidence levels. Self-testing can thus be seen as a theoretical tool that, when augmented by statistical techniques, can be transformed into a protocol for device-independent state and measurement certification.

A number of remarks are in order before proceeding to formal definitions of self-testing in the next section. First, one may wonder whether self-testing is possible using only a single device. That is, given a single device from which one observes the conditional probabilities \( p(a | x) \), is it possible to determine the quantum state \( |\psi\rangle \) inside? Notice that one possibility is that there is a pre-programmed classical computer inside the device that simply simulates the statistics \( p(a | x) \). In the device-independent scenario one cannot rule out this possibility. Thus, one cannot hope to certify any non-classical properties of \( |\psi\rangle \) with only
Bell inequality
\[ \vec{I} \cdot \vec{p} \leq \beta \]
\[ \Rightarrow |\psi\rangle, M_a|x, N_b|y \]

Figure 2: Geometric representation of self-testing. The figure is a 2-dimensional representation of the space of probability vectors \( \vec{p} = (p(00|00), p(01|00), \ldots) \). The grey area \( \mathcal{L} \) is the set of probability vectors that are obtainable with separable states (commonly called the ‘local set’). A Bell inequality (dotted line) consists of a pair \( (\vec{I}, \beta) \) such that the half space \( \vec{I} \cdot \vec{p} \leq \beta \) contains \( \mathcal{L} \). Entangled quantum states are capable of producing probability vectors that lie outside of \( \mathcal{L} \) (purple area) and thus violate Bell inequalities. Often, extremal points (e.g. \( \vec{q} \)) of the quantum set, which maximally violate some Bell inequality, admit—up to local transformations—a unique realisation in terms of a particular entangled state and measurements. Self-testing involves identifying such probability distributions and proving their unique realisation.

a single device. It is for this reason that it is crucial to move to a multipartite scenario in order to certify non-classical states. Indeed, it is precisely the phenomenon of Bell nonlocality that forbids an explanation using local pre-programmed classical devices.

Second, it is worth mentioning some practical advantages of the device-independent scenario. Suppose Alice and Bob are sold two devices that are claimed to contain a particular entangled state and perform certain measurements on it. Using self-testing, they will be able to conclude that the devices are indeed working correctly without having to understand precisely how they operate. This is clearly desirable from the perspective of Alice and Bob, especially for cryptographic applications where one would prefer not to trust the devices. From a more experimental perspective, the device-independent scenario naturally treats experimental errors at the level of the observed statistics, meaning for example, that a false positive detection of entanglement will never occur. Being the strongest form of device-independent certification, self-testing has proven to be useful for many device-independent protocols; see section 10 for more information.

Third, we note that the notion of device-independence referred to by the large majority of self-testing works assumes independent, identically distributed (i.i.d.) rounds of the experiment, i.e. the state and the measurements are assumed to be the same in each round and do not depend on past measurement results or choices. Such an assumption allows us to talk about the probability distribution \( p(a, b|x, y) \) that is valid in every experimental round. This however is not the strongest possible notion of device-independence where one further drops the i.i.d. assumption (typically used in the field of quantum cryptography). Self-testing under this more stringent notion of device-independence is little explored (although arguably relevant); see the concluding remarks in section 12 for further discussion.

Finally, we note that aside from its motivation for device-independent certification, self-testing can also be viewed as part of the general study of quantum correlations. In particular, quantum correlations are defined via the Born rule, whereby a state and measurement operators are mapped to a probability distribution. Generally, this map does not have an inverse, since many combinations of state and measurements can lead to the same probabilities. Self-testing identifies those points of the set of quantum probability distributions for which such an inverse exists, up to the local transformations defined in the next section. That is, it associates some (extremal) probability distributions with a unique realisation using a particular state and measurements. For this reason self-testing has been used to derive important results about the set of quantum correlations.

3 Definitions

3.1 Notation

We first introduce some notation. \( \mathcal{L}(\mathcal{H}) \) denotes the set of linear operators acting on Hilbert space \( \mathcal{H} \). Uppercase Roman letters denote a local party, most often either Alice, A, or Bob, B. Roman letters in either subscript or superscript denote the Hilbert space in which a state lives or on which
an operator acts e.g. $|\psi\rangle_\Lambda \in \mathcal{H}_\Lambda$. Consecutive labels denote tensor products of Hilbert spaces, e.g. $|\psi\rangle_{AB} \in \mathcal{H}_\Lambda \otimes \mathcal{H}_B$. Labels containing the same letter are implicitly assumed to refer to different local Hilbert spaces of a single subsystem, e.g. $A$ and $A'$ refer to two Hilbert spaces of Alice.

As a result, self-testing is often a useful tool in a variety of device-independent protocols (see section 10 for explicit examples). In the following section, we formalise these ideas and define precisely what it means to self-test a state and measurements.

### 3.2 The self-testing scenario

The device-independent scenario described in the previous section is commonly called a Bell test, and the probabilities $p(a, b|x, y)$ are called correlations. From quantum theory, we know that there exist measurement operators $M_{a|x} \in \mathcal{L}({\mathcal{H}}_\Lambda)$ acting on Alice’s local Hilbert space and satisfying

$$M_{a|x} \succ 0 \quad \forall x, a, \quad \sum_a M_{a|x} = 1_A \quad \forall x \quad (3)$$

describing how the outcomes $a$ are obtained given settings $x$. Similarly, there exist measurement operators $N_{b|y} \in \mathcal{L}({\mathcal{H}}_B)$ for Bob acting on his local Hilbert space. From here on, we work with a Naimark dilation of each of the measurements. The measurement operators are therefore projective:

$$M_{a|x} M_{a'|x} = \delta_{a, a'} M_{a|x} \quad \forall x, a, a', \quad N_{b|y} N_{b'|y} = \delta_{b, b'} N_{b|y} \quad \forall y, b, b'. \quad (4)$$

This can be justified if one takes the position that projective measurements and unitary evolution are the only fundamental operations in quantum theory. From this perspective, any measurement necessarily involves a projective measurement over a dilated space, the degrees of freedom of which belong to Alice and Bob’s local Hilbert spaces. For further discussion on this assumption, see Appendix B.

Now, from the Born rule, there must exist some quantum state $\rho_{AB} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B) \succ 0$, $\text{tr} \rho_{AB} = 1$ such that

$$p(a, b|x, y) = \text{tr} \left[ \rho_{AB} M_{a|x} \otimes N_{b|y} \right]. \quad (5)$$

In self-testing, one aims to infer the form of the state and the measurements in (5) from knowledge of the correlations $p(a, b|x, y)$ alone; i.e. in the device-independent scenario.

In order to write (5), one nevertheless needs to make some basic physical assumptions that constitute the definition of the device-independent scenario in most self-testing works. These are

1. The experiment admits a quantum description; i.e. there exists a quantum state and measurement operators that lead to the observed outcomes via the Born rule.

2. The laboratories of Alice and Bob are located at separate locations in space and there is no communication between the laboratories; e.g. Alice cannot send the choice of setting $x$ to Bob or vice-versa.

3. The settings $x$ and $y$ are chosen freely and independently of all other systems in the experiment. For example, the physical system used to generate $x$ does not have any correlations (quantum or classical) with the particles or the source or the laboratory of Bob.

4. Each round of the experiment is independent from all other rounds and physically equivalent to all others. That is, there exists a single density matrix and measurement operators that are valid in every round. The statistics of all the rounds are thus independent and identically distributed (i.i.d.), so that we may consider $p(a, b|x, y)$ only.

Following the majority of self-testing works, we will often choose to work with a purification $|\psi\rangle_{ABP}$ of $\rho_{AB}$, where the purification space $\mathcal{H}_P$ is external to both Alice’s and Bob’s laboratories. This will be quite convenient mathematically, since working with state vectors over density matrices will significantly shorten the length of some equations. We stress that we do not assume that the state shared between Alice and Bob is pure; indeed it is given by $\text{tr}_P[|\psi\rangle\langle\psi|_{ABP}] = \rho_{AB}$. Furthermore, our definitions of self-testing will be such that the purification space is untouched by the self-testing protocol. Hence, working with $|\psi\rangle_{ABP}$ and tracing out the purification space is equivalent to working with $\rho_{AB}$. For this reason, the definitions in the following section can be equivalently phrased using $\rho_{AB}$ by simply discarding the purification space. Using such a purification, the probabilities (5) can be written

$$p(a, b|x, y) = \langle \psi | M_{a|x} \otimes N_{b|y} \otimes 1_P | \psi\rangle_{ABP}. \quad (6)$$
where $\mathbb{I}_P$ is the identity operator on the purification space.

Now, let’s imagine we have in mind a particular pure state $|\psi\rangle_{AB}$ and projective measurements $\{M_{a|x}\}, \{N_{b|y}\}$ that we would like to self-test. We call this state and measurements the reference state and reference measurements. The state $\rho_{AB}$ and measurements $\{M_{a|x}\}, \{N_{b|y}\}$ that correspond to the actual experiment are called the physical state and physical measurements. Similarly, the realisation $\{|\psi\rangle_{AB}, \{M_{a|x}^\prime\}, \{N_{b|y}^\prime\}\}$ is called the reference experiment and $\{\rho_{AB}, \{M_{a|x}\}, \{N_{b|y}\}\}$ the physical experiment.

Note that it will not be possible to infer exactly the reference state and measurements from the correlations alone. This is for the following two reasons:

i. Due to the unitary invariance of the trace, one can reproduce the statistics of any state $|\psi\rangle$ and measurements $\{M_{a|x}^\prime\}, \{N_{b|y}^\prime\}$ by instead using the rotated state $U \otimes V |\psi\rangle$ and measurements $\{UM_{a|x}^\prime U^\dagger\}, \{VN_{b|y}^\prime V^\dagger\}$, where $U$ and $V$ are unitary transformations. Hence, one can never conclude that the state is $|\psi\rangle$ since it may in fact be $U \otimes V |\psi\rangle$.

ii. One cannot rule out additional degrees of freedom on which the measurement operators act trivially. That is, a state $|\psi\rangle \otimes |\xi\rangle$ and measurements $\{M_{a|x}^\prime \otimes 1\}, \{N_{b|y}^\prime \otimes 1\}$ (where the identity operators act on the local spaces of $|\xi\rangle$) gives the same correlations as $|\psi\rangle$ and $\{M_{a|x}^\prime\}, \{N_{b|y}^\prime\}$.

To be able to define what it means to infer a particular state in the device-independent scenario, we thus need to define an equivalence between states that takes into account the above unknowns (that is, local unitary transformations and additional unused degrees of freedom). To do this, we make use of the concept of a local isometry.

An isometry $\Phi : \mathcal{H}_A \rightarrow \mathcal{H}_A$ is a linear transformation on quantum states that preserves the inner product, and can thus be seen as a unitary operator that can increase the dimension of the space. A general isometry on a state $|\psi\rangle_A$, can be achieved by embedding $|\psi\rangle_A$ in a larger Hilbert space $\mathcal{H}_{A2}$ and performing a unitary transformation; i.e. $\Phi[|\psi\rangle_A] = U |\psi\rangle_A$. A local isometry

$\Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \rightarrow \mathcal{H}_{A^\prime} \otimes \mathcal{H}_{B^\prime}$

is an isometry that can be realised with local quantum operations; i.e. it is a tensor product of isometries acting locally. One way to implement a local isometry—often used in self-testing works—is to embed the initial state using local ancilla states $\{00\}_{A^\prime B^\prime}$, and then perform a local unitary transformation (see figure 3):

$$\Phi_A \otimes \Phi_B[|\psi\rangle_{AB}] = U_A^\prime \otimes V_B^\prime [\langle 00 |_{A' B'} |\psi\rangle_{AB}] .$$

In the following sections we will use $\Phi[|\psi\rangle]$ to denote the action of an isometry on the pure state $|\psi\rangle$ and use $\Phi[\rho]$ to denote the corresponding transformation on a density matrix. As we will see, the notion of a local isometry will be central to our definitions of self-testing.

### 3.3 Self-testing of states

We are now ready to define what it means to self-test a quantum state.

**Definition 1.** *(self-testing of pure states)*

The correlations $p(a, b|x, y)$ self-test the state $|\psi\rangle_{A'B'}$ if for any state $\rho_{AB}$ compatible with $p(a, b|x, y)$ (for some choice of local measurements) and for any purification $|\psi\rangle_{ABP}$ of $\rho_{AB}$ there exists a local isometry

$$\Phi_A \otimes \Phi_B : \mathcal{H}_A \otimes \mathcal{H}_B \rightarrow \mathcal{H}_{A' A'} \otimes \mathcal{H}_{B' B'}$$

such that

$$\Phi_A \otimes \Phi_B \otimes \mathbb{I}_P[|\psi\rangle_{ABP}] = |\psi\rangle_{A'B'} \otimes |\xi\rangle_{ABP} \quad (8)$$

for some state $|\xi\rangle_{ABP}$.

---

\(^{1}\)Self-testing of mixed states will not be possible; see section 3.5.
The above can be understood via the idea of extraction. If we trace out the purification space in (8) then we have

$$\Phi_A \otimes \Phi_B [\rho_{AB}] = |\psi'\rangle\langle\psi'|_{A'B'} \otimes \sigma_{\overline{AB}},$$  \hspace{1cm} (9)

where $\sigma_{\overline{AB}} = \text{tr}_P |\xi\rangle\langle\xi|$. Thus, if one self-tests the state $|\psi'\rangle_{A'B'}$, then there necessarily exists a local channel (given by the isometry) that allows one to extract $|\psi'\rangle_{A'B'}$ from $\rho_{AB}$ into the ancilla space. The state $|\xi\rangle_{\overline{ABP}}$ in (8), which contains everything else from $|\psi\rangle_{\overline{AB}}$ after the reference state has been extracted, is called a junk state. We note that it is not necessary that one actually perform the isometry in the laboratory; all that is needed is a mathematical proof that such a procedure exists in principle. In section 4 we will see how this is possible with an explicit example.

Definition 1 is the definition that is most commonly used in self-testing works, although the identity channel on the purification space is often left implicit. Note that it is important that the isometry does not act on the purification space since, for example, the purification of a mixed separable state would result in an entangled state, which would give the devices access to entanglement for free. Finally, the above definitions can be straightforwardly generalised to define self-testing of multipartite states; one simply uses an isometry that is local with respect to the subsystems of the multipartite state to be self-tested.

3.4 Self-testing of measurements

The correlations that are used to self-test the reference state often allow one to identify the measurements that were performed. Since our physical measurements are projective, this definition will apply to self-testing of projective measurements only. The idea is to prove that under the action of the isometry, the physical measurements map to the reference measurements acting on the reference state. More specifically,

Definition 2. (self-testing of states and measurements)

The correlations $p(a, b|x, y)$ self-test the state and measurements $|\psi\rangle_{A'B'}$, $\{M_{a|x}\}$, $\{N_{b|y}\}$ if for any state and measurements $\rho_{AB}$, $\{M_{a|x}\}$, $\{N_{b|y}\}$ compatible with $p(a, b|x, y)$ and for any purification $|\psi\rangle_{\overline{ABP}}$ of $\rho_{AB}$ there exists a local isometry $\Phi_A \otimes \Phi_B$ such that

$$\Phi_A \otimes \Phi_B [\rho_{AB}] = \frac{1}{\text{det}[\Phi_A (\rho_{AB})]} \text{tr}_P [M_{a|x} \otimes N_{b|y} \otimes \rho_{AB}],$$

for all $a, x, b, y$, for some state $|\xi\rangle_{\overline{ABP}}$.

If we wish to make a statement solely involving the measurements, we will only be able to say something about the part of the measurement that acts on the support of the physical state. Some recent works (see e.g. [Kan17]) adopt this approach and use a definition of measurement self-testing in the following spirit (for example, for Alice’s measurements).

Definition 3. (self-testing of measurements)

The correlations $p(a, b|x, y)$ self-test the measurements $\{M_{a|x}\}$ for Alice if for any measurements $\{M_{a|x}\}$ and state $\rho_{AB}$ compatible with $p(a, b|x, y)$ there exists a unitary operator $U$ such that

$$U_A [M_{a|x}] U_B^\dagger = M_{a|x} \otimes \mathbb{1}$$

for all $a, x$, where $\tilde{M}_{a|x} = \Pi M_{a|x} \Pi$ and $\Pi$ is the projector onto the support of $\rho_A = \text{tr}_B \rho_{AB}$.

3.5 Self-testing via a Bell inequality and the geometry of the set of quantum correlations

It is known that only those correlations that are extremal points of the quantum set of correlations and are achievable with finite dimensional quantum systems can be used to self-test both a state and measurements [GKW+18]. Such points can often be witnessed by the maximum violation of some Bell inequality over the set of quantum correlations. As a result, one often does not need the full set of probabilities $p(a, b|x, y)$ in order to prove self-testing statements; the maximum quantum violation of a Bell inequality may already imply the existence of the desired isometry. One can thus consider self-testing relative to a Bell inequality by replacing the observation of the correlations by the value of a Bell inequality $\mathcal{I}[p(a, b|x, y)]$ in the previous definitions. Many of the well known Bell inequalities, such as CHSH and CGLMP have been used to this effect (see section 5 for such results).

In this light one might ask if the maximal violation of every nontrivial Bell inequality, i.e. one
which can be violated in quantum theory, is also a self-test of some entangled state. Or even more generally, do all extremal points of the set of quantum correlations self-test some state? These questions are examined in [GKW+18] where it was shown that the relation between self-testing, maximisers of non-trivial Bell inequalities and the boundary of the quantum set is not as simple as one might hope for.

We also note here that bipartite mixed state correlations can always be reproduced by a pure state of the same dimension [SVW16]. This implies that self-testing of bipartite mixed states following the same spirit as definition 1 above is impossible. Since the isometry preserves the purity of the input, applying the isometry to the pure state that gives the same correlations cannot result in the desired mixed state in tensor product with a junk state.

3.6 Robust self-testing

One encounters two problems when trying to prove self-testing statements as defined above from experimental data: (i) the experiment will inevitably contain some level of noise that will dampen the correlations, and (ii) the precise values of $p(a, b|x, y)$ will be uncertain due to a finite sample size. In practice, this will mean that proving perfect self-testing of states and measurements is impossible. Point (ii) can be addressed using tools of statistical inference [LRZ+18]. Point (i) can be tackled by proving approximate self-testing statements, and is known as robust self-testing. In short, the aim of robust self-testing is to prove that if the correlations are sufficiently close to the ideal correlations, then the state and measurements must be close (in some well-defined sense) to the desired ones.

We will focus on two notions of ‘closeness’ that are frequently used in the literature. Our first definition is as follows. Imagine we have identified an isometry that allows us to prove a self-testing statement as in definition 1. If the correlations are close to the ideal then one would expect that the two vectors appearing on either side of (8) be approximately equal up to some vector norm. This leads to the following.

**Definition 4. (robust self-testing of states, vector norm)**

The correlations $p(a, b|x, y)$ self-test the state $|\psi\rangle_{A'B'}$ with distance $\delta$ in the vector norm $||\cdot||$ if for any state $\rho_{AB}$ compatible with $p(a, b|x, y)$ and for any purification $|\psi\rangle_{ABP}$ of $\rho_{AB}$ there exists a local isometry $\Phi = \Phi_A \otimes \Phi_B$ such that

$$||\Phi \otimes 1_P |\psi\rangle_{ABP} - |\psi\rangle_{A'B'} \otimes |\xi\rangle_{ABP}|| \leq \delta$$

for some state $|\xi\rangle_{ABP}$.

This definition was used as the first definition of robust self-testing.

Our second definition follows the intuition that if in the ideal case one can extract the reference state $|\psi\rangle$, then in the noisy case one should be able to extract something close to $|\psi\rangle$. Here, it is usually easiest to adopt the fidelity, defined as $F(|\psi\rangle, \rho) = \langle \psi | \rho | \psi \rangle$ as the notion of closeness. First, define $\rho^\text{EXT}_{A'B'}$ as the extracted state of the ancillas after the application of the isometry, that is,

$$\rho^\text{EXT}_{A'B'} = \text{tr}_{AB} \Phi[\rho_{AB}].$$

We then have the following definition.

**Definition 5. (robust self-testing of states, fidelity)**

The correlations $p(a, b|x, y)$ self-test the state $|\psi\rangle_{A'B'}$ with fidelity $f$ if for any state $\rho_{AB}$ compatible with $p(a, b|x, y)$ there exists a local isometry $\Phi = \Phi_A \otimes \Phi_B$ such that

$$F(|\psi\rangle_{A'B'}, \rho^\text{EXT}_{A'B'}) \geq f.$$

Ideally, one would like replace the fidelity in the above definition by the trace distance $T(\rho, \sigma) = \frac{1}{2}||\rho - \sigma||_1 = \frac{1}{2}\text{tr} \sqrt{\sqrt{\rho} - \sigma (\rho - \sigma)}$, since the trace distance is both a metric (unlike the fidelity) and relates directly to the probability of distinguishing the two states. The fact that the fidelity is much more commonly used is because its linearity in the state $\rho$ makes bounds generally much easier to compute. One can nevertheless prove an upper bound to $T(\rho^\text{EXT}_{A'B'}, |\psi\rangle_{A'B'}|\langle \psi\rangle)$ from a bound on the fidelity using the relation $T \leq \sqrt{1 - F}$ [NC18]. We point the reader to [BLM+09] where a useful discussion about appropriate figures of merit for robust self-testing can be found.

Note that a local isometry can always prepare any pure product state of the ancillas for free by simply ignoring the physical state and applying the necessary unitaries on the ancilla space.
Hence, the best bound achievable via this strategy defines a trivial bound that can always be achieved. As an example, consider the task of self-testing the state $|\psi\rangle = \cos \theta |00\rangle + \sin \theta |11\rangle$ for some $\theta \in (0, \pi/4)$. This state has fidelity $\cos^2(\theta)$ and euclidean distance $\sqrt{2 - 2 \cos \theta}$ to the state $|00\rangle$. Thus a self-tested fidelity or distance is interesting only if it surpasses the corresponding bound. Taking definition 5, this trivial fidelity is equal to the square of the largest Schmidt coefficient of the state.

3.6.1 Extractability relative to a Bell inequality

As with ideal self-testing statements, it is most common to consider robust self-testing relative to a Bell inequality $\mathcal{I}$. For example, taking definition 5 as the figure of merit, one aims to find a function $f(\beta)$ that gives a lower bound on the fidelity as a function of the Bell inequality violation $\mathcal{I}(p(a, b, x, y)) = \beta$. This can be linked to the notion of extractability of the physical state with respect to the reference state for the Bell inequality $\mathcal{I}$. Note that any CPTP map can be realised by performing an isometry and discarding some degrees of freedom [Sti55]. Thus the map $\text{tr}_{AB}\Phi[\rho_{AB}]$ is equivalent to a local CPTP map $\Lambda_A \otimes \Lambda_B : \mathcal{H}_A \otimes \mathcal{H}_B \rightarrow \mathcal{H}_A \otimes \mathcal{H}_B$ applied to $\rho_{AB}$. Given a physical state $\rho$, the extractability $\Xi$ is the maximum fidelity of $\Lambda_A \otimes \Lambda_B[\rho]$ and $|\psi\rangle$ over all CPTP maps:

$$\Xi(\rho \rightarrow |\psi\rangle) = \max_{\Lambda_A, \Lambda_B} F(\Lambda_A \otimes \Lambda_B[\rho], |\psi\rangle). \quad (12)$$

To get optimal robust self-testing statements for a given inequality $\mathcal{I}$ one therefore needs to minimise the extractability over all states compatible with $\mathcal{I} = \beta$ for all values of $\beta$. This leads to the extractability-violation trade-off function

$$Q_{\psi, \mathcal{I}} = \inf_{\rho:\text{tr}[\mathcal{I}\rho]=\beta} \Xi(\rho \rightarrow |\psi\rangle). \quad (13)$$

Finding the extractability-violation trade-off function for a given Bell inequality is a difficult task since it involves a minimisation of the fidelity over all compatible states and a maximisation over all possible CPTP maps. Moreover, the optimal CPTP map (or equivalently, isometry) generally depends on the observed violation. More commonly, one fixes a single isometry for all violations and minimises the fidelity only, leading to a sub-optimal curve.

Finally, one can use similar ideas to the above to define the robust self-testing of measurements. We do not give any definition here, but point the reader to section 8.3 for work in this direction.

3.7 Generalisations and alternative definitions

3.7.1 The issue of complex conjugation

When self-testing quantum states in the bipartite scenario, it is sufficient to consider real reference states only, i.e. states such that $|\psi\rangle = |\psi\rangle^\ast$, where $\ast$ denotes complex conjugation with respect to a fixed basis. This follows since all pure states are local unitary equivalent to a real state via the Schmidt decomposition [Pre98]. A similar argument for measurements however is not possible. As a result, definition 2 suffers from a serious drawback; it can only be used to self-test sets of measurements that are invariant under the complex conjugation of all measurement operators. To see this note that since $p(ab|xy) = (p(ab|xy))^\ast$ then (assuming a real state $|\psi\rangle$)

$$p(ab|xy) = \text{tr}[|\psi\rangle\langle\psi'| M'_{a|x} \otimes N'_{b|y}] = \text{tr}[|\psi\rangle\langle\psi'| (M'_{a|x})^\ast \otimes (N'_{b|y})^\ast]. \quad (14)$$

Thus any correlations obtained using $\{ |\psi\rangle, M'_{a|x}, N'_{b|y} \}$ can also be obtained using $\{ |\psi\rangle, (M'_{a|x})^\ast, (N'_{b|y})^\ast \}$. These two realisations are generally not equivalent under local unitary operations. In this case, one cannot self-test the set $\{ |\psi\rangle, M'_{a|x}, N'_{b|y} \}$ using definition 2 since there is always another realisation $\{ |\psi\rangle, (M'_{a|x})^\ast, (N'_{b|y})^\ast \}$ that is not related to the first via a local isometry but results in the same correlations.

A straightforward solution to this problem first proposed in [MM11] is to generalise the definition of measurement self-testing so that one self-tests the measurements $\{ M'_{a|x}, N'_{b|y} \}$ if one can show that on the support of $|\psi\rangle$, the physical measurements act as some unknown convex combination of $\{ M'_{a|x}, N'_{b|y} \}$ and $\{ (M'_{a|x})^\ast, (N'_{b|y})^\ast \}$. This is in line with the general spirit of self-testing in which one aims to certify the measurements up to all the intrinsic limitations of the device-independent scenario. See appendix A.1 for a possible definition along these lines and section 8.2.1 for an example of such a self-test.

In principle, there may be more state and measurement transformations other than com-
plex conjugation that do not affect the observed probabilities. Determining this set is still an open problem. While in the case of qubit bipartite systems one can aim at self-testing states and measurements up to local isometries and complex conjugations, it is unclear if more transformations may be present when considering higher dimensional systems or multipartite scenarios.

3.7.2 Self-testing via simulation

Another recent approach presented in [Kan16] is to adopt the philosophy that self-testing a reference state or measurements should imply that the physical state or measurements be capable to simulate the reference state or measurements. For states, this translates to finding a local quantum channel $\Lambda_A \otimes \Lambda_B$ that maps the physical state to the reference state, thus allowing the simulation of any measurement on the reference state by first applying the channel followed by the desired measurement. Note that this definition is equivalent to definition 1 since via Stinespring’s dilution theorem [Cho75, Sti55] any local channel can be realised by first applying a local isometry then tracing out any irrelevant degrees of freedom.

For measurements, one considers unital channels, i.e. quantum channels that preserve the identity (and thus map sets of measurements to sets of measurements). The idea is then (say, for Alice) that if one can find a unitary channel such that $\Lambda[M_A|x] = M_A'|x|$, then one can simulate the reference measurement $M_A'|x|$ on any state by first applying the dual quantum channel $\Lambda^\dagger$ on the state followed by the physical measurement $M_A|x|$. Since

$$\text{tr} \left[ \Lambda^\dagger[\rho]M_A|x| \right] = \text{tr} \left[ \rho \Lambda[M_A|x] \right] = \text{tr} \left[ \rho M_A'|x| \right]$$

one recovers the same statistics as making the reference measurement on any state $\rho$. This approach was used in [RKB18] to self-test the Bell state measurement (see section 8.2.4).

3.7.3 Measurement self-testing based on commutation

Yet another approach to measurement self-testing focuses on certifying that the physical measurements satisfy some desired commutation relation on the support of the physical state. This can be advantageous, as commutation relations are often the only relevant features that one is interested in, and since they are invariant under isometry maps the approach can lead to simpler proofs of self-testing. Furthermore, in the case of perfect statistics, certifying a particular commutation relation may be enough to prove full self-testing statements of the form of definition 2. This approach has been used to prove measurement self-testing statements for anti-commuting qubit observables [Kan17] and sets of mutually unbiased bases in dimension 3 [KŠT+19]. It is very close in spirit to one of the earliest self-testing statements given in [PR92]. For further discussion on this technique, see section 8.2.2.

4 A first example

The maximally entangled state of two-qubits

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle)$$

is the quantum state which is most emblematic of the significance entanglement has in quantum theory and is used in a wide class of information processing protocols [BBC+93, BW92, Eke91]. In this section we show how to prove formal self-testing statements for this state and locally anti-commuting observables. Many of the techniques used to self-test more complex states and measurements can be understood as a generalisation of those presented here. We work in a simple scenario in which Alice and Bob each have two inputs ($x, y = 0, 1$) and two outputs ($a, b = \pm 1$). We chose the convention of having $\pm 1$ valued outcomes since it will be convenient to work with the observables

$$A_x = M_{+|x} - M_{-|x}; \quad B_y = N_{+|y} - N_{-|y},$$

where $M_{a|x}, N_{b|y}$ are the physical measurements operators in (5). Note that since the physical measurement operators are projective (see section 3.2), the operators $A_x, B_y$ are by construction Hermitian and unitary. We thus have

$$A_x^\dagger = A_x; \quad A_x^2 = 1; \quad B_y^\dagger = B_y; \quad B_y^2 = 1.$$
Box 3.1: Bell Nonlocality and the CHSH inequality

Bell nonlocality is a counter-intuitive property of quantum correlations discovered by John Bell in 1964 [Bel64]. The correlations $p(a, b|x, y)$ are called local if they can be reproduced by shared classical information. To formalise this, we represent the shared information by a classical random variable $\Lambda \sim \pi(\lambda)$. Averaging over this information, the possible correlations that Alice and Bob can achieve is given by (see (a) below for the corresponding classical causal network)

$$p(a, b|x, y) = \int_\Lambda \pi(\lambda)p_A(a|x, \lambda)p_B(b|y, \lambda)d\lambda. \quad (16)$$

Notice that local measurements on any separable state $\rho = \int_\Lambda d\lambda \pi(\lambda)\sigma^A_\lambda \otimes \sigma^B_\lambda$ lead to correlations of the above form, and so (16) is precisely those correlations that can be achieved using separable states. If we collect all of the probabilities into a single vector $\mathbf{p} = (p(00|00), p(01|00), \cdots)$ then the set of local correlations forms a convex polytope, the facets of which are called Bell inequalities (see (b), below). Remarkably, if Alice and Bob share an entangled quantum system, they may produce correlations that are nonlocal, i.e. which violate a Bell inequality and therefore cannot be written in form (16). An important Bell inequality, called the CHSH Bell inequality [CHSH69], already exists in the simplest scenario in which Alice and Bob have two inputs $(x, y = 0, 1)$ and two outcomes $(a, b = \pm 1)$. It is given by

$$\beta_{\text{CHSH}} = \langle A_0 B_0 \rangle + \langle A_1 B_0 \rangle + \langle A_0 B_1 \rangle - \langle A_1 B_1 \rangle \leq 2, \quad (17)$$

where $\langle A_x B_y \rangle = \sum_{a,b} a \cdot b p(ab|xy)$ denotes the correlator for the inputs $x, y$. Measurements on the maximally entangled state $|\phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ can violate this inequality up to $\beta_{\text{CHSH}} = 2\sqrt{2}$. For a comprehensive review on Bell nonlocality, see [BCP+14].

In the following we show how to self-test $|\phi^+\rangle$ from correlations achieving the maximal violation of the CHSH inequality. The central step in the proof will be to show that Alice and Bob’s local observables anticommute on the support of their shared state, i.e. $\{A_0, A_1\} |\psi\rangle = \{B_0, B_1\} |\psi\rangle = 0$. We present two methods to achieve this; (i) a geometrical argument for anticommutativity of Bob’s observables (section 4.1), and (ii) an algebraic argument (section 4.2). Once this is achieved, the anticommuting observables can be
used to build the required local isometry that is needed to prove a formal self-testing statement (section 4.3).

4.1 Geometrical proof of anticommutativity

In this section we give a simple geometric proof that the correlations maximally violating the CHSH inequality can be achieved only by measuring locally anticommuting observables. This will require knowing all correlations, not only the observation that $\beta_{\text{CHSH}} = 2\sqrt{2}$. The ideal correlations achieving this violation are

$$
\langle \psi | A_0 B_0 | \psi \rangle = \frac{1}{\sqrt{2}}, \quad \langle \psi | A_0 B_1 | \psi \rangle = \frac{1}{\sqrt{2}}, \\
\langle \psi | A_1 B_0 | \psi \rangle = \frac{1}{\sqrt{2}}, \quad \langle \psi | A_1 B_1 | \psi \rangle = -\frac{1}{\sqrt{2}}.
$$

(21)

Let us define vectors

$$
a_0 \equiv \frac{1}{\sqrt{2}} (A_0 + A_1) | \psi \rangle, \quad a_1 \equiv \frac{1}{\sqrt{2}} (A_0 - A_1) | \psi \rangle, \\
b_0 \equiv B_0 | \psi \rangle, \quad b_1 \equiv B_1 | \psi \rangle.
$$

Equations (21) imply the following inner product values:

$$
a_0 \cdot b_0^\dagger = 1, \quad a_1 \cdot b_1^\dagger = 1. \quad (22)
$$

The Cauchy-Bunyakovski-Schwarz inequality $a \cdot b^\dagger \leq |a||b|$ implies

$$
|a_i||b_i| \geq 1 \quad \text{for} \quad i = 0, 1,
$$

where $|a_i| = \sqrt{a_i \cdot a_i^\dagger}$. Since operators $A_i$ and $B_j$ are unitary, vectors $b_0$ and $b_1$ have unit norm, which implies

$$
|a_i| \geq 1, \quad \text{for} \quad i = 0, 1. \quad (23)
$$

The norms of the vectors $a_0$ and $a_1$ satisfy

$$
|a_0|^2 + |a_1|^2 = 2 \quad (24)
$$

by construction, which together with (23) implies $|a_0| = |a_1| = 1$. Since eqs. (22) represent the saturation of the Cauchy-Bunyakovski-Schwarz inequality, vectors $a_i$ and $b_i$ for $i = 0, 1$ must be parallel, i.e. $b_i = a_i$. This implies

$$
\{B_0, B_1\} | \psi \rangle = (B_0 B_1 + B_1 B_0) | \psi \rangle
$$

(25)

$$
= \frac{(A_0 - A_1) B_0 + (A_0 + A_1) B_1}{\sqrt{2}} | \psi \rangle
$$

$$
= \frac{(A_0 - A_1) (A_0 + A_1) + (A_0 + A_1) (A_0 - A_1)}{2} | \psi \rangle = 0,
$$

and thus $B_0$ and $B_1$ anti-commute on the support of $|\psi\rangle$. Note that since the correlations are symmetric, the same result holds for Alice’s observables.

4.2 Algebraic proof of anticommutativity

In principle, it is not easy to find correlations which self-test some state and measurements. Natural candidates, however, are correlations that maximally violate a particular Bell inequality. Moreover, the structure of the Bell inequality can be useful for proving self-testing statements, especially in cases when simple geometric considerations are not possible. Here we show how one can deduce an anticommutation relation for Bob’s observables from the observation $\beta_{\text{CHSH}} = 2\sqrt{2}$. As a starting point we take the SOS decomposition of the shifted CHSH Bell operator (see box 4.1 for a summary of SOS decompositions):

$$
2\sqrt{2} \mathbb{I} - B_{\text{CHSH}} = \frac{1}{\sqrt{2}} \left[ \left( \frac{A_0 + A_1}{\sqrt{2}} - B_0 \right)^2 + \left( \frac{A_0 - A_1}{\sqrt{2}} - B_1 \right)^2 \right].
$$

(30)

which follows from the properties (20). For any state $|\psi\rangle$ leading to $\beta_{\text{CHSH}} = 2\sqrt{2}$, i.e. $\langle \psi | B_{\text{CHSH}} | \psi \rangle = 2\sqrt{2}$ we thus have

$$
\frac{A_0 \pm A_1}{\sqrt{2}} | \psi \rangle = B_0/1 | \psi \rangle, \quad (31)
$$

as explained in equation (28) in box 4.1. With these relations we can prove that $B_0$ and $B_1$ anti-commute in the same way we did in (25).

4.3 Swap gate

We now prove a formal self-testing statement for the state $|\phi^+\rangle$ in the form of definition 1. This will require proving the existence of an isometry $\Phi$ mapping the physical state $|\psi\rangle$ to our reference state $|\psi'\rangle = |\phi^+\rangle$. In the majority of self-testing proofs the isometry is explicitly constructed and in most cases it takes the form of the partial Swap gate given in figure 4.

The main idea behind this particular isometry is as follows. In the case that the physical state is a two-qubit state and the operators are $Z_A = \sigma^A_z$, $X_A = \sigma^A_x$, $Z_B = \sigma^B_z$ and $X_B = \sigma^B_x$, the action of the circuit is to swap the physical state with the state $|00\rangle$ of the registers A and B. Of course, given the device-independent scenario we cannot
Box 4.1: SOS decompositions

To every Bell inequality $\mathcal{I} = \sum_{a,b,x,y} w_{ab}^x p(a, b|x, y)$ corresponds a Bell operator

$$\mathcal{B} = \sum_{a,b,x,y} w_{ab}^x M_{a|x} \otimes N_{b|y}$$

(26)

such that the violation is obtained as $\beta = \text{tr} [\mathcal{B} \rho]$. If the maximal violation achievable by using quantum resources (i.e. the quantum bound) is $\beta_Q$ the shifted Bell operator is defined as $\beta_Q \mathbb{1} - \mathcal{B}$. Every shifted Bell operator is by construction positive semidefinite since $\langle \psi | \mathcal{B} | \psi \rangle \leq \beta_Q$ for all $|\psi\rangle$. Imagine the shifted Bell operator admits a decomposition

$$\beta_Q \mathbb{1} - \mathcal{B} = \sum_{\lambda} P_{\lambda}^\dagger P_{\lambda},$$

(27)

where each $P_{\lambda}$ is a polynomial in the operators $M_{a|x}$ and $N_{b|y}$. The decomposition (27) is called a sum of squares (SOS) decomposition of the shifted Bell operator. If the polynomials are of degree at most $n$ in either $M_{a|x}$ or $N_{b|y}$ we say the SOS decomposition is of $n$-th degree.

SOS decompositions for Bell inequalities are typically hard to find. One can use numerical methods to find SOS decompositions of various degrees via the NPA hierarchy [NPA07, NPA08] (in particular, see [PNA10] for a link to SOS decompositions).

SOS decompositions allow one to extract potentially useful information about the physical state $|\psi\rangle$ and measurements used to achieve the maximal violation of the corresponding Bell inequality. From (27) we have

$$\langle \psi | \beta_Q \mathbb{1} - \mathcal{B} | \psi \rangle = 0 \implies \sum_{\lambda} \langle \psi | P_{\lambda}^\dagger P_{\lambda} | \psi \rangle = 0 \implies \sum_{\lambda} \| P_{\lambda} | \psi \rangle \|^2 = 0 \implies P_{\lambda} | \psi \rangle = 0 \quad \forall \lambda$$

(28)

Since $P_{\lambda}$ is a function of the operators used to obtain the maximal violation, the relations of the form $\{ P_{\lambda} | \psi \rangle = 0 \}$ often represent nontrivial statements about the strategy used to maximally violate the Bell inequality under consideration. Additionally, if a non-maximal violation $\beta_Q - \epsilon$ is observed the approximate relations analogous to (28) can be obtained:

$$\langle \psi | \beta_Q \mathbb{1} - \mathcal{B} | \psi \rangle = \epsilon \implies \| P_{\lambda} | \psi \rangle \| \leq \sqrt{\epsilon} \quad \forall \lambda$$

(29)

These relations are often significant for proving robust self-testing statements.
In order for $\Phi$ to be a valid isometry, $Z_A$, $X_A$, $Z_B$ and $X_B$ must be unitary. This is automatically the case for $Z_B$ and $X_B$ (see (20)), however it is not necessarily the case for $Z_A$ and $X_A$. To deal with this problem, we need to regularise these operators so that they are unitary. Formally, to regularise a Hermitian operator $Z$, one changes all zero eigenvalues of $Z$ to 1, resulting in a new Hermitian operator $Z^*$. The regularised operator is then obtained by normalising the eigenvalues of $Z^*$, i.e. $\tilde{Z} = \frac{1}{|Z^*|}Z^*$. Note that $\tilde{Z}$ is unitary by construction. One can often show that the regularised operators act in same way as the original operators on the physical state, i.e. $\tilde{Z} |\psi\rangle = Z |\psi\rangle$. This is indeed the case for our example (see appendix A.2). From hereon we therefore take $Z_A$ and $X_A$ to be the regularised unitary operators and continue to use the substitutions (32) without problem.

After a straightforward calculation one finds that the output of the isometry is

$$\Phi[|\psi\rangle] = \frac{1}{4} \begin{bmatrix} 00 \otimes (1 + Z_A)(1 + Z_B) |\psi\rangle \\
01 \otimes (1 + Z_A)X_B(1 - Z_B) |\psi\rangle \\
10 \otimes X_A(1 - Z_A)(1 + Z_B) |\psi\rangle \\
11 \otimes X_A(1 - Z_A)X_B(1 - Z_B) |\psi\rangle \end{bmatrix},$$

or in more compact form

$$\Phi[|\psi\rangle] = \sum_{i,j \in \{0,1\}} |ij\rangle_{A'B'} \otimes f_{ij} |\psi\rangle_{ABP},$$

where

$$f_{ij} = \frac{1}{4}X_A^i(1 + (-1)^i Z_A)X_B^j(1 + (-1)^j Z_B).$$

Using (35) expressions of the form $(1 \pm Z_A)(1 \mp Z_B) |\psi\rangle$ are automatically equal to zero, setting $f_{01} |\psi\rangle = f_{10} |\psi\rangle = 0$. The expression $f_{11} |\psi\rangle$ can be simplified in the following manner:

$$f_{11} |\psi\rangle = \frac{1}{4}X_A(1 - Z_A)X_B(1 - Z_B) |\psi\rangle = \frac{1}{4}(1 + Z_A)X_A(1 + Z_B)X_B |\psi\rangle = \frac{1}{4}(1 + Z_A)(1 + Z_B) |\psi\rangle = f_{00} |\psi\rangle$$

The second line is obtained by using anticommutativity relations (33) and (34), while (35) and the unitarity of $X_B$ was used to obtain the third line. Finally we see that the output of the Swap isometry is

$$\Phi[|\psi\rangle_{ABP}] = |\phi^+\rangle_{A'B'} \otimes |\xi\rangle_{ABP},$$

where $|\xi\rangle = \sqrt{2} f_{00} |\psi\rangle$. Note that $|\xi\rangle$ is necessarily normalised since the circuit of figure 4 is unitary. We have thus self-tested the state $|\phi^+\rangle$ in the sense of definition 1. Although we have worked with a purification of the physical state, the isometry does not act on the purification space, as needed from definition 1. This is because $\Phi$ is constructed from the measurement operators themselves, which by assumption act only on the local Hilbert spaces of Alice and Bob and therefore not on the purification space of $|\psi\rangle$.

4.3.1 Partial vs full Swap gates

The partial Swap gate was used in self-testing protocols for the first time in [MYS12] and in a large number of self-testing proofs since then. The full Swap gate differs from the partial one in that it contains another controlled gate before the first Hadamard is applied to the ancillary qubit. This controlled gate can be omitted if the ancilla
4.4 Self-testing of measurements

The measurements Alice and Bob use to maximally violate the CHSH inequality can also be self-tested via the Swap isometry. Here we explicitly show how to self-test Bob’s measurement observable \( B_0 \). For that purpose we check how the partial Swap gate transforms the state \( B_0 |\psi\rangle \), which can also be written as \( Z_B |\psi\rangle \):

\[
\Phi(B_0 |\psi\rangle) = \sum_{ij \in \{0,1\}} |ij\rangle_{A'B'} \hat{f}_{ij} B_0 |\psi\rangle
\]

where

\[
\hat{g}_{ij} = \frac{1}{2} X_A^i (1 + (-1)^i Z_A) \times X_B^j (1 + (-1)^j Z_B) Z_B
\]

\[
\hat{g}_{ij} = \frac{1}{2} X_A^i (1 + (-1)^i Z_A) \times X_B^j (-1)^i 1 + Z_B)
\]

\[
\hat{g}_{ij} = (-1)^i f_{ij}.
\]

This relation implies \( g_{01} = \hat{g}_{10} = 0 \) and \( g_{11} = -\hat{g}_{00} \). Thus the output of the Swap isometry will be

\[
\Phi(B_0 |\psi\rangle) = (1 \otimes \sigma_y |\phi^+\rangle_{A'B'} \otimes |\xi\rangle_{ABP},
\]

i.e. the measurement observable acts on the support of \( |\psi\rangle \) as \( \sigma_y \). A similar method can be used to self-test all other measurement observables used for the maximal CHSH violation. Note that (38) implies a self-test of the measurement operators. Since from (19) one has \( N_{b|0} = (1 + bB_0)/2 \) it follows by linearity of \( \Phi \) that

\[
\Phi(N_{b|0} |\psi\rangle) = \left(1 \otimes \frac{1 + b\sigma_z}{2} |\phi^+\rangle_{A'B'} \right) \otimes |\xi\rangle.
\]

Combining this and the previous section, we can prove a full state and measurement self-testing statement as follows. This concludes the introductory sections of the review.

5 Self-testing of bipartite states

5.1 Self-testing of two-qubit states

5.1.1 The maximally entangled pair of qubits

The fact that the maximal violation of the CHSH inequality can be obtained only by using the maximally entangled pair of qubits or a mixture of maximally entangled qubit states corresponding to different degrees of freedom was reported already in [SW87, PR92, BMR92, Tsi93]. An alternative method to self-test the maximally entangled pair of qubits is presented in [MY04], today mostly known as the Mayers-Yao self-test. While [SW87, PR92, BMR92, Tsi93] can be considered as the avant-garde self-testing papers, [MY04] stands out as the founding work which defined self-testing as a protocol ‘on its own’ and pointed out its importance. It is worth mentioning that Mayers and Yao made a similar statement already...
in [MY98], where they called the reference correlations ‘self-checking’. In the Mayers-Yao protocol, Alice and Bob both measure three observables, $\sigma_z$, $\sigma_x$ and $(\sigma_z + \sigma_x)/\sqrt{2}$. The proof is geometric in spirit and the isometry the authors use does the same job as the Swap gate, but the authors do not make the connection to the idea of applying a swap unitary. The self-test was made robust in [MMM06]. A simplified proof of the Mayers-Yao self-test, in which Alice makes the same measurements, while Bob measures only $\sigma_z$ and $\sigma_x$ appeared in the supplementary material of [McK14].

The concept of robustness and relevant figures of merit when self-testing the maximally entangled pair of qubits were introduced in [BLM+09], alongside with some explicitly calculated robustness bounds. The first completely device-independent robust self-test of the maximally entangled pair of qubits, both CHSH and Mayers-Yao based, appeared in [MYS12]. Further inequivalent proofs for self-testing the maximally entangled pair of qubits were reported in [MS13], where the authors gave a condition for a given binary XOR game to be a robust self-test, and in [SASA16], where the chained Bell inequalities were used to self-test the maximally entangled state and an arbitrary number of real measurements. An improvement of the robustness bounds were provided numerically in [YVB14] and [BNS15], and analytically in [Kan17], which is currently the best self-test of the maximally entangled pair of qubits in terms of robustness. An important contribution to the self-testing of the maximally entangled pair of qubits is [WWS16], which characterises all the correlations that self-test the state using two dichotomic measurements per party. The robustness of these self-tests was estimated in [LWH+19].

All the results presented so far used only real measurements. The self-testing of maximally entangled pairs of qubits using $\sigma_y$ observables was introduced in [MM11] based on the chained Mayers-Yao conditions, also in [APVW16] and [ABB17] based on the elegant Bell inequality [Gis09], and in [Kan17] based on the extended version of CHSH introduced in [Slo11] (more on the issue of self-testing complex measurements will follow in section 8.2.1).

5.1.2 Self-testing of partially entangled states

All pure entangled states of two qubits admit a Schmidt decomposition

$$|\psi_0\rangle = \cos(\theta) |00\rangle + \sin(\theta) |11\rangle \quad \theta \in (0, \pi/4],$$

(40)

Such states are known as partially entangled pairs of qubits, and they maximally violate the tilted CHSH inequalities [AMP12]:

$$\alpha \langle A_0 (A_0 B_0) + \langle A_1 B_1 \rangle \rangle - \langle A_1 B_1 \rangle \leq 2 + \alpha.$$

The maximal quantum violation $\sqrt{8 + 2\alpha^2}$ is achieved by the corresponding partially entangled state (40) for $\tan 2\theta = \sqrt{2\alpha^2 - 1}/2$. To achieve the maximal violation Alice measures $A_0 = \sigma_x$ and $A_1 = \sigma_z$ while Bob measures $B_0 = \cos \mu \sigma_z + \sin \mu \sigma_x$ and $B_1 = \cos \mu \sigma_z - \sin \mu \sigma_x$, with $\tan \mu = \sin 2\theta$. The proof that the maximal violation of the tilted CHSH inequality self-tests the corresponding partially entangled pair of qubits appeared in [YN13]. It relied on an SOS decomposition of the shifted Bell operator, but the proof appeared to have an error which made the self-testing proof invalid. The work [BP15] introduced a systematic way to find SOS decompositions for arbitrary shifted Bell operators. A whole family of SOS decompositions corresponding to the tilted CHSH Bell operator is introduced which was used to show that every tilted CHSH inequality self-tests the corresponding partially entangled pair of qubits. Improved robustness bounds for self-testing partially entangled pairs of qubits through violation of the tilted CHSH inequalities were presented in [CKS19]. Two different Bell inequalities, inequivalent to the tilted CHSH inequality and useful for self-testing the partially entangled pairs of qubits appeared in [BAŠ20] and [WBSS18].

The nonlocal character of partially entangled pairs of qubits can be assessed through the Hardy test [Har92, Har93]. In [RZS12] it is proven that Hardy test can be used as a robust self-test for the following states

$$|\psi_\varphi\rangle = \alpha (|01\rangle + |10\rangle) + e^{i\varphi} \sqrt{1 - 2\alpha^2} |11\rangle,$$

where $a = \sqrt{(3 - \sqrt{5})}/2$ and $\varphi$ is a free parameter.

A recent contribution [WKB+19] presents a self-test for any partially entangled pair of qubits and all three Pauli measurements (up to complex conjugation) on Alice’s side. Bob needs to
apply six measurements. The self-test is proven from the value of three Bell inequalities; two maximally violated tilted CHSH inequalities and one non-maximally violated CHSH inequality.

5.2 Self-testing of qudit states

The self-testing of bipartite entangled states of higher local dimension (qudits) is more complicated than the self-testing of qubit states. The good understanding of the qubit case has inspired the use of methods that we call ‘subspace methods’ in which different two-qubit subspaces of the state are self-tested until enough information is gained to self-test the full state. In subsection 5.2.1 we review this approach, before focusing on more genuinely d-dimensional methods in subsections 5.2.2 and 5.2.3. Some states of local dimension 2ⁿ can be seen as a tensor product of n qubit states. In such cases the so-called parallel self-testing is often used, described in section 5.3.

5.2.1 Subspace methods

Self-testing of maximally entangled states of any dimension is discussed for the first time in [YN13]. The isometry for self-testing introduced there is a high-dimensional generalisation of the Swap gate. The authors provided a set of correlations which self-test the maximally entangled state of two qudits

\[ |\psi^+\rangle_d = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle. \] (41)

One party performs three measurements and the other four. The idea is to self-test separately maximally entangled subnormalised substates \[ |\psi_{0,1}\rangle = |00\rangle + |11\rangle, |\psi_{2,3}\rangle = |22\rangle + |33\rangle, \ldots, |\psi_{d-2,d-1}\rangle = |d-2, d-2\rangle + |d-1, d-1\rangle. \] This can be done if all of the substates maximally violate the CHSH inequality (although in [YN13] the authors used a different correlation to test the substates). For this, both parties apply the measurements which are direct sums of the ideal CHSH measurements. This step is not enough, since the mixed state \[ \frac{1}{d} \sum_{i=0}^{d-2} |\psi_{i,i+1}\rangle \langle \psi_{i,i+1}| \] could also pass the test. Another necessary step is self-testing of the substates \[ |\psi_{d-1,0}\rangle, |\psi_{1,2}\rangle \] and so on. It is clear that the mixed state given above cannot provide correlations necessary for this step, where the two parties use again measurements which are the direct sum of the ideal CHSH measurements in a shifted basis, i.e. they now self-test the states \[ |\psi_{1,2}\rangle, |\psi_{3,4}\rangle, \ldots, |\psi_{d-1,0}\rangle. \] The direct sum of \[ \sigma_z \] measurements are the same in both bases, thus one party applies three measurements in total while the other applies four.

An arbitrary pure bipartite state admits the Schmidt decomposition

\[ |\psi\rangle = \sum_{i=0}^{d-1} \lambda_i |ii\rangle. \] (42)

The generalisation of the above explained method to self-testing states of the form (42) is given in [CGS17]. In the first step the sub-states \[ |\psi_{i,i+1}\rangle = \lambda_i |ii\rangle + \lambda_{i+1} |i+1, i+1\rangle \] for \( i = 0, \ldots, d-2 \) are self-tested via the maximal violation of the tilted CHSH inequalities. The second step self-tests the shifted states \[ |\psi_{i,i+1}\rangle = \lambda_i |ii\rangle + \lambda_{i+1} |i+1, i+1\rangle \] for \( i = 1, \ldots, d-1 \). This result completed the problem of self-testing all bipartite pure states. The Bell inequalities corresponding to this type of the self-test for maximally entangled states are described in [Col18].

5.2.2 Self-testing from qudit correlations

The method for self-testing all pure bipartite entangled states presented in the previous section relied on self-testing two-qubit sub-states. The measurements used in the self-test were also block-diagonal, where all blocks were either 2 × 2 or 1 × 1. It is surprisingly difficult to prove self-testing statements about high-dimensional states without resorting to such methods. In this section we outline a few protocols for self-testing qudit states that use genuinely qudit measurements.

The first such results were proven in [BNS+15], [YVB+14] and [SAT+17] where two-qutrit states were self-tested by using the numerical Swap method (for details see section 7.1.4). In [BNS+15] and [YVB+14] the maximal violation of the CGLMP inequality [CGL+02] was used to self-test the partially entangled state of two qutrits:

\[ |\psi\rangle = \frac{1}{\sqrt{2} + \gamma^2} (|00\rangle + \gamma |11\rangle + |22\rangle) \] (43)

where \( \gamma = (\sqrt{11} - \sqrt{3})/2 \). and in [SAT+17] the SATWAP Bell inequality is introduced and used to self-test the maximally entangled pair of qutrits.
An important contribution in this direction is the analytic self-test presented in [KŠT+19]. The maximally entangled pair of qutrits is self-tested through the maximal violation of a generalised CHSH inequality. These inequalities, introduced in [KŠT+19] can be seen as a special class of Buhrman-Massar inequalities [BM05], represent good candidates for self-testing maximally entangled states in any prime dimension \( d \). Alice and Bob, both have \( d \) inputs, and the measurements necessary for the maximal violation are mutually unbiased bases. For higher dimensions, the SOS-decomposition of the shifted Bell operator is provided, but the self-testing statement is still lacking. In fact, for \( d = 5 \) and \( d = 7 \), it is proven that the maximal violation can be achieved by using inequivalent quantum realisations, however all of them involve the maximally entangled state in dimensions 5 and 7, respectively.

Another contribution to self-testing maximally entangled states of qudits in the context of nonlocal games is given in [Man14]. There, the author considers a specific type of nonlocal games, the so-called pseudo-telepathy weak projection games. A nonlocal game is called pseudo-telepathy game if it can be won with probability equal to one by using quantum finite dimensional strategy, but cannot be won by using classical strategies [BBT05]. Weak projection games belong to a sub-class of pseudo-telepathy games and [Man14] shows that every such game can be used to self-test maximally entangled states in finite dimensions.

5.2.3 Group theoretic tools

Self-testing properties of non-local games were elaborately explored in [Slo11] and [CS17b]. The common method for both works is the ‘algebraisation’ of the winning strategies in nonlocal games. The idea of relating representations of a Clifford algebra to the optimal strategies to win the CHSH game was used already in [SW87] and [Tsi87]. In [Slo11], to each XOR game \( G \) is associated a \( C^* \) algebra \( A \), such that optimal strategies to win \( G \) correspond to representations of \( A \). Furthermore, there is a relation between near-optimal strategies and approximate representations. Using these techniques a self-testing statement for high-dimensional maximally entangled states via a generalisation of the CHSH game is implicitly given in [Slo11].

In [CS17b], the authors study self-testing properties of a class of pseudo-telepathy games, known as linear-constraint system games, of which the magic square and magic pentagram games are two popular examples [Per90, Mer90b]. In these games, the players are asked for assignments to a subset of variables in a system of linear equations, and they win the game if they return consistent and valid assignments. The authors extend the representation theoretic framework of [CM14], [CLS17] and [Slo20] and obtain a generic self-testing result for linear-constraint system games of a certain kind. They apply this result to obtain a self-testing protocol for a tensor product of \( n \) maximally entangled pairs of qubits. The self-testing condition is the perfect score in either the magic square game or the magic pentagram game. It is proven in [CM14] that perfect strategy for every linear-constraint system game which is also a pseudo-telepathy game must involve a maximally entangled state. On the other side in [CLS17] it is shown that a solution group can be associated to every linear-constraint system game. Moreover, the operators used in the winning strategy must satisfy certain algebraic relations determined by the solution group. In [CS17b] the authors use these results and by exploiting algebraic properties of the solution group corresponding to the magic square and magic pentagram games prove the self-testing statement for a tensor product of maximally entangled pairs of qubits. The self-test is also proven to be robust.

5.3 Self-testing \( n \) maximally entangled pairs of qubits

In this section we outline methods and results for self-testing \( n \) copies of the maximally entangled state of two qubits (which itself is a maximally entangled state of dimension \( 2^n \)). Here, there are two main approaches; sequential self-testing and parallel self-testing.

5.3.1 Sequential self-testing

The first result relating to the self-testing of \( n \) maximally entangled pairs of qubits (here also called EPR pairs) appeared in [RUV13]. In this scheme, in each round of the experiment the devices receive inputs \( (x_i, y_i) \) for \( i = 1, \ldots, n \), labelling the measurement bases for \( i \)-th maximally entangled pair. The inputs are given to the devices sequentially: first the inputs \( (x_1, y_1) \)
are given and the outputs \((a_1, b_1)\) are returned; then the inputs \((x_2, y_2)\) are given and the outputs \((a_2, b_2)\) are returned. This process is continued until the \(n\)-th pair of outcomes is collected and it is characterised by the following transcript

\[
\begin{align*}
a_1, b_1 & \text{ given } x_1, y_1, \\
a_2, b_2 & \text{ given } x_1, x_2, b_1, y_1, y_2, \\
& \cdots, \\
a_n, b_n & \text{ given } a_1, x_1, \cdots, a_{n-1}, x_{n-1}, x_n, b_1, y_1, \\
& \cdots, b_{n-1}, y_{n-1}, y_n.
\end{align*}
\]

There is no assumption that in each round the source emits the same state and the measurement strategies in the rounds may depend on the inputs and outputs in all previous rounds. In [RU13] the authors prove that if the parties win CHSH game in \(\omega^n\) rounds, where \(\omega^n\) is the optimal probability to win the CHSH game, there is isometry mapping the state the parties shared at the beginning of the procedure to the tensor product of \(n\) EPR pairs. The result is stated in its robust form: if the parties win CHSH game in \((1 - \epsilon)\omega^n\) rounds, then at the beginning of any randomly chosen block of \(m = n^{\Omega(1)}\) rounds the state of the parties can be mapped to a state \(f(\epsilon)\)-close to the tensor product of \(m\) EPR pairs. The drawback of the work is a very low robustness, i.e. as \(\epsilon\) increases the number of rounds necessary to extract a state \(f(\epsilon)\)-close to \(m\) EPR pairs grows very fast.

### 5.3.2 Parallel self-testing

A more popular approach for self-testing \(n\) EPR pairs is parallel self-testing (see figure 5). Here, the inputs are not given sequentially but all at the same time, i.e. the devices receive input vectors \(x = (x_1, \cdots, x_n)\), \(y = (y_1, \cdots, y_n)\) and return outputs vectors \(a = (a_1, \cdots, a_n)\), \(b = (b_1, \cdots, b_n)\). In principle this makes it more difficult to prove self-testing statements than in the sequential scenario, since one assumes less structure on how the outcomes are generated. To self-test a single maximally entangled pair we saw in section 4 that it is enough to identify a pair of anticommuting observables. In the case of \(n\) pairs, one has not only to find \(n\) pairs of anticommuting observables, but also to show that observables from different pairs mutually commute. An important feature of parallel self-tests is their robustness. A parallel self-test is robust if any strategy producing correlations that are \(\epsilon\)-close to the ideal ones must use a state which is \(f(\epsilon, n)\)-close to \(|\psi^\otimes n\rangle\), where \(f(\epsilon, n)\) is a monotonically increasing function in \(\epsilon\). How quickly the function \(f(\epsilon, n)\) increases with \(\epsilon\) and \(n\) determines how good robustness is.

The first parallel self-test was proven for 2 EPR pairs in [WBMS16]. The work gives two different self-tests: one based on the optimal success in a double use of the CHSH game and the other on the optimal success in the Magic Square game [Mer90b, Per90]. The result was subsequently generalised for arbitrary \(n\): via parallel repetition of the CHSH game in [Col17] and [McK17] and via parallel repetition of the Magic Square Game in [Col17] and [CN16]. The usefulness of the Magic Pentagram game [Mer90b] for self-testing a tensor product of three EPR pairs was proven in [KM18]. Self-testing of \(n\) EPR pairs via parallel repetition of the Mayers-Yao self-test is given in [McK16b]. We also note that the self-test [CS17b] discussed in the previous section belongs to this class of parallel self-tests.

In recent years, several works appeared which also managed to self-test \(n\) EPR pairs shared by two parties but not by parallel repetition of any single self-test of a single pair. The first of such results was presented in [OV16], based on the XOR games introduced in [Slo11]. Later, self-testing of \(n\) EPR pairs with measurements performed on few of them in each round was the subject of [CRSV18]. This self-test is nondestructive: not all entanglement is consumed in the self-test, but can be used for eventual later protocols.

A combination of self-testing based on nonlocal games with the quantum version of the linearity test from [BLR93], named Pauli braiding test [NV17] led to the first self-test of \(n\) EPR pairs in which robustness does not get worse if the num-

---

**Figure 5:** Scenario for parallel self-testing. Alice and Bob both receive a list of \(n\) inputs and provide \(n\) outputs, which correspond to measurements made on \(n\) independent copies of a physical state. The aim is to prove that the statistics self-test \(n\) independent copies of the reference state.
ber of EPR pairs tested increases. Another parallel self-test keeping this desirable property is presented in [NV18]. The test can be seen as a quantum version of the classical plane-vs-point test for multivariate low-degree polynomials [RS97].

Finally, while all self-tests presented in this section certify \( n \) EPR pairs and tensor products of the observables \( \sigma_x \) and \( \sigma_z \), it is possible to extend them to involve certification of \( \sigma_y \) also. This was first done in [CGJV17] and later in [BŠCA18b], up to the uncertainty of the complex conjugation of the full \( n \)-qubit measurement operators as explained in section 3.7.

5.3.3 Overlapping qubits

A standard parallel self-test of \( n \) EPR pairs proves the existence of \( n \) pairs of anticommuting observables, where any two observables belonging to different anticommuting pairs necessarily commute. As explained in [CRSV17] each anticommuting pair of observables defines a qubit. Hence the dimension of the underlying Hilbert space in this case must be at least \( 2^n \). The main contribution of [CRSV17] is the estimation of the dimension of the underlying Hilbert space if observables from different anticommuting pairs do not commute exactly, which might happen when the self-testing conditions are approximately satisfied. This leads to the concept of ‘overlapping qubits’, which, depending on the amount of the overlap can be ‘packed’ in the Hilbert space whose dimension grows polynomially with \( n \).

6 Self-testing of multipartite states

All bipartite pure states admit a Schmidt decomposition, which simplifies the characterisation of bipartite entanglement and the self-testing of bipartite pure states. Multiparticle states do not admit such a simple characterisation, although some generalisations of the Schmidt decomposition exist in the entanglement literature [AAC+00, Kra10]. While all bipartite pure entangled states can be self-tested, when it comes to self-testing of multipartite states, only some partial results exist. Furthermore, from a ‘loophole free’ perspective, multiparticle self-testing is considerably more demanding than bipartite self-testing, since it requires space-like separation between multiple measurement devices.

In this chapter we identify four main methods for self-testing multipartite entangled states: self-testing of graph states based on the structure of their stabilizer operators (section 6.1); tailoring Bell inequalities to self-test specific states (section 6.2); reductions to bipartite methods (section 6.3); parallel self-testing of multipartite states (section 6.4); and self-testing from marginal information only (section 6.5).

6.1 Self-testing of graph states from stabilizer operators

The first multipartite states to be self-tested were graph states [McK14]. Formally, given a graph \( G \) defined by a set of vertices \( V = \{1, \cdots, N\} \) and a set of edges \( E \) (pairs of connected vertices of \( V \)), the graph state corresponding to \( G \) is given by

\[
|G\rangle = \prod_{(i,j) \in E} CZ_{i,j} |+\rangle^\otimes N,
\]

where \( CZ_{i,j} \) is the controlled-\( \sigma_z \) two-qubit unitary \( CZ = \text{diag}(1,1,1,-1) \) acting on qubits \( i \) and \( j \). Equivalently, \( |G\rangle \) can be defined as the unique state that is stabilized by (i.e. is a +1 eigenstate of) a set of \( N \) local stabiliser operators \( \sigma_x^i \otimes_{j \in N(i)} \sigma_z^j \); where \( N(i) \) is the neighbourhood of vertex \( i \); the set of vertices connected to \( i \) on \( G \).

A self-testing protocol for any graph state corresponding to a connected graph is provided in [McK14]. Note that graph states corresponding to graphs that are not connected must be separable with respect to at least one bipartition. The reference measurements needed for the self-testing are given by the stabilizer operators themselves aided by a few measurements generalising those from Mayers-Yao self-test. More specifically, for an arbitrary graph state, one party has to measure three observables: \( \sigma_x \), \( \sigma_x \) and \( (\sigma_x + \sigma_z) / \sqrt{2} \), while all the other parties measure only \( \sigma_x \) and \( \sigma_z \). The self-test is robust to small imperfections and the isometry is the multipartite generalisation of the Swap gate.

The approach from [BAS+20] can be also placed in the following subsection, but since it is intrinsically related to stabilizers we discuss it in this group. Starting from any graph state, the authors introduce a method to construct a Bell inequality that is maximally violated by the corresponding state. Moreover, the derived Bell inequality can be used to self-test the state. Each
party measures an anti-commuting pair of observables from the real plane of the Bloch sphere. Beyond graph states, the method can be used to self-test the so-called partially entangled GHZ states $\cos \theta |0\rangle^\otimes n + \sin \theta |1\rangle^\otimes n$ for any $n \geq 2$.

### 6.2 Tailoring Bell inequalities

In [PVN14] the authors introduce a method to build permutationally invariant Bell inequalities with two measurement settings per party useful for self-testing multipartite states. The method is tailored for a specific state $|\psi'\rangle$ and the measurements leading to the maximal violation are chosen from the real plane of the Bloch sphere. A linear program can be used to find a Bell operator, whose eigenstate is $|\psi'\rangle$ and maximises the ratio of the quantum and classical bound. The derived Bell inequality is just a suitable candidate for self-testing, which further must be checked by utilising the numerical Swap method technique (see section 7.1.4). Since the self-testing proof relies on the Swap method, it becomes too costly when the number of parties becomes larger than four. Examples of the successful implementation of this method involve the tripartite $W$ state, the tripartite and four-partite GHZ state and the four-qubit linear cluster state.

Another method for developing Bell inequalities, tailored for self-testing of multipartite qubit states is described in [SBWS18]. As in [PVN14], all parties can perform two different measurements and the constructed Bell inequalities are suitable candidates for self-testing applications. The starting point for choosing a Bell operator is not the permutational invariance, but the structure of the stabilizers of the state. The method can be applied to multipartite states that are not graph states, in which case these stabilizer operators will not all be tensor products of Pauli operators. Of all the Bell operators mimicking the structure of the stabilizers, constructed from the arbitrary two real measurements per party, the optimal candidate is the one whose maximum eigenvalue is the local maximum with respect to the small perturbation of the local measurement directions. The robust self-test is then checked by using semidefinite programming to find the lower bound to the fidelity of the state providing the maximal violation and $|\psi'\rangle$ (see section 7.1.3). As example the authors apply the method to self-test a family of four qubit states $CU_\phi |\phi^+\rangle \otimes |\phi^+\rangle$, where $CU_\phi = |0\rangle\langle 0| \otimes \mathbf{1} + |1\rangle\langle 1| \otimes \exp[-i\phi \sigma_x]$. Since the self-testing is proven by employing numerical methods, it becomes infeasible when the number of parties increase.

### Table 1: Comparative properties of different self-tests of $n$ EPR pairs

<table>
<thead>
<tr>
<th>Reference</th>
<th>Robustness</th>
<th>Inputs size (in bits)</th>
<th>Outputs size (in bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[BSCA18b]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[CRS18]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(\log n)$</td>
<td>1</td>
</tr>
<tr>
<td>[Col17]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[CS17b]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[CGJ17]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n \log n)$</td>
<td>2</td>
</tr>
<tr>
<td>[CN16]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[McK16b]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[McK17]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>$n$</td>
</tr>
<tr>
<td>[NV17]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(n)$</td>
<td>2</td>
</tr>
<tr>
<td>[NV18]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(poly(\log n))$</td>
<td>$\text{poly}(\log \log n)$</td>
</tr>
<tr>
<td>[OV16]</td>
<td>poly$(n, \epsilon)$</td>
<td>$O(\log n)$</td>
<td>1</td>
</tr>
</tbody>
</table>

The most important aspect of a self-test of $n$ EPR pairs when it comes to practical usefulness is its robustness to noise (or rigidity). The other relevant property is its complexity, in terms of the size of the inputs. The size of the outputs is also a relevant factor, especially in possible applications for randomness expansion. For now, the self-testing protocol presented in [NV18] has the best properties in terms of the total number of inputs (polynomial) and robustness bounds (independent on $n$). The papers use different distance measures, but all the bounds given here are in terms of the Euclidean distance using definition 4. The work [CGJ17] self-tests $\sigma_y$ measurements on each EPR pairs, and the number of inputs increases in order to deal with the issue of complex conjugation (see 3.7.1.) If one omits self-testing of $\sigma_y$ from the protocol the number of inputs is $O(n)$.
6.3 Reductions to bipartite methods

Self-testing protocols for multipartite states can be constructed by reusing self-testing protocols for bipartite states. The idea is as follows: when $n-2$ parties perform appropriate projective measurements, they might collapse the state of the remaining two parties into some pure bipartite entangled state, which in principle can be self-tested by using methods from section 5. By repeating the process of projecting and self-testing for different pairs of parties one might expect to gather enough information to self-test the whole multipartite state. An important restriction is that the measurement used by any party in the projecting part must be some of the measurement the same party uses in the self-testing part of the protocol.

The idea was first used in [WCY+14] to self-test $W$-state $|W⟩ = (|001⟩ + |010⟩ + |100⟩)/√3$. Whenever one of the parties performs the measurement in the computational basis and obtains outcome +1 the state of the remaining two parties becomes maximally entangled $∼(|01⟩ + |10⟩)$. This state can be self-tested by maximally violating the CHSH inequality, for example. The authors of [WCY+14] show that by repeating the above process twice for different parties measuring in the computational basis, the whole state can be self-tested using the Swap isometry. They also show that a similar method, based on self-testing of partially entangled two-qubit states, can be used to self-test states of the form $|W_n⟩ = (|001⟩ + |010⟩ + |100⟩)/√2 + γ^2$. The method was generalised in [SCAA18] to prove self-testing of all permutationally invariant qubit Dicke states, all qubit graph states, and all multipartite states of any local dimension admitting the Schmidt decomposition $|ψ⟩ = \sum_{i=0}^{d-1} λ_i |i, i, …, i⟩$, representing the first self-test of a high-dimensional multipartite state. Self-testing of $W$-states for any number of parties was also proven in [Wu17], and self-testing of all Dicke states was proven in [Fad17].

The self-testing of graph states whose underlying graph is a triangular lattice is shown in [HH18]. The whole graph is shared by three parties and if one party measures its qubits in the $σ_z$ basis it prepares maximally entangled pairs of qubits for the remaining two parties, which are in [HH18] self-tested through the Mayers-Yao criterion.

6.4 Parallel self-testing of multipartite states

In section 5.3 we saw many ways to self-test $n$ EPR pairs by using parallel repetition of CHSH or Magic Square game. Up to date, the only parallel self-test of some multipartite state is shown in [BKM19]. The authors use diagrammatic proofs based on categorical quantum mechanics [CK17], to prove that parallel repetition of the GHZ game robustly self-tests $n$ copies of the GHZ state.

6.5 Self-testing using only marginal information

Almost all the protocols for self-testing multipartite states presented so far require measuring full-body correlators, that is, they depend on correlations between all parties. This quickly becomes a practical problem since measuring such correlations is typically experimentally very challenging. The possibility of self-testing by measuring only few-body correlators is the subject of [LCH+18]. The authors use the numerical Swap method (see section 7.1.4) to self-test the tripartite $W$-state, a class of $W$-like states $|001⟩ + |010⟩ + γ|100⟩)/√2 + γ^2$ and the states maximally violating Bell inequalities defined in [TSV+14] by using only two-body correlators. The four-partite $W$-state is also self-tested using three-body correlators.

7 Robust self-testing of states

It is impossible to meet exactly the conditions for ideal self-testing. On one hand, experimental noise and imperfections undermine hope to reproduce exactly the reference correlations (i.e. those obtained by performing reference measurements on the reference quantum state). On the other hand, even if all noise contributions are eliminated, one must work with a finite sample size and so the precise probabilities cannot be known, but only estimated up to some statistical confidence level. In order to make self-testing protocols practically meaningful, it is therefore crucial to make them robust to deviations from the ideal case. For possible definitions of robust self-testing see Defs. 4 and 5.

The first self-testing protocol to be made robust was the Mayers-Yao self-test of maximally
entangled pair of qubits [MMMO06]. Robust self-testing through the CHSH inequality was explored in [BLM+09] and simple robust self-testing protocols based on both the Mayers-Yao and the CHSH criterion were presented in [MYS12]. The techniques presented therein remained the main tool for making self-testing protocols robust in the majority of later contributions. As more is known about ideal self-testing the focus of the research is shifting towards finding better techniques for assessing robustness. Arguably, it remains a principal challenge in the field.

In this section we review the main contributions to robust self-testing and the techniques predominantly used in the literature. In section 7.1 we identify and explain five approaches:

- An approach based on the vector norm inequalities (Section 7.1.1);
- Methods relying on the use of Jordan’s lemma (Section 7.1.2);
- An approach based on the operator inequalities (Section 7.1.3);
- The numerical Swap method (Section 7.1.4);
- An algebraic method (Section 7.1.5).

Finally, in Section 7.2 we discuss recent progress on noise-tolerant self-testing of a tensor product of many EPR pairs.

### 7.1 Robust self-testing methods

#### 7.1.1 Norm inequalities method

The bulk of self-testing protocols start from the observed probabilities $p(a, b|x, y) = \langle \psi|M_{a|x} \otimes N_{b|y}|\psi\rangle$ or the maximal violation of some Bell inequality and deduce equations of the type (see e.g. (34) and (35) from section 4)

$$f(\{M_{a|x}\}, \{N_{b|y}\})|\psi\rangle = 0,$$

(45)

where $f$ is some polynomial function in the measurement operators. Such relations can be drawn from either geometrical arguments (like in [MY04]), algebraic identities (like in [McK14]) or SOS decompositions (like in [BP15]). The relations (45) are a necessary step in proving that the appropriate isometry (the Swap gate in most cases) maps the physical state to the reference one,

$$\Phi(|\psi\rangle) = |\psi\rangle \otimes |\xi\rangle.$$

(46)

Although the self-testing proof requires that the correlations be ideal, one could hope to follow the same proof in which the exact relations are exchanged with approximate ones, leading to a noise-dependent bound on the self-tested fidelity. More precisely, when the observed probabilities are within $\varepsilon$ distance from the ideal ones (or the violation of the Bell inequality is $\varepsilon$-far from the maximal value), analogously to (45), the aim is to find approximate relations:

$$\|f(\{M_{a|x}\}, \{N_{b|y}\})|\psi\rangle\| \leq g_f(\varepsilon),$$

(47)

where $\| \cdot \|$ is a vector vector norm (usually taken to be Euclidean) and $g_f(\varepsilon)$ are some increasing functions for which $g_f(0) = 0$. One can then often guarantee that the appropriate isometry $\Phi$ satisfies

$$\|\Phi(|\psi\rangle) - |\psi\rangle \otimes |\xi\rangle\| \leq g_\Phi(\varepsilon),$$

(48)

where $g_\phi$ is obtained by propagating the uncertainties (47) through the isometry circuit.
The asymptotic behaviour of $g_{\Phi}$ for different self-testing protocols based on this method is given in Table 2.

### Table 2: Comparative properties of different robust self-tests based on vector norm inequalities.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Asymptotic behaviour of $g_{\Phi}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[MYS12]</td>
<td>$O(\epsilon^{3})$</td>
</tr>
<tr>
<td>[McK14]</td>
<td>$O(\epsilon^{3})$</td>
</tr>
<tr>
<td>[BP15]</td>
<td>$O(\epsilon^{2})$</td>
</tr>
<tr>
<td>[ŠASA16]</td>
<td>$O(\epsilon^{2})$</td>
</tr>
<tr>
<td>[WCY+14]</td>
<td>$O(\epsilon^{2})$</td>
</tr>
</tbody>
</table>

7.1.2 Utilising Jordan’s lemma

One of the main difficulties in the device-independent description of quantum experiments is related to the inability to fix the dimension of the underlying Hilbert space, which prevents the parameterisation of the measurements and states used in the experiment. This difficulty stays the prime hurdle towards calculating robust self-testing bounds. A very useful theoretical asset enabling a solution in scenarios where each party has two dichotomic measurements is the Jordan lemma [PAB+09] (see lemma 2 therein). It allows to effectively reduce an arbitrary-dimensional experiment to the one in which the local subsystems are qubit systems.

For the purposes of robust self-testing the Jordan lemma was first time used in [BLM+09] to obtain robust self-testing of the maximally entangled pair of qubits through violation of the CHSH inequality. Later, it was used in [SBWS18] for the robust self-testing of an arbitrary multipartite states using the Bell inequalities introduced therein and described in section 6.2 of this review. For simplicity, here we give a short description of the method to the bipartite scenario, while keeping in mind that, as described in [SBWS18], it can straightforwardly be applied to the multipartite case.

The Jordan lemma states that given two Hermitian matrices of finite or countably infinite dimension and with eigenvalues ±1, there exists a unitary transformation that simultaneously block diagonalises them, where each block is of size at most $2 \times 2$. Consider a self-testing protocol in which Alice and Bob each have a pair of ±1 valued observables $A_x$, $x = 0, 1$ for Alice and $B_y$, $y = 0, 1$ for Bob. It follows there is a choice of local basis in which these observables take the block structure described above. One can further assume that each of the blocks is of size $2 \times 2$, since a one-dimensional block is equivalent to a two-dimensional block where the state has support only on one of these dimensions. One can then apply additional unitary rotations to each of the blocks so that they take real values only.

Given this structure, one can parametrise the observables as follows

$$A_x = \bigoplus_i A_i = \bigoplus_i \cos \alpha_i \sigma_x + (-1)^x \sin \alpha_i \sigma_z,$$

$$B_y = \bigoplus_j B_j = \bigoplus_j \cos \beta_j \sigma_x + (-1)^y \sin \beta_j \sigma_z$$

(49)

This parameterisation covers all possibilities: $\alpha_i = 0$ implies that the observables commute in that block, whereas $\alpha_i = \pi/4$ implies anticommutation in that block. Consequently, the Bell operator can be written as $B = \oplus_{ij} B(A_i, B_j)$. Following such parametrisation the Bell violation can be written as

$$\beta = \sum_{ij} p_{ij} \text{tr}[B(A_i, B_j)\rho_{ij}]$$

(50)

where $p_{ij}\rho_{ij}$ are projections of the physical state $\rho$ onto the blocks of Alice’s and Bob’s observables. Each block can then be treated separately to achieve an expression of the form

$$F(\Lambda_A \otimes \Lambda_B(\rho_{ij}), |\psi\rangle\langle\psi|) \geq f(\beta).$$

(51)

In [SBWS18] it is proven that if $f$ is a convex function of $\beta$ there exist maps $\Lambda_A$ and $\Lambda_B$ such
that the fidelity between $\Lambda_A \otimes \Lambda_B(\rho)$ and $|\psi\rangle \langle \psi'|$ given the violation $\beta$ is lower bounded by $f(\beta)$. In [BLM+09] a similar convexity argument is used to obtain the final bound.

The remaining challenge is to obtain relations of the form (51). In [BLM+09] the problem is solved analytically and the isometry used is just the one that rotates the blocks of the observables to obtain the form given in (49). The work [SBWS18] provides a general recipe: (51) can be solved by using a nonlinear optimisation with one variable per party.

7.1.3 Operator inequalities method

An analytic approach to robust self-testi ng, introduced in [Kan16] currently gives the best robustness bounds for the self-testing of two-qubit states. It is suited for self-testing protocols based on a Bell inequality violation. The method uses the notion of extraction (see section 3.6.1) and works by proving an operator inequality of the form

$$K \geq sB_I + \mu 1,$$

for all Bell operators $B_I$ for the Bell inequality $I$ in question, where $K = \Lambda_A^\dagger \otimes \Lambda_B^\dagger (|\psi\rangle \langle \psi'|)$ and $\Lambda^\dagger$ is the dual channel of $\Lambda$ with respect to the Hilbert-Schmidt inner product. This allows one to make linear robust self-testing statements, that is, to prove the existence of real parameters $s$ and $\mu$ such that the extractability-violation trade-off defined in (13) satisfies

$$Q_{\psi,B_I}(\beta) \geq s\beta + \mu.$$  

One thus has

$$F(\Lambda_A \otimes \Lambda_B(\rho), |\psi\rangle \langle \psi'|) \geq s\beta + \mu$$

for all states $\rho$ achieving violation greater that $\beta$.

In principle it is a difficult task to prove the operator inequality (52) for all Bell operators regardless of the dimension. In [Kan16] Jordan’s lemma is exploited to derive the current best robustness bounds for self-testing the maximally entangled state of two qubits. The method uses the CHSH inequality. The local channel $\Lambda_A \otimes \Lambda_B$ appearing in (54) is as follows. First, local unitary transformations are applied to Alice and Bob’s subsystems so that via the Jordan lemma, their local observables take a block diagonal form

as in (49). Then, for each block, one applies the $\alpha$-dependent dephasing channel

$$\Lambda_\alpha(\rho) = \frac{1 + g(\alpha)}{2} \rho + \frac{1 - g(\alpha)}{2} \Gamma(\alpha) \rho \Gamma(\alpha).$$

Here $g(\alpha) = (1 + \sqrt{2})(\sin \alpha + \cos \alpha - 1)$ and

$$\Gamma(\alpha) = \begin{cases} \sigma_x & \alpha \in [0, \pi/4] \\ \sigma_z & \alpha \in (\pi/4, \pi/2]. \end{cases}$$

Bob’s channel $\Lambda_B$ is defined analogously. This choice is shown to imply the lower bound (54) to the fidelity with $s = (2 + \sqrt{2})/8$ and $\mu = -(1 + 2\sqrt{2})/4$ (see figure 7 for a plot).

In [Kan16] inequality (52) is also proven for Mermin inequalities in order to self-test the tripartite GHZ state. Moreover, the fidelity lower bound for the Mermin inequality is proven to be optimal in the sense that for any violation there always exists a state achieving that violation with the self-tested fidelity to the reference state. The method has also been used for robust self-testing of partially entangled pairs of qubits [CKS19] and to assess the performance of different self-tests of a maximally entangled pair of qubits [LWH+19].

7.1.4 Numerical Swap method

The analytic techniques presented in the previous two subsections are only useful for either small amounts of noise (norm inequalities method), or (for now) solvable in simple cases, mostly when each party applies two binary measurements (operator inequalities method). For self-testing protocols which cannot be made robust with analytic methods, one can resort to a numerical method called the Swap method, introduced in [BNS+15, YVB+14]. While its applicability is still limited to simpler protocols due to computational resource requirements, it is responsible for the majority of practically relevant robust self-testing bounds.

The Swap method uses the Swap gate isometry (see section 4.3) and makes use of the fidelity of the extracted state as a figure of merit, as in definition 5. To get a lower bound on the fidelity, one needs to minimise the fidelity between the state of the output registers of the Swap gate and the reference state, given that the input state to the Swap gate provides the violation $\beta$. For example, for two-qubit states we have seen in section

Accepted in Quantum 2020-23-08, click title to verify
In its most general form, the NPA hierarchy [NPA07, NPA08, PNA10] is a method to tackle optimisation problems involving polynomials of non-commuting variables, and as a result is suited to certain optimisation problems in quantum theory. Define optimisation problems of the form

$$\min_{|\psi\rangle, M_{a|x}, N_{b|y}} \langle \psi | P (\{M_{a|x}, N_{b|y}\}) | \psi \rangle \quad \text{subject to} \quad \langle \psi | F_i (\{M_{a|x}, N_{b|y}\}) | \psi \rangle \geq 0 \ \forall i ,$$

where $P$ and $F_i$ are polynomials in the measurement operators $\{M_{a|x}, N_{b|y}\}_{a,x,b,y}$ and where the dimension of the state and measurement can be potentially infinite. The NPA hierarchy is a sequence of convex optimisation problems that provide increasingly better lower bounds to the optimal solution of the above by relaxing the problem to a minimisation over a larger set. Each of these relaxations can be solved via a corresponding semi-definite program [BV04]. Many problems in quantum information can be cast in the above form, particularly in the device-independent setting where the state and measurements are unknown.

The NPA hierarchy works as follows. Consider a generic state and measurement operators $\{|\psi\rangle, \{M_{a|x}\}, \{N_{b|y}\}\}$. Then, define sets $S_k$ (each corresponding to a level of the hierarchy) comprised of the identity operator and all (non-commuting) products of operators $M_{a|x}, N_{b|y}$ up to degree $k$; e.g. $S_1 = \{1\} \cup a_x \{M_{a|x}\} \cup b_y \{N_{b|y}\}$, $S_{k+1} = S_k \cup i,j \{S_i^{(i)} S_j^{(j)}\}$, where $S_i^{(i)}$ is the $i$th element of $S_k$. Define the moment matrix of order $k$, $\Gamma_k$, by $\Gamma_k^i = \langle \psi | S_i^{(i)} | \psi \rangle$. For any state and measurements $\{|\psi\rangle, \{M_{a|x}\}, \{N_{b|y}\}\}$, the matrix $\Gamma_k$ is Hermitian positive definite and satisfies some linear constraints given by the orthogonality conditions of the measurement operators. One can thus tackle optimisation problems of the form (56) by minimising the corresponding elements of the matrix $\Gamma_k$, under linear constraints on $\Gamma_k$ and $\Gamma_k \geq 0$. Such a problem is an instance of a semi-definite program which, via duality theorems, can be run to obtain certified lower bounds to the optimal solution.

4, equation (36) that this fidelity is given by

$$F(\rho_{\text{swap}}, |\psi\rangle) = \sum_{i,j,k,l \in \{0,1\}} c_{ijkl}^k \langle \psi | f_{ijkl} | \psi \rangle ,$$

where $\rho_{\text{swap}} = \text{tr}_{AB} \{\Phi | \rho \}$, and $c_{ijkl}^k = \langle k| |\psi\rangle \langle \psi| j \rangle$. Note that from the definition of $f_{ijkl}$, the above is equal to $\langle \psi | P (\{A_x^k, B_y^k\}) | \psi \rangle$, where $P$ is a polynomial in the measurement operators. As a result, lower bounds to the minimum fidelity (subject to a Bell inequality violation) can be found numerically via a corresponding semi-definite program defined by the NPA hierarchy; see box 7.1 for more details.

The first applications of the method are given in [YVB+14, BNS+15], which involve the following self-testing results: the self-test of the singlet state and Bob’s measurements from the CHSH inequality, the self-test of the singlet state from the Mayers-Yao criterion, the self-test of partially entangled pair of qubits from the tilted CHSH inequality, the self-test of a pure two qutrit state maximally violating CGLMP [CGL+02] inequality and the self-test of entangling measurements. Subsequently the method has been used to devise robust self-tests of the maximally entangled pair of qutrits [SAT+17], the whole family of pure entangled qutrit states [WP+18], the three qubit W-state [WCY+14, PVN14], three- and four-qubit GHZ states and the four-qubit linear cluster state [PVN14], a family of tripartite pure states, including the W-state from only marginal information [LCH+18], and a tensor product of two singlet states [WBMS16]. It has also been used to compare the performance of different types of self-tests of the singlet state, presented in [WWS16].

7.1.5 Algebraic method

In section 5.2.2 we discussed self-testing through the ‘algebratisation’ of the winning strategies in nonlocal games. Let us briefly recall that the crux of the method is associating an algebraic
The correspondence between the group representations and winning strategies then allows for the use of techniques from group theory to prove self-testing statements.

In [CS17b] this reasoning is taken one step further: a quantum strategy winning the generalised magic square game with high probability allows to extract an approximate representation of the solution group, or equivalently, a mapping between the group elements and unitary operators which is approximately a homomorphism. The closeness between the approximate and the exact representation is then used to make a robust self-testing statement.

The most important ingredient for constructing robust self-tests in this way is the stability theorem for approximate representations from [GH17]. It states that for any approximate $n$-dimensional representation $f$ of a finite group $G = \{g_i\}$ there exists an exact unitary $m$-dimensional representation $h$ such that the Hilbert-Schmidt distance between $f(g_i)$ and $h(g_i)$ is small and $m$ is close to $n$. The distance between these two representations is related to the score the physical strategy gained in the LCS game under consideration. The full robustness statement is obtained through the use of the van Kampen diagrams [Kam33].

7.2 Robust certification of large entanglement

In this section we discuss few contributions dealing with robustly certifying large amounts of entanglement without explicitly stating any self-testing result. There are two main reasons why such results merit attention in a review like this. The first is that for many purposes they could be used instead of robust self-testing protocols and moreover the self-testing statement is implicitly present for noiseless correlations. The second reason is the possible influence they could have on future approaches to robust self-testing.

A difference between robustness and noise tolerance when it comes to self-testing a tensor product of $n$ entangled pairs $|\psi^n\rangle$ is emphasised in works [AFB19] and [AFY18]. The known self-testing protocols are robust in the sense that any strategy producing correlations that are $\epsilon$-close to the ideal ones must use a state which is $f(\epsilon, n)$-close to $|\psi^n\rangle$, where $f(\epsilon, n)$ scales as $an^b c^c$ for appropriate constants $a, b$ and $c$. This is, however, not the same as noise-tolerance since noisy source producing the state $\rho^{\otimes n}$, where $\rho$ is $\epsilon$-close to $|\psi^n\rangle$ is not $f(\epsilon, n)$-close to $|\psi^n\rangle$. The fidelity of such state with $|\psi^n\rangle$ drops exponentially with $n$, so there is very little hope to make any non-trivial self-testing statement about such highly entangled state. Instead of self-testing $|\phi^n\rangle^{\otimes n}$ [AFY18] designs a one-shot test which is able to certify states whose entanglement of formation [BDSW96] is $\Omega(n)$. This certification method is noise-tolerant in the sense that the states $\rho^{\otimes n}$ are able to pass the test with high probability. A method to bound the one-shot distillable entanglement [BD10] of the states produced by an uncharacterised source is presented in [AFB19]. The protocol is operationally useful since not all entanglement is consumed for certification. Both these results are implicit self-tests since the maximal score in the introduced games implies that the state produced by a source must be $|\phi^n\rangle^{\otimes n}$.
8 Self-testing of measurements

In many cases the correlations which self-test a quantum state also self-test the applied measurements. As a result, many of the state self-testing results presented in the previous sections are accompanied by a corresponding statement for the measurements. In this section we give an overview of such results. In section 8.1 we review the known self-testing results for various sets of measurements, and in section 8.2 we discuss the different methods that have been used to achieve these results. We end the section with an overview of robustness techniques in measurement self-testing in section 8.3.

8.1 Measurement self-testing results

8.1.1 Qubit measurements

The simplest set of incompatible qubit measurements is given by a pair of Pauli observables $\sigma_x$ and $\sigma_z$. Self-testing of these measurements (together with their rotated versions $(\sigma_x \pm \sigma_z)/\sqrt{2}$ for the other party) can be achieved through the maximum violation of the CHSH Bell inequality (see Section 4) or related self-tests. Such self-testing statements can be found in [MYS12, Kan17, BNS⁺15, WCY⁺14]. Self-testing of the set of local observables $\{\sigma_x, \sigma_z, (\sigma_x \pm \sigma_z)/\sqrt{2}\}$ can be achieved through the so called ‘Mayers-Yao’ self test and its generalisations [MY04, McK14, MYS12]. A method to self-test large sets of qubit observables that are equally spaced on the equator of the Bloch sphere was given in [MYS12⁺]. A method to self-test large sets of measurements, and in section 8.2 we discuss the different methods that have been used to achieve these results. We end the section with an overview of robustness techniques in measurement self-testing in section 8.3.

8.1.2 Qudit measurements

Self-testing results for measurements of dimension larger than two are much less common. The only self-test of mutually unbiased bases in a prime dimension higher that 2 was given in [KST⁺19] for dimension $d = 3$. Self-testing of the Bell state measurement was first achieved analytically in [RKB18, BSS18] (see section 8.2.4 for an outline of the method). Self-tests of sets of measurements in high dimension can be achieved using the same techniques as in parallel self-testing of states (section 5.3). In this way, $n$-fold tensor products of the measurements $\{\sigma_x, \sigma_z\}$ and $\{\sigma_x, \sigma_y, \sigma_z\}$ in dimension $2^n$ have been achieved [BSČA18b, CGJV17, WBMS16, McK17, Col17, NV17, CN16, BKM19, KM18, CS17b, NV18].

8.1.3 Non-projective measurements

Although definition 2 of measurement self-testing assumes that the physical measurements are projective, one can nevertheless aim to prove that on the support of the reduced state of the self-tested state they act as some desired POVM. More specifically, suppose we have self-tested the reference state $|\psi\rangle$. Since the trace is invariant under isometry maps, the correlations can be written

$$p(a, b|x, y) = \text{tr} \left[ |\psi\rangle\langle\psi'| A'B' \otimes \sigma_{AB} M_{a|x} \otimes N_{b|y} \right] ,$$

where the local measurements are projective and may act on both the primed and bared spaces. Taking the trace over the barred spaces we have

$$p(a, b|x, y) = \text{tr} \left[ |\psi\rangle\langle\psi'| \tilde{M}_{a,b|x,y} \right] ,$$

(57)

where

$$\tilde{M}_{a,b|x,y} = \text{tr}_A \left[ I_{A'B'} \otimes \sigma_{AB} M_{a|x} \otimes N_{b|y} \right] .$$

(58)

To ‘self-test’ non-projective measurements, one aims to show that $M_{a,b|x,y} = M'_{a|x} \otimes N'_{b|y}$, where now the reference measurements can be non-projective. Essentially, one is self-testing a Stinespring dilation [Sti55] of the non-projective measurement.

In this manner, a self-test of the ‘tetrahedral’ qubit POVM first appeared in [APVW16], with rigorous proofs appearing later in [ABB⁺17] and [ABDC18], and an experimental demonstration presented in [SMN⁺20]. These results were proven using the method of ‘post-hoc’ self-testing, that we describe in 8.2.3. To self-test

$$\text{tr} \left[ |\psi\rangle\langle\psi'| A'B' \otimes \sigma_{AB} M_{a|x} \otimes N_{b|y} \right] .$$

(58)

where

$$\tilde{M}_{a,b|x,y} = \text{tr}_A \left[ I_{A'B'} \otimes \sigma_{AB} M_{a|x} \otimes N_{b|y} \right] .$$

(58)

To ‘self-test’ non-projective measurements, one aims to show that $M_{a,b|x,y} = M'_{a|x} \otimes N'_{b|y}$, where now the reference measurements can be non-projective. Essentially, one is self-testing a Stinespring dilation [Sti55] of the non-projective measurement.

In this manner, a self-test of the ‘tetrahedral’ qubit POVM first appeared in [APVW16], with rigorous proofs appearing later in [ABB⁺17] and [ABDC18], and an experimental demonstration presented in [SMN⁺20]. These results were proven using the method of ‘post-hoc’ self-testing, that we describe in 8.2.3. To self-test
measurements which are neither projective nor rank-one POVMs [WBSS18] use the approach developed by the same authors for the self-testing of quantum channels, described here in section 9.1.

8.2 Methods in measurement self-testing

In this section we outline some of the methods that have been used to prove measurements self-testing statements.

8.2.1 Phase kickback method for self-testing complex measurements

As mentioned in section 3.7, definition 2 is not suitable for self-testing complex-valued measurement operators. Take for example the problem of self-testing $|\phi^+\rangle$, the maximally entangled state of dimension 2, and $\{\sigma_z, \sigma_x, \sigma_y\}$, the three Pauli observables for say Alice. In section 4, we have seen how one can self-test the state $|\phi^+\rangle$ and $\{\sigma_x, \sigma_z\}$. Here, the issue of complex conjugation is not a problem since there exists a local basis in which the measurements and state are both real. However, there is no local basis in which the observables $\{\sigma_z, \sigma_x, \sigma_y\}$ are all real. Thus, we have two distinct possibilities for Alice’s measurements, $\{\sigma_z, \sigma_x, \sigma_y\}$ and $\{\sigma_z^*, \sigma_x^*, \sigma_y^*\}=\{\sigma_z, \sigma_x, -\sigma_y\}$, both of which are compatible with the observed correlations.

A natural question to ask is, given this uncertainty, what is the strongest possible self-testing statement that one could hope to prove? This question was first tackled by [MM11], see also [CGS17, BŠCA18b]. The basic idea is as follows. Consider a self-testing scenario in which Alice has (at least) three measurements given by the observables $A_0, A_1, A_2$. Take a known self-testing protocol for the state $|\phi^+\rangle$ and observables $\{\sigma_z, \sigma_x\}$ for Alice. Use this self-testing protocol three times for the pairs $\{A_0, A_1\}$, $\{A_0, A_2\}$, $\{A_1, A_2\}$, introducing new measurements for Bob and Alice if necessary. Since this proves that each pair $A_i, A_j$ anti-commute, one proves that the observables $\{A_0, A_1, A_2\}$ pairwise anti-commute and should essentially be $\{\sigma_z, \sigma_x, \sigma_y\}$ or $\{\sigma_z, \sigma_y, -\sigma_x\}$. More precisely, one introduces a pair or local ancillas $|00\rangle_{A' B'}$ for Alice and another pair $|00\rangle_{B' B''}$ for Bob and proves the existence of an isometry $\Phi$ such that

\[
\Phi[|\psi\rangle] = |\phi^+\rangle_{A' B'} \otimes |\xi\rangle
\]

\[
\Phi[A_0 |\psi\rangle] = (\sigma_z \otimes 1 |\phi^+\rangle_{A' B'} \otimes |\xi\rangle
\]

\[
\Phi[A_1 |\psi\rangle] = (\sigma_x \otimes 1 |\phi^+\rangle_{A' B'} \otimes |\xi\rangle
\]

\[
\Phi[A_2 |\psi\rangle] = (\sigma_y \otimes 1 |\phi^+\rangle_{A' B'} \otimes \sigma_z A' |\xi\rangle
\]

where the state $|\xi\rangle$ has the form

\[
|\xi\rangle = |\xi_0\rangle_{A B} \otimes |00\rangle_{A' B'} + |\xi_1\rangle_{A B} \otimes |11\rangle_{A' B''}.
\]

In (59) the additional $\sigma_z$ measurement on the junk state acts as an effective ‘controlled conjugation’ for the measurement of $\sigma_y$ on $|\phi^+\rangle$, where the probability to perform the conjugation is given by $\langle \xi_1 |\xi_1\rangle$, which remains unknown. Note that if we consider only the space $\mathcal{H}_A \otimes \mathcal{H}_B$ then the action of $A_2$ is to perform some unknown convex combination of $\sigma_y$ and $-\sigma_y$, as expected. Similar statements can be proven for Bob, where the register $B'$ acts as a control for a possible conjugation of his measurements. Note that given the form of (60), Alice and Bob will conjugate their measurement operators in a correlated fashion, as required from (14). The isometry, introduced in [MM11] and later used in [JMS20] to prove the above self-testing statement is an extension to the Swap isometry (see figure 4) introduced in section 4. The full isometry consists of the regular Swap isometry, followed by two extra ‘phase kickback’ controlled unitaries; see figure 8.

8.2.2 Self-testing measurements based on commutation

The majority of self-testing protocols first prove self-testing of the state, and then move to self-test the form of measurements on the support of the self-tested state. An alternative approach for binary observables was used already in [PR92] and revived in [Kan17]. In this approach the violation of a Bell inequality is directly related to the commutation properties of measurement observables without the need to prove a statement of the form of definition 2. In the two input, two output scenario the figure of merit (for Alice’s observables) is directly relatable to the maximal violation of the CHSH inequality and is given by

\[
t_{01} = \frac{1}{2} \text{tr}(||A_0 A_1|| \rho_A),
\]

where $\rho_A = \text{tr}_B \rho_{AB}$ is Alice’s reduced state of the physical state. The maximal CHSH violation implies $t_{01} = 1$, which further can be used to infer
anticommutativity of $A_0$ and $A_1$ on $\rho_A$. Beyond CHSH, it is also proved that the maximal violation of the Mermin-Ardehali-Belinskii-Klyshko inequalities [Mer90a, Ard92, BK93] implies the parties use anticommuting observables to achieve the maximal violation.

The method is also applied to self-test a set of three mutually anti-commuting observables. Note that there exists a basis in which any two anticommuting observables $A_0$ and $A_1$ can be written

$$A_0 = \sigma_z \otimes \mathbb{1}, \quad A_1 = \sigma_x \otimes \mathbb{1}. \quad (62)$$

If an inequality involving three observables $A_0, A_1, A_2$ can be used to certify $t_{01} = t_{02} = t_{12} = 1$ then besides relation (62) the following relation can be extracted

$$A_2 = \sigma_y \otimes A_Y, \quad (63)$$

where $A_Y$ is a Hermitian $\pm 1$-eigenvalue operator. Hence, when measuring $A_2$ the measurement result for $A_Y$ on the ‘junk’ Hilbert spaces takes care of the complex conjugation issue in the same way as (59).

Once the form (62) is extracted for measurement observables of all parties it is possible to also make statements about the underlying state. Furthermore, in [Kan17] this method is useful to certify observables which are not anticommuting but maximally violate weighted CHSH inequalities introduced in [LLP10]. Going beyond binary measurements, the extension of this method is applied also in [KŠT+19] to self-test mutually unbiased bases in the dimension $d = 3$.

8.2.3 Post-hoc self-testing of measurements

Once a state and sufficiently many measurements have been self-tested, further measurements can often be self-tested ‘for free’ through a method that we call post-hoc self-testing. As an example, the maximal violation of the CHSH inequality assures that up to an isometry the shared state is $|\phi^\perp\rangle$, Alice’s measurement observables are $\sigma_z$ and $\sigma_x$, and Bob’s measurement observables $(\sigma_x \pm \sigma_z)/\sqrt{2}$. The self-testing protocol can be extended to certify any other real qubit observable applied by either Alice or Bob. Assume Bob uses another measurement observable $B_2$ which we want to self-test as $\cos \theta \sigma_z + \sin \theta \sigma_x$ on his half of the maximally entangled state. If Bob performs this measurement, we will observe the correlations $\langle \psi | A_0 B_2 | \psi \rangle = \cos \theta$ and $\langle \psi | A_1 B_2 | \psi \rangle = \sin \theta$. Now, the maximal violation of the CHSH inequality implies that $A_0 |\psi\rangle$ and $A_1 |\psi\rangle$ are orthogonal. If we take these two states to be the first two states in an orthonormal basis of the full space, it must be that $B_2 |\psi\rangle = \cos \theta A_0 |\psi\rangle + \sin \theta A_1 |\psi\rangle$. Applying the isometry to this, we have

$$\Phi[B_2 |\psi\rangle] = \Phi[\cos \theta A_0 |\psi\rangle + \sin \theta A_1 |\psi\rangle] = (\cos \theta \sigma_z \otimes \mathbb{1} + \sin \theta \sigma_x \otimes \mathbb{1} |\phi^+\rangle) \otimes |\xi\rangle = (\mathbb{1} \otimes \cos \theta \sigma_z + \mathbb{1} \otimes \sin \theta \sigma_x |\phi^+\rangle) \otimes |\xi\rangle,$$

as required, where we have used the property $M \otimes \mathbb{1} |\phi^+\rangle = \mathbb{1} \otimes M^T |\phi^+\rangle$.

This technique can be understood from the perspective of measurement tomography. Given a set of linearly independent pure states that are tomographically complete on some subspace, one can infer the form of any measurement in this

![Diagram](image-url)
8.2.4 Self-testing of entangling measurements

An entangling measurement is one whose measurement operators are non-separable with respect to some bipartition of the Hilbert space. Two recent works [RKB18, BSS18] have presented analytic methods to robustly self-test the Bell state measurement (BSM), the entangling measurement whose eigenvectors are the four maximally entangled Bell states \( \{|\phi^+\rangle, |\phi^-\rangle, |\psi^+\rangle, |\psi^-\rangle\} \). Here the reference scenario is an entanglement swapping protocol: Bob possesses two particles, one maximally entangled with Alice’s particle and the other with Carmela’s. Bob performs the BSM on his two particles, and depending on his outcome projects the particles of Alice and Carmela onto one of the Bell states. These four Bell states can be self-tested by maximally violating the four different CHSH inequalities (mutually related by relabelling) conditioned on the outcome of Bob. The idea for self-testing is simple: if Alice and Carmela maximally violate all CHSH inequalities then it must be that Bob’s measurement is the BSM. Importantly, to maximally violate each of the four different Bell inequalities, Alice and Carmela use the same measurements.

In [RKB18, BSS18], the entangling swapping scenario is used to self-test the BSM using the notion of measurement self-testing via simulation described in section 3.7.2. Here, one necessarily needs to identify some well defined local Hilbert spaces for Bob in order to define entanglement. This is achieved by assuming that there are two independent sources between Alice and Bob, and between Bob and Carmela. Entanglement is then defined with respect to the Hilbert spaces of these two sources. A self-testing protocol for entangling measurements whose eigenvectors are partially entangled pairs of qubits, or GHZ states is also presented in [RKB18]. In [BSS18] the techniques for self-testing quantum channels is used to self-test the BSM. For more details on these techniques see section 9.1.

It is also worth mentioning here that in [RHC+11], a protocol to device-independently certify the existence of an entangling measurement is given, which was later shown to be robust in [BNS+15] via the use of the numerical Swap method (section 7.1.4). Both these works however only prove the existence of some entangling measurement but do not provide a self-testing statement for a particular measurement. In the context of witnessing irreducible dimension a way to certify entangling measurements is also presented in [CCBS17] alongside with a figure of merit quantifying how entangled the measurement operators are.

8.3 Robust measurement self-testing

In this section we give an overview of the approaches to robust self-testing of measurements. As we saw in section 3.6 there are a number of valid figures of merit one could consider when robustly self-testing a state. Given the increased complexity of measurements compared to states, the self-testing of measurements is even more diversified. In what follows we discuss a few approaches to quantify how close a physical measurement \( \{M_x\} \) is to some reference measurement \( \{M_x\} \).

The straightforward approach, in accordance with definition 4 uses the same methods as in the self-testing of states. In robust self-testing of states one defines a figure of merit that captures the closeness between the states \( \Phi[\rho_{AB}] \) and \( |\psi'\rangle\langle\psi'|_{AB} \otimes \sigma_{AB} \). For measurement self-testing, one can simply use the same figure of merit between the sub-normalised states \( \Phi \left[ M_{a|x} \otimes 1 \rho_{AB} M_{a|x}^\dagger \otimes 1 \right] \) and \( (M'_{a|x} \otimes 1 |\psi'\rangle\langle\psi'|_{AB} M_{a|x}^\dagger \otimes 1 ) \otimes \sigma_{AB} \). This will
mean that there will be a different value associated to each of the measurement operators; a single figure of merit can be obtained by, for example, taking the average or maximum of these values. Such an approach is used, for example, in [MYS12, McK14, BP15, ŠASA16].

The Swap method [YVB+14, BNS+15] can also be used to define a figure of merit for robust measurement self-testing. Taking the CHSH example, to estimate the closeness of Alice’s measurements to the reference measurements, the Swap gate is applied only on her system (i.e. only Alice’s local branch of the Swap gate is used) and the ancilla is initiated in one of the eigenstates \( |\varphi_{\pm}\rangle \) of her reference observable \( A'_i \). In the ideal case, if Alice measures the reference observable \( A'_i \) after the Swap gate is applied she will deterministically obtain the outcome \( \pm 1 \). The probability that this measurement gives the result \(+1\) is then used as a figure of merit to assess the closeness of the measurement, however it is not proven to give a distance measure. As with the estimation of the fidelity with the reference state, one can use the NPA hierarchy to lower bound this quantity with respect to the given CHSH violation.

A different figure of merit, analogous to the notion of state extractability (see section 3.6), was suggested in different self-testing contexts [TKV+18, RKB18, BSS18, TSV+20, WBSS18]. If \( D \) is a suitably chosen distance measure on the set of measurements, one can define the distance \( D \) between the physical measurement \( M_{a|z} \) and the reference one \( M'_{a|z} \) as

\[
D(M_{a|z}, M'_{a|z}) = \frac{1}{c_a} \max_{\Lambda} \sum_{a} D(\Lambda(M_{a|z}), M'_{a|z}),
\]

where \( c_a \) is a normalisation factor and the maximisation is taken over all completely positive and unital maps \( \Lambda : \mathcal{L}(H_{a}) \rightarrow \mathcal{L}(H_{a}') \). Depending on the type of reference measurements \( D \) can be chosen to be the overlap (as in [RKB18, TKV+18, TSV+20]), Uhlmann fidelity (as in [BSS18, WBSS18]) or any other distance measure. While the physical state \( \rho \) does not explicitly appear in (64), the map \( \Lambda \) has to depend in some way on it. In [WBSS18, BSS18] measurements are self-tested through their action on the maximally entangled pair of qudits and the state appears explicitly in the distance measure.

9 Extensions of self-testing to other scenarios

In this section we cover three extensions to the standard scenario of self-testing. In section 9.1 we cover works that self-test the action of a quantum gate in a device-independent manner. In section 9.2 we focus on the so-called semi-device-independent approaches. In section 9.2.3 we focus on self-testing via contextuality.

9.1 Self-testing of quantum gates and circuits

The paradigm of self-testing can be useful in scenarios going beyond the certification of states and measurements. Anticipating usefulness in the certification of devices for quantum computing, one may ask if it is possible to certify quantum gates, i.e. unitary transformations. First answers to this question came already in the early days of self-testing with two contributions devoted to the task of self-testing of quantum gates or quantum circuits, [vDMMS07] and [MMMO06]. Although [vDMMS07] has the phrase ‘self-testing’ in its title, it does not correspond to the fully device-independent scenario; the certification relies on several assumptions such as knowledge of the dimension of the system, which shifts it to the landscape of semi-device-independent scenarios.

The first protocol providing a recipe to self-test quantum gates acting on an arbitrary number of qubits is presented in [MMMO06]. Denote the physical implementation of the gates Alice and Bob use with \( G_A \) and \( G_B \). For the protocol to work, Alice and Bob must have access to the same gate, that is, \( G_A = G_B \). The core of the protocol is the Mayers-Yao self-test of the maximally entangled pair of qubits. To self-test a one-qubit unitary gate \( G'_A \) acting on her system, Alice has to share a maximally entangled pair of qubits. For self-test a one-qubit unitary gate \( G'_A \) acting on her system, Alice has to share a maximally entangled pair of qubits with Bob. As usual, the state shared between Alice and Bob is \( |\psi\rangle \) and they perform measurements \( \{M_{a|x}\} \) and \( \{N_{b|y}\} \). The aim is to show that there is a local isometry \( \Phi \) such that

\[
\Phi \otimes \mathbb{1}_P \left[ G_A M_{a|x} \otimes 1 |\psi\rangle_{ABP} \right] = G'_A M'_{a|x} \otimes 1 |\phi^+\rangle_{AB} \otimes |\xi\rangle_{ABP}
\]

where \( \{M'_{a|x}\} \) are the reference measurements for the Mayers-Yao self-test. The protocol consists of three parts:
- The Mayers-Yao self-test on the input state $|\psi\rangle$,

- The Mayers-Yao self-test on the output state $G_A \otimes G_B |\psi\rangle$,

- A check that $G_A \otimes I_B |\psi\rangle_{AB}$ reproduces the statistics of $G'_A \otimes I_{B'} |\phi^+\rangle_{A'B'}$ with respect to the Mayers-Yao measurements.

The first two steps serve for self-testing the underlying state and ensure that $G$ is a unitary gate (at this step potentially the identity gate). The third step can be seen as a tomography of $G$, since the measurements and state are already self-tested in the first two steps. Note that this means one can only self-test gates having real coefficients with respect to the self-tested measurements. The method is also extended to many-qubit gates. Each of Alice’s qubit on which a gate acts is maximally entangled with another qubit of Bob, and the tensor product structure of Alice’s and Bob’s Hilbert spaces is assumed. The procedure repeats as in the case of one-qubit qubits, with three steps involving the self-test of the input and output states which are now tensor products of many maximally entangled pairs of qubits and tomography of the corresponding multi-qubit gate. By using this method one can self-test the whole quantum circuit by self-testing each gate in sequence according to the recipe given above. This self-test is also proven to be robust.

Another self-test of quantum gates with simpler structure and significantly better robustness bounds is given in [SBWS18]. It is more general than [MMM06] since it provides a framework to lower bound the fidelity with an arbitrary quantum channel $\Gamma$. For Alice to self-test the channel $\Gamma'$ she again has to share a maximally entangled pair of qubits with Bob, but now Bob does not perform any channel to his system. Let the physical implementation of the channel be denoted as $\Gamma$. The aim is to find the fidelity between the reference channel $\Gamma'$ and the physical one $\Gamma$. The protocol consists of two steps:

- The self-test of the input state $|\psi\rangle$. The result provides a lower bound to the input fidelity $F_I$ between $\Lambda_A' \otimes \Lambda_B |\psi\rangle$ and $|\phi^+\rangle$, where $\Lambda_A'$ and $\Lambda_B$ are the CPTP maps as explained in 3.7.2.

- The appropriate self-test which finds a lower bound of the output fidelity $F_O$ between the state $\Lambda'_A \otimes \Lambda_B (\Gamma_A \otimes I_B |\psi\rangle_{AB})$ and the reference output state $\Gamma'_A \otimes I_{B'} |\phi^+\rangle_{A'B'}$.

The fidelity of the physical channel to the reference channel is proven to be lower bounded by $\cos(\arccos(F_I) + \arccos(F_O))$. The main challenge is to find the appropriate self-test necessary for the second step. The paper gives the solution for unitary channels (i.e. quantum gates), generalises the protocol for many-qubit channels and provides the explicit solution for arbitrary two-qubit controlled gate of the form $CU_\varphi = |0\rangle \langle 0| \otimes 1 + |1\rangle \langle 1| \otimes e^{-i\varphi}s$. Since such gates are necessary and sufficient for universal quantum computing (together with a set of single qubit gates) the toolkit represents an important contribution to the self-testing of all the building-blocks of a quantum computer.

Aside from this, the paper [SBWS18] is also valuable for two contributions independent of quantum channel self-testing. One is a method to self-test multipartite states, described in 6.2 and the other a technique useful for robust self-testing, described in 7.1.3. Furthermore, by defining the gates in terms of their Krauss representation, the techniques from [SBWS18] have been generalised in [WBSS18] to allow for self-testing of measurements other than rank-one POVMs.

### 9.2 Semi-device-independent scenarios

A number of works have investigated extensions of self-testing to so-called semi-device-independent (SDI) scenarios. In the device-independent scenario, all devices are treated as black boxes and one hence imposes minimal assumptions on the states and measurements. In the SDI scenario, some additional assumptions are added, without assuming a full characterisation of the entire set-up. As such, the SDI scenario can be seen as a weaker version of the DI scenario, intermediate between the scenarios of full device independence and full characterisation. Moving to the SDI scenario can be advantageous for at least three reasons. First, the additional assumptions can overcome some of the mathematical difficulties of the DI scenario and make statements easier to prove and results more tolerant to noise; second, for some scenarios it may actually be necessary to move to SDI scenario in order to make any non-trivial statements (see 9.2.2), and third, the additional assumptions...
may be very natural given a particular experimental set-up or level of trust in some devices. In this section we give an overview of three extensions of self-testing to the SDI scenario, namely one-sided device-independent self-testing, commonly known as the EPR steering scenario (section 9.2.1), self-testing in prepare-and-measure scenarios (section 9.2.2), and self-testing based on noncontextuality inequalities (section 9.2.3).

9.2.1 One sided device-independent self-testing (EPR steering)

The one-sided device-independent scenario (also commonly referred to as the EPR steering scenario), is equivalent to the standard self-testing scenario, with the additional assumption that there is one trusted party (here Bob) whose device is fully characterised, that is, his measurement operators are known. Thus, Bob is able to apply any quantum measurement and can in principle perform quantum state tomography of his half of the state. Alice, as in the self-testing scenario, receives classical input $x$ to her device and outputs classical output $a$. The subnormalised state of Bob conditioned on Alice’s input $x$ and output $a$ is given by

$$\sigma_{a|x} = \text{tr}_A \left[ M_{a|x} \otimes \mathbb{I} \rho_{AB} \right]$$

(65)

The set $\{\sigma_{a|x}\}_{a,x}$ is called an assemblage. It is said that the assemblage $\{\sigma_{a|x}\}_{a,x}$ admits a local hidden state model (LHS) if it admits a decomposition

$$\sigma_{a|x} = \int d\lambda q(\lambda) p_{a|x,\lambda} \rho_{\lambda}, \quad \forall a, x, \quad (66)$$

where $q(\lambda)$ is a normalised probability density and $\rho_{\lambda}$ is a normalised density operator acting on the local Hilbert space of Bob. If the assemblage $\{\sigma_{a|x}\}_{a,x}$ is incompatible with a LHS model one says that it demonstrates steering. The existence of a LHS model can be refuted by violation of steering inequalities, which take into account the correlations between Alice’s outputs and the outputs of known measurements performed by Bob. Another way to prove that the assemblage $\{\sigma_{a|x}\}_{a,x}$ demonstrates steering is by using simple SDP optimisations [WJD07],[CS17a],[UCNG19].

The decomposition (66) demonstrates that the shared state must be entangled. A natural question is in which cases can we go beyond witnessing entanglement and recover the shared state $\rho_{AB}$. This task was introduced in [SH16] and [GKW15] under the name one-sided device-independent (1SDI) self-testing. In [SH16] the authors are mostly interested in the robustness of 1SDI self-testing and how it compares to the robustness of standard self-testing, while the authors of [GKW15] are also interested in the application to delegated quantum computing protocols (see Section 10.4). Two types of 1SDI self-testing are introduced: correlation based, which draws conclusions only from the violations of steering inequalities, and assemblage-based, which works with the full assemblage. An interesting conclusion is that in the case of the simplest self-test of the maximally entangled pair of qubits, the asymptotic behaviour of the self-tested fidelity as a function of noise is the same in both the 1SDI and DI scenarios. How general this statement is remains as an open question. The 1SDI scenario is also very useful for self-testing a tensor product of many EPR pairs. In the standard self-testing of such states the main difficulty is establishing a tensor product structure, while in 1SDI scenario this comes for free due to the fact that Bob’s device is characterised. Numerical techniques, similar to the Swap method, for robust self-testing in the 1SDI scenario were also presented in work [SH16]. 1SDI self-testing of all pure two-qubit states is presented in [GBD+18].

Figure 9: Semi-device-independent scenarios. (a) The one-sided device-independent scenario. One of the parties (here Bob) is assumed to have a trusted measurement device (i.e. his measurement operators are known). (b) The prepare-and-measure scenario. Alice sends a quantum state to Bob, conditioned on her input.
9.2.2 Self-testing in the prepare-and-measure scenario

A recent series of works have adapted the self-testing scenario to the prepare-and-measure scenario. Here, Alice sends one of a number of states $|\psi_x\rangle$, labelled by $x$, to Bob, who measures $\{M_{b|y}\}$ conditioned on input $y$ and obtains outcome $b$. The statistics of the experiment are therefore given by

$$p(b|x,y) = \text{tr}[|\psi_x\rangle\langle\psi_x|M_{b|y}]$$

(67)

In analogy to the case of self-testing entangled states and measurements, one aims to infer from the statistics that the preparations and measurements $\{|\psi_x\rangle, M_{b|y}\}$ are equal to some reference set $\{|\psi_x\rangle', M_{b|y}'\}$ up to some unknown isometry.

In contrast to the Bell scenario, non-trivial statements can only be made if one places additional assumptions on the experiment. To see this, note that any statistics $p(b|x,y)$ can be reproduced by sending the label $x$ to Bob, i.e. with the preparations $|\psi_x\rangle = |x\rangle$, and Bob simply outputting $b$ with probability $p(b|x,y)$. Thus, self-testing any set of preparations that are not diagonal in the same basis is impossible without some additional assumptions. As is common in the prepare-and-measure scenario, a number of recent works [TKV+18, TSV+20, MP19, FK19] overcome this by assuming an upper bound on the Hilbert space dimension of the preparations and measurements. A similar assumption has also been studied in the Bell scenario [BLM+09, GBS16] where the state and measurements are assumed to be qubit systems, making self-testing statements significantly easier to prove.

A convenient figure of merit in prepare-and-measure scenario is the success in a game called a random access code (RAC) [ANTSV99, Nay99, THMB15]. In a $n^d \to 1$ RAC, Alice receives $n$ dits $x = (x_1, x_2, \cdots, x_n)$ and Bob receives $y = 1, 2, \cdots, n$. The aim is to maximise the average probability that Bob correctly guesses the input bit $x_y$, i.e. to maximise the expression

$$A_{n^d\to 1} = \frac{1}{2d^n} \sum_{x,y} p(b = x_y|x,y).$$

(68)

In [TKV+18], the authors study as a figure of merit the $2^2 \to 1$ RAC. It is proven that the value $A_{2^2\to 1}$ can be used to perform robust self-testing of the preparations $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ and measurements given by the qubit observables $(\sigma_x \pm \sigma_z)\sqrt{2}$ under the assumption of qubit preparations and measurements, using a technique based on that in [Kan16] (similar statements also appeared in [WLP13, BQB14] in the context of SDI quantum key distribution protocols and dimension witnesses). This is then generalised to a self-test of any pair of non-commuting qubit observables by adding an input-dependent bias to (68). The authors also use $A_{2^2\to 1}$ to self-test a non-trivial set of qutrit preparations and measurements, and implement an adaptation of the numerical Swap method (see Section 7.1.4) to deal with the prepare-and-measure scenario. In [FK19], the authors study the $2^d \to 1$ RAC. It is proven that this game provides a robust self-test of a pair of measurements that correspond to two mutually unbiased bases in dimension $d$. Further to this, the authors show how the score of the RAC can also be used to bound both the incompatibility robustness [HKR15] of the pair and the randomness of the measurement outputs. It is worth emphasising that self-testing claims in this scenario can be made for arbitrary dimension-bounded communication games, i.e. it is not restricted to RACs.

Several works have investigated self-testing of non-projective measurements in the dimension-bounded prepare-and-measure scenario. In [MP19] the authors start from the $3^2 \to 1$ RAC to develop a self-test of the extremal ‘tetrahedral’ qubit POVM. In [TSV+20], a general method is given to self-test any extremal qubit POVM, given a self-test of a set of preparations with opposite Bloch vectors to the POVM on the Bloch sphere. In a similar fashion in [TRR19] the authors prove a self-test of $d$-dimensional SIC POVM (whenever it exists). To prove the ideal self-testing all works use the same idea; if there is one outcome of a measurement that never occurs for a given preparation, it follows that the corresponding POVM element must be opposite to the preparation on the Bloch sphere. From this one can use a self-test of preparations to effectively tomograph the POVM measurement in a similar manner to that in Section 8.2.3. In [TSV+20] the protocol is made robust by introducing a convenient distance measure which is used in the experimental demonstration reported therein. All three works also discuss certification of the non-projective nature of a measurement, which is a weaker form of certification than ro-
bust self-testing since it does not discuss closeness to any particular POVM.

The recent work [MBP19] investigates the self-testing of non-projective measurements in a prepare-and-measure scenario involving a sequence of measurements, using figures of merit that are closely linked to the $2^2 \rightarrow 1$ RAC. To achieve the optimal success probability in the game, one party needs to perform a so-called Lüders instrument, which corresponds to a non-projective measurement. The authors then derive bounds on the maximal eigenvalue of the corresponding measurement operators given an observed score in the game, using a numerical approach based on an adaptation of the NPA hierarchy to the prepare-and-measure scenario.

### 9.2.3 Self-testing through noncontextuality inequalities

One of the important features of quantum theory is contextuality, first noticed via the Kochen-Specker theorem [KS67]. In general, in any deterministic hidden variable model reproducing quantum correlations, the outcome of a measurement $M$ must depend on its context, i.e. the set of compatible measurements one may perform alongside $M$. Contextuality can be detected through the violation of noncontextuality inequalities, of which the simplest is the Klyshko-Can-Binicioglu-Shumovsky (KCBS) inequality [KCBS08]. It corresponds to the scenario in which five projective binary measurements $\{M_i = (M_{0|i}, M_{1|i})\}_{i=1}^5$ can be performed and each pair $M_i, M_{i+1}$ is compatible, i.e.

$$[M_i, M_{i+1}] = 0 \quad (69)$$

and exclusive i.e. $\text{tr}(M_{0|i}M_{0|i+1}) = 0$ (the labels of measurements are taken modulo 5). If we denote $p_i = \text{tr}[M_{0|i}\rho]$, where $\rho$ is the measured state the KCBS inequality reads:

$$\sum_{i=1}^5 p_i \leq 2.$$

The inequality is satisfied in all outcome deterministic noncontextual theories, while quantum measurements achieve the value $c_q = 5\cos(\pi/5)/(1 + \cos(\pi/5))$. The maximal quantum value is achieved by measuring a pure state $\rho = |v_0⟩⟨v_0|$ and measurements $M_{0|i} = |v_i⟩⟨v_i|$ where

$$|v_0⟩ = (1, 0, 0)^T,$$

$$|v_i⟩ = (\cos \theta, \sin \theta \sin \phi_i, \sin \theta \cos \phi_i)^T,$$

where $\cos \theta = \cos(\pi/5)/(1 + \cos(\pi/5))$ and $\phi_i = i\pi(4/5)$.

Note that since this scenario consists of making measurements on a single quantum system, the statistics can always be simulated classically and it is thus impossible to achieve self-testing in a fully device-independent manner. In [BRV+19] the authors overcome this by assuming compatibility relations such as (69), thus rendering the scheme semi-device independent. Using this, they give a general scheme to self-test quantum states using noncontextuality inequalities and prove that the maximal quantum violation of the KCBS inequality robustly self-tests the above strategy. In other words, for any state $\rho$ and measurements $\{|v_i⟩⟨v_i|\}$ which can be used to achieve the violation $c_q - \epsilon$ of KCBS inequality there exist a unitary $U$ such that

$$\|U |v'_i⟩ - |v_i⟩\| \leq O(\sqrt{\epsilon}), \quad \forall i.$$

The proof holds also for the generalisations of KCBS inequality given in [BBC+11, AQB+13, LSW11]. The crux of the proof is an equivalence between the optimal strategy for violating the KCBS inequality and the solution to a certain type of SDP optimisation known as the Lovász theta number of an odd cycle graph [Lov79].

### 10 Applications of self-testing

The birth of self-testing is usually associated to the Mayers-Yao paper [MY04] from 2004. It set the terminology and formalism, including the first usage of the term self-testing in this context and identifying local isometries as relevant transformations. A similar main result as in [MY04] was presented in [MY98], although in the context of whether untrusted sources can be pertinent for cryptographic tasks. This earlier paper used the term ‘self-checking’ instead of ‘self-testing’. Moreover, [MY98] is at the same time one of the pioneering works in device-independent cryptography, indicating importance of self-testing for the development of device-independent protocols. Since then self-testing has been scrutinised as a task of twofold significance:
• purely theoretical, related to exploring the conditions for a probability distribution to determine a specific quantum state and/or measurements, and proving that such statements also hold approximately. This theoretical aspect was reviewed in Sections 4, 5, 6, 7, 8 and 9.2. As a result, ideas and results from self-testing can lead to progress in related theoretical areas.

• practical, relevant for creating new device-independent or semi-device-independent protocols for different tasks. In the Mayers-Yao paper [MY04] the authors say ‘We hope that it will have application in different areas of quantum information processing’. Fifteen years later we can observe that this hope is fulfilled.

In this section we give an overview of the applications of self-testing during the first fourteen years after the technique has been formally introduced. On the practical side, we cover the relation of self-testing with device-independent randomness generation in section 10.1, device-independent quantum cryptography in section 10.2, and device-independent entanglement certification in section 10.3. In section 10.4 we describe the applications to delegated quantum computing 10.4. Finally, from the theoretical side, in section 10.5 we describe the influence self-testing has had in understanding the structure of the set of quantum correlations. In particular in section 10.5.1 we highlight a link to the study of quantum correlations produced by finite vs infinite quantum systems, and in section 10.5.2 how self-testing has been used to prove an inequivalence between uncertainty relations and Bell nonlocality.

10.1 Device-independent randomness generation

The probabilistic nature of quantum mechanics can be exploited for generation of random numbers. In the simplest example, measuring in the computational basis a qubit in the state $|+\rangle$ results in a perfectly random output bit. However, the certification of the random nature of bits obtained this way relies on the exact characterisation of both the quantum state and the measurement performed. The device-independent scenario offers much less stringent requirements for randomness certification, by qualitatively relating randomness with nonlocality. By treating her devices as non-communicating black boxes Alice can certify some amount of randomness by observing the violation of a Bell inequality. The first results in this direction show that the maximal violation of the Mermin [Col06, CK11b] and the CHSH inequality [PAM+10] can be used in this way. For more information on certification of quantum randomness see [AM16] and references therein.

Here, we comment on the relation between self-testing and randomness. A self-testing protocol proves the existence of a pure entangled state and a certain set of measurements acting on it. Once this conclusion is made, certified random bits come for free since local measurements on a pure entangled state necessarily produce random outcomes. As a result, ideas from self-testing are often either implicitly present in device-independent randomness works or are explicitly used as tools to prove randomness lower bounds.

In the pioneering works of [Col06, CK11b] a self-testing statement is implicitly present, where the authors prove that only an orthogonal sum of GHZ states can maximally violate the Mermin inequality. In [CY14], the sequential self-testing of $n$ EPR-pairs proven in [RUV13] is used as a sub-protocol for infinite randomness expansion with a constant number of devices.

Simple symmetry-based arguments are used in [DPA13] to prove that the violation of some Bell inequalities can be used to certify the presence of genuine randomness. A necessary condition is that there exists a unique probability distribution maximally violating the Bell inequality. One way to prove such uniqueness is through self-testing: if the maximal violation of the Bell inequality is a self-test the maximally violating probability distribution has to be unique. Furthermore, incomplete results from [DPA13], were proven to be true by using self-testing techniques in [ŠASA16].

The results on self-testing properties of binary XOR games from [MS13] were expanded in [MS16] and used to devise protocols for exponential randomness expansion. More recently, the authors of [BMP18] directly use robust self-testing bounds for the tilted-CHSH inequality [BP15] to lower bound the randomness generated in their protocol. Self-testing techniques are also
used in [APVW16, ABDC18, WKB+19] to prove that two bits of local randomness can be certified from a two-qubit entangled state.

10.2 Device-independent quantum cryptography

10.2.1 Quantum key distribution

Quantum key distribution (QKD) is the most widely studied quantum cryptographic protocol in which two parties, Alice and Bob, use quantum resources to generate a shared private key which can later be used for encryption and decryption of messages. The security of a standard QKD protocol relies on the correct characterisation of all devices, which can be difficult to achieve in practice and far from ideal from a security perspective. An alternative approach comes from device-independent quantum key distribution (DIQKD), where security is based only on the observation of the correlations, and can be proven even if the constituent devices are treated as black boxes. DIQKD is intimately related to DI randomness generation; whereas in randomness generation one aims to have random outcomes, in a DIQKD protocol one aims to have random outcomes that are also correlated between Alice and Bob (thus ensuring a shared private key). As with randomness generation, the security of DIQKD is often measured against the violation of some Bell inequality.

An indication of a close relation between DIQKD and self-testing is their common root in the Mayers-Yao work [MY98]. It discusses self-testing as a protocol for the first time (under the name self-checking) and recognises that it can help to use untrusted devices in cryptographic setting. [MY98] consider the BB84 protocol [BB84] in which Alice certifies an untrusted source she wants to use. The source is supposed to emit EPR pairs with Alice keeping one particle and measuring it and sending the other one to Bob. The untrusted source can be self-tested using the Mayers-Yao self-testing criterion, as explained in section 5. The protocol is later discussed in [MT02] in the context of the Ekert QKD protocol [Eke91]. The second Mayers-Yao paper [MY04], improving the first one by characterising the measurements (and introducing the phrase ‘self-testing’) also discussed the relation of self-testing with the BB84 protocol.

Ever since then self-testing and DIQKD have been intertwined. A certification of some quantum resource is implicitly present in every DIQKD security proof, however in some works the relation between self-testing, as the strongest form of certification, and DIQKD was explicitly examined. The effect of the inability to self-test complex measurements on the security of cryptographic tasks has been the subject of [MM11]. The authors prove that the 6-state QKD protocol [BBBW84, Bru98] can be secure even if the devices are untrusted, despite the issue with complex conjugation. Similarly like in the protocols for randomness expansion, the self-testing properties of XOR binary nonlocal games explored in [MS13] were used in [MS16] to prove the security of certain class of QKD protocols. Finally, the concept of parallel DIQKD developed analogously to parallel self-testing was first introduced in [JMS17] and the security proof relied on the rigidity of the magic square game [WBMS16] allowing for parallel self-testing of two singlets. A simplified proof appeared in [Vid17].

10.2.2 Cryptography beyond quantum key distribution

Bit commitment—Bit commitment is a cryptographic primitive in which Alice chooses a bit $b$ that she wants to first commit, then later reveal, to Bob. The protocol should be both binding (Alice should not be able to change her choice of $b$ after the commit step) and hiding (Bob should not be able to know $b$ until Alice chooses to reveal). In a classical protocol, either Alice or Bob can cheat with probability 1 without being caught, i.e. either Alice can alter the bit after committing or Bob can learn it before it is revealed. Although unconditionally secure bit commitment is known to be impossible even using quantum resources [May97, LC97], there exist quantum protocols in which either party’s probability to successfully cheat is strictly smaller than 1 [CK11a]. [SCA+11] introduces the idea of a DI quantum bit commitment protocol in which besides not trusting each other, Alice and Bob do not trust their equipment either. The security of the protocol is based on the self-testing fact that the maximal violation of the Mermin inequality can only be achieved by measuring the GHZ state. A version of DI quantum bit commitment based on the
violation of the CHSH inequality is presented in [AMPS16], in which the security can be seen as consequence of the self-testing properties of the CHSH inequality. Similarly, the violation of the CHSH inequality has been used to prove the security of DI relativistic bit commitment in [AK15], ruling out the possibility of location attacks in which devices are able to track their own space and time coordinates. None of the works explicitly relate their results to the corresponding self-testing protocol however.

Weak string erasure— Weak string erasure (WSE) [KWW12] is a primitive which can be used in two-party cryptographic protocols in which no large scale reliable quantum storage is available to the cheating party. WSE provides a random bit string \((b_1, \ldots, b_n)\) to Alice, while sending a randomly chosen substring \((b_1, \ldots, b_k)\) to Bob, together with the set \((i_1, \ldots, i_k)\) specifying the location of substring bits. WSE is secure against Bob if he cannot learn much about the full string given to Alice, while it is secure against Alice if she cannot learn the location of Bob’s bits. A DI version of this protocol useful for bit commitment or oblivious transfer is introduced in [KW16] and the security is related to the self-testing properties of the CHSH inequality.

Position verification— Finally, we briefly mention the position verification primitive, useful in position-based cryptography in which the parties have to convince the (honest) verifiers that they are located at a particular location. Protocols for position verification that improve the security by using quantum communication have been proposed in [KMS11, Mal10]. The DI security of position verification against adversaries with no quantum memory is proven in [RTK+18], and can also be traced to the self-testing properties of the CHSH inequality.

10.3 Entanglement detection

One of the most basic tasks in quantum information is that of detecting entanglement of a bipartite quantum system via local measurements on its subsystems. Device-independent entanglement detection considers this problem in the device-independent scenario, i.e. where all local measurement devices are treated as black boxes. Since the observation of Bell nonlocal correlations necessarily implies that the underlying state is entangled, the standard approach to DI entanglement detection involves violating a Bell inequality. However, since there exist entangled mixed states that do not violate any Bell inequality\(^2 [Wer89, Bar02, ADA14, BHQB16, JHA+15, BFF+16, HQB+16]\), this method cannot be used for all entangled states. A partial solution to this problem, allowing for the entanglement detection of all entangled states, was given in [Bus12] (see also [BRLG13]) using the concept of a ‘semi-quantum game’. Here, the classical inputs in a Bell test are replaced by ‘quantum inputs’ \(|\psi_x\rangle, |\psi_y\rangle\), that is, a set of known quantum states that are sent to the measurement device instead of the classical labels \(x\) and \(y\). This scenario is semi-device-independent since although the measurement devices are treated as black boxes, the quantum input states must be trusted.

In [BŠCA18a, BŠCA18b], tools from self-testing and semi-quantum games were used to construct fully DI protocols for the entanglement detection of all entangled mixed states. The idea is as follows. If one achieves a self-test of a particular state and local measurements for Alice, then this certifies (up to a local isometry) the reduced states of Bob conditioned on a particular choice of input/output for Alice. In this way one can certify an ensemble of state preparations (conditioned on Alice’s input/output) on Bob’s local Hilbert space. These preparations can then be used as quantum inputs in a semi-quantum game. Since (i) the quantum inputs are now certified device-independently, (ii) the semi-quantum games scenario can be applied to all entangled states, the two can be combined to construct a fully device-independent protocol that works for all entangled states. Specifically, one needs to consider a network scenario in which the state of interest is augmented with two auxiliary bipartite states that are used to prepare the quantum inputs. Here, tools from parallel self-testing as well as the issue of complex conjugation become important for the general proof.

\(^2\)At least in the original Bell scenario in which Alice and Bob can perform any number of non-sequential local measurements on a single copy of the state. In more complex measurement scenarios (see [CASA11, SDSB+05, Pal12, TRC19, BRGP12, HQBB13, Pop95]) it is generally unknown if such states exist.
Delegated quantum computing

Delegated computation is a protocol in which a party, usually called a verifier, delegates a computational task to another party, usually called a prover. The verifier aims to solve difficult computational tasks, but does not have enough computational resources. The prover, on the other side, has a very powerful computer and is able to solve any task the verifier is interested in. When one talks about delegated quantum computation (DQC) the prover possesses a quantum computer, while the verifier has either only classical computing resources or limited quantum resources but wants to solve a problem intractable for classical computing devices. For a concise review on the existing approaches in DQC see [GKK18].

There are two desirable properties of a DQC protocol: verifiability and blindness. The protocol is said to be verifiable if the verifier can be convinced that the solution provided by the prover(s) is correct. This is non-trivial, since the verifier is unable to solve the problem. Blindness of the protocol is related to the secrecy of the computation. It is ensured when the prover(s) cannot learn anything about the computational task the verifier wants to perform. It is very difficult to construct a DQC protocol with a fully classical verifier and a single prover which is verifiable and blind. The first protocol that achieves this, under computational assumptions, is [Mah18]. In principle, it is easier to achieve both verifiable and blind protocol with a classical verifier when there is more than one prover. In this case the provers are entangled and forbidden to communicate. Verifiability is proven if the verifier can be convinced that the two or more non-communicating provers are performing the prescribed sequence of measurements. This, of course, requires that the verifier be able to test that the provers perform measurements from a set that is universal for quantum computing. The latter is exactly a self-testing task. Thus it is no surprise, that self-testing is useful for such a delegation protocol, yet orchestrating such a computation in a verifiable fashion is a delicate task. Delegation protocols with two or more provers based on self-testing typically achieve information theoretic security.

The first such protocol was presented by Reichardt, Unger and Vazirani (RUV) in [RUV13]. The protocol involves two provers sharing a tensor product of many EPR pairs and the computation model is quantum computation by teleportation [GC99]. The provers are able to convince the verifier that they possess \(N\) EPR pairs by obtaining the optimal score in sequential playing of the CHSH game. The complexity of RUV protocol in terms of time and the number of EPR pairs needed is extremely large. The protocol was subsequently improved in [JMS20] where each Bell pair is shared by two provers, making the number of provers increase significantly at the expense of reducing the overall complexity. The advantage of parallel self-testing of \(N\) EPR pairs instead of sequential was exploited in [NV17] and [CGJV17]. The latter obtains a protocol with an almost optimal overhead (in the size of the computation) in terms or resources used, by exploiting parallel self-tests with robustness independent of the number of EPR pairs tested, discussed previously in subsection 5.3. The work [GKW15] shows that RUV protocol can be significantly improved if the verifier is actually quantum and wants to be convinced that they share with the prover \(n\) EPR pairs. In this case the self-testing of \(n\) EPR pairs through steering becomes a relevant sub-protocol.

Another example of self-testing incorporated into a delegated quantum computing protocol is [McK16a]. The DQC protocol involves many provers which share a graph state. The model of computation is measurement-based quantum computing (MBQC) [RB01, RBB03]. The protocol is made verifiable by using the self-testing of graph-states mostly based on [McK14]. Based on a similar idea a significantly simplified protocol appeared in [HH18]: since the triangular lattice graph state is universal for MBQC [MP13] the number of provers can be reduced to three and the number of necessary copies of the graph states is also considerably smaller than in [McK16a].

Structure of the set of quantum correlations

Here we highlight two ways in which self-testing as furthered our understanding of the set of quantum correlations.

Correlations from finite vs infinite dimensional quantum strategies

It is said that the correlations \(\{p(a, b|x, y)\}_{a, b, x, y}\) admit a quantum strategy if there exists a state...
\( |\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B \) and projective measurements \( \{M_{a|x}\}_{a,x}, \{N_{b|y}\}_{b,y} \) such that
\[
p(a, b|x, y) = \langle \psi | M_{a|x} \otimes N_{b|y} |\psi\rangle.
\]
This strategy is also called a tensor-product strategy due to the tensor product between Alice’s and Bob’s spaces. Different sets of quantum correlations arise when certain conditions are imposed on the state \( |\psi\rangle \) and/or measurements \( \{M_{a|x}\}_{a,x}, \{N_{b|y}\}_{b,y} \). The correlations obtained through a tensor-product strategy on finite-dimensional Hilbert spaces \( \mathcal{H}_A \) and \( \mathcal{H}_B \) constitute the set denoted as \( C_q \). If the Hilbert spaces \( \mathcal{H}_A \) and \( \mathcal{H}_B \) can also be infinite dimensional the set is denoted as \( C_{qs} \). The closure of the set \( C_q \) is called \( C_{qa} \). Additionally, a commuting-operator strategy is one in which no tensor product structure is imposed but instead all measurement operators of Alice commute with all those of Bob, i.e.
\[
p(a, b|x, y) = \langle \psi | M_{a|x} N_{b|y} |\psi\rangle.
\]
with \( [M_{a|x}, N_{b|y}] = 0 \). The commuting-operator model is used in algebraic quantum field theory and all such correlations are denoted by the set \( C_{qc} \). The set inclusion relation defines a hierarchy among these sets [PT15]:
\[
C_q \subseteq C_{qs} \subseteq C_{qa} \subseteq C_{qc}.
\] (70)
Whether \( C_{qc} \) is equivalent to either \( C_{qs} \) or \( C_{qa} \) are problems known as Tsirelson’s problems [Tsi93]. Recently, it has been proven that \( C_{qs} \neq C_{qc} \) [Slo20], and also that \( C_{qs} \neq C_{qa} \) in [Slo19]. Whether \( C_{qa} \subseteq C_{qc} \) remains an open problem.

Self-testing techniques have also inspired proving a strict inclusion of \( C_q \) in \( C_{qs} \). The inequivalence between these two sets was proven in [CS17c] in cases when either the number of inputs or the number of outputs is infinite. The separation between \( C_q \) and \( C_{qs} \) for finite input or output alphabets was proven by the same authors in [CS18]. The separation is demonstrated by explicitly giving the correlations which can be obtained with infinite-dimensional quantum strategies, but not with any finite-dimensional ones. The proof is inspired by the protocol for self-testing infinite-dimensional bipartite pure states, described in section 5.2.1. In the finite case, the self-test of a bipartite state can be interpreted as involving one of two different protocols depending on the parity of the dimension. An infinite-dimensional state does not have a defined parity and can thus be self-tested by either protocol. The authors use this fact as the theoretical basis to prove the separation.

Following Slofstra’s proof of non-closure of the quantum set of correlations [Slo19] alternative proofs appeared in [DPP19] and [MR19]. All these proofs rely on the representation theory of \( C^* \)-algebras. A relatively simpler proof, using embezzling entanglement [vDH03] and self-testing, is presented in [Col20].

10.5.2 Uncertainty relations and Bell nonlocality

Self-testing has also been used as a tool to understand the relationship between uncertainty relations, steering and Bell nonlocality. In [OW10] it was shown that in the optimal quantum strategy for a large class of Bell inequalities called XOR games, the steered states of Bob after a measurement by Alice are always such that they saturate a fine-grained uncertainty relation defined by the Bell inequality itself. This was then extended to more Bell inequalities in [ZGZ16]. Whether such a link was true for generic Bell inequalities remained an open question.

In [RM+18] (building on [ZGZ+16]) the authors answer this in the negative by providing specific examples of Bell inequalities for which the relationship does not hold. To achieve this, self-testing statements for the Bell inequalities are proven, from which the form of the steered states can be determined and checked against the corresponding uncertainty relations. The self-testing statements are proven using Jordan’s lemma (see section 7.1.2) to reduce the problem to essentially a two-qubit strategy, which simplifies the analysis significantly.

11 Experiments

The bulk of self-testing procedures are still only theoretical recipes. This is understandable since the majority of robust self-testing protocols have fidelity bounds that decrease rapidly with noise and serve only as a proof of principle. However, in the last years there has been an increasing number of self-testing protocols robust to realistic amounts of noise (see section 7). Here we mention the few experimental realisations of such protocols.
The biggest experimental hurdle towards fully device-independent protocols is simultaneously closing detection and locality loopholes. [BRS++18] reports the self-testing of a Bell state distributed over 398 meters through the violation of the CHSH inequality free of both detection and locality loopholes and furthermore free of the i.i.d. assumption. In this light it is the first fully device-independent self-testing protocol to be implemented in practice. Entanglement between the distant atoms is generated by entangling the spin of each atom with polarisation of a single photon. The obtained fidelity is 55.54% at a confidence level of 99%. By applying the same theoretical tools to analyse the data obtained in the loophole-free Bell test presented in [HKB++16] no fidelity higher than the trivial 50% could be found.

The remainder of the experimental self-testing contributions are not based on a completely loophole-free Bell tests. [TWE++17] reports a high violation of the CHSH inequality by a pair of 9Be+ ions. The violation is used to make a self-testing statement, based on [Kan16]. At the 95% confidence level the pair of 9Be+ ions has 0.958 fidelity with the maximally entangled pair of qubits. An overview of the inferred self-testing bounds from some previous works reporting CHSH violations is also presented. In [ZCP++18] the operator inequalities for robust self-testing from [Kan16] are tested for a large number of bipartite and tripartite qubit states encoded in photon polarisation degrees of freedom.

Experimental robust self-testing of partially entangled pairs of qubits using the Swap method is presented in [ZCY++19]. The systems under consideration are polarisation entangled photons. Self-testing of partially entangled qubit pairs is also used to heuristically estimate the fidelity of a produced ququart state with a given reference state. Robust self-testing of partially entangled pairs of qubits encoded in photon polarisation degrees of freedom was reported in [GMM++19]. The self-testing was done through the violation of the tilted CHSH inequality and the robustness bounds were estimated by using the numerical results from [CKS19]. In [GPL++19] the authors explore the certification of partially entangled pairs of photons encoded in the polarisation degree of freedom. The fidelity of the entangled pair with the corresponding partially entangled pair of qubits is estimated in two ways: by using standard tomographic methods and self-testing. The obtained fidelities have ratio $\approx 0.998$, implying that for the case of qubit states self-testing can be used to achieve almost the same conclusions as tomography. It is argued that self-testing may have an advantage over tomography even when the detection and locality loopholes are not closed since it avoids characterisation of measurements and assumptions about dimension and, in principle, requires estimating fewer average values.

In [WPD++18] various two-qutrit entangled states are self-tested using photons entangled in the mode degree of freedom of the waveguides in a silicon based integrated optical chip. Fidelity bounds were obtained numerically via the Swap method. The self-tested states are the maximally entangled pair of qutrits (estimated fidelity 0.799), the state maximally violating the CGLMP inequality [CGL++02] (estimated fidelity 0.68) and a state maximally violating an extension of one of the SATWAP inequalities [SAT++17] (estimated fidelity 0.832).

Experimental self-testing in the steering scenario is the subject of [LLW++19]. The fidelity of the underlying physical state with the GHZ state is estimated based on the violation of the Mermin’s steering inequality [CHR++11]. The fidelity lower bound is estimated to be 0.7866, while the tomographically retrieved fidelity is $0.8725 \pm 0.0034$.

Finally, in [ZCP++19] the authors report an experimental realisation of a robust self-test of a Bell state measurement based on the entanglement swapping protocols of [SBWS18, RKB18] (see section 9.1). Photon pairs that are hyperentangled in the spatial and polarisation degrees of freedom are used to encode the two maximally entangled pairs of qubits that are needed for the entanglement swapping protocol.

12 Concluding remarks and open questions

Recent years have seen an increased interest in device-independent self-testing, accompanied by the plethora of self-testing protocols and methods presented in this review. However, there are still many important unresolved questions.
Without aiming to exhaust the list, we name some open questions and research directions which we believe worthy of attention.

**Analytic methods for dimension larger than 2**—The majority of known self-testing protocols are either built to self-test multi-qubit states and measurements, or apply existing qubit protocols to the self-testing of higher dimensional systems. The self-testing of states and measurements using methods that exploit the genuine $d$ dimensional nature of quantum systems, is however still a very unexplored area. For instance, it has been known for a long time that via the numerical Swap method, the CGLMP inequality self-tests the two-qutrit state of equation (43). However, a corresponding analytic proof of this statement is still lacking, despite the relative simplicity of the inequality. Similarly desirable are analytic proofs for the self-testing of the maximally entangled states in dimension $d$ using the SATWAP inequalities [SAT+17]. With respect to high dimensional measurement self-testing, one important open question is to extend the analytic self-test of a set of qutrit mutually unbiased basis measurements and the maximally entangled state of [KŠT+19] to higher dimensions. Moving beyond high dimensional systems to continuous variable systems, essentially nothing is known and there exist no protocols to self-test such states.

**Multipartite methods**—In a similar vein, techniques for self-testing general multipartite states are also needed, since current methods are only known for restricted classes such as graph states. One potential line of research in this direction would be to develop methods to self-test multi-qubit hypergraph states [RHBM13], which exhibit a richer structure than graph states but still admit a useful description in terms of Clifford group stabiliser operators. One would thus need a general method to construct Bell inequalities for such states, as was done for graph states [GTHB05, BAS+20]. This appears more complicated for hypergraph states however since the nonlocal nature of the hypergraph stabilizer operators means they do not have an obvious interpretation as local measurement observables (although some progress has been made [GBG16]).

**Identifying the set of undetectable transformations**—Part of the challenge in going beyond two qubit methods is to identify the set of local transformations defining the equivalence classes of self-testable states and measurements in higher dimensions. As we have seen, the standard definitions presented in section 3 need to be adapted in order to self-test complex valued measurements, stemming from the invariance of quantum correlations under complex conjugation of the state and measurement operators. In higher dimensions, it is still unknown whether there exist more state and measurement transformations that leave correlations invariant. If such transformations exist, an all-encompassing definition of what it means to self-test a state and measurements in general dimension is therefore still missing.

**Self-testing of a state or measurements only**—In many self-testing works, a self-testing statement for the state is accompanied by an additional self-testing statement for the measurements. One interesting problem would be to find situations where the correlations allow one to identify the state, but not the measurements, even allowing for the freedom of complex conjugation of the measurements in the definition of measurement self-testing. Similarly, it would be interesting to find correlations that allow one to self-test the measurements, but not the state.

**Self-testing in non-i.i.d. scenarios**—Throughout this review we have made the assumption that each round of the experiment is independent and identical to all others. It would be interesting and practically relevant to attempt to remove this assumption from self-testing protocols. One way to achieve this would be to build sequential self-testing protocols akin one in [RUV13] or by leveraging some recently introduced techniques [BRS+18, DFR16, AFRV19, AF18] for such scenarios. Methods could be borrowed also from the protocols for delegated quantum computation, i.e. [CGJV17]. Since the state and measurements can now depend on settings and outcomes in previous rounds, one would also need to adapt the definition of self-testing to apply to such scenarios, as well as derive corresponding confidence bounds from finite statistics. For example, the aim would be to show that, to high statistical confidence, the source is producing something close (by some measure) to $n$ independent copies of $|\psi\rangle$, where $n$ is the number of experimental rounds.
Improved robustness methods—Finally, methods to improve the robustness bounds of general self-testing protocols are much in need. In practice, the applicability of the majority of self-testing protocols is hindered by very poor tolerance to noise. Significant improvements have been achieved for simple scenarios [Kan16], however it is not clear if these methods can be extended to scenarios with more inputs and outputs due to their dependence on Jordan’s lemma. Finding a good robustness bound involves a difficult maximisation over all local isometries, and as a result nearly all methods use one of the few Swap isometries that are known to give good results in the well-explored simple cases. Thus, knowing more useful isometries and understanding which work well for particular classes of states would likely lead to improved robustness bounds.

Acknowledgements

We are grateful to Alexia Salavrakos, Jean-Daniel Bancal, Andrea Coladangelo, Jed Kaniewski, Armin Tavakoli, Erik Woodhead, Alejandro Pozas-Kerstjens, Aleksandra Dimić, Marc-Olivier Renou, Felix Hüber, Sébastien Designeolle, Nicolas Brunner, Yeong Cherng Liang, Nikolai Miklin and Rotem Arnon-Friedman for suggestions while preparing the manuscript. We acknowledge funding from the ERC CoG QITBOX, the Spanish MINECO (QIBEQI FIS2016-80773-P, Severo Ochoa SEV-2015-0522), Fundacio Cellex, Generalitat de Catalunya (SGR 1381 and CERCA Programme). IŚ acknowledges funding from SNSF (Starting grant DIAQ). JB acknowledges funding from the Juan de la Cierva-formación grant the AXA chair in quantum information science.

A Appendix

A.1 Self-testing complex measurements

Here we give a possible definition of self-testing of states and complex valued measurements.

Definition 6. (self-testing of states and complex measurements)

We say that the correlations \( p(a,b|x,y) \) self-test the state and measurements \( |\psi\rangle_{A'B'}, \{M'_{a|x}\}, \{N'_{b|y}\} \) if for all states and measurements \( \rho_{AB}, \{M_{a|x}\}, \{N_{b|y}\} \) compatible with \( p(a,b|x,y) \) there exists a local isometry \( \Phi = \Phi_A \otimes \Phi_B \) such that for any purification \( |\psi\rangle_{ABP} \) of \( \rho_{AB} \) there exists some state \( |\xi\rangle_{ABP} \) such that

\[
\Phi \otimes I_P \left[ M_{a|x} \otimes N_{b|y} \otimes I_P |\psi\rangle_{ABP} \right] = \tilde{M}_{a|x} \otimes \tilde{N}_{b|y} \otimes I_P \left( |\psi\rangle_{A'B'} \otimes |\xi\rangle_{ABP} \right),
\]

for all \( a, x, b, y \), and where

\[
\begin{align*}
\tilde{M}_{a|x} &= M_{a|x}^* \otimes S_0^A + (M_{a|x}^*)^* \otimes S_0^A, \\
\tilde{N}_{b|y} &= N_{b|y}^* \otimes T_0^B + (N_{b|y}^*)^* \otimes T_0^B, \\
S_0 + S_1 I_A, &
T_0 + T_1 = I_B, \\
\langle \xi | (S_0 \otimes T_0 + S_1 \otimes T_1) \otimes I_P |\xi\rangle &= 1.
\end{align*}
\]

Here, the \( S_i \) and \( T_i \) part of the measurements are acting as effective controlled complex conjugations of the reference measurements. The final condition ensures that this conjugation is performed in a correlated fashion as implied from (14). The probability that the conjugation is performed depends on the (unknown) junk state and is thus unknown. Note that if one traces out all but the \( AB \) space, one obtains some convex combination of the reference and conjugated measurements acting on the reference state.

A.2 Regularisation trick

In this appendix we give more details about the so-called regularisation trick. It refers to the case when one of the operators used to build the Swap gate (see figure 4) is not unitary. This happens already in the case described in Chapter 4 where the operators \( Z_A \) and \( X_A \) from equation (32) might have some zero eigenvalues. Let us focus on \( Z_A = (A_0 + A_1)/\sqrt{2} \). The first step in the regularisation procedure is to change all the zero eigenvalues of \( Z_A \) to 1, resulting in a new operator \( Z_A^* \). In the second step, all eigenvalues are normalized, i.e. the new operator defined as \( \tilde{Z}_A = Z_A^*/|Z_A^*| \) is unitary by construction. However, one has to prove that \( \tilde{Z}_A \) acts on the physical state in the same way as \( Z_A \). For that the following series of inequalities can be used (Note that \( Z_A^* \) acts on \( |\psi\rangle \) in the same way as \( Z_A \) since it can be seen as \( Z_A + P \) where \( P \) is the projector...
The inequality follows from the operator inequality $A \leq |A|$. The last line stems from equation (31). The key ingredient necessary for regularisation is exactly equation (31). In general, the regularisation of any operator $A$ can be done if there is a unitary $U$ such that $A \otimes 1 |\psi\rangle = 1 \otimes U |\psi\rangle$.

A.3 Swap isometries

In this appendix we provide further comments on the different Swap isometries used in the self-testing protocols. In section 4.3 we mentioned that the partial Swap gate given on figure 4 is appropriate only if the ancillas are initiated in the state $|0\rangle$. In the case that the ancillas are in a different state the correct isometry to use is the full Swap gate, given in figure 10.

A generalisation of the Swap gate, given on figure 11, can be used for self-testing of bipartite qudit states $|\psi\rangle = \sum_{j=0}^{d-1} \lambda_j |jj\rangle$, where $\lambda_j$ are positive real numbers. The gate $F$ is the Fourier transform defined as:

$$F \left| j \right\rangle = \frac{1}{\sqrt{d}} \sum_{k=0}^{d-1} \omega^{jk} \left| k \right\rangle,$$

where $d$ is the local dimension of the reference state, and $\omega$ is the $d$-th root of the unity. The controlled gates $CZ$ and $CX$ are defined as follows:

$$CX \left| j \right\rangle = \left| j \right\rangle \left( X^{(j)} \right) |\psi\rangle$$

$$CZ \left| j \right\rangle = \left| j \right\rangle \left( Z^{(j)} \right) |\psi\rangle.$$

For the gate on figure 11 to work as an effective Swap gate the operators $X$ and $Z$ have to satisfy certain conditions, mimicking anticommutativity\footnote{Note that $\{X^{(j)}\}$ are $j$ different operators, while $\{Z^{(j)}\}$ are $j$-th powers of the operator $Z$.}

from the qubit case. In [YN13] the authors give the recipe: operators $\tilde{Z}_A$ and $\tilde{Z}_B$ have to satisfy

$$\sum_{j=0}^{d-1} \omega^{ja} \tilde{Z}_A^j \otimes 1 |\psi\rangle = 1 \otimes \sum_{j=0}^{d-1} \omega^{ja} \tilde{Z}_B^j |\psi\rangle,$$

for all $a \in \{1, \cdots, d\}$. In addition, operators $\tilde{X}_A^{(k)}$ and $\tilde{X}_B^{(k)}$ must satisfy

$$\lambda_0 \tilde{X}_A^{(k)} \otimes \sum_{j=0}^{d-1} \omega^{jk} \tilde{Z}_B^j |\psi\rangle$$

$$= \lambda_k \sum_{j=0}^{d-1} \omega^{jd} \tilde{Z}_A^j \otimes \left( \tilde{X}_B^{(j)} \right)^\dagger |\psi\rangle, \quad \forall k. \quad (72)$$

It can be proven that the output state of the isometry $\Phi_d$ given on figure 11 and built from the operators satisfying the conditions (71) and (72) has the form $\sum_k \lambda_k |kk\rangle_{AB} \otimes |\xi\rangle_{AB}$ where $|\xi\rangle$ is some normalised state. For more details about the qudit Swap isometry see [YN13, CGS17, ŠCAA18].

A.4 Localising matrices in the Swap method

In section 7.1.4 we presented the numerical Swap method used in robust self-testing protocols. The
isometry used in the Swap method is the Swap gate and as we discussed in Appendix A.2 when one of the operators $X_A$, $X_B$, $Z_A$ or $Z_B$ is defined as a sum or difference of physical measurement observables, the Swap isometry might not be unitary. Let us, for simplicity, focus the CHSH case and the operator $Z_A = (A_0 + A_1)/\sqrt{2}$ and $X_A = (A_0 - A_1)/\sqrt{2}$ used to build the Swap isometry. In Appendix A.2 we showed how to regularize such operators for the purposes of the ideal self-testing.

The procedure to solve this problem when using the Swap isometry in robust self-testing protocols is introduced in [BNS+15, YVB+14]. In the context of the Swap method one solves the problem by introducing two new operators $A_2$ and $A_3$ which are unitary and not too different from $(A_0 + A_1)/\sqrt{2}$ and $(A_0 - A_1)/\sqrt{2}$, respectively. That way the isometry built by defining $Z_A = A_2$ and $X_A = A_3$ is necessarily unitary. One way to impose the proximity of $A_2$ and $A_3$ to $(A_0 + A_1)/\sqrt{2}$ and $(A_0 - A_1)/\sqrt{2}$, respectively, is by imposing the relaxation

$$A_2 (A_0 + A_1)/\sqrt{2} \geq 0 \quad (73a)$$

$$A_3 (A_0 - A_1)/\sqrt{2} \geq 0. \quad (73b)$$

This can be enforced by introducing two new moment matrices called the localizing matrices. The condition (73a) can be imposed as a requirement that the moment matrix defined as

$$\Gamma_{i,j}^k (A_2) = \langle \psi | S^{(i)} | \psi \rangle A_2 \frac{A_0 + A_1}{\sqrt{2}} S^{(j)} | \psi \rangle, \quad (74)$$

where $S = \{1, A_0, A_1, A_2\}$, is positive semi-definite. An analogous constraint can be made to enforce the condition (73b).

## B State and measurement assumptions

Here we overview the assumptions made about the state and measurements in the definitions of self-testing.

### B.1 State

In the definitions of section 3 we use a purification $|\psi\rangle_{ABP}$ of the potentially mixed physical state $\rho_{AB}$. One does not assume that the state shared between Alice and Bob is pure however, since it is given by

$$\rho_{AB} = \text{tr}_P [|\psi\rangle_{ABP} \langle \psi |_{ABP}]. \quad (75)$$

Suppose we have a self-testing statement in the form of definition 1:

$$\Phi_A \otimes \Phi_B [\rho_{AB}] = [\langle \psi' |_{A'B'} \langle \xi |_{\bar{A}\bar{B'}}].$$

If we trace over the purification space on both sides of the above we obtain

$$\Phi_A \otimes \Phi_B [\rho_{AB}] = [\langle \psi' |_{A'B'} \otimes \sigma_{AB}].$$

where $\sigma = \text{tr}_P [|\xi \rangle \langle \xi |]$. Thus a self-testing statement of the form of definition 1, implies that the same isometry maps $\rho_{AB}$ to the reference state. This is possible because the isometry map acts trivially on the purification space. This follows from the fact that the isometry is constructed from the measurement operators, which by assumption themselves act non-trivially only on $\rho$.

### B.2 Measurements

We do, however, assume that the physical measurements are projective. This can lead to some confusion, since naturally one may want to repeat an analogous argument to the above that would allow one to treat the measurements as general POVM measurements, and simply make use of a Naimark dilation for mathematical convenience. We elaborate on this difficulty in doing this below.

Let us drop the projective assumption, so that the physical measurements are in general POVM measurements. Imagine we aim to self-test a set of measurements $M'_{\alpha|x}$ for Alice and the reference state $|\psi'\rangle$ in the sense of definition 2. This immediately poses a problem for the majority of self-testing proofs. For example in (30), one needs the physical measurements to be projective to guarantee $A_2^2 = 1$ in order to prove anti-commutativity of the observables. In order to proceed to use the standard proofs of measurement self-testing, of which almost all assume the projective nature of the measurements, one has three options:

1. Prove that even if treating the physical measurements as POVM measurements, the only physical measurements that are compatible with the observed correlations are projective measurements.
2. Prove a general theorem that states that if one has an isometry mapping any Naimark dilation of the physical measurements to the reference measurements, this implies a (possibly different) isometry mapping the POVM physical measurements to the reference measurements in the sense of definition 2.

3. Argue that projective measurements are the only fundamental measurements in quantum theory. That is, POVM measurements can only be physically realised via a projective measurement on a dilated space.

Option 1 has been done very rarely; perhaps the only example can be found in [Kan16] when self-testing qubit Pauli measurements. Option 2 would be an analogue for the ability to use a purification of the state as explained above. However, such a theorem has not been proven to the best of our knowledge, and my may not be possible. We comment on option 3 at the end of this section.

Let us focus further on option 2. Define the physical measurements for Alice as usual by $M_{a|x}$, which may now be POVM. Furthermore, define a Naimark dilation of these measurements by $\hat{M}_{a|x}$. The vast majority of self-testing works prove an isometry mapping the dilated measurements $\hat{M}_{a|x}$ to the reference measurements:

$$\Phi \left[ \hat{M}_{a|x} \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{0}\rangle \right] = M_{a'|x} \otimes \mathbb{1}_{B'} |\psi'\rangle \otimes |\xi\rangle ,$$

(76)

where we have explicitly written the ancilla state $|\tilde{0}\rangle \in \mathcal{H}_A$ used for the dilation, and the purification space of the state is left implicit as in section 4. Now consider the physical (potentially POVM) measurements acting on the physical state.

$$M_{a|x} \otimes \mathbb{1}_B |\psi\rangle .$$

(77)

At this point, it is tempting to define an isometry $\Omega$ that maps this measurement to it’s Naimark dilation:

$$\Omega [M_{a|x} \otimes \mathbb{1}_B |\psi\rangle] = \hat{M}_{a|x} \otimes \mathbb{1}_{B'} |\psi\rangle \otimes |\tilde{0}\rangle ,$$

(78)

One could then use the standard proof of self-testing by concatenating isometries. Note, however, that fixing an input $x$, the vectors on the right hand side of (78) are orthogonal for different $a$ since for $a \neq a'$

$$\langle \psi \otimes |\tilde{0}\rangle |M_{a'|x} \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{0}\rangle \rangle = \langle \psi \otimes |\tilde{0}\rangle |\hat{M}_{a'|x} \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{0}\rangle \rangle = 0 .$$

(79)

The corresponding inner product between the vectors inside the isometry in (78) is

$$\langle \psi | M_{a'|x} \otimes \mathbb{1}_B |\psi\rangle ,$$

(80)

which is generally not equal to zero if the measurements are POVM. Thus, (78) is generally not valid, since isometry maps conserve the inner product between vectors. This means that one cannot simply consider a Naimark dilation ‘for free’ since the map which takes the POVM measurements to the dilation is not an isometry.

Nevertheless, one may hope to prove that given a self-test of the dilated measurements (76), one could recover the corresponding map for the physical measurements by somehow discarding the ancilla degrees of freedom used in the dilation, as was done for the purification space of the state. The difficulty here however is that—unlike the purification space—the dilated space is explicitly used by the isometry, since it is constructed from the dilated measurements.

Consider a general state $\rho$ on $\mathcal{H}_A$. For any Naimark dilation $\tilde{M}_{a|x}$ of $M_{a|x}$ one has

$$\text{tr}[\tilde{M}_{a|x} \rho \otimes |\tilde{0}\rangle \langle \tilde{0}|] = \text{tr}[M_{a|x} \rho] .$$

(81)

Tracing over the ancilla space only we have

$$\text{tr}[\mathbb{1} \otimes |\tilde{0}\rangle \langle \tilde{0}| |M_{a|x} \mathbb{1} \otimes |\tilde{0}\rangle \rho] = \text{tr}[M_{a|x} \rho]$$

(82)

and since this holds for all $\rho$ we must have

$$M_{a|x} = \mathbb{1} \otimes |\tilde{0}\rangle \langle \tilde{0}| \tilde{M}_{a|x} \mathbb{1} \otimes |\tilde{0}\rangle .$$

(83)

Returning to (76), we insert the identity $\mathbb{1}_A \otimes \sum_k |\tilde{k}\rangle \langle \tilde{k}|_A$ inside the isometry, giving

$$\Phi \left[ (\mathbb{1}_A \otimes \sum_k |\tilde{k}\rangle \langle \tilde{k}|_A ) \tilde{M}_{a|x} \mathbb{1}_B |\psi\rangle \otimes |\tilde{0}\rangle \right]$$

$$= M'_{a|x} \otimes \mathbb{1}_{B'} |\psi'\rangle \otimes |\xi\rangle .$$

(84)

Taking the sum outside we find

$$\sum_k \Phi \left[ \tilde{M}^k_{a|x} \otimes \mathbb{1}_A \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{k}\rangle \right]$$

$$= M'_{a|x} \otimes \mathbb{1}_{B'} |\psi'\rangle \otimes |\xi\rangle .$$

(85)

Where

$$\tilde{M}^k_{a|x} = \mathbb{1}_A \otimes |\tilde{k}\rangle (\tilde{M}_{a|x}) \mathbb{1}_A \otimes |\tilde{0}\rangle .$$

(86)

From (83) we have $\tilde{M}^0_{a|x} = M_{a|x}$ and so

$$\Phi \left[ M_{a|x} \otimes \mathbb{1}_A \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{0}\rangle \right]$$

$$= M'_{a|x} \otimes \mathbb{1}_B |\psi'\rangle \otimes |\xi\rangle - \sum_k \Phi \left[ \tilde{M}^k_{a|x} \otimes \mathbb{1}_B |\psi\rangle \otimes |\tilde{k}\rangle \right].$$

(87)
If the sum in the right hand side of the above is zero then we have proven that the isometry $\Phi$ indeed maps the physical measurements to the reference measurements. However, it is not clear that there exists a Naimark extension such that this is always the case. Showing whether this is or is not possible would be a valuable contribution to the field.

Finally we comment on option 3. This position can be justified on the basis that all other operations in quantum theory (CPTP maps, non-projective measurements) can be seen as the result of combining unitary evolution and projective measurement, and so there is no logical problem that arises with this stance. For some, this argument is not convincing however, since they consider POVMs to be as ‘real’ as projective measurements, and thus deserve a place in the ontology of the theory.

An alternative way to argue the projective-only assumption is from a information theoretic perspective. More precisely, the measurement update rule in quantum theory can be understood as a process by which an observer updates their description of a quantum state given new information. This new information is given by the outcome of a measurement, and is classical information from the perspective of the observer (i.e. it is the classical information that the observer reads from the macroscopic degrees of freedom of the measurement device). Thus it is onto these degrees of freedom (which may have become entangled with the quantum system during the measurement procedure) that the measurement operators act. Since these degrees of freedom are perfectly distinguishable, the measurement operators are described by orthogonal projectors. If one performs another measurement, this amounts to applying a different unitary Schrodinger evolution to the state before observing the measurement procedure, thus mapping the measurement to a different projective measurement. The POVM update rule can then be understood as the effective measurement operator applied to the state that is induced by the observation of this classical information. In this sense, any measurement operator—which by definition describes the effect of obtaining classical information—is described by a projector.

References


[AFY18] Rotem Arnon-Friedman and Henry Yuen. Noise-Tolerant Testing of High Entanglement of Formation. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Samella, editors, 45th International Colloquium on Automata, Languages, and Pro-


[CGS17] Andrea Coladangelo, Koon Tong Goh, and Valerio Scarani. All pure bipartite entangled states can be self-tested. Nature Communications, 8:15485, may 2017. doi:10.1038/ncomms15485.


[WLP13] Erik Woodhead, Charles Ci Wen Lim, and Stefano Pironio. Semi-device-independent QKD based on BB84 and a CHSH-type estimation. In Kazuo Iwama, Yasuhiro Kawano,


