

---

# Complex System and Untrusted Device Certification from Bell's Inequality

---

## Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät  
der Universität Basel

von

**Sebastian Wagner**

Basel, 2019

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel

<http://edoc.unibas.ch>



This work is licensed under a Creative Commons  
Attribution-NonCommercial-NoDerivatives 4.0 International License.

The complete text may be viewed here:

<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät  
auf Antrag von

Prof. Dr. Nicolas Sangouard

Prof. Dr. Philipp Treutlein

Prof. Dr. Jędrzej Kaniewski

Basel, den 19. Februar 2019

Prof. Dr. Martin Spiess  
Dekan

# Acknowledgements

Before we start the main part of this thesis, I would like to dedicate a few lines to the people who accompanied and supported me over the last three years.

First of all I thank Nicolas Sangouard for giving me the opportunity to work in his group, proposing me to work on interesting subjects and of course for his excellent supervision. The advice he has given me regarding quantum physics and life in general is invaluable and I appreciate it greatly. With respect to this thesis, I thank him for his many helpful remarks and suggestions.

Secondly I thank Philipp Treutlein for agreeing to be my second supervisor and for being a member of the committee.

I also thank Jędrzej Kaniewski for agreeing to be the external referee. I highly appreciate the time he sacrifices for reading this manuscript and for attending the defence.

A big thank-you is devoted to the other members of the group, namely to Azadeh, Enky, Jean-Daniel, Melvyn, Pavel and Xavier. Their friendly, humorous and supportive nature made it very enjoyable to work in this group. I also enjoyed our lunch discussions on quantum healing, the supremacy of the French croissant and many essential topics more.

Among the members of the group, I would especially like to thank Jean-Daniel Bancal and Pavel Sekatski. Jean-Daniel already supervised me during my Masters studies and taught me to treat my results critically. Pavel joined the group during my PhD and introduced me to various interesting mathematical concepts. I was fortunate enough to work with these two brilliant researchers on many projects and it is due to them and Nicolas that I am in the position to finish my PhD. They taught me many things about quantum mechanics, mathematics and programming.

Last but not least I want to thank my family and my girlfriend. They did not always find me in a good mood but nevertheless always encouraged and supported me. I also thank Ella Fitzgerald and Frank Sinatra for providing the soundtrack to my studies.



# CONTENTS

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>  | <b>7</b>  |
| 1.1      | Context . . . . .  | 7         |
| 1.2      | Bell's Inequality . . . . .  | 10        |
| 1.2.1    | Local Models . . . . .   | 10        |
| 1.2.2    | The Clauser-Horne-Shimony-Holt Inequality . . . . .                    | 10        |
| 1.3      | Finding Bell Inequalities . . . . .                                    | 12        |
| 1.4      | Self-Testing . . . . .   | 13        |
| 1.4.1    | Definition . . . . .   | 13        |
| 1.4.2    | Example: Self-Testing of a Bell State . . . . .                        | 15        |
| 1.5      | Outline . . . . .  | 16        |
| <b>2</b> | <b>Bell-Correlations in Large Systems</b>                              | <b>19</b> |
| 2.1      | Multipartite Bell Inequalities . . . . .                               | 19        |
| 2.2      | Bell Witness . . . . .   | 19        |
| 2.3      | Finite Statistics . . . . .  | 20        |
| <b>3</b> | <b>Variational Method for Tailoring Bell Inequalities</b>              | <b>41</b> |
| 3.1      | Eigenvalue Perturbation . . . . .                                      | 41        |
| 3.2      | A New Self-testing Inequality for Partially-Entangled States . . . . . | 43        |
| <b>4</b> | <b>Self-Testing of Quantum Channels</b>                                | <b>49</b> |
| <b>5</b> | <b>Self-Testing of Quantum Measurements</b>                            | <b>69</b> |
| <b>6</b> | <b>Conclusions and Perspectives</b>                                    | <b>79</b> |
| 6.1      | Conclusions . . . . .  | 79        |
| 6.2      | Perspectives . . . . .   | 81        |
| <b>7</b> | <b>References</b>  | <b>85</b> |

|   |           |
|---|-----------|
| <b>8 Appendix</b>   | <b>89</b> |
| 8.1 Dephasing Channels for "Non-Orthogonal Operators" . . . . . | 89        |

## INTRODUCTION

### 1.1 Context

---

In physics, we use fundamental theories to describe and explain phenomena occurring in nature. Two of the most prominent theories are classical mechanics, pioneered in the 17th century by Sir Isaac Newton, and quantum mechanics which arose in the beginning of the 20th century. While classical and quantum mechanics dissent in various aspects, the most pronounced difference is called *entanglement*.

Entanglement is a purely quantum mechanical phenomenon and hence does not have a classical counterpart. We characterise it by defining what it means for a state to be separable: A state collectively describing two or more subsystems is called separable if all the information is encoded in the subsystems individually. Mathematically this means that a state  $\rho$  acting on a Hilbert space  $\mathcal{H} = \bigotimes_k \mathcal{H}_k$  is denoted separable if and only if one can write the state as  $\rho = \sum_i p_i \bigotimes_k \rho_k^i$ , where  $\rho_k^i$  acts on  $\mathcal{H}_k$  and  $\sum_i p_i = 1$ . If the state is not separable we refer to it as being entangled. This means that if two systems, e.g. two photons, are entangled their state is more than the product of individual states; we have to describe their state as a whole. This implies that the measurement outcomes for experiments on the two entangled systems - irrespective of their spatial separation - can be correlated with potentially stronger correlations than the ones observed with separable states.

One might think entanglement is a weak point because it complicates the description, but on the contrary, entanglement provides us with powerful tools that allow us to perform tasks which are not possible by classical means [1]. Quantum computing studies the possible use of quantum principles for computational tasks. For example the factorisation of large numbers is infeasible with classical computers but can be done efficiently with Shor's algorithm [2]. The power of quantum computers can be intuitively understood by realizing that its basic unit - called qubit (quantum bit) and being a state in  $\mathbb{C}^2$  - can encode infinitely many states via the

superposition principle.

Another significant application of quantum mechanics is quantum key distribution (QKD) [3, 4, 5]. The goal is to create two identical strings of random bits, called a key, at two spatially-separated locations. This key is then used to encrypt a message allowing for secret communication. Secret communication is essential in our modern society, in which we use the internet to manage our bank account, buy products in online stores and send personal messages to friends. We communicate over great distances and desire that this communication be secure. The current classical cryptographic protocols use complex mathematical problems such as factorization to create secure keys. The downside of this procedure is that the security is based on the complexity of mathematical problems and relies on assumptions about the computational power of the person who wants to break the cryptographic system. Hence there is every chance that a potential eavesdropper hacks the key - especially if he has access to quantum computers. On the other hand, a key can be obtained by performing appropriate measurements on an entangled state. This provides the means to actually create secret keys with provable security.

In order to achieve long-distance QKD, we envision quantum networks whose purpose is to transmit entanglement between two arbitrary parties on earth. A network consists of various quantum mechanical devices, including sources for creating quantum information, memories which allow for the storage of it, as well as quantum gates and projective measurements for processing the information.

Quantum networks and quantum computers sound very appealing. However, with the benefits of quantum mechanics there also come great challenges. A central challenge we want to tackle in this thesis is how to certify that one indeed works with quantum mechanical devices. The subtlety here lies in the fact that we humans are classical and thus cannot directly observe quantum features such as entanglement. As a consequence, we desire to employ certification schemes which do not overburden our classical competences. The need for such certifications becomes apparent when considering the following scenario: Basic quantum machines are already available commercially, for example true random number generators [6, 7]. If we purchase such a device that promises to prepare entangled states or act as a quantum gate, we aim at verifying that the promise actually holds. We want to do this without breaking or opening the device since we would lose the warranty or anyway be overstrained by the complexity of the physics involved. At best, the certification should be such that even an unqualified user can conduct it. In this thesis we will discuss how this can be achieved.

Every certification requires assumptions. Therefore we dedicate a small paragraph to the role assumptions play in every scientific work. First of all, we would like to emphasise that assumptions are crucial and necessary, for a lack of presumptions may result in the problem being infeasible. On the other hand, this might lead to one being tempted to circumnavigate theoretical and/or experimental issues by simply adding presumptions that handle these issues. In this sense, an excess of assumptions reduces the significance of the results. What is more, in case of invalid assumptions, this might even lead to false-positive results. It is thus advisable to aim for a minimal set of assumptions.

Following the previous discussion, there are multiple ways to approach the challenge of certification. One way is - as is the case in many device-dependent certification schemes - to simply trust the measurements by assuming one has full knowledge about them. This assumes that the measurements are perfectly calibrated. Trusting the measurements then allows one to do tomography on the state. But what happens if the measurements are not perfectly calibrated? Rosset *et al.* showed that even slight misconfigurations in the settings can result in false-positive results when one wants to witness entanglement [8]. Another device-dependent approach is to include assumptions about the dimensions of the tested state. While this might seem a reasonable and harmless presumption, Acín *et al.* showed how carelessly utilizing a witness for state certifications may lead to false-positive certification outcomes [9]: In the space of two-qubit states and measurements, the inequality

$$\langle \sigma_x^A \otimes \sigma_x^B \rangle + \langle \sigma_z^A \otimes \sigma_z^B \rangle \leq 1 \quad (1.1)$$

holds for any separable state and is maximally violated uniquely by the maximally-entangled state  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . If one cannot ensure the Hilbert space dimension, however, then the measurements could be  $\sigma_x^A = \sigma_x^{A_1} \otimes \mathbb{1}_2^{A_2}$ ,  $\sigma_z^A = \mathbb{1}_2^{A_1} \otimes \sigma_z^{A_2}$ , and analogously for the second party. But then the product state  $|+, 0\rangle_A \otimes |+, 0\rangle_B$  also achieves a maximal witness value of 2. What is more, with a separable state one can also achieve a value of 2 and even reproduce the marginal probabilities [10]. The correlations measured in Eq. (1.1) are the ones relevant for the BB84 QKD protocol [3] and the example before was used to show that it is secure only under the assumption that qubits are used. This is because a separable state allows a potential eavesdropper to create a copy of the state shared by Alice and Bob. The eavesdropper thus also shares their information. These two examples related to calibration of the measurements and dimension of the tested state emphasise the importance of a careful choice and handling of assumptions.

In our opinion, the preferred path to certification therefore must be a device-independent one not involving assumptions on the proper calibration of the measurements or on the dimension of the tested state. We only want to trust our capabilities to collect and process statistics. Based on these statistics, the aim is to be able to make qualitative as well as quantitative statements about states and measurements. For certain entangled states and measurements this is indeed possible. We demonstrate in section 1.4 how to certify device-independently a two-qubit maximally-entangled state.

All our certification schemes are based on Bell's Theorem. The next section is dedicated to this cornerstone of modern quantum physics.

---

## 1.2 Bell's Inequality

---

Einstein, Podolsky and Rosen (EPR) found the correlations in the outcomes of measurements on entangled states not to be compatible with their understanding and expectations of a reasonable physical theory. They concluded that quantum mechanics must be incomplete and proposed a refined theory of local *hidden variables* [11]. In this theory, the assumption is that the individual subsystems share locally hidden parameters which cannot be accessed experimentally but which determine the measurement results deterministically. However, the local hidden variable approach proved itself to be wrong.

### 1.2.1 Local Models

The cornerstone for the insight that some phenomena can only be described non-locally was set by the famous physicist John Bell. In his reply to the theory proposed by EPR, he derived an inequality which sets a constraint on scenarios that allow for a local description [12]. From this point on in the thesis, every inequality of this form is denoted a Bell inequality. There is experimental evidence that nature violates such inequalities [13, 14, 15, 16] implying that it indeed requires non-local theories to describe our world - at least on an atomic scale.

Let us briefly introduce the concept of locality by considering the following scenario. Two parties share a pair of particles which they measure using the settings  $X$  and  $Y$  to get outcomes  $a$  and  $b$ , respectively. In the hidden-variable picture, the particles share a program that determines the measurement results prior to the separation of the particles. This program is represented by the classical variable  $\lambda$ . In this description, the observed correlations are determined locally in the sense that joint probabilities are expressed in terms of marginal probabilities via

$$P(a, b|X, Y) = \int d\lambda p(\lambda)p(a|X, \lambda)p(b|Y, \lambda) , \quad (1.2)$$

where  $p(\lambda)$  is the (unknown) probability distribution of the classical variable. This means that on Alice's side, the outcome  $a$  depends on the choice of setting  $X$  and the shared program  $\lambda$  but it does not depend on Bob's measurement choices. This type of correlations is called local in the sense of Bell. When testing whether certain observed correlations can be explained by a local program  $\lambda$ , we test if they admit a decomposition of the form (1.2). If no decomposition of this form exists, we say that the correlations are non-local.

Rather than explaining relevant features of nonlocality with the help of the original inequality by Bell, we will now introduce the very well-known Clauser-Horne-Shimony-Holt inequality.

### 1.2.2 The Clauser-Horne-Shimony-Holt Inequality

The most relevant Bell inequality to this date clearly is the one derived by Clauser, Horne, Shimony and Holt (CHSH) [17]. Because of its relevance in quantum science, we will quickly

summarise the result. This also serves the purpose of introducing the notation that is used throughout this thesis.

Two parties denoted *Alice* and *Bob* individually perform measurements on a joint system. In the CHSH scenario, Alice can choose between the measurements  $X_0$  and  $X_1$  with *binary* outcomes  $a_0, a_1 = \pm 1$ . An analogous statement holds for Bob. Under these conditions, it then follows that any model allowing for a local program has to fulfil the CHSH inequality

$$\beta = \langle X_0 Y_0 \rangle + \langle X_0 Y_1 \rangle + \langle X_1 Y_0 \rangle - \langle X_1 Y_1 \rangle \leq \beta_L = 2, \quad (1.3)$$

where  $\langle \cdot \rangle$  is the expectation value of outcomes, that is the difference between the probabilities of correlated and anti-correlated results, i.e.  $\langle X_i Y_j \rangle = p(a=b|X_i, Y_j) - p(a \neq b|X_i, Y_j)$ . We denote  $\beta_L = 2$  the *local bound*, that is, the maximum obtained with correlations of the form (1.2). We will not go through the proof in detail (we direct an interested reader to Ref. [18]). However, there is an important fact when dealing with local theories that is worth mentioning. Since the set of local strategies is a convex polytope in the space of probability distributions, the extrema correspond to deterministic strategies. A deterministic strategy is characterised by the fact that each measurement setting only allows for one outcome, that is the outcome appears with probability one. The local bound can thus be obtained by computing the value of  $\beta$  for these  $2^4 = 16$  different strategies.

In the quantum formalism, the violation of a Bell inequality highlights the presence of entanglement. The reason for this is that separability provides a local model and consequently separable states can at best achieve the local bound. But we can conclude even more:

A violation of a Bell inequality indicates the presence of a non-local state, also referred to as a Bell-correlated state. Bell correlation is a stronger form of quantum correlation than entanglement in the sense that the set of Bell-correlated states is a subset of entangled states. This means that any non-local state is entangled, but that there are certain states which are entangled but do not violate any Bell-type inequality [19].

In case of a maximal violation, we can conclude even more than just entanglement and Bell correlation. Perfect statistics allow us to identify the state and the corresponding measurements. This is discussed in detail in section 1.4. Before we will define remaining terminology that is used in the thesis.

For a joint system of two quantum bits, i.e. in the Hilbert space  $\mathbb{C}^2 \otimes \mathbb{C}^2$ , there are four commonly used states which achieve a maximal CHSH-violation. We thus refer to them as being maximally entangled, and denote them *Bell states*. They are

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), \quad |\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle), \quad (1.4)$$

where we used the Dirac notation of states and the convention  $|ab\rangle = |a\rangle_A \otimes |b\rangle_B$ . Additionally, the states  $|0\rangle$  and  $|1\rangle$  are the (+1)- and (-1)-eigenstates of the Pauli matrix  $\sigma_z$ , the three Pauli matrices being defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (1.5)$$

The identity matrix was added for completeness. The corresponding eigenbases are  $\{|+\rangle, |-\rangle\}$ ,  $\{|R\rangle, |L\rangle\}$  and  $\{|0\rangle, |1\rangle\}$ . The  $z$ -basis is also called the computational basis because of its resemblance to the classical bits 0 and 1. At this point we would also like to introduce the term *Bell operator*. A Bell operator is the quantum mechanical equivalent of a Bell inequality. One arrives at it by identifying the measurement variables with operators such as the ones of Eq. (1.5). The two-qubit CHSH operator can be written as

$$\mathcal{B}_{\text{CHSH}} = \hat{X}_0 \otimes (\hat{Y}_0 + \hat{Y}_1) + \hat{X}_1 \otimes (\hat{Y}_0 - \hat{Y}_1), \quad (1.6)$$

with  $\hat{X}_i = \mathbf{x}_i \cdot \boldsymbol{\sigma}$  and  $\hat{Y}_i = \mathbf{y}_i \cdot \boldsymbol{\sigma}$ , where  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  is the vector of Pauli matrices. A state  $\rho$  then achieves the Bell value  $\beta = \text{Tr}(\mathcal{B}_{\text{CHSH}}\rho)$ .

The Bell states of Eq. (1.4) together with the appropriate measurements achieve a maximal CHSH-value of  $2\sqrt{2}$ , denoted the *quantum bound* [20]. As hinted at earlier, an experimental observation of the quantum bound allows us to conclude that the state at hand is one of the Bell states. This is referred to as self-testing and introduced in section 1.4. Also, in the imperfect case quantitative statements about the state and measurement fidelities can be made.

### 1.3 Finding Bell Inequalities

---

Once we have a Bell inequality we can use it for many interesting things such as state certification or the detection of nonlocality. However, we often meet the problem of first having to derive new Bell inequalities. In case one wants to certify a state which does not maximally violate any known Bell inequality, one has to tailor a new Bell inequality to this state. A method serving this purpose is introduced in chapter 3. Also in other branches the derivation of new Bell inequalities is a relevant issue - for example when aiming at detecting quantum features in large systems (chapter 2) for which only few inequalities are known.

Once one has a promising candidate for a Bell inequality, there are different tasks to conduct. The most difficult part is computing the local bound. This can be done in multiple ways, one path being the brute-force method of simply trying out all deterministic strategies. This approach, however, is not suitable for systems involving many parties and/or many measurement settings per party and/or many possible measurement outcomes. This is because the number of deterministic strategies for a system of  $N$  parties,  $m$  measurement settings and  $\ell$  possible outcomes per measurement is  $\ell^{mN}$ , which in the simplest many-body scenario for  $m = \ell = 2$  is larger than  $10^{30}$  for  $N \geq 100$ . Computing Bell values for such a number of strategies is infeasible. In cases like this, one computes the local bound for small values of  $N$  that do not overburden our computational competences. Then one tries to find a pattern and generalize the findings to the arbitrary- $N$  case. Often, we also simplify our task of finding Bell inequalities for example by restricting the search to those inequalities that are symmetric under exchange

of parties or by only considering few-body correlators. A suitable approach for multipartite scenarios is discussed in chapter 2.

## 1.4 Self-Testing

### 1.4.1 Definition

As was already mentioned in chapter 1, assumptions play a crucial part in every scientific study. An excess of assumptions (for example on the dimension of the state and measurements) is detrimental for the significance of the result. We therefore strive to minimize the number of assumptions. Self-testing [21] allows us to certify states and measurements from the statistics only and without assumptions on the dimension of the Hilbert space or the calibration of the measurements.

Note that with self-testing, we can only certify states and measurements up to local isometries [22]: Firstly, it is impossible to deduce from the statistics the dimension of the actual state and measurements, because in principle there can always be dimensions of the state on which the measurements act trivially:

$$\mathrm{Tr}[(M_A \otimes M_B)\rho_{AB}] = \mathrm{Tr}[(M_A \otimes \mathbb{1}_{A'} \otimes M_B \otimes \mathbb{1}_{B'})\rho_{AB} \otimes \rho_{A'B'}], \quad (1.7)$$

where  $M_A$  and  $M_B$  are the local measurements of Alice and Bob. In addition to this dimensional equivalence class, local unitaries cannot be detected:

$$\begin{aligned} \mathrm{Tr}[(M_A \otimes M_B)\rho_{AB}] &= \mathrm{Tr}[(M'_A \otimes M'_B)\rho'_{AB}], \quad \text{where} \\ M'_A &= U_A M_A U_A^\dagger, \quad M'_B = U_B M_B U_B^\dagger, \quad \rho'_{AB} = (U_A \otimes U_B)\rho_{AB}(U_A^\dagger \otimes U_B^\dagger), \end{aligned} \quad (1.8)$$

with  $U_A$  and  $U_B$  local unitaries.

The fact that we can only certify states and measurements up to these local isometries is not an issue. The two classes of isometries do not affect the feature we are interested in - nonlocality. The first class implies that Alice and Bob share the state  $\rho_{AB}$  in some subsystems; where exactly we do not know, but we also do not care since we can be sure that also the measurements act in these subsystems as modelled by  $M_A$  and  $M_B$ . The second class of isometries simply refers to a local change of frame. This leads us to the formal definition of self-testing.

**Definition 1** *Self-Testing* – We say that a Bell violation  $\beta$  self-tests the  $N$ -partite state  $\bar{\rho}$  if the violation implies the existence of local extraction channels  $\Phi_i$ ,  $i = 1, \dots, N$  extracting the target state  $\bar{\rho}$  from the experimental state  $\rho$ :

$$\exists \Phi_i : \bigotimes_{i=1}^N \Phi_i[\rho] = \bar{\rho} \otimes \rho_{ext},$$

where  $\rho_{ext}$  is an irrelevant state also denoted junk or extra state.

Analogously we say that a Bell violation self-tests the measurements  $\overline{M}_i$ ,  $i = 1, \dots, N$  if the **same** extraction channels  $\Phi := \bigotimes_{i=1}^N \Phi_i$  achieve

$$\Phi \left[ \left( \bigotimes_{i=1}^N M_i \right) \rho \left( \bigotimes_{i=1}^N M_i \right) \right] = \left( \bigotimes_{i=1}^N \overline{M}_i \right) \overline{\rho} \left( \bigotimes_{i=1}^N \overline{M}_i \right) \otimes \rho_{ext}.$$

This is remarkable. It means that a single number allows for conclusions about the state and the measurements. Of course, in practice this single number needs many experimental runs to assess and what is more, we would like to emphasise that perfect statistics can never be observed. Hence the above definition is not useful in practice, implying the need for a measure for the distance of the experimental state  $\rho$  to our target state  $\overline{\rho}$ .

**Definition 2** *Overlap of States* – The overlap of an  $N$ -party state  $\rho$  extracted by the channels  $\Lambda_i$ ,  $i = 1, \dots, N$  with a target state  $\overline{\rho}$  is

$$\mathcal{O}(\bigotimes_i \Lambda_i[\rho], \overline{\rho}) = \text{Tr} \left( \bigotimes_i \Lambda_i[\rho] \overline{\rho} \right).$$

The extraction channels are completely positive trace-preserving (CPTP) maps also tracing out irrelevant subsystems, that is  $\Lambda[\rho] = \text{Tr}_{ext}(\Phi[\rho])$ .

The overlap can be directly interpreted as the self-testing *fidelity* of the experimental state  $\rho$  with the target state  $\overline{\rho}$  [23]. In this thesis, we also use a different definition of the fidelity between states, namely the Uhlmann fidelity.

**Definition 3** *Uhlmann fidelity* – The Uhlmann fidelity of two states  $\rho$  and  $\sigma$  is given by

$$F(\rho, \sigma) = \left[ \text{Tr} \left( \sqrt{\sqrt{\sigma} \rho \sqrt{\sigma}} \right) \right]. \quad (1.9)$$

**Lemma 1** *For pure states the Uhlmann and self-testing fidelities are related via  $F(\rho, \sigma) = \sqrt{\mathcal{O}(\rho, \sigma)}$ .*

This is shown in the appendix of our paper on the certification of building blocks of quantum computers (chapter 4).

We would like to emphasise that, in the generic case, self-testing is an immensely difficult problem because the dimensional freedom allows for an infinite set of parameters to optimize over. The task of self-testing hence is not straightforward and may not be executable in general. There is, however, a scenario in which we can reduce the dimensional complexity. Because of its importance to the approach we follow, we will briefly introduce it here. If the Bell test involves at most *two binary* measurements per party, we can make use of Jordan's lemma [22]:

**Lemma 2** *Jordan's Lemma* – Let  $X$  and  $Z$  be two Hermitian operators with eigenvalues  $-1$  and  $+1$ . Then there exists a basis in which both operators are block-diagonal, with blocks of dimension  $2 \times 2$  at most.

This directly implies that if a Bell operator  $\mathcal{B}$  corresponding to the Bell test consists of at most two binary observables per party then it can be written as

$$\mathcal{B} = \bigoplus_{\alpha_1} \bigoplus_{\alpha_2} \dots \bigoplus_{\alpha_N} \mathcal{B}_{\alpha_1 \dots \alpha_N}. \quad (1.10)$$

In this expression,  $\mathcal{B}_{\alpha_1 \dots \alpha_N}$  is an  $N$ -qubit Bell operator, and the indices  $\alpha_1, \dots, \alpha_N$  denote the block of parties  $1, \dots, N$ , respectively. This allows one to reduce the device-independent certification of  $N$ -qubit states to an  $N$ -qubit problem.

Apart from the dimensionality, a difficult problem is related to the extraction channels. These are also unknown, so in order to achieve the best fidelity, we would have to optimize over all possible CPTP maps, that is we would like to get  $\left[ \max_{\Lambda_i} \min_{\rho} F(\bigotimes_i \Lambda_i[\rho], \bar{\rho}) \right]$ . However, most of the time we content ourselves with lower bounds on the state fidelity. In this way, we can simply fix the extraction channels beforehand and interpret the resulting fidelities as relevant lower bounds.

#### 1.4.2 Example: Self-Testing of a Bell State

A prominent example for the self-testing of states and measurements is provided by the CHSH-test. The maximal CHSH-violation self-tests the maximally-entangled two-qubit state  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  up to local isometries. As a reminder, the CHSH inequality is given as  $\langle X_0(Y_0 + Y_1) + X_1(Y_0 - Y_1) \rangle \leq 2$  (see also Eq. (1.3)). Quantum mechanics allows for the violation of the inequality and the maximal quantum violation of  $2\sqrt{2}$  can be achieved by the Bell states of Eq. (1.4).  $|\phi^+\rangle$  achieves the quantum bound when choosing the measurements

$$\hat{X}_0 = \sigma_x, \quad \hat{X}_1 = \sigma_z, \quad \hat{Y}_0 = \frac{\sigma_x + \sigma_z}{\sqrt{2}}, \quad \hat{Y}_1 = \frac{\sigma_x - \sigma_z}{\sqrt{2}}.$$

As is nicely explained in Ref. [22], a perfect violation implies the existence of extraction channels filtering out the target state  $|\phi^+\rangle$  as well as identifying the subspaces in which the experimental measurements act on  $|\phi^+\rangle$  as modelled by the observables listed above. This can also be seen by conducting the following semi-definite program (SDP):

$$\begin{aligned} \min_{\rho, a, b} \{ \mathcal{O}[(\Lambda_a \otimes \Lambda_b)[\rho], |\phi^+\rangle\langle\phi^+|] \} \quad & s.t. \quad \text{Tr}(\mathcal{B}_{\text{CHSH}}\rho) = 2\sqrt{2}, \\ & \rho \geq 0, \\ & \text{Tr}(\rho) = 1, \\ & \rho^\dagger = \rho, \end{aligned} \quad (1.11)$$

where the extraction channels  $\Lambda_a$  and  $\Lambda_b$  are the ones of Ref. [23]. Only the state  $\rho = |\phi^+\rangle\langle\phi^+|$  satisfies all constraints and hence the target state is certified. Note that we only certify states up to local isometries. Since the four Bell states are LU-equivalent, this means a CHSH-value of  $2\sqrt{2}$  "only" certifies that the state is maximally entangled. Also, in case of imperfect statistics, i.e.  $\text{Tr}(\mathcal{B}_{\text{CHSH}}\rho) = \beta < 2\sqrt{2}$ , one can conclude about the fidelity of the tested state with respect to the Bell state. Indeed, one observes that the certification is very robust to noise in the sense that even for small violations  $\beta \gtrsim 2.11$ , the resulting lower bound on the fidelity is larger than the trivial fidelity of 0.5 [23].

The self-testing of states is the essential ingredient in our works on the device-independent certification of building blocks of quantum computers (chapter 4) and the self-testing of quantum measurements (chapter 5).

## 1.5 Outline

---

The aim of this thesis is to show device-independent certifications beyond two-qubit maximally-entangled states and projective qubit measurements.

In chapter 2 we derive a new class of Bell inequalities involving an arbitrary number of parties and measurement settings. These inequalities are symmetric under exchange of parties and only involve one- and two-body correlators. From these inequalities we derive witnesses whose purpose is to detect Bell correlation in large systems of spin- $\frac{1}{2}$  particles. We use the data of previous experiments to conclude about Bell correlation in a Bose-Einstein condensate of about 500 Rubidium atoms. We also study the effect finite statistics have on the significance of the results and provide quantitative bounds on the number of experimental runs needed as a function of the number of parties to reach conclusive results.

In chapter 3 we introduce a new method for designing Bell inequalities that are tailored to given target states. The starting point for the approach are stabilising operators of the state of interest. We then shape a Bell operator such that its maximal eigenvalue drops rapidly when departing from the perfect settings and the target state. Our method is designed specifically for the rough conditions present in self-testing. We show that even in the well-studied case of partially-entangled two-qubit states, this approach results in new inequalities that provide unprecedented robustness to noise.

In chapter 4 the goal is to device-independently certify basic building blocks of quantum computers. These building blocks among others are quantum memories and two-qubit gates. We achieve this goal by combining state fidelities: One first certifies a maximally-entangled input state (prior to implementing the gate) and then the state in case the building block is present. This is possible because the action of a quantum gate is fully determined by its action on half a maximally-entangled state. Our certification schemes are very robust to noise.

In chapter 5 we study the device-independent certification of non-projective quantum meas-

---

urements. We first characterise a measurement by giving the corresponding Kraus operators. Then we look at the action of this measurement on a maximally-entangled state and attribute the possible outcomes to a register. Alongside each outcome, there is also a post-measurement state. We self-test the maximally-entangled input state as well as each of the possible output states. This allows us to bound the fidelity of each Kraus operator. Eventually we lower-bound the measurement fidelity by combining the individual Kraus operator fidelities.

Globally, our results represent essential steps towards the device-independent certification of complex systems and untrusted devices.



## BELL-CORRELATIONS IN LARGE SYSTEMS

### 2.1 Multipartite Bell Inequalities

---

While there exist well-studied Bell inequalities for states involving multiple parties [24, 25], the study of Bell correlations in many-body systems is still in its infancy. This is mostly due to the inadequacy of known multipartite Bell inequalities which rely on expectation values involving many parties and require an individual addressing of each party. Tura *et al.* were the first to derive Bell inequalities for systems of arbitrary size with only one- and two-body correlation functions [26]. In this chapter we will add a new class of two-body correlator Bell inequalities. This class is also suitable for systems of arbitrary size and is symmetric under exchange of parties. Additionally it allows for an arbitrary number of settings per party. We will show how they can be used to witness strong quantum correlations in many-body systems.

### 2.2 Bell Witness

---

A Bell inequality is a device-independent test to detect quantum correlations. It does not rely on assumptions neither on the underlying state nor on the measurements performed. On the other hand, the purpose of a Bell correlation witness, just as in the case of an entanglement witness, is a device-dependent one: It serves to detect strong quantum correlations with assumptions on the Hilbert space dimension and/or the functioning of the measurement device. Here we assume that our Bell inequalities are tested using projective measurements performed on a system of spin- $\frac{1}{2}$  particles. Our Bell inequalities then reduce to witnesses revealing a stronger form of correlation than entanglement, namely Bell correlation. Bell correlated states are those

violating a Bell inequality which is not a property shared by all entangled states. The Bell correlation witnesses we derive from our class of Bell inequalities allow for the detection of a larger set of states compared to previously-known witnesses.

## 2.3 Finite Statistics

---

One of the problems arising with every Bell inequality (e.g. Ineq. (1.3)) is the issue of finite statistics. The subtlety lies in the fact that statements are made about expectation values and not about outcomes of individual experimental runs. In practice it is of course infeasible to perform the infinite number of experiments required to assess the average value. Therefore it is necessary to study the consequences finite statistics have on the significance of the results. We would like to emphasise that the study of finite statistics must not be confused with error bars and/or standard deviations. Latter quantities only shed light on the accuracy of the measurements performed but not on the probability that the observed Bell violation can be reproduced by classical events.

While in few-party scenarios such as CHSH, the role of finite statistics can be secondary in some experiments, its importance increases as the size of the system grows. For the number of parties considered in this chapter, it is one of the key factors when deciding how many runs are necessary for a sufficiently-small  $p$ -value. Here, we will derive quantitative bounds lower-bounding the number of experimental runs needed for a small probability that classical events reproduce the observed violation of a Bell correlation witness. We show in particular that the bounds increase linearly with the number of parties.

# Paper No. 1

## Bell Correlations in a Many-Body System with Finite Statistics

Sebastian Wagner, Roman Schmied, Matteo Fadel, Philipp Treutlein, Nicolas Sangouard, and Jean-Daniel Bancal

Physical Review Letters, 119, 170403 (2017)



## Bell Correlations in a Many-Body System with Finite Statistics

Sebastian Wagner,<sup>1</sup> Roman Schmied,<sup>2</sup> Matteo Fadel,<sup>2</sup> Philipp Treutlein,<sup>2</sup> Nicolas Sangouard,<sup>1</sup> and Jean-Daniel Bancal<sup>1</sup>

<sup>1</sup>*Quantum Optics Theory Group, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

<sup>2</sup>*Quantum Atom Optics Lab, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

(Received 10 February 2017; published 27 October 2017)

A recent experiment reported the first violation of a Bell correlation witness in a many-body system [Science **352**, 441 (2016)]. Following discussions in this Letter, we address here the question of the statistics required to witness Bell correlated states, i.e., states violating a Bell inequality, in such experiments. We start by deriving multipartite Bell inequalities involving an arbitrary number of measurement settings, two outcomes per party and one- and two-body correlators only. Based on these inequalities, we then build up improved witnesses able to detect Bell correlated states in many-body systems using two collective measurements only. These witnesses can potentially detect Bell correlations in states with an arbitrarily low amount of spin squeezing. We then establish an upper bound on the statistics needed to convincingly conclude that a measured state is Bell correlated.

DOI: [10.1103/PhysRevLett.119.170403](https://doi.org/10.1103/PhysRevLett.119.170403)

*Introduction.*—Bell nonlocality, as revealed by the violation of a Bell inequality, constitutes one of the strongest forms of nonclassicality [1,2]. However, its demonstration has long been restricted to systems involving few particles [3–7]. Recently, the discovery of multipartite Bell inequalities that only rely on one- and two-body correlators opened up new possibilities [8]. Although these inequalities have not yet lead to the realization of a multipartite Bell test, they have been used to derive witnesses able to detect Bell correlated states, i.e., states capable of violating a Bell inequality [9,10].

These witnesses have triggered two experiments [9,11] which successfully detect the presence of Bell correlations in a many-body system under the assumption of Gaussian statistics [12,13]. The witness used in Refs. [9,11] involves one- and two-body correlation functions and takes the form  $\mathcal{W} \geq 0$ , where the inequality is satisfied by measurements on states that are not Bell correlated. Observation of a negative value for  $\mathcal{W}$  then leads to the conclusion that the measured system is Bell correlated. However, reaching such a conclusion in the presence of finite statistics requires special care [14,15]. In particular, an assessment of the probability with which a non-Bell-correlated state could be responsible for the observed data is required before concluding about the presence of Bell correlations without further assumptions.

Concretely, the witness of Refs. [9] has the property of admitting a quantum violation lower bounded by a constant  $\mathcal{W}_{\text{opt}} < 0$ , while the largest possible value  $\mathcal{W}_{\text{max}} > 0$  is achievable by a product state and increases linearly with the size of the system  $N$ . These properties imply that a small number of measurement rounds on a state of the form

$$\rho = (1 - q)|\psi\rangle\langle\psi| + q(|\uparrow\rangle\langle\uparrow|)^{\otimes N}, \quad (1)$$

where  $\mathcal{W}(|\psi\rangle) = \mathcal{W}_{\text{opt}}$ ,  $\mathcal{W}(|\uparrow\rangle^{\otimes N}) = \mathcal{W}_{\text{max}}$  and  $q$  is small, is likely to produce a negative estimate of  $\mathcal{W}$ , even though

the state is not detected by the witness in the limit of infinitely many measurement rounds [9]. This state thus imposes a lower bound on the number of measurement rounds required to exclude, through such witnesses, all non-Bell-correlated states with high confidence. Contrary to other assessments, this lower bound increases with the number of particles involved in the many-body system. Therefore, it is not captured by the standard deviation of one- and two-body correlation functions (which on the contrary decreases as the number of particles increases).

For small systems, this dependence of the number of measurement rounds on the size of the measured system merely represents a technical overhead: a conclusion may still be obtained at the price of performing few more measurements. For large systems, however, any bound on the number of measurements that can be performed imposes a hard limit on the maximal size of systems on which a reliable conclusion can be drawn. The question of statistical significance thus constitutes a fundamental question for many-body systems.

It is worth noting that states of the form (1) put similar bounds on the number of measurement rounds required to perform any hypothesis tests in a many-body system satisfying the conditions above. This includes in particular tests of entanglement [16–19] based on the entanglement witnesses of Ref. [20–22].

In this Letter, we address this statistical problem in the case of Bell correlation detection by providing a number of measurement rounds sufficient to exclude non-Bell-correlated states from an observed witness violation. Let us mention that in Refs. [9,11], this finite statistics issue is circumvented by the addition of an assumption on the set of local states being tested. This has the effect of reducing the scope of the conclusion: the data reported in Refs. [9,11], are only able to exclude a subset of all non-Bell-correlated

states (as pointed out in the references). Here, we show that such additional assumptions are not required in experiments on many-body systems, and thus argue that they should be avoided in the future.

In order to minimize the amount of statistics required to reach our conclusion, we start by investigating improved Bell correlation witnesses. For this, we first derive Bell inequalities with two-body correlators and an arbitrary number of settings. This allows us to obtain Bell correlation witnesses that are more resistant to noise compared to the one known to date [9]. We then analyse the statistical properties of these witnesses and provide an upper bound on the number of measurement rounds needed to rule out all local states in a many-body system. We show that this upper bound is linear in the number of particles, hence demonstrating the possibility of reliable detection of Bell correlations in systems with a large number of particles.

*Symmetric two-body correlator Bell inequalities with an arbitrary number of settings.*—Multipartite Bell inequalities that are symmetric under exchange of parties and which involve only one- and two-body correlators have been proposed in scenarios where each party uses two measurement settings and receives an outcome among two possible results [8]. Similar inequalities were also obtained for translationally invariant systems [23], or based on Hamiltonians [24]. Here, we derive a similar family of Bell inequalities that is invariant under arbitrary permutations of parties but allows for an arbitrary number of measurement settings per party.

Let us consider a scenario in which  $N$  parties can each perform one of  $m$  possible measurements  $M_k^{(i)}$  ( $k = 0, \dots, m-1$ ;  $i = 1, \dots, N$ ) with binary outcomes  $\pm 1$ . We write the following inequality:

$$I_{N,m} = \sum_{k=0}^{m-1} \alpha_k S_k + \frac{1}{2} \sum_{k,l} S_{kl} \geq -\beta_c, \quad (2)$$

where  $\alpha_k = m - 2k - 1$ ,  $\beta_c$  is the local bound, and the symmetrized correlators are defined as

$$S_k := \sum_{i=1}^N \langle M_k^{(i)} \rangle, \quad S_{kl} := \sum_{i \neq j} \langle M_k^{(i)} M_l^{(j)} \rangle. \quad (3)$$

Let us show that Eq. (2) is a valid Bell inequality for  $\beta_c = \lfloor (m^2 N / 2) \rfloor$ , where  $\lfloor x \rfloor$  is the largest integer smaller or equal to  $x$ . Below, we assume that  $m$  is even; see Appendix A in the Supplemental Material [25] for the case of odd  $m$ .

Since  $I_{N,m}$  is linear in the probabilities and local behaviors can be decomposed as a convex combination of deterministic local strategies, the local bound of Eq. (2) can be reached by a deterministic local strategy [1]. We thus restrict our attention to these strategies and write

$$\langle M_k^{(i)} \rangle = x_k^i = \pm 1 \quad \Rightarrow \quad S_{kl} = S_k S_l - \sum_{i=1}^N x_k^i x_l^i, \quad (4)$$

where  $x_k^i$  is the (deterministic) outcome party  $i$  produces when asked question  $k$ . This directly leads to the following decomposition:

$$I_{N,m} = \sum_{k=0}^{\frac{m}{2}-1} \alpha_k (S_k - S_{m-k-1}) + \frac{1}{2} B^2 - \frac{1}{2} C \geq -\beta_c, \quad (5)$$

with  $B := \sum_{k=0}^{m-1} S_k$  and  $C := \sum_{i=1}^N (\sum_{k=0}^{m-1} x_k^i)^2$ . Because of the symmetry under exchange of parties of this Bell expression, it is convenient to introduce, following Ref. [8], variables counting the number of parties that use a specific deterministic strategy:

$$\begin{aligned} a_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = -1 \text{ iff } k \in \{j_1, \dots, j_n\}\}, \\ \bar{a}_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = +1 \text{ iff } k \in \{j_1, \dots, j_n\}\}, \\ n &\leq \frac{m}{2}, \quad \bar{a}_{j_1, \dots, j_{\frac{m}{2}}} \equiv 0, \end{aligned} \quad (6)$$

where  $\#$  denotes the set cardinality. Since each party has to choose a strategy, the variables sum up to  $N$ :

$$\sum_{\text{all variables}} = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1, \dots, j_n} + \bar{a}_{j_1, \dots, j_n}) = N. \quad (7)$$

The correlators can now be expressed as

$$S_k = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1, \dots, j_n} - \bar{a}_{j_1, \dots, j_n}) y_k^{j_1, \dots, j_n}, \quad (8)$$

with  $y_k^{j_1, \dots, j_n} = -1$  if  $k \in \{j_1, \dots, j_n\}$ , and  $+1$  otherwise.

The first term of (5) concerns the difference between two correlators. Let us see how this term decomposes as a function of the number of indices present in its variables. From Eq. (8), it is clear that a variable with  $n$  indices only appears in the difference  $S_k - S_l$  if  $y_k^{j_1, \dots, j_n} \neq y_l^{j_1, \dots, j_n}$ . But the corresponding strategy only has  $n$  differing outcomes and each correlator in this term only appears once, so a variable with  $n$  indices appears in at most  $n$  of these differences. Moreover, if it appears, it does so with a factor  $\pm 2$ . The coefficient in front of a variable with  $n$  indices in the first sum of Eq. (5) thus cannot be smaller than  $-2 \sum_{k=0}^{n-1} \alpha_k = 2n(n-m)$ .

The second term of Eq. (5) can be bounded as  $B^2 \geq 0$ , while the third one can be expressed as

$$C = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1, \dots, j_n} + \bar{a}_{j_1, \dots, j_n}) (m - 2n)^2. \quad (9)$$

Putting everything together and using property (7), we arrive at

$$\begin{aligned} I_{N,m} &\geq \sum_{k=0}^{\frac{m}{2}-1} \alpha_k (S_k - S_{m-k-1}) - \frac{1}{2} C \\ &\geq -\frac{m^2}{2} \sum_{\text{all variables}} = -\frac{m^2 N}{2} = -\beta_c, \end{aligned} \quad (10)$$

which concludes the proof.

Note that this bound is achieved for  $a_{01\dots(m/2)-1} = N$ , i.e., when for each party exactly the first half of the  $m$  measurements yields the result  $-1$ . Note also that the Bell inequality (2) does not reduce to Eq. (6) of Ref. [8] when  $m = 2$ . Indeed, while none of these inequalities is a facet of the local polytope, the latter one is a facet of the symmetrized 2-body correlator local polytope [8,30].

*From Bell inequalities to Bell-correlation witnesses.*— Let us now derive a set of Bell-correlation witnesses assuming a certain form for the measurement operators. Here, no assumptions are made on the measured state.

Following Ref. [9], we start from inequality (2) and introduce spin measurements along the axes  $\vec{d}_k$ ,  $k=0, \dots, m-1$ , as well as the collective spin observables  $\hat{S}_k$ :

$$M_k^{(i)} = \vec{d}_k \cdot \vec{\sigma}^{(i)}, \quad \hat{S}_k = \frac{1}{2} \sum_{i=1}^N M_k^{(i)}, \quad (11)$$

where  $\vec{\sigma}$  is the Pauli vector acting on a spin- $\frac{1}{2}$  system. The correlators can be expressed in terms of these total spin observables and the measurement directions [8]:

$$S_k = 2\langle \hat{S}_k \rangle, \quad S_{kl} = 2[\langle \hat{S}_k \hat{S}_l \rangle + \langle \hat{S}_l \hat{S}_k \rangle] - N\vec{d}_k \cdot \vec{d}_l. \quad (12)$$

This defines the Bell operators

$$\hat{W}_{N,m} := 2 \sum_{k=0}^{m-1} \alpha_k \hat{S}_k + 2 \sum_{k,l} \hat{S}_k \hat{S}_l - \frac{N}{2} \sum_{k,l} \vec{d}_k \cdot \vec{d}_l + \left[ \frac{m^2 N}{2} \right], \quad (13)$$

whose expectation values are positive for states that are not Bell correlated. Note that the expectation value of these operators need not be negative for all Bell correlated states and every choice of measurement directions, though. A negative value may only be achieved for specific choices of states and measurement settings.

We now consider measurement directions  $\vec{d}_k = \vec{a} \cos(\vartheta_k) + \vec{b} \sin(\vartheta_k)$  lying in a plane spanned by two orthonormal vectors  $\vec{a}$  and  $\vec{b}$ , with the antisymmetric angle distribution  $\vartheta_{m-k-1} = -\vartheta_k$ . Note that the coefficients  $\alpha_k$  share the same antisymmetry. Defining  $\mathcal{W}_m := \langle \hat{W}_{N,m} / (2\hat{N}) \rangle$  for even  $m$ , we arrive at the following family of witnesses:

$$\mathcal{W}_m = C_b \sum_{k=0}^{\frac{m}{2}-1} \alpha_k \sin(\vartheta_k) - (1 - \zeta_a^2) \left( \sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k) \right)^2 + \frac{m^2}{4}, \quad (14)$$

with  $\mathcal{W}_m \geq 0$  for states that are not Bell correlated. These Bell correlation witnesses depend on  $m/2$  angles  $\vartheta_k$  and involve just two quantities to be measured: the scaled collective spin  $C_b := \langle \hat{S}_b / (\hat{N}/2) \rangle$  and the scaled second moment  $\zeta_a^2 := \langle \hat{S}_a^2 / (\hat{N}/4) \rangle$ .

The tightest constraints on  $C_b$  and  $\zeta_a^2$  that allow for a violation of  $\mathcal{W}_m \geq 0$  are obtained by minimizing  $\mathcal{W}_m$  over

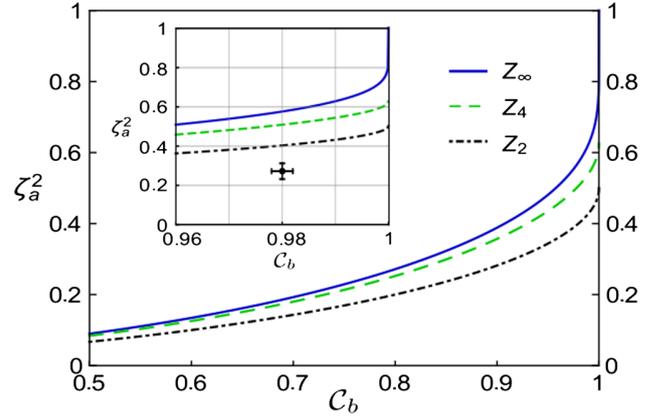


FIG. 1. Plots of the critical lines  $Z_2$ ,  $Z_4$ , and  $Z_\infty$ . The witness obtained from the Bell inequality with 4 settings already provides a significant improvement over the case of 2 settings. The black point in the inset shows the data point from Ref. [9], with  $N = 476 \pm 21$ .

the angles  $\vartheta_k$ . Solving  $\partial \mathcal{W}_m / \partial \vartheta_k = 0$  yields (see Appendix B in the Supplemental Material [25]):

$$\vartheta_k = -\arctan[\lambda_m(m-2k-1)], \quad (15)$$

$$\frac{C_b}{2\lambda_m(1-\zeta_a^2)} = \sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k). \quad (16)$$

Equation (16) is a self-consistency equation for  $\lambda_m$  that has to be satisfied in order to minimize  $\mathcal{W}_m$ .

Using these parameters, we can rewrite our witness in terms of the physical parameters  $C_b$  and  $\zeta_a^2$  only. For two measurement directions ( $m = 2$ ), we find that states which are not Bell correlated satisfy

$$\zeta_a^2 \geq Z_2(C_b) = \frac{1}{2} \left( 1 - \sqrt{1 - C_b^2} \right). \quad (17)$$

This recovers the bound obtained from a different inequality in Ref. [9]. Note that in the present case, the argument is more direct since it does not involve  $C_a$ , the first moment of the spin operator in the  $a$  direction.

Increasing the number of measurement directions allows for the detection of Bell correlations in additional states. In the limit  $m \rightarrow \infty$ , we find (see Appendix B in the Supplemental Material [25])

$$\zeta_a^2 \geq Z_\infty(C_b) = 1 - \frac{C_b}{\operatorname{artanh}(C_b)}. \quad (18)$$

Figure 1 shows the two witnesses (17) and (18) together with the one obtained similarly for  $m = 4$  settings in the  $C_b - \zeta_a^2$  plane. The curve  $Z_\infty$  reaches the point  $C_b = \zeta_a^2 = 1$ , therefore allowing, in principle, for the detection of Bell correlations in presence of arbitrarily low squeezing. It is known, however, that some values of  $C_b$  and  $\zeta_a^2$  can only be reached in the limit of a large number of spins [31]. For any fixed  $N$ , a finite amount of squeezing is thus necessary in

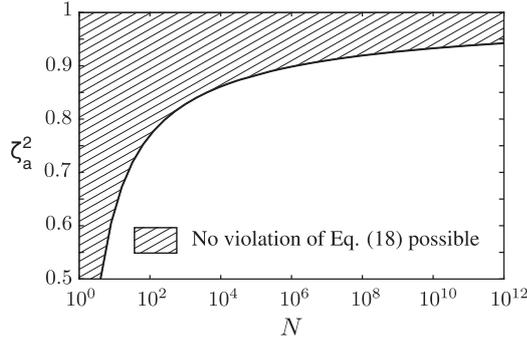


FIG. 2. Upper bound on the value of  $\zeta_a^2$  required to see a violation of the Bell correlation witness (18). The bound depends on the number of particles  $N$ .

order to allow for the violation of our witness (see Appendix C in the Supplemental Material [25]). The corresponding upper bound on  $\zeta_a^2$  is shown in Fig. 2.

Points below the curve  $Z_m$  in Fig. 1 indicate a violation of the witness  $\mathcal{W}_m \geq 0$  obtained from the corresponding  $m$ -settings Bell inequality. Violation of any such bound reveals the presence of a Bell correlated state. However, as discussed in the introduction, conclusions in the presence of finite statistics have to be examined carefully, since in practice, one can never conclude from the violation of a witness that the measured state is Bell correlated with 100% confidence. The point shown in the inset of Fig. 1 corresponds to the data reported in Ref. [9] from measurements on a spin-squeezed Bose-Einstein condensate. This point clearly violates the witnesses for  $m = 2, 4, \infty$  by several standard deviations, although the number of measurement rounds is too small to guarantee that the measured state is Bell correlated without further assumptions [9].

*Finite statistics.*—In this section, we put a bound on the number of experimental runs needed to exclude with a given confidence that a measured state is not Bell correlated. Note that such a conclusion does not follow straightforwardly from the violation of the witness by a fixed number of standard deviations. Indeed, standard deviations inform on the precision of a violation, but fail at excluding arbitrary local models [15], including, e.g., models which may show non-Gaussian statistics with rare events. We thus look here for a number of experimental runs which is sufficient to guarantee a  $p$  value lower than a given threshold for the null hypothesis “The measured state is not Bell correlated.” Since we are concerned with the characterization of physical systems in the absence of an adversary, we assume that the same state is prepared in each round (i.i.d. assumption).

For this statistical analysis, let us consider a different Bell correlation witness than Eq. (18). Indeed, we derived this inequality in order to maximize the amount of violation for given data, but here we rather wish to maximize the statistical evidence of a violation. For this, we take Eq. (14) and consider the representation of the angles given in

Eq. (15), but without taking Eq. (16) into account. In the limit of infinitely many measurement settings, we find (see Appendix B in the Supplemental Material [25])

$$\mathcal{W}_{\text{stat}} = -C_b \Delta_\nu - (1 - \zeta_a^2) \Lambda_\nu^2 + \frac{1}{4} \geq 0, \quad \text{with} \quad (19)$$

$$\Delta_\nu = \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2}, \quad \Lambda_\nu = \frac{\text{arsinh}(\nu)}{2\nu}, \quad (20)$$

where  $\nu = \lim_{m \rightarrow \infty} \lambda_m m$  is a free parameter that fully specifies the set of measurement angles.

In order to model the experimental evaluation of  $\mathcal{W}_{\text{stat}}$ , we introduce the following estimator:

$$\mathcal{T} = \frac{\chi(Z=0)}{q} X + \frac{\chi(Z=1)}{1-q} Y + \left( \frac{1}{4} - \Delta_\nu - \Lambda_\nu^2 \right). \quad (21)$$

Here,  $\chi$  denotes the indicator function and the binary random variable  $Z$  accounts for the choice between the measurement of either  $C_b$  or  $\zeta_a$ . Each measurement round thus allows for the evaluation of the corresponding random variables  $X = \Delta_\nu(1 - C_b)$  or  $Y = \Lambda_\nu^2 \zeta_a^2$ . Assuming that  $Z$  is independent of  $X$  and  $Y$  and choosing  $q = P[Z=0]$  guarantees that  $\mathcal{T}$  is a proper estimator of  $\mathcal{W}_{\text{stat}}$ , i.e.,  $\langle \mathcal{T} \rangle = \mathcal{W}$ .  $q$  then corresponds to the probability of performing a measurement along the  $b$  axis. We choose  $q = (1 + (\Lambda_\nu^2 N / 2\Delta_\nu))^{-1}$  so that the contributions of both measurement choices to  $\mathcal{T}$  have the same magnitude; i.e., the maximum values of  $X/q$  and  $Y/(1-q)$  are equal within the domain  $|C_b| \leq 1$  and  $\zeta_a^2 \in [0, N]$ . This also guarantees that the spectrum of  $\mathcal{T}$  matches that of  $\mathcal{W}_{\text{stat}}$ .

Suppose the measured state is non-Bell correlated, i.e., that its mean value  $\mu = \langle \mathcal{T} \rangle = \mathcal{W}_{\text{stat}} \geq 0$ . We are now interested in the probability that after  $M$  experimental runs the estimated value  $T = (1/M) \sum_{i=1}^M \mathcal{T}_i$  of the witness  $\mathcal{W}_{\text{stat}}$  falls below a certain value  $t_0 < 0$ , with  $\mathcal{T}_i$  being the value of the estimator in the  $i$ th run.

In statistics, concentration inequalities deal with exactly this issue. In Appendix D in the Supplemental Material [25], we compare four of these inequalities, namely, the Chernoff, Bernstein, Uspensky, and Berry-Esseen ones [25] and show explicitly that in the regime of interest the tightest and, therefore, preferred bound results from the Bernstein inequality:

$$P[T \leq t_0] \leq \exp \left( - \frac{(\mu - t_0)^2 M}{2\sigma_0^2 + \frac{2}{3}(t_u - t_l)(\mu - t_0)} \right) \leq \varepsilon. \quad (22)$$

Here,  $t_0$  is the experimentally observed value of  $T$  after  $M$  measurement rounds,  $t_l = \frac{1}{4} - \Delta_\nu - \Lambda_\nu^2$  and  $t_u = \frac{1}{4} + \Delta_\nu + \Lambda_\nu^2(N+1)$  are lower and upper bounds on the random variable  $\mathcal{T}$ , respectively, and  $\sigma_0^2$  is its variance for a local state.

We show in Appendix D in the Supplemental Material [25] that the largest  $p$  value is obtained by setting  $\mu = 0$  and  $\sigma_0^2 = -t_l t_u$ . A number of measurement rounds

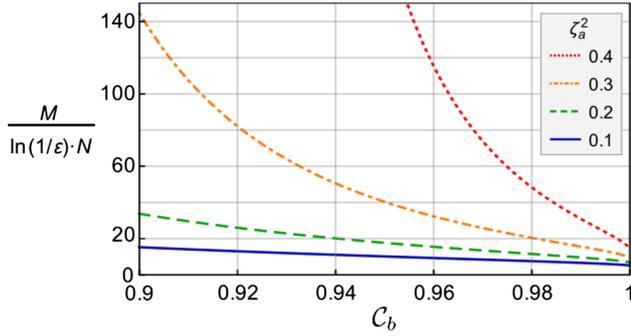


FIG. 3. Number of experimental runs per spins required to rule out non-Bell-correlated states with a confidence of  $1 - \epsilon$  as a function of  $C_b$  and  $\zeta_a^2$ . For  $C_b = 0.98$  and  $\zeta_a^2 = 0.272$  (as reported in Ref. [9]), approximately  $17 \ln(100) \approx 80$  runs per spin are sufficient to reach a confidence level of 99%.

sufficient to exclude the null hypothesis with a probability larger than  $1 - \epsilon$  is then given by

$$M \geq \frac{-2t_l t_u - \frac{2}{3}(t_u - t_l)t_0}{t_0^2} \ln\left(\frac{1}{\epsilon}\right). \quad (23)$$

This quantity can be minimized by choosing the free parameter  $\nu$  appropriately. As shown in Appendix D in the Supplemental Material [25], optimizing  $\nu$  at this stage allows us to reduce the number of measurement rounds by  $\sim 30\%$ . It is thus clearly advantageous not to consider the witness (18) when evaluating statistical significance.

The number of runs in Eq. (23) depends linearly on  $t_l$  and, therefore, also linearly on  $N$ . The ratio  $(M/N)$  thus tends to a constant for large  $N$  (see Appendix D in the Supplemental Material [25] for more details). This implies that a number of measurement rounds growing linearly with the system size is both necessary and sufficient to reliably conclude that the measured state is Bell correlated [9].

Figure 3 depicts the required number of measurement rounds per spin as a function of the scaled collective spin  $C_b$  and the scaled second moment  $\zeta_a^2$ . For a confidence level of  $1 - \epsilon = 99\%$ , between 20 and 500 measurement rounds per spin are required in the considered parameter region.

**Conclusion.**—In this Letter, we started by introducing a class of multipartite Bell inequalities involving two-body correlators and an arbitrary number of measurement settings. Assuming collective spin measurements, these inequalities give rise to the witness (18), which can be used to determine whether Bell correlations can be detected in a many-body system. This criterion detects states that were not detected by the previously known witness [9].

We then discussed the role of finite statistics in experiments involving many-body systems. We provided a bound, Eq. (23), on the number of measurement rounds that allows one to detect Bell correlated states without further assumptions. This bound shows that all non-Bell-correlated

states can be convincingly ruled out at the cost of performing a number of measurement rounds that grows linearly with the system size. This puts the detection of quantum correlations in many-body systems on firm grounds and opens the way for a possible use of many-body systems in the context of device-independent quantum information processing.

We thank Baptiste Allard, Remik Augusiak, and Valerio Scarani for helpful discussions. This work was supported by the Swiss National Science Foundation (SNSF) through Grants No. PP00P2-150579, No. 20020-169591, and No. NCCR QSIT. N.S. acknowledges the Army Research Laboratory Center for Distributed Quantum Information via the project SciNet.

- [1] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (2014).
- [2] V. Scarani, *Acta Phys. Slovaca* **62**, 347 (2012).
- [3] B. P. Lanyon, M. Zwerger, P. Jurcevic, C. Hempel, W. Dür, H. J. Briegel, R. Blatt, and C. F. Roos, *Phys. Rev. Lett.* **112**, 100403 (2014).
- [4] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, and H. Weinfurter, *Phys. Rev. Lett.* **90**, 200403 (2003).
- [5] Z. Zhao, T. Yang, Y.-A. Chen, A.-N. Zhang, M. Żukowski, and J.-W. Pan, *Phys. Rev. Lett.* **91**, 180401 (2003).
- [6] J. Lavoie, R. Kaltenbaek, and K. J. Resch, *New J. Phys.* **11**, 073051 (2009).
- [7] B. Hensen, H. Bernien, A. E. Drau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. L. Vermeulen, R. N. Schouten, C. Abelln, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiu, and R. Hanson, *Nature (London)* **526**, 682 (2015); L. K. Shalm *et al.*, *Phys. Rev. Lett.* **115**, 250402 (2015); M. Giustina *et al.*, *Phys. Rev. Lett.* **115**, 250401 (2015); W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter, *Phys. Rev. Lett.* **119**, 010402 (2017).
- [8] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, *Science* **344**, 1256 (2014).
- [9] R. Schmied, J.-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, and N. Sangouard, *Science* **352**, 441 (2016).
- [10] S. Pelisson, L. Pezzè, and A. Smerzi, *Phys. Rev. A* **93**, 022115 (2016).
- [11] N. J. Engelsen, R. Krishnakumar, O. Hosten, and M. A. Kasevich, *Phys. Rev. Lett.* **118**, 140401 (2017).
- [12] D. C. Bailey, *R. Soc. Open Sci.* **4**, 160600 (2017).
- [13] W. J. Youden, *Technometrics* **14**, 1 (1972).
- [14] R. Gill, *Stat. Sci.* **29**, 512 (2014).
- [15] Y. Zhang, S. Glancy, and E. Knill, *Phys. Rev. A* **84**, 062118 (2011).
- [16] C. Gross, T. Zibold, E. Nicklas, J. Estève, and M. K. Oberthaler, *Nature (London)* **464**, 1165 (2010).
- [17] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, and P. Treutlein, *Nature (London)* **464**, 1170 (2010).
- [18] B. Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, and C. Klempt, *Phys. Rev. Lett.* **112**, 155304 (2014).

- [19] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, [arXiv:1609.01609](https://arxiv.org/abs/1609.01609).
- [20] A. S. Sørensen, L.-M. Duan, J. I. Cirac, and P. Zoller, *Nature (London)* **409**, 63 (2001).
- [21] P. Hyllus, L. Pezzè, A. Smerzi, and G. Tóth, *Phys. Rev. A* **86**, 012337 (2012).
- [22] G. Tóth, C. Knapp, O. Gühne, and H. J. Briegel, *Phys. Rev. A* **79**, 042334 (2009).
- [23] J. Tura, A. B. Sainz, T. Vértesi, A. Acín, M. Lewenstein, and R. Augusiak, *J. Phys. A* **47**, 424024 (2014).
- [24] J. Tura, G. de las Cuevas, R. Augusiak, M. Lewenstein, A. Acín, and J. I. Cirac, *Phys. Rev. X* **7**, 021005 (2017).
- [25] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.119.170403> for a comparison of the four inequalities, which includes Refs. [26–29].
- [26] <http://cseweb.ucsd.edu/~klevchen/techniques/chernoff.pdf>.
- [27] G. Bennett, *J. Am. Stat. Assoc.* **57**, 33 (1962).
- [28] J. V. Uspensky, *Introduction to Mathematical Probability* (McGraw-Hill, New York, 1937).
- [29] A. C. Berry, *Trans. Am. Math. Soc.* **49**, 122 (1941).
- [30] J.-D. Bancal, N. Gisin, and S. Pironio, *J. Phys. A* **43**, 385303 (2010).
- [31] A. S. Sørensen and K. Mølmer, *Phys. Rev. Lett.* **86**, 4431 (2001).

# Supplementary material: Bell correlations in a many-body system with finite statistics

Sebastian Wagner,<sup>1</sup> Roman Schmied,<sup>2</sup> Matteo Fadel,<sup>2</sup> Philipp Treutlein,<sup>2</sup> Nicolas Sangouard,<sup>1</sup> and Jean-Daniel Bancal<sup>1</sup>

<sup>1</sup>*Quantum Optics Theory Group, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

<sup>2</sup>*Quantum Atom Optics Lab, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

(Dated: May 17, 2017)

## SM I. PROOF OF THE BELL INEQUALITIES

In this appendix, we expand on the proof of Ineq. (2) given in the main text, and cover the case of odd numbers of [measurement choices](#).

### A. Symmetric Bell inequality for $m$ [measurement settings](#)

We consider local measurements on  $N$  parties. For each party, one can choose between  $m$  [measurement settings](#)  $M_k^{(i)}$ , where  $k \in \{0, 1, \dots, m-1\}$  and  $i \in \{1, \dots, N\}$ . Each measurement has the two possible outcomes  $\pm 1$ . We are interested in Bell inequalities, i.e. inequalities every local theory has to obey [1]. We only consider one- and two-body mean values, so the general form of such an inequality is

$$I_{N,m} = \sum_{k=0}^{m-1} \sum_{i=1}^N \alpha_k^i \langle M_k^{(i)} \rangle + \sum_{k,l} \sum_{i < j} \beta_{kl}^{ij} \langle M_k^{(i)} M_l^{(j)} \rangle \geq -\beta_c, \quad (\text{SM 1})$$

where  $\langle M_k^{(i)} \rangle = \sum_{a \in \{-1,1\}} a \text{Prob}(M_k^{(i)} = a)$  and  $\langle M_k^{(i)} M_l^{(j)} \rangle = \sum_{a,b \in \{-1,1\}} ab \text{Prob}(M_k^{(i)} = a, M_l^{(j)} = b)$ .

We now restrict ourselves to Bell inequalities which are symmetric under exchange of parties, i.e.  $\alpha_k^i = \alpha_k$  and  $\beta_{kl}^{ij} = \beta_{kl}$ . After defining the symmetrized correlators

$$S_k = \sum_{i=1}^N \langle M_k^{(i)} \rangle, \quad S_{kl} = \sum_{i \neq j} \langle M_k^{(i)} M_l^{(j)} \rangle, \quad (\text{SM 2})$$

symmetric inequalities can be expressed as

$$I_{N,m} = \sum_{k=0}^{m-1} \alpha_k S_k + \frac{1}{2} \sum_{k,l}^{m-1} \beta_{kl} S_{kl} \geq -\beta_c. \quad (\text{SM 3})$$

We are interested in cases for which the coefficients are  $\alpha_k = m - 2k - 1$  ( $k = 0, \dots, m-1$ ) and  $\beta_{kl} = 1$ . We note that  $\alpha_{m-k-1} = -\alpha_k$  and claim that local theories have to fulfill the Bell inequalities

$$I_{N,m} = \sum_{k=0}^{m-1} (m - 2k - 1) S_k + \frac{1}{2} \sum_{k,l}^{m-1} S_{kl} \geq - \left\lfloor \frac{m^2 N}{2} \right\rfloor = -\beta_c, \quad (\text{SM 4})$$

where  $\lfloor x \rfloor$  is the largest integer smaller or equal to  $x$ .

### B. Computation of the local bound

In this section we prove the above claim, i.e. the validity of Ineq. (SM 4). One of the most important properties of a local theory is its equivalence to a mixture of deterministic local theory. That is why, by considering only deterministic theories, there is no loss of generality. We can therefore assume that a measurement  $M_k^{(i)}$  will lead to

an outcome  $x_k^i = \pm 1$  with probability 1, i.e.  $\langle M_k^{(i)} \rangle = x_k^i$ . The two-body correlators  $S_{kl}$  can thus be expressed as  $S_{kl} = S_k S_l - \sum_{i=1}^N x_k^i x_l^i$ . By also taking the antisymmetry of  $\alpha_k$  into account, and introducing the quantities

$$A = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} (m - 2k - 1)(S_k - S_{m-k-1}), \quad B = \sum_{k=0}^{m-1} S_k, \quad C = \sum_{i=1}^N \left[ \sum_{k=0}^{m-1} x_k^i \right]^2. \quad (\text{SM } 5)$$

we arrive at

$$I_{N,m} = A + \frac{1}{2}B^2 - \frac{1}{2}C. \quad (\text{SM } 6)$$

### 1. Strategy variables

We want to rewrite  $I_{N,m}$  further. Therefore we introduce variables counting the strategies chosen by the parties. Because there are  $m$  measurements with binary outcomes, the number of possible strategies per party is  $2^m$ . We define the following  $2^m$  variables:

$$\begin{aligned} a_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = -1 \text{ iff } k \in \{j_1, \dots, j_n\}\} \quad \text{for } n \leq \left\lfloor \frac{m}{2} \right\rfloor, \\ \bar{a}_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = +1 \text{ iff } k \in \{j_1, \dots, j_n\}\} \quad \text{for } n \leq \left\lfloor \frac{m-1}{2} \right\rfloor, \\ \bar{a}_{j_1, \dots, j_{\lfloor \frac{m}{2} \rfloor}} &\equiv 0, \end{aligned} \quad (\text{SM } 7)$$

where  $\#$  denotes the set cardinality. For example,  $a_j$  counts the parties  $k$  whose outcomes are  $x_k^{j'} = 1 - 2\delta_{jj'}$ .  $\bar{a}_j$  is the number of parties following the opposite strategy. Variables with  $n$  indices thus correspond to a strategy for which either exactly  $n$  of the  $m$  outcomes are  $+1$  or exactly  $n$  of the outcomes are  $-1$ , i.e.  $n$  outcomes differ from the rest. Note that the conjugate variables in the case of  $\frac{m}{2}$  indices are set to zero for the case of even  $m$  in order to prevent strategies from being counted twice. Since every party has to choose one strategy, the variables sum up to  $N$ , i.e.

$$a + \bar{a} + \sum_{j=0}^{m-1} (a_j + \bar{a}_j) + \dots = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}) = N. \quad (\text{SM } 8)$$

Note that  $S_k$  can be expressed in terms of the strategy variables as follows:

$$S_k = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) y_k^{j_1 \dots j_n}, \quad \text{with} \quad (\text{SM } 9)$$

$$y_k^{j_1 \dots j_n} = \begin{cases} -1 & \text{if } k \in \{j_1, \dots, j_n\} \\ +1 & \text{else} \end{cases}. \quad (\text{SM } 10)$$

### 2. Decomposition of $A$ and $C$ in terms of the strategy variables

Equation (SM 9) results in the following representation of  $S_k - S_l$ :

$$S_k - S_l = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) (y_k^{j_1 \dots j_n} - y_l^{j_1 \dots j_n}). \quad (\text{SM } 11)$$

Clearly, a variable only appears in this expression if  $y_k \neq y_l$ , i.e. if the strategy is such that the outcome of the  $k^{\text{th}}$  measurement differs from the  $l^{\text{th}}$ . This, for example, cannot be the case if the number of indices is zero, i.e. if all measurement outcomes are the same. So  $a$  and  $\bar{a}$  do not show up in  $S_k - S_l$ .

With the help of the introduced strategy variables, we can express  $A$  as

$$A = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} (m - 2k - 1) \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) \left( y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right). \quad (\text{SM } 12)$$

and  $C$  as

$$\begin{aligned} C &= m^2(a + \bar{a}) + (m - 2)^2 \sum_{j=0}^{m-1} (a_j + \bar{a}_j) + (m - 4)^2 \sum_{j_1 < j_2} (a_{j_1 j_2} + \bar{a}_{j_1 j_2}) + \dots \\ &= \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} (m - 2n)^2 \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}). \end{aligned} \quad (\text{SM } 13)$$

In other words, we notice that if a variable has  $n$  indices it contributes to  $C$  with a factor  $(m - 2n)^2$ .

### 3. A bound independent of the number of indices

In this section, we study the contributions of  $A$  and  $C$  to  $I_{N,m}$ . For this, we make use of the following theorem:

**Theorem 1.** *A strategy with  $n$  equal outcomes satisfies the inequality*

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq 2n.$$

*Proof.* First we note that the summation is such that no  $y_k^{j_1 \dots j_n}$  appears twice. Also, we know that since we consider binary outcomes,  $|y_k - y_{m-k-1}|$  is either 0 or 2. We thus have

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| = 2l, \quad l \in \mathbb{N}.$$

Assume now that the above inequality is violated, i.e.  $l > n$ .

$\Leftrightarrow y_k^{j_1 \dots j_n} \neq y_{m-k-1}^{j_1 \dots j_n}$  for  $l > n$  values of  $k$ .

$\Leftrightarrow$  The strategy  $(j_1, \dots, j_n)$  has  $l$  differing outcomes.

This is a contradiction to the definition of the strategy. Therefore the assumption must be wrong and the inequality holds for all strategies.  $\square$

**Corollary 1.1.** *A function  $f(k)$  which is monotonically decreasing with  $k$  satisfies*

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} f(k) \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq 2 \sum_{k=0}^{n-1} f(k).$$

*Proof.* Theorem 1 implies that  $y_k \neq y_{m-k-1}$  for at most  $n$  values of  $k$ . Taking into account that  $f(k)$  is monotonically decreasing, we find that

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} f(k) \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq \sum_{\substack{n \text{ values} \\ \text{of } k}} f(k) \cdot 2 \leq 2 \sum_{k=0}^{n-1} f(k).$$

$\square$

With the help of Corollary 1.1, we rewrite the quantity  $A$  as

$$\begin{aligned} A &= \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \alpha_k (y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n}) \\ &\geq \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) (-2) \sum_{k=0}^{n-1} \alpha_k = -2 \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} n(m-n) \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}). \end{aligned} \quad (\text{SM 14})$$

Making use of Eq. (SM 13) and (7), we then find that

$$\begin{aligned} A - \frac{1}{2}C &\geq \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \left[ -2n(m-n) - \frac{(m-2n)^2}{2} \right] \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) \\ &= -\frac{m^2}{2} \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) = -\frac{m^2}{2}N \end{aligned} \quad (\text{SM 15})$$

#### 4. Putting the pieces together

In order to conclude the proof, we now only miss the contribution of the term  $B$ . For this, we look at the case of even and odd  $m$  separately. When  $m$  is even,  $B = \sum_k S_k$  is also even. We thus find that  $B^2 \geq 0$ . This means that

$$I_{N,m} \geq A - \frac{1}{2}C \geq -\frac{m^2 N}{2}. \quad (\text{SM 16})$$

If  $m$  is odd,  $B$  shares the parity of  $N$ . That is why we have  $B^2 \geq 0$  for even  $N$ , and  $B^2 \geq 1$  for odd  $N$ , resulting in

$$I_{N,m} \geq \begin{cases} -\frac{m^2 N}{2} & \text{for even } N \\ -\frac{m^2 N}{2} + \frac{1}{2} & \text{for odd } N \end{cases}. \quad (\text{SM 17})$$

In general, the classical bound is thus  $\beta_c = \lfloor \frac{m^2 N}{2} \rfloor$ .

## SM II. OPTIMIZATION OF THE WITNESSES

In this appendix, we optimize the witnesses  $\mathcal{W}_m$  as given in Eq. (14) of the main text over the measurement angles. Let us remind the form of  $\mathcal{W}_m$ :

$$\mathcal{W}_m = \mathcal{C}_b \sum_{k=0}^{\frac{m}{2}-1} \alpha_k \sin(\vartheta_k) - (1 - \zeta_a^2) \left[ \sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k) \right]^2 + \frac{m^2}{4}. \quad (\text{SM 18})$$

We do this optimization by searching for those angles leading to the minimum of  $\mathcal{W}_m$ . This is equivalent to solving the system of equations arising from  $\frac{\partial \mathcal{W}_m}{\partial \vartheta_k} = 0$ :

$$\frac{\partial \mathcal{W}_m}{\partial \vartheta_k} = (m - 2k - 1) \mathcal{C}_b \cos(\vartheta_k) + 2 \sin(\vartheta_k) (1 - \zeta_a^2) \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = 0. \quad (\text{SM 19})$$

We eventually want to find angles such that  $\mathcal{W}_m$  is negative. To achieve this, the last term of Eq. (SM 18) must be compensated. Since  $\zeta_a^2 \geq 0$ , the second term of Eq. (SM 18) is bounded by  $-\frac{m^2}{4}$  and thus cannot be sufficient for a negative  $\mathcal{W}_m$ . On the other hand, the first term is bounded by  $-\frac{m^2}{4} + 1$  due to the fact that  $|\mathcal{C}_b| \leq 1$ . So we find that in order to reach  $\mathcal{W}_m < 0$ , we need  $\sin(\vartheta_k)$ ,  $\cos(\vartheta_k)$  and  $\mathcal{C}_b$  to differ from zero and  $\zeta_a^2 < 1$ . In the following studies, we assume these necessary constraints, allowing us to rewrite Eq. (SM 19) as

$$\frac{2(1 - \zeta_a^2)}{\mathcal{C}_b} \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = -(m - 2k - 1) \frac{\cos(\vartheta_k)}{\sin(\vartheta_k)} \quad \forall k. \quad (\text{SM 20})$$

Since the left side of Eq. (SM 20) does not explicitly depend on  $k$ , this can only be achieved if both sides are equal to a constant. The assumptions about  $\zeta_a^2$ ,  $\mathcal{C}_b$  and the angles, as reasoned above, allow us to write

$$\frac{2(1 - \zeta_a^2)}{\mathcal{C}_b} \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = \frac{1}{\lambda_m} = -(m - 2k - 1) \frac{\cos(\vartheta_k)}{\sin(\vartheta_k)}, \quad (\text{SM 21})$$

where  $\lambda_m$  is a constant depending for given  $\mathcal{C}_b$  and  $\zeta_a^2$  only on  $m$ . We find the optimal angles

$$\vartheta_k = -\arctan(\lambda_m(m - 2k - 1)). \quad (\text{SM 22})$$

For a minimal  $\mathcal{W}_m$ , the constants  $\lambda_m$  have to fulfill the self-consistency equations

$$\frac{\mathcal{C}_b}{2\lambda_m(1 - \zeta_a^2)} = \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = \sum_{l=0}^{\frac{m}{2}-1} \frac{1}{\sqrt{1 + \lambda_m^2(m - 2l - 1)^2}}. \quad (\text{SM 23})$$

Here, we used the fact that  $\cos(\arctan(x)) = \frac{1}{\sqrt{1+x^2}}$ . For further steps, we note that  $\sin(\arctan(x)) = \frac{x}{\sqrt{1+x^2}}$  and define the following functions:

$$\Lambda_m(\lambda_m) := \sum_{k=0}^{\frac{m}{2}-1} \frac{1}{\sqrt{1 + \lambda_m^2(m - 2k - 1)^2}} = \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \lambda_m^2(2k - 1)^2}}, \quad (\text{SM 24})$$

$$\Delta_m(\lambda_m) := \sum_{k=0}^{\frac{m}{2}-1} \frac{\lambda_m(m - 2k - 1)^2}{\sqrt{1 + \lambda_m^2(m - 2k - 1)^2}} = \sum_{k=1}^{\frac{m}{2}} \frac{\lambda_m(2k - 1)^2}{\sqrt{1 + \lambda_m^2(2k - 1)^2}}. \quad (\text{SM 25})$$

If we assume the representation of the angles given in Eq. (SM 22), the witnesses can be expressed as

$$\mathcal{W}_m = -\mathcal{C}_b \Delta_m(\lambda_m) - (1 - \zeta_a^2) \Lambda_m^2(\lambda_m) + \frac{m^2}{4} \geq 0, \quad (\text{SM 26})$$

which holds for non-Bell-correlated states. Note that in this expression, we only assume the *arctan*-angle-distribution, without taking the self-consistency equations into account, i.e. without optimizing the actual differences between angles.

For the case of  $m \rightarrow \infty$ , we need to rewrite the above witnesses, since  $\mathcal{W}_m$  diverges in this limit. We define

$$\mathcal{W}'_m := \frac{\mathcal{W}_m}{m^2} = -\mathcal{C}_b \frac{\Delta_m(\lambda_m)}{m^2} - (1 - \zeta_a^2) \left( \frac{\Lambda_m(\lambda_m)}{m} \right)^2 + \frac{1}{4} \geq 0. \quad (\text{SM 27})$$

We also have to rewrite the constant  $\lambda_m$ . We define  $\lambda_m = \frac{\nu_m}{m}$  and rewrite Eq. (SM 23) as

$$\frac{\mathcal{C}_b}{2\nu_m(1 - \zeta_a^2)} = \frac{\Lambda_m(\frac{\nu_m}{m})}{m}. \quad (\text{SM 28})$$

If we define  $s_k = \frac{2k-1}{m}$ , we see that  $\frac{1}{m}$  can be expressed as  $\frac{s_{k+1} - s_k}{2}$ . Note that for  $m \rightarrow \infty$ ,  $s_1 \rightarrow 0$  and  $s_{m/2} \rightarrow 1$ . Using the convention  $\nu_\infty = \nu$ , we find

$$\begin{aligned} \Lambda_\nu &:= \lim_{m \rightarrow \infty} \frac{\Lambda_m(\nu_m/m)}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \nu_m^2 \frac{(2k-1)^2}{m^2}}} \cdot \frac{1}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \nu^2 s_k^2}} \frac{s_{k+1} - s_k}{2} \\ &= \frac{1}{2} \int_0^1 \frac{1}{\sqrt{1 + \nu^2 s^2}} ds = \frac{\text{arsinh}(\nu)}{2\nu} \end{aligned} \quad (\text{SM 29})$$

$$\begin{aligned} \Delta_\nu &:= \lim_{m \rightarrow \infty} \frac{\Delta_m(\nu_m/m)}{m^2} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{\nu_m \frac{(2k-1)^2}{m^2}}{\sqrt{1 + \nu_m^2 \frac{(2k-1)^2}{m^2}}} \cdot \frac{1}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{\nu s_k^2}{\sqrt{1 + \nu^2 s_k^2}} \frac{s_{k+1} - s_k}{2} \\ &= \frac{1}{2} \int_0^1 \frac{\nu s^2}{\sqrt{1 + \nu^2 s^2}} ds = \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2}. \end{aligned} \quad (\text{SM 30})$$

This yields the witness

$$\mathcal{W}'_\infty = -\mathcal{C}_b \Delta_\nu - (1 - \zeta_a^2) \Lambda_\nu + \frac{1}{4} \quad (\text{SM 31})$$

$$= -\mathcal{C}_b \left( \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2} \right) - (1 - \zeta_a^2) \frac{\text{arsinh}^2(\nu)}{4\nu^2} + \frac{1}{4} \geq 0. \quad (\text{SM 32})$$

We now search for those points in the  $\mathcal{C}_b$ - $\zeta_a^2$ -plane that allow for a violation of the correlation witnesses of Ineq. (SM 26) and (SM 32). For this purpose, we assume the optimal angles given in Eq. (SM 23) and (SM 28) respectively. We define  $Z_m(\mathcal{C}_b)$  to be the scaled second moment, as a function of the scaled collective spin, such that  $\mathcal{W}_m$  vanishes.

### 1. $m = 2$

In the case of  $m = 2$  measurement settings, we have to solve the following system of equations in order to find  $Z_2$ :

$$0 = \mathcal{W}_2 = -\mathcal{C}_b \frac{\lambda_2}{\sqrt{1 + \lambda_2^2}} - (1 - \zeta_a^2) \frac{1}{1 + \lambda_2^2} + 1, \quad (\text{SM 33})$$

$$\frac{\mathcal{C}_b}{2\lambda_2(1 - \zeta_a^2)} = \frac{1}{\sqrt{1 + \lambda_2^2}}. \quad (\text{SM 34})$$

From these, we find the critical line  $Z_2$  and therefore the following condition, satisfied by every non-Bell-correlated state:

$$\zeta_a^2 \geq Z_2(\mathcal{C}_b) = \frac{1}{2} \left( 1 - \sqrt{1 - \mathcal{C}_b^2} \right). \quad (\text{SM 35})$$

### 2. Limit $m \rightarrow \infty$

The critical line  $Z_\infty$  is determined by solving the following equation for  $\zeta_a^2$

$$0 = \mathcal{W}'_\infty = -\mathcal{C}_b \left( \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2} \right) - (1 - \zeta_a^2) \frac{\text{arsinh}^2(\nu)}{4\nu^2} + \frac{1}{4}, \quad \text{where } \nu = \sinh \left( \frac{\mathcal{C}_b}{1 - \zeta_a^2} \right). \quad (\text{SM 36})$$

We find that any non-Bell-correlated state satisfies

$$\zeta_a^2 \geq Z_\infty(\mathcal{C}_b) = 1 - \frac{\mathcal{C}_b}{\text{artanh}(\mathcal{C}_b)}. \quad (\text{SM 37})$$

## SM III. SQUEEZING REQUIREMENT

Here, we find a bound on the amount of squeezing that is needed as a function of the number of spins  $N$  in order to violate the Bell correlation witness (18) described in the main text. Due to the structure of spin systems, the first moment  $\mathcal{C}_b$ , the second moment  $\zeta_a^2$  and the number of spins  $N$  satisfy the following constraints [2]:

$$\zeta_a^2 \geq 1 - \frac{N}{2} \left[ \sqrt{(1 - \mathcal{C}_b^2) \left[ \left( 1 + \frac{2}{N} \right)^2 - \mathcal{C}_b^2 \right]} + \mathcal{C}_b^2 - 1 \right] \quad (\text{SM 38})$$

(notice that there is an error in the expression given in the reference). For any number of spins  $N$ , equating the right-hand side of this constraint with the right-hand side of Eq. (18) gives the maximum value of  $\mathcal{C}_b$  under which a violation of the witness (18) is possible. The corresponding maximum value of  $\zeta_a^2$  is plotted as a function of the number of spins  $N$  in Fig. 2 of the main text. For large  $N$ , this function can be expanded as

$$Z_N^* = 1 - \frac{1}{\omega} - \frac{1}{2\omega^3} - \frac{3}{4\omega^4} - O(\omega^{-5}) \quad (\text{SM 39})$$

where  $\omega = W_{-1} \left( -\frac{1}{2\sqrt{N+1}} \right)$  and  $W_{-1}$  is the lower branch of the Lambert  $W$  function.

## SM IV. FINITE STATISTICS

In this appendix, we introduce four concentration inequalities and determine their bound on the p-value for generic non-Bell-correlated states. We also compare these p-values and choose the optimal one to estimate a number of experimental runs sufficient to exclude non-Bell-correlated states with a confidence  $1 - \varepsilon$ . Eventually we minimize this number of runs by optimizing the measurement angles.

### A. Concentration inequalities

In statistics, concentration inequalities bound the probability that a random variable  $X$  exceeds or falls below a certain value. In what follows, we recall the definitions of some of these inequalities. We then discuss some of their properties in view of our problem in the following section.

#### 1. Chernoff bound

The following version of the Chernoff bound was proven by Van Vu at the University of California, San Diego [3]. It was done for discrete, independent random variables. However, the bound also applies in the case of continuous random variables.

**Theorem 2.** *Let  $X_1, \dots, X_M$  be independent random variables with  $|X_i| \leq 1$  and expectation values  $E[X_i] = 0$  for all  $i$ . Let  $X = \sum_{i=1}^M X_i$  and  $\sigma^2$  be the variance of  $X$ . Then*

$$\begin{aligned} P[X \leq -\lambda\sigma] &\leq \exp\left(-\frac{\lambda^2}{4}\right), & \text{for } 0 \leq \lambda \leq 2\sigma, \\ P[X \leq -x_0] &\leq \exp\left(-\frac{x_0^2}{4\sigma^2}\right), & \text{for } 0 \leq x_0 \leq 2\sigma^2. \end{aligned}$$

**Corollary 2.1.** *Let  $X_1, \dots, X_M$  be independent random variables with  $E[X_i] = \mu$  and  $a \leq X_i \leq b$  for all  $i$ . Let  $X = \frac{1}{M} \sum_{i=1}^M X_i$ ,  $\sigma_i^2 = \text{Var}[X_i]$  and  $\sigma_0^2 = \max_i \{\sigma_i^2\}$ . Then*

$$P[X \leq x_0] \leq \exp\left(-\frac{(\mu - x_0)^2 M}{4\sigma_0^2}\right), \quad \text{for } \mu \geq x_0 \geq \mu - \frac{\sum_i \sigma_i^2}{M(b-a)}.$$

#### 2. Bernstein inequality

The following expression known as Bernstein inequality, was proven by Bernstein in 1927, but we refer to the work of George Bennett [4]. The inequality is valid under certain restrictions for the absolute moments. Since our random variables are bounded, we can be sure these restrictions to be fulfilled.

**Theorem 3.** *Let  $X_1, \dots, X_M$  be independent random variables with  $E[X_i] = 0$  and  $|X_i| \leq \xi$  for all  $i$ . Also let  $X = \frac{1}{M} \sum_{i=1}^M X_i$  and  $\sigma^2 = \frac{1}{M} \sum_{i=1}^M \text{Var}[X_i]$ . Then*

$$\begin{aligned} P[X \geq x_0] &\leq \exp\left(-\frac{x_0^2 M}{2\sigma^2 + \frac{2}{3}\xi x_0}\right), & \forall x_0 > 0, \\ P[X \leq -x_0] &\leq \exp\left(-\frac{x_0^2 M}{2\sigma^2 + \frac{2}{3}\xi x_0}\right), & \forall x_0 > 0. \end{aligned}$$

**Corollary 3.1.** Let  $X_1, \dots, X_M$  be independent random variables with  $E[X_i] = \mu$ ,  $\sigma_i^2 = \text{Var}[X_i]$  and  $a \leq X_i \leq b$  for all  $i$ . Also let  $X = \frac{1}{M} \sum_{i=1}^M X_i$  and  $\sigma^2 = \frac{1}{M} \sum_{i=1}^M \sigma_i^2 \leq \sigma_0^2 = \max_i \{\sigma_i^2\}$ . Then

$$\begin{aligned} P[X \leq x_0] &\leq \exp\left(-\frac{(\mu - x_0)^2 M}{2\sigma^2 + \frac{2}{3}(b-a)(\mu - x_0)}\right) \\ &\leq \exp\left(-\frac{(\mu - x_0)^2 M}{2\sigma_0^2 + \frac{2}{3}(b-a)(\mu - x_0)}\right), \quad \forall x_0 < \mu. \end{aligned}$$

### 3. Uspensky inequality

Uspensky stated in [5] an inequality for a stochastic variable:

**Theorem 4.** Let  $X$  be a random variable with mean value  $E[X] = 0$  and  $a \leq X \leq b$ . Additionally  $b \geq |a|$ . Let  $\sigma^2 = \text{Var}[X]$ . Then for  $x_0 \leq 0$

$$P[X \leq x_0] \leq \frac{\sigma^2}{\sigma^2 + x_0^2}.$$

**Corollary 4.1.** Let  $X_1, \dots, X_M$  be independent random variables with mean values  $E[X_i] = \mu$  and  $a \leq X_i \leq b$  for all  $i$ . Additionally  $b \geq |a|$ . Let  $\sigma_i^2 = \text{Var}[X_i]$  and  $\sigma_0^2 = \max_i \{\sigma_i^2\}$ . Then  $\sigma^2 = \text{Var}[X] = \text{Var}\left[\frac{1}{M} \sum_{i=1}^M X_i\right] = \frac{1}{M^2} \sum_{i=1}^M \sigma_i^2 \leq \frac{1}{M} \sigma_0^2$  so that for  $x_0 \leq \mu$

$$P[X \leq x_0] \leq \frac{\sigma_0^2}{\sigma_0^2 + (x_0 - \mu)^2 M}.$$

### 4. Berry-Esseen inequality

Andrew C. Berry and Carl-Gustav Esseen proved the following theorem [6]:

**Theorem 5.** Let  $X_1, \dots, X_M$  be independent identically distributed random variables with  $E[X_i] = \mu$  for all  $i$  and with  $X = \frac{1}{M} \sum_{i=1}^M X_i$ . Additionally, the variance  $\sigma^2 = \text{Var}[X_i] = E[(X_i - \mu)^2]$  and the third absolute moment  $\rho = E[|X_i - \mu|^3]$  are finite. Then there exists a constant  $C$  such that for all  $x$

$$|F_M(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3 \sqrt{M}}, \quad (\text{SM } 40)$$

where  $F_M(x) = P\left[\sqrt{\frac{M}{\sigma^2}}(X - \mu) \leq x\right]$  and  $\Phi(x) = \frac{1}{2}(1 + \text{erf}(x))$ .

Esseen also proved that the constant  $C$  has to fulfill  $C \geq \frac{\sqrt{10+3}}{6\sqrt{2\pi}}$ . A good estimate for  $C$  follows from

$$|F_M(x) - \Phi(x)| \leq \frac{0.33554(\rho + 0.415\sigma^3)}{\sigma^3 \sqrt{M}}. \quad (\text{SM } 41)$$

**Corollary 5.1.** A direct consequence of Theorem 5 is

$$\begin{aligned} P[X \leq x_0] &= P\left[\sqrt{\frac{M}{\sigma^2}}(X - \mu) \leq \sqrt{\frac{M}{\sigma^2}}(x_0 - \mu)\right] \\ &\leq \Phi\left(\sqrt{\frac{M}{\sigma^2}}(x_0 - \mu)\right) + \frac{C\rho}{\sigma^3 \sqrt{M}}. \end{aligned}$$

### B. Largest p-value of the concentration inequalities

In this section we determine the largest p-value for the concentration inequalities listed above, under the assumption that  $x_0 < 0$  and  $\mu \geq 0$ . We denote with  $X$  the random variable  $\frac{1}{M} \sum_{i=1}^M X_i$ , with  $x_l \leq X_i \leq x_u$  and  $x_l < 0$ .

The crucial point here is that we have no information about the random variable apart from its non-negative mean value. This lack of information directly disqualifies the Berry-Esseen inequality of Corollary 5.1 as a potential tight bound. Indeed, the Berry-Esseen bound does not result in a tighter restriction than the trivial bound  $P[X \leq x_0] \leq 1$ . To see this, we consider the following distribution with three peaks:

Consider the case, for which all  $X_i$  satisfy the following probability distribution:

$$P[X_i = x] = \begin{cases} p_l & \text{for } x = x_l \\ p_u & \text{for } x = x_u \\ 1 - p_l - p_u & \text{for } x = 0 \\ 0 & \text{else} \end{cases}, \quad (\text{SM 42})$$

where  $p_l, p_u < 1$ . Additionally, we demand  $E[X_i] = 0$  leading to the condition  $-p_l x_l = p_u x_u$ . We thus arrive at the variance and third absolute moment:

$$\sigma^2 = p_u x_u (x_u - x_l), \quad (\text{SM 43})$$

$$\rho = p_u x_u (x_u^2 + x_l^2) \quad (\text{SM 44})$$

$$\Rightarrow \frac{\rho}{\sigma^3} = \frac{x_u^2 + x_l^2}{\sqrt{p_u x_u (x_u - x_l)^3}}. \quad (\text{SM 45})$$

In this expression,  $p_u$  remains as a parameter scaling the weight on the edges compared to the weight at  $x = 0$ . Note that for  $p_u \rightarrow \frac{-x_l}{x_u - x_l}$ , we arrive at a binomial distribution while for  $p_u \rightarrow 0$  we have a delta distribution. From Eq. (SM 45), we see that in the limit  $p_u \rightarrow 0$ ,  $\frac{\rho}{\sigma^3} \rightarrow \infty$ . Thus, we can write for the Berry-Esseen bound

$$P[X \leq x_0] \leq \Phi \left( \sqrt{\frac{M}{\sigma^2}} (x_0 - \mu) \right) + \frac{C\rho}{\sigma^3 \sqrt{M}} \leq \frac{C\rho}{\sigma^3 \sqrt{M}} = \frac{C}{\sqrt{M}} \frac{x_u^2 + x_l^2}{\sqrt{p_u x_u (x_u - x_l)^3}} \xrightarrow{p_u \rightarrow 0} \infty. \quad (\text{SM 46})$$

Since we have no information on the actual probability distribution, the largest p-value of the Berry-Esseen bound is 1. We thus restrict our interests to the Chernoff, Bernstein and Uspensky bounds.

The inequalities of Corollaries 2.1, 3.1 and 4.1 have certain properties in common: They all depend on the variance  $\sigma_0^2$  and the mean value  $\mu$ . More explicitly, the dependence on  $\sigma_0^2$  in all three cases is such that if one increases the variance, the bounds are also increased. We therefore use the following strategy to determine the largest p-values: We increase the variance of an arbitrary random variable in a way leaving the mean value unaffected. We eventually arrive at an easy-to-handle probability distribution with a maximal variance. The bounds resulting from this distribution then serve as upper bounds for all distributions of the same mean value. The bounds given by the concentration inequalities will then only depend on the mean value, so that we can optimize over  $\mu$ .

Eventually, we show that the largest p-value results from the binomial distribution centered around  $\mu = 0$ .

**Theorem 6.** *Let  $X$  be a random variable in the interval  $[a, b]$  with an arbitrary probability distribution and with  $E[X] = \mu$ . Let  $X_{bi}$  be a binomially distributed random variable with peaks at the edges  $a$  and  $b$ . Furthermore,  $X_{bi}$  has the same mean value as  $X$ , i.e.  $P[X_{bi} = a] = \frac{\mu - b}{a - b}$ ,  $P[X_{bi} = b] = \frac{a - \mu}{a - b}$  and  $P[X_{bi} = x] = 0$  otherwise. Additionally, let  $a \leq 0$ . Then*

$$\text{Var}[X_{bi}] \geq \text{Var}[X].$$

*Proof.* We define a third random variable  $Y$  satisfying

$$P[Y = x] = \begin{cases} P[X = x] & \text{if } x \notin dx_0 \cup \{a, b\} \\ 0 & \text{if } x \in dx_0 \\ P[X = a] + qP[X \in dx_0] & \text{if } x = a \\ P[X = b] + (1 - q)P[X \in dx_0] & \text{if } x = b \end{cases}. \quad (\text{SM 47})$$

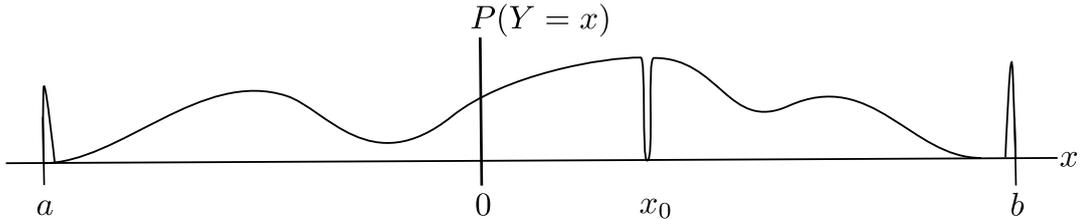


FIG. 1: Sketch of the probability distribution function of  $Y$  in Theorem 6.

Here,  $dx_0$  is an infinitesimal set around  $x_0 \in ]a, b[$ . So in other words, the probability distribution function of  $Y$  is almost the same as the one of  $X$ . The only difference is that the set  $dx_0$  is "cut out" and the probabilities at  $a$  and  $b$  are increased (see Fig. 1). They are increased in a way which leaves the mean value unaffected, i.e.  $q$  is chosen such that  $E[Y] = E[X] = \mu$ :

$$\begin{aligned} E[Y] &= E[X] - P[X \in dx_0]x_0 + qP[X \in dx_0]a + (1-q)P[X \in dx_0]b \\ &= \mu + P[X \in dx_0][-x_0 + q(a-b) + b] = \mu \\ \Leftrightarrow \quad q &= \frac{x_0 - b}{a - b}. \end{aligned} \tag{SM 48}$$

With this and  $a \leq 0$ , we show that  $\text{Var}[Y] \geq \text{Var}[X]$ :

$$\begin{aligned} \text{Var}[Y] &= \text{Var}[X] + P[X \in dx_0] [-x_0^2 + q(a^2 - b^2) + b^2] = \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0(a+b) - ab] \\ &\geq \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0(a+x_0) - ax_0] \geq \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0^2] = \text{Var}[X]. \end{aligned} \tag{SM 49}$$

So we find that  $\text{Var}[Y] \geq \text{Var}[X]$ , with the equal sign only if  $P[X \in dx_0] = 0$ . By induction, one can gradually "cut out" all the other points in  $]a, b[$ . During this process, the variance is constantly increased while the mean value remains unchanged. Eventually, one arrives at the random variable  $X_{bi}$  with its binomial distribution.  $\square$

By applying the different bounds and using Theorem 6, we find

$$P[X \leq x_0] \leq \begin{cases} \exp\left(-\frac{(\mu-x_0)^2 M}{4\sigma_{bi}^2(\mu)}\right) =: p_C(\mu, M) & \text{Chernoff} \\ \exp\left(-\frac{(\mu-x_0)^2 M}{2\sigma_{bi}^2(\mu) + \frac{2}{3}(x_u-x_l)(\mu-x_0)}\right) =: p_B(\mu, M) & \text{Bernstein ,} \\ \frac{\sigma_{bi}^2(\mu)}{\sigma_{bi}^2(\mu) + (x_0-\mu)^2 M} =: p_U(\mu, M) & \text{Uspensky} \end{cases} \tag{SM 50}$$

where  $\sigma_{bi}^2(\mu) = (x_u - \mu)(\mu - x_l)$ . As stated above,  $\mu \in [0, x_u]$ . Restricted to this interval, the three functions  $p_C$ ,  $p_B$  and  $p_U$  are strictly monotonous decreasing with  $\mu$  and therefore take their maximal values at  $\mu = 0$ . This allows us to write

$$P[X \leq x_0] \leq \begin{cases} \exp\left(-\frac{x_0^2 M}{4\sigma_{bi}^2}\right) =: p_C(M) & \text{Chernoff} \\ \exp\left(-\frac{x_0^2 M}{2\sigma_{bi}^2 - \frac{2}{3}(x_u-x_l)x_0}\right) =: p_B(M) & \text{Bernstein ,} \\ \frac{\sigma_{bi}^2}{\sigma_{bi}^2 + x_0^2 M} =: p_U(M) & \text{Uspensky} \end{cases} \tag{SM 51}$$

with  $\sigma_{bi}^2 := \sigma_{bi}^2(0) = -x_l x_u$ .

We identify  $M$  with the number of experimental runs, and determine the minimum number of runs required to

have  $P[X \leq x_0] \leq \varepsilon$ . This corresponds to solving  $p_i(M) \leq \varepsilon$  to  $M$ , where  $i = C, B, U$ . We find

$$M \geq \begin{cases} M_C := \frac{4\sigma_{bi}^2}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) & \text{Chernoff} \\ M_B := \frac{2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) & \text{Bernstein} \\ M_C := \frac{\sigma_{bi}^2(1-\varepsilon)}{\varepsilon x_0^2} & \text{Uspensky} \end{cases} \quad (\text{SM } 52)$$

### C. Comparison of the bounds

We now compare the three bounds stated in Ineq. (SM 52) and show that in the regime of interest, the Bernstein bound is the tightest bound. By comparing  $M_B$  to  $M_C$ , one finds

$$\begin{aligned} M_B \cdot \frac{x_0^2}{\ln(1/\varepsilon)} &= 2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0 < 2\sigma_{bi}^2 - (x_u - x_l)x_0 \\ &< -2x_lx_u - x_u x_l - x_lx_u = -4x_lx_u = 4\sigma_{bi}^2 = M_C \cdot \frac{x_0^2}{\ln(1/\varepsilon)} \\ &\Rightarrow M_B < M_C. \end{aligned} \quad (\text{SM } 53)$$

This means the bound resulting from Bernstein's inequality is tighter than the Chernoff bound for all values of  $x_l$ ,  $x_u$  and  $x_0$ .

Now we compare  $p_B$  to  $p_U$ . Plotting both functions reveals that Uspensky's inequality is better for small numbers of experimental runs, whereas Bernstein's is better for larger numbers of runs. We now want to estimate for which  $\varepsilon$  both approximations require the same number of runs  $M$ . We therefore set  $M_B = M_U$ :

$$\frac{2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) = \frac{\sigma_{bi}^2(1-\varepsilon)}{\varepsilon x_0^2} \Leftrightarrow 2 + \frac{2}{3} \frac{(x_u - x_l)(-x_0)}{\sigma_{bi}^2} = \frac{1-\varepsilon}{\varepsilon \ln\left(\frac{1}{\varepsilon}\right)}. \quad (\text{SM } 54)$$

The function on the right side of Eq. (SM 54) is strictly monotonous decreasing with  $\varepsilon$ . The left side is just a number. So the minimal  $\varepsilon$  possible is achieved if the left side is maximized. We thus make the following estimation:

$$\frac{(x_u - x_l)|x_0|}{\sigma_{bi}^2} = \frac{(x_u + |x_l|)|x_0|}{|x_l|x_u} \leq \frac{x_u|x_l| + |x_l|x_u}{|x_l|x_u} = 2.$$

For this value, Eq. (SM 54) yields  $\varepsilon \approx 0.127$ . This means that the Uspensky bound can only be better than the Bernstein bound for  $\varepsilon \geq 0.127$ . Since we are interested in probabilities  $\varepsilon \leq 0.1$ , the Bernstein bound remains the preferred one.

### D. Statistical optimization

Let us now present the result of numerical studies on the number of measurement runs allowing one to rule out non-Bell-correlated states. Here we rely on the Bernstein bound presented in Ineq. (SM 52).

#### 1. Choice of settings

First, we study the choice of measurement settings, i.e. the choice of  $\nu$  in Eq. (SM 32), which allow for the strongest statistical claim. For this, we minimize the number of experimental runs sufficient to conclude about the presence of a Bell-correlated state with given confidence  $1 - \varepsilon$  over the choices of  $\nu$ . We then compare this number of **measurement rounds**  $M$  to the one obtained when choosing  $\nu = \sinh\left(\frac{\mathcal{C}_b}{1-\zeta_a^2}\right)$ , i.e. for the settings which maximize the witness value (see Eq. (SM 36)).

The result of this comparison is presented in Figure 2 for a particular choice of  $\mathcal{C}_b$  and  $\zeta_a^2$ . Clearly, it is advantageous to reoptimize the measurement setting in order to maximize the statistical evidence, and thus a different witness should be considered in this case. In the rest of this appendix, measurement settings are optimized in order to maximize statistical evidence.

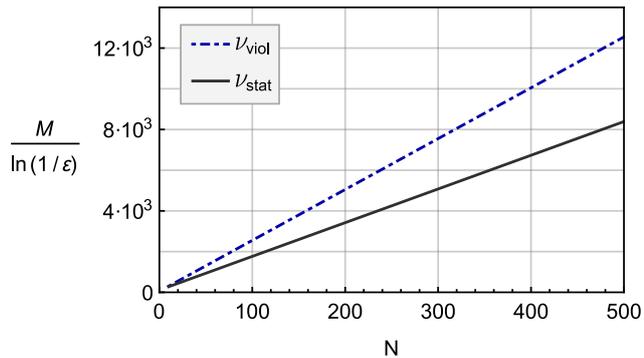


FIG. 2: Plot of the required number of runs as a function of the number of atoms, with  $C_b = 0.98$  and  $\zeta_a^2$ . The blue dash-dotted line results from the  $\nu$  which maximizes the violation of witness (19) whereas the solid black line corresponds to the  $\nu$  minimizing the number of runs. As one can see, the optimization of  $\nu$  in terms of statistics reduces the number of runs by a factor of approximately  $\frac{2}{3}$ .

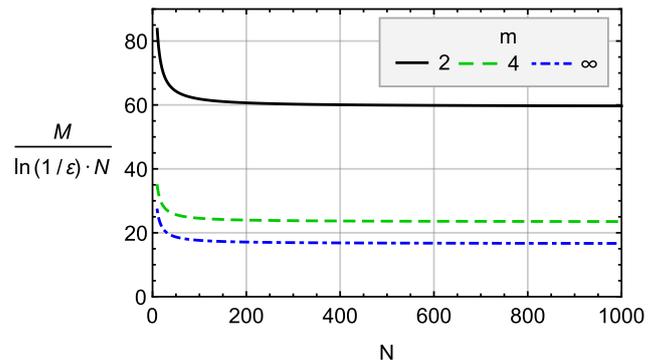


FIG. 3: Plots of the required number of experimental runs  $M$  per spin for a confidence of  $\varepsilon$  in units of  $\ln(\frac{1}{\varepsilon})$ , as a function of the number of spins  $N$ . The plots are for  $m = 2, 4, \infty$ , with  $\zeta_a^2 = 0.272$  and  $C_b = 0.98$ . The ratios tend to constants for larger systems.

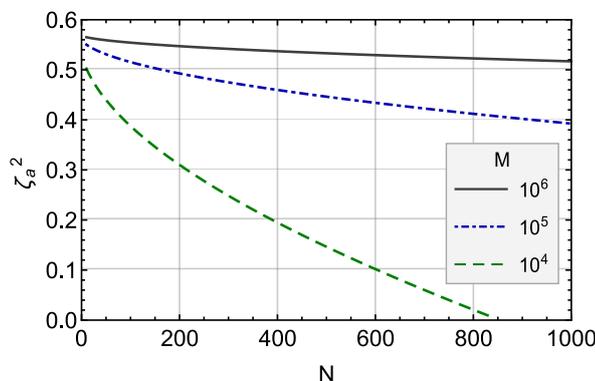


FIG. 4: Plots of the required scaled second moment  $\zeta_a^2$  as a function of the number of atoms  $N$ , with  $C_b = 0.98$  and  $\varepsilon = 0.1$ . The different curves represent different numbers of experimental runs ( $M = 10^4, 10^5, 10^6$ ).

### 2. Linear relation between the number of measurement runs and the number of spins

Figure 3 depicts the number of measurement runs per spin needed in order to reach a confidence of  $1 - \varepsilon$ . The plots are for the values of  $C_b$  and  $\zeta_a^2$  stated in [7], and for  $m = 2, 4, \infty$  settings. As explained in the main text, we observe that the ratio  $\frac{M}{N}$  tends to a constant as  $N$  increases. This plot also illustrates the gain obtained in using the newly derived Bell inequality with  $m$  settings.

### 3. Minimum squeezing requirement for finite $M$

As discussed in the main text, violation of our witness requires a finite amount of squeezing for systems of finite size  $N$ . Here we study this requirement when finite number of measurement runs are taken into account. In Fig. 4 we show the upper bound on  $\zeta_a^2$  as a function of the number of spins  $N$  and the number of measurement runs  $M$  for the case  $C_b = 0.98$ ,  $\varepsilon = 0.1$ . Although Fig. 2 in the main text shows that a violation with  $N = 1000$  spins is possible when  $\zeta_a^2 \lesssim 0.8$ , we see here that a squeezing of  $\zeta_a^2 \lesssim 0.5$  is needed if the number of measurement rounds is limited to  $10^6$ .

- [2] A. S. Sørensen, K. Mølmer, Phys. Rev. Lett. **86**, 4431 (2001).
- [3] <http://cseweb.ucsd.edu/~klevchen/techniques/chernoff.pdf>
- [4] G. Bennett, Journal of the American Statistical Association, **57**, 297 (1962).
- [5] J. V. Uspensky, Introduction to Mathematical Probability. No. 519.1 U86. 1937.
- [6] A. C. Berry, Trans. Amer. Math. Soc., **49**, 122-136 (1941).
- [7] R. Schmied, J-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, N. Sangouard, Science, **352**, 6284 (2016).



## VARIATIONAL METHOD FOR TAILORING BELL INEQUALITIES

In this chapter, we want to introduce a variational method which is very convenient for finding new Bell inequalities tailored to the certification of specific target states. We will do this in two steps; first we will introduce basic concepts and then we will explain our approach directly by applying it to partially-entangled two-qubit states. Our approach is an essential part in the works on the self-testing of quantum channels (chapter 4) and quantum measurements (chapter 5). A manuscript on the variational method is in preparation [27].

### 3.1 Eigenvalue Perturbation

---

Suppose we have an operator  $\mathcal{B}(\vec{\delta})$  depending on a parameter  $\vec{\delta}$ . We denote with  $\mathcal{B}_0$  the operator  $\mathcal{B}(\vec{\delta} = 0)$  and we know that this operator has a maximal eigenstate  $|\phi_0\rangle$ . Maximal eigenstate means that  $|\phi_0\rangle$  is an eigenstate of  $\mathcal{B}_0$  and that the corresponding eigenvalue  $\lambda_0$  is the largest in the spectrum of  $\mathcal{B}_0$ . Now consider small perturbations to the operator, the state and the

eigenvalues. Including second order terms, this can be written as

$$\mathcal{B}(\vec{\delta}) = \mathcal{B}_0 + \sum_{i=0}^{d-1} \delta_i \mathcal{B}'_i + \frac{1}{2} \sum_{i,j=0}^{d-1} \delta_i \mathcal{B}''_{ij} \delta_j = \mathcal{B}_0 + \vec{\delta}^\top \vec{\mathcal{B}}' + \frac{1}{2} \vec{\delta}^\top \mathcal{B}'' \vec{\delta}, \quad (3.1)$$

$$|\phi\rangle_{\vec{\delta}} = |\phi_0\rangle + \sum_{i=0}^{d-1} \delta_i |\phi'_i\rangle + \frac{1}{2} \sum_{i,j=0}^{d-1} \delta_i |\phi''_{ij}\rangle \delta_j = |\phi_0\rangle + \vec{\delta}^\top \vec{\phi}' + \frac{1}{2} \vec{\delta}^\top \phi'' \vec{\delta}, \quad (3.2)$$

$$\lambda(\vec{\delta}) = \lambda_0 + \sum_{i=0}^{d-1} \delta_i \lambda'_i + \frac{1}{2} \sum_{i,j=0}^{d-1} \delta_i \lambda''_{ij} \delta_j = \lambda_0 + \vec{\delta}^\top \vec{\lambda}' + \frac{1}{2} \vec{\delta}^\top \lambda'' \vec{\delta}, \quad (3.3)$$

where  $\vec{\mathcal{B}}'$  ( $\mathcal{B}''$ ) is a vector (matrix) of operators, and analogously for  $\vec{\phi}'$ ,  $\phi''$ ,  $\vec{\lambda}'$  and  $\lambda''$ . We recall that  $\mathcal{B}_0|\phi_0\rangle = \lambda_0|\phi_0\rangle$  and force  $\mathcal{B}(\vec{\delta})|\phi\rangle_{\vec{\delta}} = \lambda(\vec{\delta})|\phi\rangle_{\vec{\delta}}$ , resulting in

$$(\mathcal{B}_0 + \vec{\delta}^\top \vec{\mathcal{B}}' + \frac{1}{2} \vec{\delta}^\top \mathcal{B}'' \vec{\delta})(|\phi_0\rangle + \vec{\delta}^\top \vec{\phi}' + \frac{1}{2} \vec{\delta}^\top \phi'' \vec{\delta}) = (\lambda_0 + \vec{\delta}^\top \vec{\lambda}' + \frac{1}{2} \vec{\delta}^\top \lambda'' \vec{\delta})(|\phi_0\rangle + \vec{\delta}^\top \vec{\phi}' + \frac{1}{2} \vec{\delta}^\top \phi'' \vec{\delta}). \quad (3.4)$$

To first order in the perturbation parameter  $\vec{\delta}$ , Eq. (3.4) can be expressed as

$$\begin{aligned} \mathcal{B}_0 \vec{\delta}^\top \vec{\phi}' + \vec{\delta}^\top \vec{\mathcal{B}}' |\phi_0\rangle &= \lambda_0 \vec{\delta}^\top \vec{\phi}' + \vec{\delta}^\top \vec{\lambda}' |\phi_0\rangle, \\ \text{or equivalently: } \delta_i (\mathcal{B}_0 |\phi'_i\rangle + \mathcal{B}'_i |\phi_0\rangle) &= \delta_i (\lambda_0 |\phi'_i\rangle + \lambda'_i |\phi_0\rangle) \quad \forall i. \end{aligned} \quad (3.5)$$

We can express the primed states in the basis of the unperturbed operator ( $\{|\phi_k\rangle\}$ ) by writing  $|\phi'_i\rangle = \sum_{k=0}^{d-1} \mathcal{E}_{ik} |\phi_k\rangle$ . Multiplying Eq. (3.5) from the left with  $\langle\phi_k|$  leads to:

$$\lambda'_i = \langle\phi_0|\mathcal{B}'_i|\phi_0\rangle = \langle\phi_0|\frac{\partial \mathcal{B}}{\partial \delta_i}|\phi_0\rangle, \quad \mathcal{E}'_{i,k} = \frac{\langle\phi_k|\mathcal{B}'_i|\phi_0\rangle}{\lambda_0 - \lambda_k} \quad \forall k > 0. \quad (3.6)$$

The second order terms in Eq. (3.4) result in

$$\delta_i \delta_j (\mathcal{B}_0 |\phi''_{ij}\rangle + 2\mathcal{B}'_i |\phi'_j\rangle + \mathcal{B}''_{ij} |\phi_0\rangle) = \delta_i \delta_j (\lambda_0 |\phi''_{ij}\rangle + 2\lambda'_i |\phi'_j\rangle + \lambda''_{ij} |\phi_0\rangle). \quad (3.7)$$

Again, we express the doubly-primed states in terms of the unperturbed states. We write  $|\phi''_{ij}\rangle = \sum_{\ell=0}^{d-1} \alpha_{ij}^{(\ell)} |\phi_\ell\rangle$  and find by multiplying Eq. (3.7) with  $\langle\phi_0|$ , and making use of Eqs. (3.6) that

$$\lambda''_{ij} = \langle\phi_0|\mathcal{B}''_{ij}|\phi_0\rangle + 2 \sum_{\ell=1}^{d-1} \frac{\langle\phi_\ell|\mathcal{B}'_i|\phi_0\rangle \langle\phi_0|\mathcal{B}'_j|\phi_\ell\rangle}{\lambda_0 - \lambda_\ell} =: \mu_{ij} + \nu_{ij}, \quad (3.8)$$

with  $\mathcal{B}'_i = \frac{\partial \mathcal{B}}{\partial \delta_i}$  and  $\mathcal{B}''_{ij} = \frac{\partial^2 \mathcal{B}}{\partial \delta_i \partial \delta_j}$ . As a reminder, the eigenvalues are  $\lambda_\ell = \langle\phi_\ell|\mathcal{B}_0|\phi_\ell\rangle$ .

## 3.2 A New Self-testing Inequality for Partially-Entangled States

A partially-entangled two-qubit state can always be represented in the following way:

$$|\phi_\theta\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle, \quad (3.9)$$

where  $\theta \in (0, \frac{\pi}{4}]$  quantifies the degree of entanglement, with  $\theta = \frac{\pi}{4}$  corresponding to the case of a maximally-entangled state. In the following studies we will use the short notation  $c_x := \cos(x)$ ,  $s_x := \sin(x)$ .

An inequality suitable for self-testing  $|\phi_\theta\rangle$  is given by the so-called tilted-CHSH Bell inequality [28]:

$$\langle \alpha(\theta)A_0 + A_0(B_0 + B_1) + A_1(B_0 - B_1) \rangle \leq \beta_L(\theta), \quad (3.10)$$

where  $\alpha(\theta) = 2/\sqrt{1 + 2\tan(2\theta)^2}$  and the local bound is  $\beta_L(\theta) = 2 + \alpha(\theta)$ . The operators achieving the quantum bound of  $\beta_Q(\theta) = \sqrt{8 + 2\alpha(\theta)^2}$  (together with  $|\phi_\theta\rangle$ ) are

$$A_0 = \sigma_z, \quad B_0(\theta) = c_{\mu(\theta)}\sigma_z + s_{\mu(\theta)}\sigma_x, \quad (3.11)$$

$$A_1 = \sigma_x, \quad B_1(\theta) = c_{\mu(\theta)}\sigma_z - s_{\mu(\theta)}\sigma_x, \quad (3.12)$$

where  $\tan(\mu) = s_{2\theta}$ .

In this section we will derive a new inequality which is symmetric under exchange of parties and which allows for more noise-tolerant self-testing. Let us start by introducing the notion of stabilising operators. In group theory, a stabilizer is a subgroup of transformations that leave a certain element fixed. One such transformation from the set we will denote a stabilising operator:

**Definition 4** *Stabilising Operator* – An operator  $S$  is called a stabilising operator of the state  $|\varphi\rangle$  if it fulfils the condition:

$$S|\varphi\rangle = |\varphi\rangle. \quad (3.13)$$

With this definition, we find two stabilising operators of  $|\phi_\theta\rangle$ :

$$S_1 = \sigma_z \otimes \sigma_z, \quad (3.14)$$

$$S_2 = \frac{1}{2}c_{2\theta}(\sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z) + s_{2\theta}\sigma_x \otimes \sigma_x. \quad (3.15)$$

Their eigenvalues are  $\{1, 1, -1, -1\}$  and  $\{1, s_{2\theta}, -s_{2\theta}, -1\}$ , respectively. Note that there are also other stabilising operators. The crucial feature, however, is the fact that one can find a common eigenbasis since  $[S_1, S_2] = 0$ . If one includes other operators, this is no longer the case. By constructing the operator

$$\mathcal{B}_S = (1 - p)S_1 + pS_2, \quad (3.16)$$

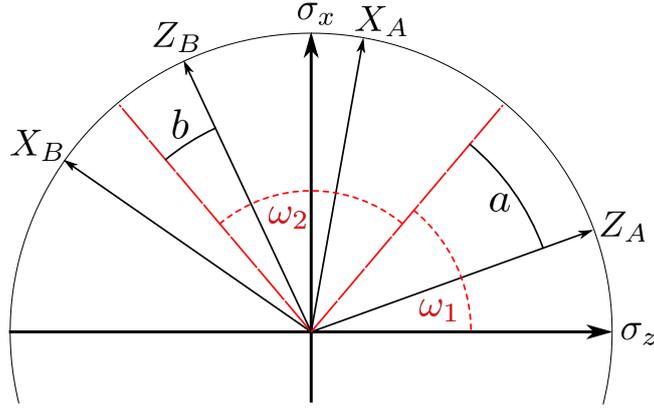


Figure 3.1: Scheme illustrating how  $X_A, Z_A, X_B, Z_B$  are constructed from the observables of the stabilising operators.

we can be sure that for any  $p \in (0, 1)$ , the maximal eigenvalue is 1 and exclusively achieved by the state  $|\phi_\theta\rangle$ . Therefore  $\mathcal{B}_S$  is a good starting point for our search for a self-testing inequality. The difficult part is constructing a Bell inequality from the operator Eq. (3.16). In order to do so, we introduce the following operators:

$$X_A = \cos(\omega_1 + a)\sigma_z + \sin(\omega_1 + a)\sigma_x, \quad (3.17)$$

$$Z_A = \cos(\omega_1 - a)\sigma_z + \sin(\omega_1 - a)\sigma_x, \quad (3.18)$$

$$X_B = \cos(\omega_1 + \omega_2 + b)\sigma_z + \sin(\omega_1 + \omega_2 + b)\sigma_x, \quad (3.19)$$

$$Z_B = \cos(\omega_1 + \omega_2 - b)\sigma_z + \sin(\omega_1 + \omega_2 - b)\sigma_x. \quad (3.20)$$

Here,  $\omega_1, \omega_2 \in [0, 2\pi)$ ,  $\omega_1 + \omega_2 < 2\pi$  and  $a, b \in (0, \frac{\pi}{2})$ . These operators are depicted in Fig. 3.1. By introducing them, we now have five parameters  $(p, \omega_1, \omega_2, a, b)$  to optimize over, in order to find a suitable Bell operator  $\mathcal{B}_S(p, \omega_1, \omega_2, a, b)$ . To simplify our life, we fix the reference direction on Alice's side by choosing  $\omega_1 = \frac{\pi}{4}$ . The question now arises what the optimal choice for the remaining variables is and how to characterize "optimality". We do this by making use of the variational method introduced earlier.

### Applying Eigenvalue Perturbation

Our variational method relies on the following idea: One introduces small perturbations to the operators  $X_{A(B)}$  and  $Z_{A(B)}$  and studies how the maximal eigenvalue of the Bell operator  $\mathcal{B}_S$  changes as one departs from the perfect settings. In case of a Bell operator unsuitable for self-testing, the maximal eigenvalue of the perturbed Bell operator would *not* change or at least not significantly. This would imply that unintended settings and non-target states can reproduce the Bell violation of the perfect scenario. Hence one could not certify the state device-independently. What we aim for is that the maximal eigenvalue drops virtually equally

fast in every perturbation direction. Then Bell violations close to the quantum bound can only be achieved by states and measurements close to the desired ones; hence the state fidelity is high.

The perturbations are introduced as follows:

$$X_{A(B)} \rightarrow X_{A(B)} + \delta_{1(3)} X_{A(B)}^\perp - \frac{1}{2} \delta_{1(3)}^2 X_{A(B)}, \quad (3.21)$$

$$Z_{A(B)} \rightarrow Z_{A(B)} + \delta_{2(4)} Z_{A(B)}^\perp - \frac{1}{2} \delta_{2(4)}^2 Z_{A(B)}, \quad (3.22)$$

where  $X_B^\perp = \cos(\frac{\pi}{4} + \omega_2 + b)\sigma_x - \sin(\frac{\pi}{4} + \omega_2 + b)\sigma_z$ ,  $Z_B^\perp = \cos(\frac{\pi}{4} + \omega_2 - b)\sigma_x - \sin(\frac{\pi}{4} + \omega_2 - b)\sigma_z$ , and analogously for  $X_A^\perp$ ,  $Z_A^\perp$ .  $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3, \delta_4)$  is the vector of infinitesimal perturbation strengths.

In order for the desired settings and  $|\phi_\theta\rangle$  to achieve a local maximum with respect to the maximal eigenvalue of  $\mathcal{B}_S$ , the following conditions must be fulfilled (compare to Eqs (3.6) and (3.8)):

$$0 \stackrel{!}{=} \langle \phi_\theta | \left. \frac{\partial \mathcal{B}_S}{\partial \delta_i} \right|_{\boldsymbol{\delta} \rightarrow 0} | \phi_\theta \rangle, \quad i = 1, 2, 3, 4, \quad (3.23)$$

$$0 \stackrel{!}{\succeq} \gamma := \mu + \nu, \quad (3.24)$$

$$\mu_{ij} = \langle \phi_\theta | \left. \frac{\partial^2 \mathcal{B}}{\partial \delta_i \partial \delta_j} \right|_{\boldsymbol{\delta} \rightarrow 0} | \phi_\theta \rangle, \quad (3.25)$$

$$\nu_{ij} = 2 \sum_{k=1}^3 \frac{\langle \phi_\theta | \mathcal{B}_i^{(1)} | \phi_k \rangle \langle \phi_k | \mathcal{B}_j^{(1)} | \phi_\theta \rangle}{1 - \lambda_k}, \quad (3.26)$$

where  $|\phi_1\rangle = s_\theta|00\rangle - c_\theta|11\rangle$ ,  $|\phi_2\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$  and  $|\phi_3\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ . To be more precise, two of the perturbation directions are trivial and hence always two eigenvalues are zero. Thus for a true maximum, two of the four eigenvalues of  $\gamma$  must be negative.

The first order perturbations (3.23) for the observables of Alice ( $i = 1, 2$ ) thus result in the condition

$$0 = \sin(\frac{\pi}{4} - a)[1 - p(1 - \sin(2\theta)^2 - \frac{1}{2} \cos(2\theta)^2)]. \quad (3.27)$$

Since we require the weighting to satisfy  $0 < p < 1$ , the condition can only be satisfied if  $a = \frac{\pi}{4}$ . Taking this into account, the first order perturbations on Bob's observables ( $i = 3, 4$ ) result in the condition

$$0 = (c_{2b} + s_{2\omega_2})[1 - p + \frac{p}{2}c_{2\theta}^2] + (c_{2b} - s_{2\omega_2})ps_{2\theta}^2. \quad (3.28)$$

We know that in the limit  $\theta \rightarrow \frac{\pi}{4}$ , a perfect choice for  $\omega_2$  is  $\frac{\pi}{4}$ . Therefore we assume  $\omega_2 = \frac{\pi}{4}$ , which together with the constraint  $b < \frac{\pi}{2}$  results in the equation:

$$0 = (1 - p + \frac{p}{2}c_{2\theta}^2) - \tan^2(b)ps_{2\theta}^2 \quad (3.29)$$

$$\iff_{\theta, p, b > 0} b = \arctan \sqrt{\frac{1 - p + \frac{p}{2}c_{2\theta}^2}{ps_{2\theta}^2}}. \quad (3.30)$$

With these results,  $\gamma$  takes the following form:

$$\gamma = \frac{1}{2} \begin{pmatrix} \gamma_1 & \gamma_1 & 0 & 0 \\ \gamma_1 & \gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_2 & -\gamma_2 \\ 0 & 0 & -\gamma_2 & \gamma_2 \end{pmatrix}, \quad (3.31)$$

where  $\gamma_1$  and  $\gamma_2$  are lengthy expressions in  $p$  and  $\theta$ . The eigenvalues of  $\gamma$  are thus  $\{0, 0, \gamma_1, \gamma_2\}$ . Our studies show that if we choose  $p = \frac{1}{2}$ , i.e. an equal mixing of the two stabilising operators, then

$$\gamma_1 = \frac{(-7 + c_{4\theta})s_{2\theta}^2}{17 + c_{4\theta}}, \quad (3.32)$$

$$\gamma_2 = -\frac{(17 + c_{4\theta})s_{2\theta}^2}{8(5 + c_{4\theta})}. \quad (3.33)$$

This is a good result since  $|\gamma_1 - \gamma_2| < 0.01$  (see Fig. 3.2), meaning that the maximal eigenvalue drops in every non-trivial perturbation direction virtually equally fast. Also, importantly,  $\gamma_2 \leq \gamma_1 < 0$  for  $0 < \theta \leq \frac{\pi}{4}$  implying that  $\mathcal{B}_S$  has a local maximum for the desired settings together with  $|\phi_\theta\rangle$  for all partially-entangled states. Therefore we found a new Bell operator which is highly suitable for the task of device-independent state certification. It is given as

$$\mathcal{B}_S = \frac{1}{4} \left[ \frac{A_0(B_0 - B_1)}{\sin(b_\theta)} + \frac{\sin(2\theta)}{\cos(b_\theta)} A_1(B_0 + B_1) + \cos(2\theta) \left( A_0 + \frac{B_0 - B_1}{2 \sin(b_\theta)} \right) \right], \quad (3.34)$$

where  $b_\theta = \arctan \sqrt{\frac{1 + \frac{1}{2}c_{2\theta}^2}{s_{2\theta}^2}}$  and the quantum bound of 1 is achieved only when choosing the operators

$$A_0 = \sigma_z, \quad B_0 = \cos(b_\theta)\sigma_x + \sin(b_\theta)\sigma_z, \quad (3.35)$$

$$A_1 = \sigma_x, \quad B_1 = \cos(b_\theta)\sigma_x - \sin(b_\theta)\sigma_z. \quad (3.36)$$

The Bell inequality corresponding to Eq. (3.34) is denoted  $\mathcal{I}_\theta$  and we find that its local bound is achieved by the deterministic local strategy  $\{A_0 = A_1 = B_0 = 1, B_1 = -1\}$  and given as

$$\beta_L^{(S)} = \frac{1}{4} \left[ c_{2\theta} + (2 + c_{2\theta}) \sqrt{\frac{7 - c_{4\theta}}{5 + c_{4\theta}}} \right]. \quad (3.37)$$

The local bound of Eq. (3.37) is compared to the (normalized) local bound of the tilted-CHSH operator in Fig. 3.2. The plot reveals that  $\beta_L^{(T)} \leq \beta_L^{(S)}$ . However, our later studies in chapter 5 will show that the symmetric Bell operator allows for a more noise-tolerant state certification. For completeness, the individual state overlaps resulting from violations of the tilted-CHSH and our newly-derived symmetric inequality are compared in Fig. 3.3. The reason for the

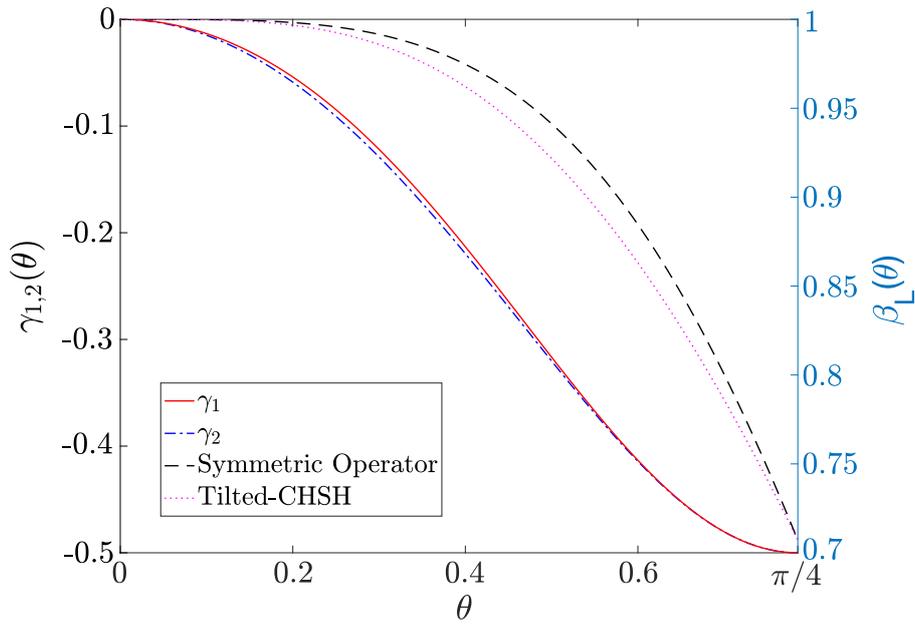


Figure 3.2: Plot of the eigenvalues of  $\gamma$  and of the local bounds  $\beta_L$  of  $\mathcal{B}_S$  and  $\mathcal{B}_T$ .

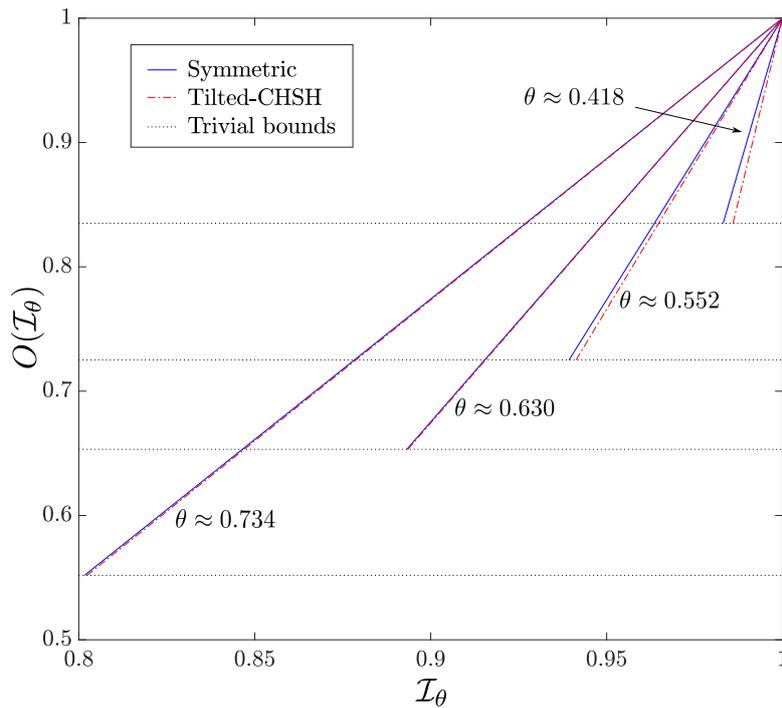


Figure 3.3: The overlap resulting from a violation of the tilted-CHSH inequality is compared to the overlap resulting from a violation of our new inequality. Despite the larger local bound (see Fig. 3.2), the symmetric inequality nevertheless achieves a larger overlap.

counter-intuitive result is the following: Local bounds only provide information on the distance between the target state (in our case  $|\phi_\theta\rangle$ ) and deterministic strategies (parallel measurement settings and product states). In self-testing, on the other hand, we need to bound the distance of the target state to arbitrary states for arbitrary measurements. Therefore it is crucial that the performance worsens drastically when departing from the perfect scenario. This essential condition can be imposed by following the perturbative approach described above.

Hence our method of eigenvalue perturbation is an ideal tool for the shaping of self-testing inequalities. It is easily generalized to more complex systems and it was a constitutive ingredient in our works on channel and measurement certification (chapters 4 and 5).

## SELF-TESTING OF QUANTUM CHANNELS

Quantum channels are inevitable in all quantum computational operations. They are present whenever information is transferred, stored or processed. Hence their certification is crucial for any computation.

In this chapter we demonstrate how this certification can be conducted device-independently. The central idea is formulated as follows: The action of a quantum channel is completely determined by its action on half a maximally-entangled state. Hence one can compare an experimental channel  $\mathcal{E}$  to a target channel  $\bar{\mathcal{E}}$  by comparing the two output states that result when applying the channels to such a maximally-entangled state. If these output states are the same, we can be sure the experimental channel works as modelled by the target channel.

However, there is an important detail we neglected so far. In order to certify the channel device-independently, we are not allowed to assume that we start with a maximally-entangled state but we have to self-test this state as well. We will show how two state certifications, namely of the maximally-entangled input state and the output state in case the channel is applied, can be combined to provide a lower bound on the fidelity of the channel. The lower bound is robust to noise - in the input as well as the output state.

We will illustrate our scheme by studying two relevant cases. The first one is the identity qubit-channel that is used to model quantum memories and transmission lines. The second one is a two-qubit controlled-unitary channel which is important for entangling operations. For both scenarios, we will provide robust lower bounds on the channel fidelity.



## Paper No. 2

# Certifying the Building Blocks of Quantum Computers from Bell's Theorem

Pavel Sekatski, Jean-Daniel Bancal, Sebastian Wagner, and Nicolas Sangouard

Physical Review Letters 121, 180505 (2018)



## Certifying the Building Blocks of Quantum Computers from Bell's Theorem

Pavel Sekatski,<sup>1,2,\*</sup> Jean-Daniel Bancal,<sup>1,\*</sup> Sebastian Wagner,<sup>1</sup> and Nicolas Sangouard<sup>1</sup>

<sup>1</sup>*Quantum Optics Theory Group, Universität Basel, Klingelbergstraße 82, CH-4056 Basel, Switzerland*

<sup>2</sup>*Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria*

 (Received 23 February 2018; published 2 November 2018)

Bell's theorem has been proposed to certify, in a device-independent and robust way, blocks either producing or measuring quantum states. In this Letter, we provide a method based on Bell's theorem to certify coherent operations for the storage, processing, and transfer of quantum information. This completes the set of tools needed to certify all building blocks of a quantum computer. Our method distinguishes itself by its robustness to experimental imperfections, and so could be used to certify that today's quantum devices are qualified for usage in future quantum computers.

DOI: 10.1103/PhysRevLett.121.180505

Experimental research on quantum computing is progressing at an unprecedented rate [1]. Five-qubit quantum computations combining around a dozen of quantum logical gates can nowadays be performed with a mean gate fidelity of  $\sim 98\%$  using trapped ions [2] or superconducting circuits [3]. However, for implementing large-scale quantum computation, it is crucial to proceed in a scalable way and certify that each new component is qualified for use in a quantum computer, independently of the purpose for which that larger device is used.

Such a certification must be device independent, that is, it cannot rely on a physical description of the actual implementation. Indeed, an exhaustive model of the setup is challenging, if not impossible, to establish. Relying on any particular model therefore amounts to making assumptions about the functioning of blocks. But seemingly harmless assumptions can have dramatic consequences when they are not perfectly satisfied. An assumption on the Hilbert space dimension, for example, can completely corrupt the security guarantees of a network of small quantum computers used to communicate securely [4,5]. Blocks certified in a device-dependent way thus cannot be used safely for arbitrary purposes.

Bell's theorem [6] has led to device-independent certification schemes for components either producing quantum states or performing quantum measurements [7–18]. But these are just some of the elementary blocks needed to build a quantum computer (see Fig. 1). In particular, a device-independent method that can be used in present-day experiments for assessing the quality of components in charge of the transfer, processing, and storage of quantum information is still missing. Together with existing techniques, such a method would, in principle, allow for the certification of all kinds of elementary building blocks needed in a quantum computer.

Here, we show how to certify a trace preserving quantum channel acting on one or several systems, that is, a general

transformation taking quantum states and returning other quantum states. Our approach involves no description of the internal functioning of either the tested channel or the certification setup, but relies on the device-independent characterization of two entangled states, the first one serving as input to the channel, the second one being the output state. Interestingly, we can use state certifications that are robust to experimental imperfections to certify channels robustly.

Our goal is in sharp contrast with a line of research aiming to certify quantum computations [21–25]. Our work addresses elementary blocks of a quantum computer and certifies that they are qualified for use in future larger quantum devices. It builds on the work of Magniez *et al.* [26], but differs (i) in its formulation, (ii) methodology, and (iii) robustness. In particular, (i) we show how to use the device to be certified to perform the desired operation between well-identified subspaces and subsystems with

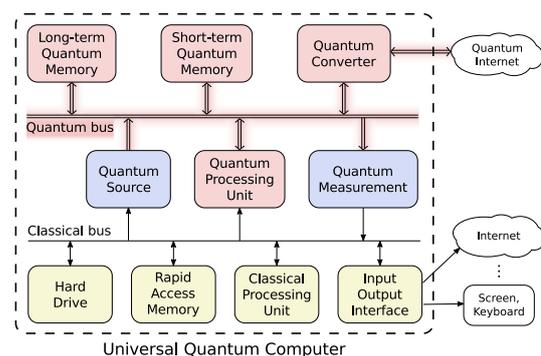


FIG. 1. Possible architecture of a future universal quantum computer (see also Refs. [19,20]). Elements in yellow are classical, and thus well characterized. Blue elements already admit device-independent certification schemes. Here, we demonstrate how to certify the components in red. In practice, several blocks may be merged into a single physical unit.

predefined Hilbert space dimensions, (ii) our recipe does not require two copies of the box to be certified and (iii) the robustness of our results is compatible with current technological capabilities. In opposition to Ref. [27], we provide lower bounds on the quality of the blocks. Detailed recipes are given to certify the unitarity of one-qubit channels as well as two-qubit entangling operations. These recipes could be used in present-day experiments to certify transmission lines between processing and storage areas, storage devices, converters between various information carriers, and arbitrary two-qubit controlled-unitary gates independently of the details and imperfections of the actual implementation.

*Device-independent certification of a quantum channel.*— We start by providing a definition of the device-independent certification of quantum channels. For this, we consider a scenario with two sides  $\mathcal{A}$  and  $\mathcal{B}$ , each side containing potentially several parties depending on the channel to be certified. Each party performs measurements on one part of a shared state  $\rho \in L(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}})$  and records the result of each experimental run. In addition, the parties on side  $\mathcal{A}$  have the freedom to decide whether or not to apply the channel to be certified  $\mathcal{E}$ , an endomorphism on states in  $\mathcal{H}_{\mathcal{A}}$ , before performing the measurements [see Figs. 3(a) and 3(c)]. The sources preparing the initial state, the measurement devices, and the channel are treated as black boxes and the parties do not communicate with each other. The partial state prepared by the source at side  $\mathcal{A}$  is denoted  $\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}}\rho$ .

We say that the channel  $\mathcal{E}$  is certified device independently if the sole knowledge of the results given the measurement choices implies the existence of local isometries  $\Phi_i: \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i \rightarrow \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i^{\text{ext}}$  and  $\Phi_o: \mathcal{H}_{\mathcal{A}} \rightarrow \mathcal{H}_o \otimes \mathcal{H}_o^{\text{ext}}$ , such that

$$\begin{aligned} & (\Phi_o \circ \mathcal{E} \circ \Phi_i \otimes \mathbb{1})[\rho_{\mathcal{A}} \otimes |\phi^+\rangle\langle\phi^+|] \\ &= (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|] \otimes \rho_{\text{ext}}^{(i,o)}, \end{aligned}$$

where  $\bar{\mathcal{E}}$  is the reference channel mapping states from Hilbert space  $\mathcal{H}_i$  to the Hilbert space  $\mathcal{H}_o$ . Here,  $|\phi^+\rangle$  is a maximally entangled state in  $\mathcal{H}_i \otimes \mathcal{H}_i$ , and  $\rho_{\text{ext}}^{(i,o)}$  is some irrelevant residual state on  $\mathcal{H}_i^{\text{ext}} \otimes \mathcal{H}_o^{\text{ext}}$ . We emphasize that in device-independent certification, assumptions are made neither on the system's state on which  $\mathcal{E}$  operates, nor on the dimension of the underlying Hilbert space. The local isometries  $\Phi_i$  and  $\Phi_o$  identify subspaces and subsystems in which the channel  $\mathcal{E}$  acts exactly as the reference channel  $\bar{\mathcal{E}}$ .

When the above equality does not hold exactly we quantify the relation between the channels  $\mathcal{E}$  and  $\bar{\mathcal{E}}$  through the following fidelity

$$\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}}) = \max_{\Lambda_i, \Lambda_o} F((\Lambda_o \circ \mathcal{E} \circ \Lambda_i \otimes \mathbb{1})[\rho_{\mathcal{A}} \otimes |\phi^+\rangle\langle\phi^+|], \bar{\rho}),$$

$$(1) \quad F^o = F((\Lambda_o^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}})[(\mathcal{E} \otimes \mathbb{1})[\rho]], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]). \quad (3)$$

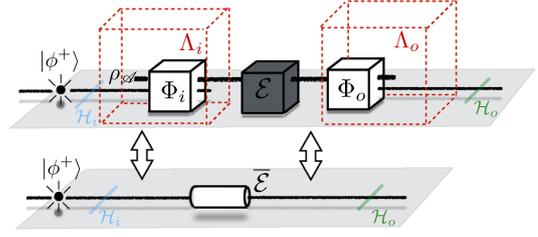


FIG. 2. Comparison between an unknown channel  $\mathcal{E}$  and a reference channel  $\bar{\mathcal{E}}$  operating on a Hilbert space  $\mathcal{H}_i$ . Half a maximally entangled state belonging to  $\mathcal{H}_i \otimes \mathcal{H}_i$  is presented to  $\mathcal{E}$  by a local map  $\Lambda_i$ , which can also act on the initial quantum state  $\rho_{\mathcal{A}}$ . Degrees of freedom that are not transmitted to the channel at this point are discarded. A local map  $\Lambda_o$  is then used at the output of the channel  $\mathcal{E}$  to remove extra systems and extract the state of a subsystem to be compared with the Choi state  $\bar{\rho} = (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]$  of the reference channel. The channel fidelity  $\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})$  is then obtained by maximizing the overlap between  $\bar{\rho}$  and the channel output over all possible input isometries and output maps.

Here,  $F(\rho, \sigma) = \text{Tr}(\sqrt{\sigma^{1/2}\rho\sigma^{1/2}})$  is the Uhlmann fidelity.  $\mathbb{1}$  acts on the second half of  $|\phi^+\rangle$ .  $\Lambda_i[\cdot] = \text{Tr}_{\mathcal{H}_i^{\text{ext}}}(\Phi_i[\cdot])$  traces out all degrees of freedom which are not in the preimage of  $\mathcal{E}$  while  $\Lambda_o[\cdot] = \text{Tr}_{\mathcal{H}_o^{\text{ext}}}(\Phi_o[\cdot])$  traces out all degrees of freedom which are not in the image of  $\bar{\mathcal{E}}$ .  $\bar{\rho} = (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]$ . (See Fig. 2 and Supplemental Material A.1 for details [28].)

This fidelity, which is optimized over all maps, can be understood as an extension of the Choi fidelity to device-independent scenarios. It guarantees that the channel  $\mathcal{E}$  can be used to play the role of  $\bar{\mathcal{E}}$  in any circumstance with fidelity  $\mathcal{F}$ . The maps achieving this fidelity describe the recipe for how to do that. Furthermore, the fidelity  $\mathcal{F}$  of Eq. (1) can be used to bound the distance between the two channels through the diamond norm, which informs us on the highest probability to distinguish the two channels in a single shot upon acting on arbitrary states [31]; see the Supplemental Material A.3 [28].

In the case where the target channel  $\bar{\mathcal{E}}$  acts on several parties, we distinguish these parties  $\{A^{(k)}\}$  on the side  $\mathcal{A}$ . The input and output Hilbert spaces then have a tensor structure  $\mathcal{H}_{i/o} = \otimes_{k=1}^n \mathcal{H}_{i/o}^{(k)}$  and the same is required from the maps  $\Lambda_i$  and  $\Lambda_o$ , as spelled out in Supplemental Material A.2 [28].

*A practical device-independent bound on the channel fidelity.*— We show that a channel certificate can be obtained by combining two certifications, one for the state serving as input of the channel and one for the output state, that is,

$$F^i = F((\tilde{\Lambda}_i^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}})[\rho], |\phi^+\rangle\langle\phi^+|), \quad (2)$$

$F^i$  corresponds to the fidelity of the input state  $\rho$  with respect to the maximally entangled state  $|\phi^+\rangle$ .  $F^o$  is the fidelity of the output state with respect to the image of  $|\phi^+\rangle$  under the reference channel. As before, the role of the maps  $\tilde{\Lambda}_i^{\mathcal{A}}$ ,  $\Lambda_o^{\mathcal{A}}$ , and  $\Lambda^{\mathcal{B}}$  is to identify subspaces where the system states and the reference states can be compared, and the underlying isometries are enforced to have a product structure with respect to the partition of  $\mathcal{A}$  into separate parties.

The triangle and processing inequalities for the fidelity as well as properties of the isometries in Eqs. (2)–(3) allow one to show that the device independent Choi fidelity given in Eq. (1) can be bounded by

$$\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}}) \geq \cos[\arccos(F^i) + \arccos(F^o)]. \quad (4)$$

Importantly, the bound holds for channels acting on several parties, in which case the states in Eqs. (2)–(3) are multipartite and the maps  $\tilde{\Lambda}_i^{\mathcal{A}}$  and  $\Lambda_o^{\mathcal{A}}$  are products of local maps for each party, see Supplemental Material B [28].

Formula (4) shows how two channels can be compared even though they operate on Hilbert spaces with (possibly unknown) different dimensions. This relation is made possible by the fact that the map  $\Lambda^{\mathcal{B}}$  is identical in both equations Eqs. (2) and (3). One way to guarantee that the map is the same is to obtain certificates for both states with the same measurement boxes on side  $\mathcal{B}$ . If this is fulfilled, a robust bound on the channel fidelity is obtained as soon as the input and output states are certified robustly. Interestingly, there are several known results and methods for state certification that are robust to noise [11–16,18].

We show how Eq. (4) can be used for the robust certification of (i) a one-qubit unitary and (ii) two-qubit quantum logical gates.

*Device-independent certification of a single-qubit unitary channel.*—Memories such as hard drives and RAM units, transmission lines between different units of a computer, and converters between different information carriers are elements mappings input to output qubits, either separated in time or space or carried by different physical systems, which are all ideally modeled by the identity channel. Applying the formalism presented earlier to  $\bar{\mathcal{E}} = \mathbb{1}$  in dimension two, involves ideally a maximally entangled two-qubit state as input state [see Fig. 3(a)]. As the reference channel does not alter the input, we assess the fidelities of both input and output with the Clauser-Horne-Shimony-Holt (CHSH) test [32]. The condition that  $\Lambda^{\mathcal{B}}$  is identical in both situation is then naturally satisfied. Given the CHSH values  $\beta^{i/o}$ , it is possible to bound the state fidelity as [14]

$$F^{i/o} \geq F_{\text{CHSH}} = \sqrt{\frac{1}{2} \left( 1 + \frac{\beta^{i/o} - \beta^*}{2\sqrt{2} - \beta^*} \right)}, \quad (5)$$

where  $\beta^* = [2(8 + 7\sqrt{2})/17] \approx 2.11$ . Inserting these fidelities into Eq. (4), yields a robust device-independent

certification of one-qubit unitaries depicted in Fig. 3(b). Examples confirming the robustness can be found in Supplemental Material C.

Note that testing the input state is not necessary for the certification of a unitary channel. Indeed one can see the channel itself as part of the local isometry. Hence, it is always possible to define  $\tilde{\Lambda}_i^{\mathcal{A}}$  such that the fidelity of the input state is at least as large as the output fidelity, i.e.,  $F^i \geq F^o$ . This relation together with Eq. (4) give a bound on the channel fidelity  $\mathcal{F} \geq 2(F^o)^2 - 1$  in terms of the output fidelity alone.

*Device-independent certification of two-qubit entangling channels.*—Entangling gates are necessary for any non-trivial manipulation and sufficient to enable universal quantum computation [33]. We present a setup that allows for the certification of an arbitrary two-qubit controlled-unitary gate. Such a gate can be put in the form

$$CU_\varphi = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes e^{-i\varphi Y}. \quad (6)$$

$CU_{(\pi/2)}$  is equivalent to the controlled-NOT gate while  $CU_0$  is the two-qubit identity channel.

In order to bound the fidelity of an actual gate with the bipartite  $CU_\varphi$  gate, we need to split side  $\mathcal{A}$  into two parties  $A^{(1)}$  and  $A^{(2)}$ . Similarly, we also split side  $\mathcal{B}$  into  $B^{(1)}$  and  $B^{(2)}$  so that sharing a maximally entangled state of dimension 4 between  $\mathcal{A}$  and  $\mathcal{B}$  amounts to sharing two-qubit maximally entangled states  $|\phi_2^+\rangle$  between  $A^{(1)}$  and  $B^{(1)}$  and between  $A^{(2)}$  and  $B^{(2)}$ ; cf. Fig. 3(c). As we show now, four-partite statistics obtained after parties  $A^{(1)}$  and  $A^{(2)}$  jointly decide to use the device which supposedly performs the  $CU_\varphi$  gate on their systems or not can lead to the certification of this gate.

The first step consists of deriving a new family of Bell inequalities suitable for the certification of the input  $|\phi_2^+\rangle^{\otimes 2}$  and output state  $(CU_\varphi \otimes \mathbb{1}_{\mathcal{B}})|\phi_2^+\rangle^{\otimes 2}$ , that is, for an arbitrary state of the form  $|\xi_\varphi\rangle = (CU_\varphi \otimes \mathbb{1}_{\mathcal{B}})|\phi_2^+\rangle^{\otimes 2}$ . We consider the case where each party has a measurement box with two inputs and two outcomes. Let  $B_\varphi$  be a family of Bell expressions, i.e., a weighted sum of expectation values of measurement outcomes whose coefficients depend on  $\varphi$ . Let  $B_\varphi^Q$  be the operator obtained by replacing the inputs in  $B_\varphi$  by quantum observables corresponding to projections into directions such that the unique  $\max_\rho \text{Tr}(B_\varphi^Q \rho)$  is obtained for  $\rho = |\xi_\varphi\rangle\langle \xi_\varphi|$ . To construct  $B_\varphi^Q$ , we consider a set of Hermitian operators having the state  $|\xi_\varphi\rangle$  for unique maximal eigenstate. These operators are obtained by applying a gate LU equivalent to  $(CU_\varphi \otimes \mathbb{1}_{\mathcal{B}})$  on convex sums of stabilizers of the state  $|\phi_2^+\rangle^{\otimes 2}$ . We find  $B_\varphi$ , i.e., the proper correspondence between the measurement inputs and the Pauli matrices, by requesting the maximum eigenvalue of the operator to be a local maximum with respect to small

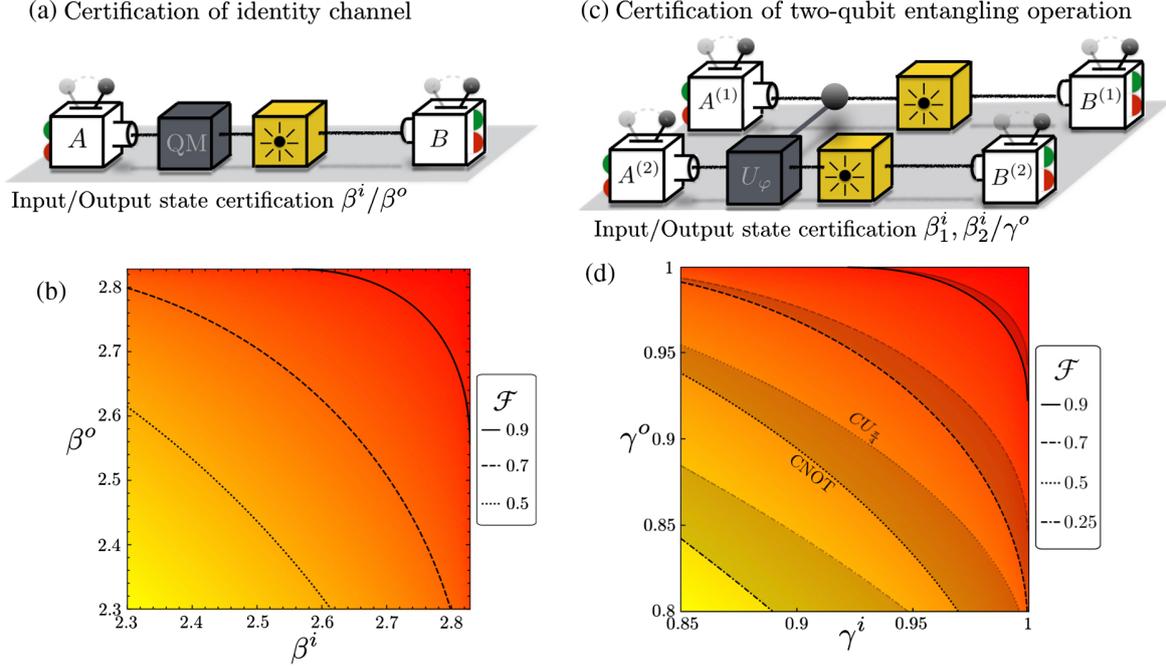


FIG. 3. Certification of the one-qubit identity channel [(a) and (b)] and two-qubit entangling operation [(c) and (d)]. (a) The certification of the identity in dimension 2 uses a source (yellow box) producing ideally a maximally two-qubit entangled state. The measurement devices (white boxes)  $A$  and  $B$  are used to perform two CHSH tests, with and without the tested device (black box). The sole knowledge of two CHSH values  $\beta^i$  and  $\beta^o$  gives a bound on the fidelity  $\mathcal{F}$  of the tested device with respect to the identity. (b) Robustness of the qubit identity certification as a function of the two CHSH values (color is a guide for the eye). (c) The certification of a two-qubit entangling operation uses a source (here represented with two yellow boxes) ideally producing two maximally entangled two-qubit states. Four measurement devices are used to perform Bell tests with and without the gate to be certified. The two Bell values  $\beta_1^i$  and  $\beta_2^i$  obtained to certify the two states produced by the source and the one obtained at the output of the gate (black box)  $\gamma^o$  are used to bound the fidelity of any two-qubit controlled-unitary gates. (d) Robustness of the certification of two-qubit controlled-unitary operations (color is a guide for the eye). The best robustness is obtained for a class of gates including the controlled-NOT gate (CNOT). The gray lines show the worst case. The greenish area thus includes all two-qubit controlled-unitary gates.

perturbations of the measurement inputs, see Supplemental Material D.4 [28].

The resulting Bell expressions  $B_\varphi$  have two inputs and two outputs per party. This allows us to make use of Jordan's Lemma in order to quantify their self-testing property, that is, we look for bounds on the fidelity assuming that qubit measurements are performed locally. If the extraction isometries only depend on local measurement settings and the square of the obtained fidelity bounds are convex functions of the mean value of the Bell operator, they automatically hold independently of the dimension [34]; see the Supplemental Material D.2 [28].

We find such bounds by using the isometries proposed in Ref. [14], which are known to provide very robust results for the singlet state. To do so we look for the state and measurement settings that minimize the fidelity of the extracted four qubit state with respect to  $|\xi_\varphi\rangle$  ( $|\phi_2^+\rangle^{\otimes 2}$ ) while keeping a fixed expectation value  $\gamma^o$  ( $\gamma^i$ ) of the Bell operator  $B_\varphi$  ( $B_0$ ), cf. Supplemental Material D.3 [28]. The resulting bound on the fidelity is given by

$$F^{i/o} \geq \sqrt{\frac{1}{2} \left( 1 + \frac{\gamma^{i/o} - \gamma^*}{1 - \gamma^*} \right)}, \quad (7)$$

where  $\gamma^*$  is a constant that could, in principle, depend on the gate to be tested. This constant is upper bounded by 0.85 for all  $\varphi$ , cf. Supplemental Material D.5 [28]. Note that our approach to find Bell inequalities and deduce the corresponding robust fidelity bounds is applicable to other  $N$ -qubit states.

Given the bounds on the fidelity  $F^i$  of the initial state and on the fidelity  $F^o$  of the output state, and checking that they have been obtained with common measurements for parties  $B^{(1)}$  and  $B^{(2)}$ , we get from Eq. (4) a bound on the fidelity between the actual gate  $\mathcal{E}$  and the reference gate  $\tilde{\mathcal{E}} = CU_\varphi$ . The result is shown in Fig. 3(d) as a function of the observed Bell values. Examples illustrating the robustness can be found in Supplemental Material C [28].

In analogy with the one-qubit identity certification, it is possible to prove that the actual two-qubit gate acts

as a global unitary on side  $\mathcal{A}$  from  $F^o$  only using  $\mathcal{F} \geq 2(F^o)^2 - 1$ . This information alone is, however, not sufficient to identify the gates  $CU_\varphi$  up to local isometries without additional assumptions, because the final state  $|\xi_\varphi\rangle$  could be directly prepared by the source and merely transmitted by the device to be certified.

*Discussions.*—We have introduced a framework for the device-independent certification of quantum channels. We applied our methods to individual elements of quantum computers, namely, single qubit identity channels and two qubit controlled unitary operations. Our technique does not certify the proper functioning of composite circuits but is the first necessary verification step and the relevant one given the status of on-going experiments. This is also relevant in the long term as our technique could be used to identify the elements causing the failure of a quantum computation.

This work was supported by the Swiss National Science Foundation (SNSF) through Grants No. PP00P2-179109, No. P300P2-167749, and No. 200021-175527. We also acknowledge the Army Research Laboratory Center for Distributed Quantum Information via the project SciNet.

\*P. S. and J.-D. B. contributed equally to this work.

- [1] T. D. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. L. O'Brien, *Nature (London)* **464**, 45 (2010).
- [2] S. Debnath, N. M. Linke, C. Figgatt, K. A. Landsman, K. Wright, and C. Monroe, *Nature (London)* **536**, 63 (2016).
- [3] R. Barends *et al.* *Nature (London)* **508**, 500 (2014).
- [4] A. Acín, N. Gisin, and L. Masanes, *Phys. Rev. Lett.* **97**, 120405 (2006).
- [5] L. Lydersen, C. Wiechers, C. Wittmann, D. Elser, J. Skaar, and V. Makarov, *Nat. Photonics* **4**, 686 (2010).
- [6] J. S. Bell, *Physics* **1**, 195 (1964).
- [7] S. Popescu and D. Rohrlich, *Phys. Lett. A* **169**, 411 (1992).
- [8] S. L. Braunstein, A. Mann, and M. Revzen, *Phys. Rev. Lett.* **68**, 3259 (1992).
- [9] D. Mayers and A. Yao, *Quantum Inf. Comput.* **4**, 273 (2004).
- [10] M. Tomamichel and E. Hänggi, *J. Phys. A* **46**, 055301 (2013).
- [11] C.-E. Bardyn, T. C. H. Liew, S. Massar, M. McKague, and V. Scarani, *Phys. Rev. A* **80**, 062327 (2009).
- [12] C. Miller and Y. Shi, *8th Conference on the Theory of Quantum Computation, Communication and Cryptography, TQC 2013, Guelph, Ontario, Canada*, Vol. 22 (Schloss Dagstuhl- Leibniz-Zentrum für Informatik GmbH, Dagstuhl Publishing, 2013), p. 254.
- [13] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués, *Phys. Rev. Lett.* **113**, 040401 (2014).
- [14] J. Kaniewski, *Phys. Rev. Lett.* **117**, 070402 (2016).
- [15] C. Bamps and S. Pironio, *Phys. Rev. A* **91**, 052111 (2015).
- [16] A. Natarajan and T. Vidick, *Proc. of STOC* **17**, 1003 (2017).
- [17] A. Coladangelo, K. T. Goh, and V. Scarani, *Nat. Commun.* **8**, 15485 (2017).
- [18] I. Šupić, A. Coladangelo, R. Augusiak, and A. Acín, *New J. Phys.* **20**, 083041 (2018).
- [19] M. Mariantoni, H. Wang, T. Yamamoto, M. Neeley, R. C. Bialczak, Y. Chen, M. Lenander, E. Lucero, A. D. O'Connell, D. Sank, M. Weides, J. Wenner, Y. Yin, J. Zhao, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, *Science* **334**, 61 (2011).
- [20] G. K. Brennen, D. Song, and C. J. Williams, *Phys. Rev. A* **67**, 050302(R) (2003).
- [21] M. McKague, *Theory Comput.* **12**, 1 (2016).
- [22] B. W. Reichard, F. Unger, and U. Vazirani, *Nature (London)* **496**, 456 (2013).
- [23] M. Hajdušek, C. A. Pérez-Delgado, and J. F. Fitzsimons, *arXiv:1502.02563*.
- [24] J. F. Fitzsimons, M. Hajdušek, and T. Morimae, *Phys. Rev. Lett.* **120**, 040501 (2018).
- [25] A. Coladangelo, A. Grilo, S. Jeffery, and T. Vidick, *arXiv:1708.07359*.
- [26] F. Magniez, D. Mayers, M. Mosca, and H. Ollivier, in *Proceedings of the 33rd International Colloquium on Automata*, Lang Lecture Notes in Computer Science No. 4052, edited by M. Bugliesi (Springer, Berlin, Heidelberg, 2006), p. 72.
- [27] M. Dall'Arno, S. Brandsen, and F. Buscemi, *Proc. R. Soc. A* **473**, 20160721 (2017).
- [28] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.121.180505> for how to certify channels device independently and give all the proofs absent in the main text. We also show the extension to multipartite scenarios and demonstrate how one can bound the channel fidelity from state fidelities. We investigate the robustness of these state fidelities to noise by devising Bell tests tailored to the states, which includes Refs. [29,30].
- [29] P. Sekatski *et al.* (to be published).
- [30] P. E. M. F. Mendonca, R. d. J. Napolitano, M. A. Marchioli, C. J. Foster, and Y.-C. Liang, *Phys. Rev. A* **78**, 052330 (2008).
- [31] Avraham Ben-Aroya and Amnon Ta-Shma, *Quantum Inf. Comput.* **10**, 77 (2010).
- [32] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
- [33] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [34] V. Scarani, *Acta Phys. Slovaca* **62**, 347 (2012).

# Supplemental Material : Certifying the building blocks of quantum computers from Bell's theorem

Pavel Sekatski,<sup>1,2,\*</sup> Jean-Daniel Bancal,<sup>1,\*</sup> Sebastian Wagner,<sup>1</sup> and Nicolas Sangouard<sup>1</sup>

<sup>1</sup>Quantum Optics Theory Group, Universität Basel,  
Klingelbergstraße 82, CH-4056 Basel, Switzerland

<sup>2</sup>Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 21a, A-6020 Innsbruck, Austria

(Dated: October 8, 2018)

**Appendix A** *Device-independent certification of quantum channels*– In this appendix we give a thorough description of the device-independent channel certification introduced in the main text. There are three sections. Section A.1 provides a detailed definition of the device-independent certification of quantum channels and comments on Eq. (1) of the main text. Section A.2 addresses the extension to multi-party scenarios. Section A.3 shows how the Choi fidelity bounds the diamond norm between the two channels.

**Appendix A.1** *Formal definition of device-independent channel certification*– In full generality, a quantum channel  $\mathcal{E}$  is a completely positive trace preserving map between linear operators on two Hilbert spaces

$$\mathcal{E} : L(\mathcal{H}) \rightarrow L(\mathcal{H}'). \quad (1)$$

It maps a quantum state  $\rho$  to

$$\mathcal{E}[\rho] = \sum_{i=1}^I K_i \rho K_i^\dagger, \quad (2)$$

where the Kraus operators  $\{K_i\}_{i=1\dots I}$  are represented by  $\dim(\mathcal{H}') \times \dim(\mathcal{H})$  complex matrices satisfying the relation  $\sum_i K_i^\dagger K_i = \mathbb{1}$ . In the following we will assume that the input and output Hilbert spaces are the same  $\mathcal{H} = \mathcal{H}' = \mathcal{H}_{\mathcal{A}}$ . However, all the results can be straightforwardly generalized to the case where they do not match. In a bi-partite Bell-type scenario with measurement observables  $M_{a|x}$  and  $M_{b|y}$  corresponding to input  $x$  on side  $\mathcal{A}$  and  $y$  on side  $\mathcal{B}$  and outcome  $a$  and  $b$  respectively, we say that a behavior  $P$  certifies device-independently a reference channel  $\bar{\mathcal{E}} : L(\mathcal{H}_i) \rightarrow L(\mathcal{H}_o)$  if, for every quantum realization  $(\rho, \{M_{a|x}, M_{b|y}\}, \mathcal{E})$  compatible with  $P$ , there exist two local isometries

$$\Phi_i : \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i \rightarrow \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i^{\text{ext}}, \quad (3)$$

$$\Phi_o : \mathcal{H}_{\mathcal{A}} \rightarrow \mathcal{H}_o \otimes \mathcal{H}_o^{\text{ext}} \quad (4)$$

such that

$$\begin{aligned} \text{Tr}_{\text{ext}} \left( (\Phi_o \circ \mathcal{E} \circ \Phi_i \otimes \mathbb{1}) [\rho_{\mathcal{A}} \otimes |\phi^+\rangle\langle\phi^+|] \right) \\ = (\bar{\mathcal{E}} \otimes \mathbb{1}) [|\phi^+\rangle\langle\phi^+|], \end{aligned} \quad (5)$$

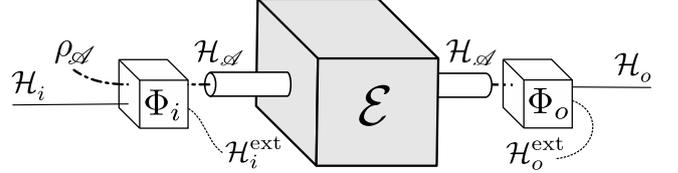


FIG. 1. The isometries  $\Phi_i$  and  $\Phi_o$  give a recipe of how the channel  $\mathcal{E}$  can be used in order to perform the desired operation  $\bar{\mathcal{E}}$  between  $\mathcal{H}_i$  and  $\mathcal{H}_o$ .

where the trace of over all the external subsystems, i.e. on  $\mathcal{H}_i^{\text{ext}} \otimes \mathcal{H}_o^{\text{ext}}$ . Here,  $|\phi^+\rangle$  is the maximally entangled state in  $\mathcal{H}_i \otimes \mathcal{H}_i$ . The maps are applied on the first Hilbert space and the identities on the second one.  $\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}} \rho$  is the unknown state prepared by the source on side  $\mathcal{A}$ . The role of the isometries is to identify a subspace for the channel input and a subsystem for the channel output between which the channel  $\mathcal{E}$  acts as desired. To put it differently, the isometries define a recipe

$$\begin{aligned} \mathcal{E}_{i-o} : L(\mathcal{H}_i) \rightarrow L(\mathcal{H}_o) \\ \varrho \mapsto \text{Tr}_{\text{ext}} \left( (\Phi_o \circ \mathcal{E} \circ \Phi_i) [\rho_{\mathcal{A}} \otimes \varrho] \right) \end{aligned} \quad (6)$$

of how the channel  $\mathcal{E}$  can be used in order to perform the desired operation, i.e.  $\mathcal{E}_{i-o} = \bar{\mathcal{E}}$ , as depicted in Fig. 1.

When the equality presented before does not hold exactly, we would naturally define the distance between the actual channel and the reference one through

$$\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}}) = \max_{\Phi_o, \Phi_i} F \left( (\mathcal{E}_{i-o} \otimes \mathbb{1}) [|\phi^+\rangle\langle\phi^+|], \bar{\rho} \right) \quad (7)$$

where  $F(\rho, \sigma) = \text{Tr} \left( \sqrt{\sigma^{1/2} \rho \sigma^{1/2}} \right)$  is the Uhlmann fidelity and  $\bar{\rho} = (\bar{\mathcal{E}} \otimes \mathbb{1}) [|\phi^+\rangle\langle\phi^+|]$  is the target state. Note that the fidelity is symmetric.

In Proposition .2 below, we prove that the fidelity (7) can be related to a Uhlmann fidelity computed on the full Hilbert space, including the external systems  $\rho_{\text{ext}} \in L(\mathcal{H}_i^{\text{ext}} \otimes \mathcal{H}_o^{\text{ext}})$ :

$$\begin{aligned} \mathcal{F}(\mathcal{E}, \bar{\mathcal{E}}) \geq \\ \max_{\Phi_o, \Phi_i, \rho_{\text{ext}}} F \left( (\Phi_o \circ \mathcal{E} \circ \Phi_i) \otimes \mathbb{1} [\rho_{\mathcal{A}} \otimes |\phi^+\rangle\langle\phi^+|], \bar{\rho} \otimes \rho_{\text{ext}} \right), \end{aligned} \quad (8)$$

with equality when the target channel  $\bar{\mathcal{E}}$  is unitary (i.e. the target state  $\bar{\rho}$  is pure). This last expression is closer to the original definition of quantum state certification [1]. When the target channel  $\bar{\mathcal{E}}$  is not unitary, however, inequality (8) is not tight, and the expressions in (7) better captures the relation between channels  $\mathcal{E}$  and  $\bar{\mathcal{E}}$ . To see this, consider the example where the reference channel  $\bar{\mathcal{E}}$  is a totally depolarizing single-qubit channel. The channel to be certified  $\mathcal{E}$  can implement this depolarizing channel by entangling the system qubit with an external qubit. While  $\mathcal{E}$  and  $\bar{\mathcal{E}}$  operate identically on the system,  $\mathcal{E}$  outputs a pure global state. Moreover, the latter is entangled and the maximal fidelity cannot be obtained by optimizing it over a product state  $\bar{\rho} \otimes \rho_{\text{ext}}$ .

To shorten the notation we define the injection map

$$\Lambda_i : L(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i) \rightarrow L(\mathcal{H}_{\mathcal{A}}) \quad (9)$$

$$\varrho \mapsto \text{Tr}_{\mathcal{H}_i^{\text{ext}}}(\Phi_i[\varrho]) \quad (10)$$

and extraction map

$$\Lambda_o : L(\mathcal{H}_{\mathcal{A}}) \rightarrow L(\mathcal{H}_o) \quad (11)$$

$$\varrho \mapsto \text{Tr}_{\mathcal{H}_o^{\text{ext}}}(\Phi_o[\varrho]) \quad (12)$$

which allow to simply write  $\mathcal{E}_{i-o}[\bullet] = \Lambda_o \circ \mathcal{E} \circ \Lambda_i[\rho_{\mathcal{A}} \otimes \bullet]$ .

**Lemma .1.** *Given three quantum states  $\rho \in L(\mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{ext}})$ ,  $\bar{\rho} = |\psi\rangle\langle\psi| \in L(\mathcal{H}_{\text{sys}})$  and  $\sigma \in L(\mathcal{H}_{\text{ext}})$  the following relation holds:*

$$F(\rho, \bar{\rho} \otimes \sigma) = F(\rho_{\text{sys}}, \bar{\rho}) F(\varrho_{\text{ext}}, \sigma), \quad (13)$$

where  $\rho_{\text{sys}} = \text{Tr}_{\text{ext}}(\rho)$  and  $\varrho_{\text{ext}} = \frac{\text{Tr}_{\text{sys}}(\rho \bar{\rho} \otimes \mathbf{1})}{\text{Tr}(\rho \bar{\rho} \otimes \mathbf{1})}$ .

*Proof.* Let us first note that  $|\psi\rangle\langle\psi|^{\frac{1}{2}} = |\psi\rangle\langle\psi|$  and expand the fidelity:

$$\begin{aligned} F(\rho, \bar{\rho} \otimes \sigma) &= \text{Tr} \left( (|\psi\rangle\langle\psi| \otimes \sigma)^{\frac{1}{2}} \rho (|\psi\rangle\langle\psi| \otimes \sigma)^{\frac{1}{2}} \right)^{\frac{1}{2}} \\ &= \text{Tr} \left( (\mathbf{1} \otimes \sqrt{\sigma}) (|\psi\rangle\langle\psi| \otimes \mathbf{1}) \rho (|\psi\rangle\langle\psi| \otimes \mathbf{1}) (\mathbf{1} \otimes \sqrt{\sigma}) \right)^{\frac{1}{2}} \\ &= \text{Tr} \left( (\mathbf{1} \otimes \sqrt{\sigma}) (|\psi\rangle\langle\psi| \otimes \text{Tr}_{\text{sys}}(\rho |\psi\rangle\langle\psi| \otimes \mathbf{1})) (\mathbf{1} \otimes \sqrt{\sigma}) \right)^{\frac{1}{2}} \\ &= \text{Tr}_{\text{sys}}(|\psi\rangle\langle\psi|)^{\frac{1}{2}} \text{Tr}_{\text{ext}} \left( \sqrt{\sigma} \text{Tr}_{\text{sys}}(\rho |\psi\rangle\langle\psi| \otimes \mathbf{1}) \sqrt{\sigma} \right)^{\frac{1}{2}} \\ &= \sqrt{\text{Tr}(\rho |\psi\rangle\langle\psi| \otimes \mathbf{1})} \text{Tr}_{\text{ext}} \left( \underbrace{\sqrt{\sigma} \frac{\text{Tr}_{\text{sys}}(\rho |\psi\rangle\langle\psi| \otimes \mathbf{1})}{\text{Tr}(\rho |\psi\rangle\langle\psi| \otimes \mathbf{1})} \sqrt{\sigma}}_{\varrho_{\text{ext}}} \right)^{\frac{1}{2}} \\ &= \sqrt{\text{Tr}_{\text{sys}}(\text{Tr}_{\text{ext}}(\rho) |\psi\rangle\langle\psi|)} F(\varrho_{\text{ext}}, \sigma). \\ &= \sqrt{\text{Tr}_{\text{sys}}(\rho_{\text{sys}} |\psi\rangle\langle\psi|)} F(\varrho_{\text{ext}}, \sigma). \end{aligned} \quad (14)$$

Finally, it is easy to see that the term with the square root equals to  $F(\rho_{\text{sys}}, \bar{\rho})$ . To this end expand

$$\begin{aligned} F(\rho_{\text{sys}}, \bar{\rho}) &= \text{Tr} \sqrt{|\psi\rangle\langle\psi| \rho_{\text{sys}} |\psi\rangle\langle\psi|} = \\ &= \text{Tr} (|\psi\rangle\langle\psi|) \sqrt{\text{Tr}(\rho_{\text{sys}} |\psi\rangle\langle\psi|)} = \sqrt{\text{Tr}(\rho_{\text{sys}} |\psi\rangle\langle\psi|)} \end{aligned}$$

which completes the proof.  $\square$

**Proposition .2.** *Given a target state  $\bar{\rho} \in L(\mathcal{H}_{\text{sys}})$  and any state  $\Phi[\rho] \in L(\mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{ext}})$  with  $\Lambda[\rho] = \text{Tr}_{\text{ext}}(\Phi[\rho]) \in L(\mathcal{H}_{\text{sys}})$ , the following relation holds*

$$F(\Lambda[\rho], \bar{\rho}) \geq \max_{\rho_{\text{ext}}} F(\Phi[\rho], \bar{\rho} \otimes \rho_{\text{ext}}).$$

*Moreover, when the target state  $\bar{\rho} = \bar{\rho}^2$  is pure, the maximum is attained for the state  $\rho_{\text{ext}} = \frac{\text{Tr}_{\text{sys}}(\Phi[\rho] \bar{\rho} \otimes \mathbf{1})}{\text{Tr}(\Phi[\rho] \bar{\rho} \otimes \mathbf{1})}$  and the inequality is saturated.*

*Proof.* First note that the processing inequality implies  $F(\Phi[\rho], \bar{\rho} \otimes \rho_{\text{ext}}) \leq F(\Lambda[\rho], \bar{\rho})$  – tracing out subsystems can only increase the fidelity. This implies the inequality. Second, when  $\bar{\rho}$  is pure, we can apply Lemma .1 to the fidelity  $F(\Phi[\rho], \bar{\rho} \otimes \rho_{\text{ext}})$ , which gives

$$F(\Phi[\rho], \bar{\rho} \otimes \rho_{\text{ext}}) = F(\Lambda[\rho], \bar{\rho}) F(\varrho_{\text{ext}}, \rho_{\text{ext}}), \quad (15)$$

with  $\rho_{\text{ext}} = \frac{\text{Tr}_{\text{sys}}(\Phi[\rho] \bar{\rho} \otimes \mathbf{1})}{\text{Tr}(\Phi[\rho] \bar{\rho} \otimes \mathbf{1})}$ . Hence, the equality  $F(\Lambda[\rho], \bar{\rho}) = F(\Phi[\rho], \bar{\rho} \otimes \rho_{\text{ext}})$  is obtained for the choice  $\rho_{\text{ext}} = \varrho_{\text{ext}}$ .  $\square$

## Appendix A.2 Extension to multi-partite scenarios

In the framework of the certification of channel with several inputs, e.g. two-qubit gates, the channel  $\mathcal{E}$  supposedly implements an interaction between a number of physically distinct subsystems, which can be clearly identified at the input and the output of the tested device. In this case, the side  $\mathcal{A}$  is composed of  $n$  parties<sup>1</sup>  $\{A^{(k)}\}_{k=1}^n$  carrying one subsystem each, but also the reference channel  $\bar{\mathcal{E}}$  comes with a product structure for the Hilbert spaces  $\mathcal{H}_{i/o} = \bigotimes_{k=1}^n \mathcal{H}_{i/o}^{(k)}$ . In this context, a certification up to global isometries on side  $\mathcal{A}$  is not sufficient, e.g. this would not distinguish an entangling two-qubit gate (a CNOT gate for example) with the identity. Hence, in order to allow for the device-independent certification of channels acting on several systems, the the maps  $\Lambda_i$  and  $\Lambda_o$  must be local with respect to the partition of  $\mathcal{A}$ , as we now describe.

A channel  $\mathcal{E} : L(\mathcal{H}_{\mathcal{A}}) \rightarrow L(\mathcal{H}_{\mathcal{A}})$  describing the tested gate operates on an Hilbert space  $\mathcal{H}_{\mathcal{A}}$  of unknown dimension. Nevertheless, the gate acts on several subsystems

<sup>1</sup> Note that more generally the number of input and output subsystems can differ, but all the formalism straightforwardly generalizes to this case.

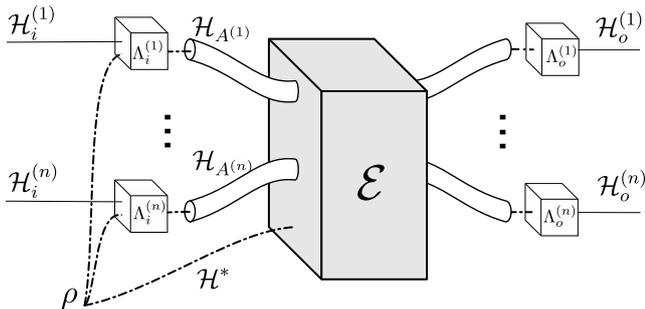


FIG. 2. Device independent certification of multipartite channels: decomposition of the input and output of the channel to be certified  $\mathcal{E}$  into Hilbert spaces belonging to the different parties  $A^{(k)}$ , and the role of injection/extraction maps. The induced map from  $\mathcal{H}_i$  to  $\mathcal{H}_o$  can be compared to the target channel  $\bar{\mathcal{E}}$ .

that can be unambiguously separated. Hence,  $\mathcal{H}_{\mathcal{A}}$  comes with a partition

$$\mathcal{H}_{\mathcal{A}} = \left( \bigotimes_{k=1}^n \mathcal{H}_{A^{(k)}} \right) \otimes \mathcal{H}^*, \quad (16)$$

where each  $\mathcal{H}_{A^{(k)}}$  carries the state of the subsystem  $k$  on which the gate acts, and  $\mathcal{H}^*$  can carry external systems that define the initial state of the tested device. For example, the on/off button on the device can be such an external subsystem proper to the device. The Hilbert space for each physical subsystem  $\mathcal{H}_{A^{(k)}}$  has an unspecified dimension. We extend it with the input Hilbert space  $\mathcal{H}_i^{(k)}$  of the specified dimension, which will carry part of the maximally entangled state  $|\phi^+\rangle \in \mathcal{H}_i = \bigotimes_{k=1}^n \mathcal{H}_i^{(k)}$ .

Then each local injection map sends state of this extended Hilbert space

$$\Lambda_i^{(k)} : L\left(\mathcal{H}_{A^{(k)}} \otimes \mathcal{H}_i^{(k)}\right) \rightarrow L\left(\mathcal{H}_{A^{(k)}}\right) \quad (17)$$

into the physically relevant Hilbert space  $\mathcal{H}_{A^{(k)}}$  on which the gate  $\mathcal{E}$  can act. With the overall map  $\Lambda_i = \left( \bigotimes_{k=1}^n \Lambda_i^{(k)} \right) \otimes \mathbb{1}^*$ , where  $\mathbb{1}^*$  simply says that the internal state of the device is untouched. Similarly, the local extraction map

$$\Lambda_o^{(k)} : L\left(\mathcal{H}_{A^{(k)}}\right) \rightarrow L\left(\mathcal{H}_o^{(k)}\right) \quad (18)$$

maps the state of the physical output of the gate into the output Hilbert space of the ideal channel  $\bar{\mathcal{E}}$  such that the global map reads  $\Lambda_o = \text{Tr}_{\mathcal{H}^*} \left( \bigotimes_{k=1}^n \Lambda_o^{(k)} \right) \otimes \mathbb{1}^*$ .

With these definitions, the channel composed with the maps takes the form

$$\Lambda_o \circ \mathcal{E} \circ \Lambda_i : L\left(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i\right) \rightarrow L\left(\mathcal{H}_o\right). \quad (19)$$

By setting the state of the subsystem  $\mathcal{H}_{\mathcal{A}}$  to be the one prepared by the source  $\rho_{\mathcal{A}}$ , we finally get the desired

recipe

$$\mathcal{E}_{i-o} : L\left(\mathcal{H}_i\right) \rightarrow L\left(\mathcal{H}_o\right) \quad (20)$$

$$\varrho \mapsto \Lambda_o \circ \mathcal{E} \circ \Lambda_i[\rho_{\mathcal{A}} \otimes \varrho] \quad (21)$$

that allows to compare  $\mathcal{E}$  to the target channel  $\bar{\mathcal{E}}$ , as depicted in Fig. 2.

Note that the state  $\rho_{\mathcal{A}}$  also contains the initial internal state of the device itself. This allows to address eccentric scenarios where the source can be entangled with the device. In practice, one can often assume that the state produced by the source is independent from the device, in which case the state  $\rho_{\mathcal{A}}$  decomposes as  $\rho_{\mathcal{A}} \otimes \rho^*$ , where  $\mathcal{A}$  only refers to the source. Whenever this is possible, we can forget about the internal state of the gate  $\rho^*$  and the Hilbert space  $\mathcal{H}^*$ , absorbing it in the definition of the channel  $\mathcal{E}$ . We assume that this is the case for the rest of this section.

As a particular application of the above definition, let us now consider the case where the state in the experiment is produced by  $n$  independent sources, each of which distributes an entangled state to  $A^{(k)}$  and  $B^{(k)}$ . Under this assumption the marginal state of  $\mathcal{A}$  takes the form

$$\rho_{\mathcal{A}} = \bigotimes_{k=1}^n \rho_{A^{(k)}}. \quad (22)$$

Because each auxiliary state  $\rho_{A^{(k)}}$  can be created locally, one can absorb them in newly defined injection maps

$$\Lambda_i^{\prime(k)} : L\left(\mathcal{H}_i^{(k)}\right) \rightarrow L\left(\mathcal{H}_{A^{(k)}}\right) \quad (23)$$

$$\varrho \mapsto \Lambda_i^{\prime(k)}[\rho_{A^{(k)}} \otimes \varrho]. \quad (24)$$

With the maps defined this way, the explicit dependence on the state of the source  $\rho_{\mathcal{A}}$  disappears, that is

$$\mathcal{E}_{i-o} = \left( \bigotimes_{k=1}^n \Lambda_o^{(k)} \right) \circ \mathcal{E} \circ \left( \bigotimes_{\ell=1}^n \Lambda_i^{\prime(\ell)} \right), \quad (25)$$

as depicted in Fig. 3.

It is important to realize that if the sources are not independent such a formulation of the device-independent certification of a gate is impossible because one can easily imagine channels which can only perform the desired operation when the source provides an entangled state in some auxiliary degrees of freedom. For example, imagine that the source distributes two auxiliary maximally entangled states  $|\phi^+\rangle_{A^{(k)}-A^{(\ell)}}$  to each pair of subsystems  $A^{(k)}$  and  $A^{(\ell)}$ , and that the channel  $\mathcal{E}$  first teleports all the inputs on one subsystem, say  $A^{(1)}$ , performs the desired operation  $\bar{\mathcal{E}}$  locally at  $A^{(1)}$ , and then teleports the resulting states back to the respective parties. This sequence of operations would allow the channel to act as desired, but only provided that these singlets are actually distributed by the source. In other words, it would fail to work when provided independent auxiliary

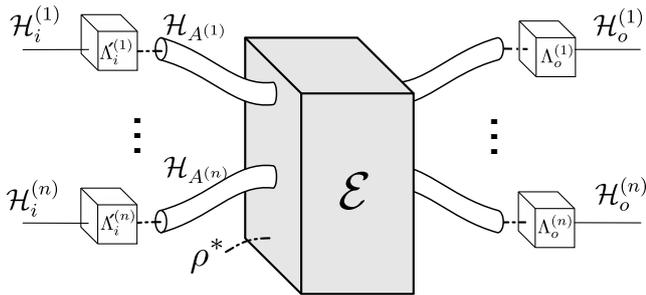


FIG. 3. Device independent certification of multipartite channels with  $n$  independent sources. The local injection maps  $\Lambda_i^{(k)}$  can be optimized independently of the system's state  $\rho$ .

states. The recipe defined in Eq. (25) would then fail to act as the target channel  $\bar{\mathcal{E}}$  and only the procedure defined in Eq. (21), which includes the partial state of the source  $\rho_{\mathcal{A}}$  produced in experiment, would perform the desired operation.

**Appendix A.3 Relation between the Uhlmann channel fidelity and the diamond norm**— We show how the Uhlmann fidelity between the Choi state of two channels can be used to bound the diamond norm between these two channels.

Let us first recall that the diamond norm between two channels  $\mathcal{E}_{i-o}$  and  $\bar{\mathcal{E}}$ , acting between the same Hilbert spaces  $L(\mathcal{H}_i) \rightarrow L(\mathcal{H}_o)$ , is defined as<sup>2</sup>

$$\|\mathcal{E}_{i-o} - \bar{\mathcal{E}}\|_{\diamond} = \sup_{|\zeta\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i} \|(\mathcal{E}_{i-o} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|] - (\bar{\mathcal{E}} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|]\|_1. \quad (26)$$

This expression has an operational meaning: It directly relates to the maximal probability to discriminate the two channels in a single measurement, when comparing the images of some state  $|\zeta\rangle$ .

We are interested to compare the reference channel  $\bar{\mathcal{E}}$  and the channel  $\mathcal{E}_{i-o} = \Lambda_o \circ \mathcal{E} \circ \Phi_i$ . In particular, we show that the fidelity  $\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})$  of Eq. (1) of the main text can be used to bound the diamond norm between  $\mathcal{E}_{i-o}$  and  $\bar{\mathcal{E}}$ .

**Proposition .3.** *The diamond norm  $\|\mathcal{E}_{i-o} - \bar{\mathcal{E}}\|_{\diamond}$  between two quantum channels*

$$\mathcal{E}_{i-o} \text{ and } \bar{\mathcal{E}} : L(\mathcal{H}_i) \rightarrow L(\mathcal{H}_o) \quad (27)$$

is upper bounded by

$$\|\mathcal{E}_{i-o} - \bar{\mathcal{E}}\|_{\diamond} \leq 2 \dim(\mathcal{H}_i) \sqrt{1 - \mathcal{F}^2(\mathcal{E}_{i-o}, \bar{\mathcal{E}})}, \quad (28)$$

<sup>2</sup> Remark that one can sometimes see the definition with the maximization running over states  $|\zeta\rangle$  on an arbitrarily extended Hilbert space  $\mathcal{H}_i \otimes \mathcal{H}_{\text{ext}}$ . But because of the Schmidt decomposition any such state only has support on a subspace of  $\mathcal{H}_{\text{ext}}$  of dimension  $\dim(\mathcal{H}_i)$  at most. Hence it is sufficient to consider states  $|\zeta\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i$ .

where

$$\mathcal{F}(\mathcal{E}_{i-o}, \bar{\mathcal{E}}) = F((\mathcal{E}_{i-o} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]) \quad (29)$$

is the Choi fidelity between the two channels defined with the maximally entangled state  $|\phi^+\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i$ .

*Proof.* We start by showing that any state  $|\zeta\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i$  can be probabilistically prepared from the maximally entangled state  $|\phi^+\rangle$  by solely acting on the second Hilbert space. Given the Schmidt decomposition of the state  $|\zeta\rangle = \sum_i \zeta_i |i\rangle |i'\rangle$  there exists a local unitary  $\mathbb{1} \otimes U$  such that

$$|\phi^+\rangle = \mathbb{1} \otimes U |\phi^+\rangle = \frac{1}{\sqrt{\dim(\mathcal{H}_i)}} \sum_i |i\rangle |i'\rangle. \quad (30)$$

Furthermore, the operator  $K = \sum_i \zeta_i |i'\rangle\langle i'|$  applied on this state results into

$$\mathbb{1} \otimes K |\phi^+\rangle = \frac{1}{\sqrt{\dim(\mathcal{H}_i)}} |\zeta\rangle. \quad (31)$$

$K$  is a valid Kraus operator  $0 \leq K^\dagger K \leq 1$  and hence together with the unitary  $U$  can be implemented as one branch of some probabilistic announced two-branch protocol  $\mathcal{P}$  acting on the second Hilbert space (a two-outcome POVM). From the initial state  $|\phi^+\rangle$  this protocol prepares the state  $|\zeta\rangle$  with probability  $p = \frac{1}{\dim(\mathcal{H}_i)}$  (successful branch), and some other state  $\varrho$  with probability  $(1-p)$  (failure branch).

Next recall that the Uhlmann fidelity between any two states  $\rho$  and  $\sigma$  upper bounds the trace-distance between the same states as

$$D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 \leq \sqrt{1 - F^2(\rho, \sigma)}. \quad (32)$$

Hence, from Eq. (29) one gets

$$D((\mathcal{E}_{i-o} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]) \leq \sqrt{1 - \mathcal{F}^2}. \quad (33)$$

The trace distance  $D$  satisfies the processing inequality, meaning that it cannot increase whatever common operation is performed on the two states. In particular, this holds for the protocol  $\mathcal{P}$ , which, in addition, commutes with the channels and can therefore be applied before them, leading to

$$D((\mathcal{E}_{i-o} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]) \geq \quad (34)$$

$$pD((\mathcal{E}_{i-o} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|]) + (1-p)D'. \quad (35)$$

$D'$  stands for the trace distance between the states prepared in the failure branch and satisfies  $(1-p)D' \geq 0$ . Combining with the previous inequality yields

$$D((\mathcal{E}_{i-o} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\zeta\rangle\langle\zeta|]) \leq \frac{1}{p} \sqrt{1 - \mathcal{F}^2}. \quad (36)$$

Since this holds for any state  $|\zeta\rangle$  and  $p = \frac{1}{\dim(\mathcal{H}_i)}$ , the proposition is proven.  $\square$

**Appendix B Bounding the channel fidelity with state fidelities**– In this Appendix, we prove Eq. (4) of the main text.

**Proposition .4.** *Let  $\rho$  be a quantum state shared between the sides  $\mathcal{A}$  and  $\mathcal{B}$ ,  $\mathcal{E}$  and  $\bar{\mathcal{E}}$  two channels acting on side  $\mathcal{A}$  with  $\bar{\mathcal{E}} : L(\mathcal{H}_i) \rightarrow L(\mathcal{H}_o)$  a reference channel, and  $|\phi^+\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i$  a maximally entangled state. Given the local maps  $\tilde{\Lambda}_i^{\mathcal{A}}[\cdot] = \text{Tr}_{\text{ext}}(\tilde{\Phi}_i^{\mathcal{A}}[\cdot])$ ,  $\Lambda_o^{\mathcal{A}}[\cdot] = \text{Tr}_{\text{ext}}(\Phi_o^{\mathcal{A}}[\cdot])$  and  $\Lambda^{\mathcal{B}}[\cdot] = \text{Tr}_{\text{ext}}(\Phi^{\mathcal{B}}[\cdot])$  acting on  $\mathcal{A}$  and  $\mathcal{B}$  respectively with the corresponding isometries  $\tilde{\Phi}_i^{\mathcal{A}}$ ,  $\Phi_o^{\mathcal{A}}$  and  $\Phi^{\mathcal{B}}$ , the following two fidelities*

$$F^i = F\left(\tilde{\Lambda}_i^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}}[\rho], |\phi^+\rangle\langle\phi^+|\right)$$

$$F^o = F\left((\Lambda_o^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}})[(\mathcal{E} \otimes \mathbb{1}_{\mathcal{B}})[\rho]], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]\right),$$

lead to the following bound

$$\arccos(\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})) \leq \arccos(F^i) + \arccos(F^o).$$

*Proof.* First note that the same map  $\Lambda_{\mathcal{B}}$  appears in both equations. Hence we get rid of all the external systems on the side  $\mathcal{B}$  and introduce the state  $\rho' = (\mathbb{1}_{\mathcal{A}} \otimes \Lambda_{\mathcal{B}})[\rho] \in L(\mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_i)$  where the dimension of  $\mathcal{H}_{\mathcal{A}}$  is unspecified but the output subsystems on side  $\mathcal{B}$  is already identified. Expressing the two fidelities  $F^i$  and  $F^o$  with  $\rho'$  gives

$$F^i = F\left(\tilde{\Lambda}_i^{\mathcal{A}} \otimes \mathbb{1}[\rho'], |\phi^+\rangle\langle\phi^+|\right) \quad (37)$$

$$F^o = F\left((\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \otimes \mathbb{1})[\rho'], (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]\right). \quad (38)$$

Using Proposition .2 for  $\mathcal{H}_{\text{sys}} = \mathcal{H}_i$ , we can express the first fidelity as

$$F^i = F\left((\tilde{\Phi}_i^{\mathcal{A}} \otimes \mathbb{1})[\rho'], |\phi^+\rangle\langle\phi^+| \otimes \rho_{\text{ext}}^{\mathcal{A}}\right), \quad (39)$$

where the auxilliary state on Alice side  $\rho_{\text{ext}}^{\mathcal{A}}$  is by definition given by

$$\rho_{\text{ext}}^{\mathcal{A}} = \text{Tr}_{(\mathcal{H}_{\mathcal{A}}^{\mathcal{A}} \otimes \mathcal{H}_i^{\mathcal{B}})}(\tilde{\Phi}_i \otimes \mathbb{1}[\rho']) = \text{Tr}_{\mathcal{H}_i^{\mathcal{A}}} \tilde{\Phi}_i[\rho_{\mathcal{A}}] \quad (40)$$

with  $\rho_{\mathcal{A}} = \text{Tr}_{\mathcal{B}} \rho$  as defined in the main text.

An isometry  $\tilde{\Phi}_i^{\mathcal{A}}$  is a unitary embedding of the state into an Hilbert space of a larger dimension. As such, it can be decomposed as  $(\tilde{\Phi}_i^{\mathcal{A}} \otimes \mathbb{1})[\rho'] = (\tilde{U}_i \otimes \mathbb{1})[\rho' \otimes |\widetilde{\text{ext}}\rangle\langle\widetilde{\text{ext}}|_i^{\mathcal{A}}]$ . Plugging this into Eq. (39) gives

$$\begin{aligned} F^i &= F\left((\tilde{U}_i \otimes \mathbb{1})[\rho' \otimes |\widetilde{\text{ext}}\rangle\langle\widetilde{\text{ext}}|_i^{\mathcal{A}}], |\phi^+\rangle\langle\phi^+| \otimes \rho_{\text{ext}}^{\mathcal{A}}\right) = \\ &= F\left(\rho' \otimes |\widetilde{\text{ext}}\rangle\langle\widetilde{\text{ext}}|_i^{\mathcal{A}}, (\tilde{U}_i^\dagger \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\text{ext}}^{\mathcal{A}}]\right) \leq \\ &= F\left(\rho', \text{Tr}_{\text{ext}}(\tilde{U}_i^\dagger \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\text{ext}}^{\mathcal{A}}]\right) \end{aligned} \quad (41)$$

where we used the invariance of fidelity under unitary transformations and the fact that it can only increase when subsystems are traced out. Note that the right term in the last fidelity can be simply written as

$$\begin{aligned} &\text{Tr}_{\text{ext}}(\tilde{U}_i^\dagger \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+| \otimes \text{Tr}_{\mathcal{H}_i^{\mathcal{A}}} \tilde{\Phi}_i[\rho_{\mathcal{A}}]] = \\ &\Lambda_i^{\mathcal{A}} \otimes \mathbb{1}[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\mathcal{A}}] \end{aligned} \quad (42)$$

defining the injection map  $\Lambda_i^{\mathcal{A}}$ . Hence, we have

$$F\left(\rho', \Lambda_i^{\mathcal{A}} \otimes \mathbb{1}[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\mathcal{A}}]\right) \geq F_i. \quad (43)$$

Now we apply the map  $(\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \otimes \mathbb{1})$  on both states<sup>3</sup> in the last equation. The processing inequality ensures that the fidelity can only increase by doing so, therefore

$$\begin{aligned} &F\left((\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \otimes \mathbb{1})[\rho'], (\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \circ \Lambda_i^{\mathcal{A}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\mathcal{A}}]\right) \\ &\geq F^i. \end{aligned} \quad (44)$$

Next, we use the equivalence of the triangle inequality for the Unlmann fidelity

$$\arccos(F(\varrho_1, \varrho_3)) \leq \arccos(F(\varrho_1, \varrho_2)) + \arccos(F(\varrho_2, \varrho_3)) \quad (45)$$

for the states

$$\varrho_1 = (\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \circ \Lambda_i^{\mathcal{A}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+| \otimes \rho_{\mathcal{A}}] \quad (46)$$

$$\varrho_2 = (\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \otimes \mathbb{1})[\rho'] \quad (47)$$

$$\varrho_3 = (\bar{\mathcal{E}} \otimes \mathbb{1})[|\phi^+\rangle\langle\phi^+|]. \quad (48)$$

Eq. (38) directly gives  $F(\varrho_2, \varrho_3) = F^o$  while Eq. (44) gives the bound  $F(\varrho_1, \varrho_2) \geq F_i$ , which in turn implies  $\arccos(F(\varrho_1, \varrho_2)) \leq \arccos(F_i)$ . This leads to

$$\arccos(F(\varrho_1, \varrho_2)) \leq \arccos(F^i) + \arccos(F^o).$$

Finally, noticing that

$$\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}}) = F(\varrho_1, \varrho_2)$$

implies

$$\arccos(\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})) \leq \arccos(F^i) + \arccos(F^o). \quad \square$$

In the proof of the previous proposition, we have explicitly constructed the injection map  $\Lambda_i^{\mathcal{A}}$  identifying input subspaces on which the channel  $\mathcal{E}_{i-o}$  acts. This map is constructed from the extraction map  $\tilde{\Lambda}_i^{\mathcal{A}}$  in the certificate of the input state Eq. (2) of the main text. We now show that a similar result holds in the multipartite case where the channel  $\mathcal{E}$  implements an interaction between several physical systems.

**Proposition .5.** *Let  $\rho$  be a quantum state shared between the sides  $\mathcal{A}$  and  $\mathcal{B}$ ,  $\mathcal{A}$  consisting of several parties  $\{A^{(k)}\}_{k=1}^n$ . Let also  $\mathcal{E}$  be a channel acting jointly on all the parties on side  $\mathcal{A}$ ,  $\bar{\mathcal{E}} : L\left(\bigotimes_{k=1}^n \mathcal{H}_i^{(k)}\right) \rightarrow L\left(\bigotimes_{k=1}^n \mathcal{H}_o^{(k)}\right)$*

<sup>3</sup> Remark that both  $\mathcal{E}$  and  $\Lambda_o^{\mathcal{A}}$  only act nontrivially on the "physical" system  $\rho'$ , the additional subsystems added by the isometries are traced out by the composed map  $\Lambda_o^{\mathcal{A}} \circ \mathcal{E}$ .

a reference channel, and  $|\phi^+\rangle \in \mathcal{H}_i \otimes \mathcal{H}_i$  a maximally entangled state. Given the maps  $\tilde{\Lambda}_i^{A^{(k)}}[\cdot] = \text{Tr}_{\widetilde{\text{ext}}_k}(\tilde{\Phi}_i^{A^{(k)}}[\cdot])$ ,  $\Lambda_o^{A^{(k)}}[\cdot] = \text{Tr}_{\text{ext}_k}(\Phi_o^{A^{(k)}}[\cdot])$  acting locally on the respective parties  $A^{(k)}$  and  $\Lambda^{\mathcal{B}}[\cdot] = \text{Tr}_{\text{ext}}(\Phi^{\mathcal{B}}[\cdot])$  acting on  $\mathcal{B}$ , and the product maps

$$\begin{aligned}\tilde{\Lambda}_i^{\mathcal{A}} &= \tilde{\Lambda}_i^{A^{(1)}} \otimes \dots \otimes \tilde{\Lambda}_i^{A^{(n)}} \\ \Lambda_o^{\mathcal{A}} &= \Lambda_o^{A^{(1)}} \otimes \dots \otimes \Lambda_o^{A^{(n)}},\end{aligned}$$

the following two equations

$$\begin{aligned}F^i &= F\left(\left(\tilde{\Lambda}_i^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}}\right)[\rho], |\phi^+\rangle\langle\phi^+|\right) \\ F^o &= F\left(\left(\Lambda_o^{\mathcal{A}} \otimes \Lambda^{\mathcal{B}}\right)\left[(\mathcal{E} \otimes \mathbf{1}_{\mathcal{B}})[\rho]\right], (\bar{\mathcal{E}} \otimes \mathbf{1})\left[|\phi^+\rangle\langle\phi^+|\right]\right),\end{aligned}$$

imply the following bound on the fidelity between the channels  $\mathcal{E}$  and  $\bar{\mathcal{E}}$

$$\arccos(\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})) \leq \arccos(F^i) + \arccos(F^o). \quad (49)$$

In addition, the injection and extraction maps  $\Lambda_i$  and  $\Lambda_o$  in the definition of  $\mathcal{F}(\mathcal{E}, \bar{\mathcal{E}})$  (see Eq. (4) of the main text) also have a local structure

$$\Lambda_i^{\mathcal{A}} = \Lambda_i^{(1)} \otimes \dots \otimes \Lambda_i^{(n)} \quad (50)$$

$$\Lambda_o^{\mathcal{A}} = \Lambda_o^{(1)} \otimes \dots \otimes \Lambda_o^{(n)}, \quad (51)$$

as described in Appendix A.2.

*Proof.* The proof is a simple modification of the proof of Proposition .4. It is similar until Eq. (39)

$$F^i = F\left(\left(\tilde{\Phi}_i^{\mathcal{A}} \otimes \mathbf{1}\right)[\rho'], |\phi^+\rangle\langle\phi^+| \otimes \rho_{\text{ext}}^{\mathcal{A}}\right). \quad (52)$$

At this point we use the product structure of the isometry  $\tilde{\Phi}_i^{\mathcal{A}} = \tilde{\Phi}_i^{A^{(1)}} \otimes \dots \otimes \tilde{\Phi}_i^{A^{(n)}}$ , where each isometry acting on the party  $A^{(k)}$  can be written as  $\tilde{\Phi}_i^{A^{(k)}}[\cdot] = \tilde{U}_i^{(k)}[(\cdot) \otimes |\widetilde{\text{ext}}\rangle\langle\widetilde{\text{ext}}|_i^{(k)}]$ . Hence the unitary in (41) also has a product structure  $\tilde{U}_i = \tilde{U}_i^{(1)} \otimes \dots \otimes \tilde{U}_i^{(n)}$ , which is inherited by the injection map  $\Lambda_i^{\mathcal{A}}$  defined in Eq. (42). The map  $\Lambda_o^{\mathcal{A}}$  have a product structure by definition. The rest simply follows the proof of proposition .4.  $\square$

To conclude, we notice that in the case where the quantum state  $\rho$  is produced by independent sources it has the form

$$\rho = \rho_{A^{(1)}-B^{(1)}} \otimes \dots \otimes \rho_{A^{(n)}-B^{(n)}} \otimes \rho^*, \quad (53)$$

where each state  $\rho_{A^{(k)}-B^{(k)}}$  is shared between the parties  $A^{(k)}$  and  $B^{(k)}$ , and  $\rho^*$  is the initial state of the device. Consequently, the marginal state of Alice is product with respect to all subsystems

$$\rho_{\mathcal{A}} = \bigotimes_{k=1}^n \rho_{A^{(k)}}, \quad (54)$$

with  $\rho_{A^{(k)}} = \text{Tr}_{B^{(k)}} \rho_{A^{(k)}-B^{(k)}}$ . This allows to conveniently absorb each state into the local injection map  $\Lambda_i^{(i)}$  performed locally by the corresponding subsystem

$$\Lambda_i^{A^{(\ell)}}[\cdot] = \Lambda_i^{A^{(\ell)}}[\cdot \otimes \rho_{A^{(\ell)}}] \quad \forall \ell = 1, \dots, n \quad (55)$$

Hence, in the proof of the above proposition the state of the source  $\rho_{\mathcal{A}}$  can be absorbed in this novel definition of the injection map  $\Lambda_i^{\mathcal{A}}$ , except for the degrees of freedom that describe the initial state of the measurement device itself  $\rho^*$  (which may specify, for example, that the device is plugged in). Considering the initial internal state of the device  $\rho^*$  as part of the channel  $\mathcal{E}$ , we can write the following corollary

**Corollary .5.1.** *Given all the condition of the proposition .5, if in addition the sources for each input subsystem of Alice are independent, it follows that*

$$\arccos(\mathcal{F}'(\mathcal{E}, \bar{\mathcal{E}})) \leq \arccos(F^i) + \arccos(F^o).$$

where

$$\mathcal{F}'(\mathcal{E}, \bar{\mathcal{E}}) = F\left(\left(\Lambda_o^{\mathcal{A}} \circ \mathcal{E} \circ \Lambda_i^{\mathcal{A}}\right) \otimes \mathbf{1}\left[|\phi^+\rangle\langle\phi^+|\right], \bar{\mathcal{E}} \otimes \mathbf{1}\left[|\phi^+\rangle\langle\phi^+|\right]\right) \quad (56)$$

and the maps  $\Lambda_o^{\mathcal{A}}$  and  $\Lambda_i^{\mathcal{A}}$  have a tensor product structure with respect to subsystems of  $\mathcal{A}$ .

*Proof.* We just presented the proof.  $\square$

As discussed in Appendix A.2, this result would not be possible without the assumption of independent sources.

**Appendix C Robustness of channel certification in presence of white noise**– Here we illustrate the robustness of the bounds obtained for the fidelity of the single qubit unitary channel and the two qubit CNOT gate. To this end, we consider a simple model where each element suffers from white noise. In particular, each source produces the two qubit mixed state

$$\rho(\epsilon_S) = (1 - \epsilon_S) |\phi_2^+\rangle\langle\phi_2^+| + \epsilon_S \frac{1}{4} \mathbb{1}, \quad (57)$$

that is, with probability  $\epsilon_S$ , it fails to output the maximally entangled two-qubit state  $|\phi_2^+\rangle$  and produces a totally depolarized two-qubit state  $\frac{1}{4} \mathbb{1}$  instead. The measurement devices output a random result with probability  $\epsilon_M$ , which corresponds to replacing each Pauli by  $X(Z) \rightarrow (1 - \epsilon_M)X(Z)$  in the Bell operator. Similarly, the channel  $\mathcal{E}$  fails to perform the desired operation with probability  $\epsilon_C$ , in which case it outputs a totally depolarized state of the corresponding dimension.

Within such a model, the calculation of expected Bell values is straightforward and gives

$$\beta^i = 2\sqrt{2}(1 - \epsilon_S)(1 - \epsilon_M)^2 \quad (58)$$

$$\beta^o = 2\sqrt{2}(1 - \epsilon_S)(1 - \epsilon_M)^2(1 - \epsilon_C) \quad (59)$$

$$\begin{aligned} \gamma^i = & 1 - 2.4\epsilon_M^3(1 - \epsilon_S)^2 + \epsilon_M(\epsilon_S(5.6 - 2.4\epsilon_S) - 3.2) \\ & + 0.6\epsilon_M^4(1 - \epsilon_S)^2 + \epsilon_M^2(\epsilon_S(3.6\epsilon_S - 7.6) + 4) \\ & + (0.6\epsilon_S - 1.6)\epsilon_S \end{aligned} \quad (60)$$

$$\begin{aligned} \gamma^o = & 1 - 3.2\epsilon_M + 3.9\epsilon_M^2 - 2.2\epsilon_M^3 + 0.5\epsilon_M^4 - 1.6\epsilon_S \\ & + 5.5\epsilon_M\epsilon_S - 7.2\epsilon_M^2\epsilon_S + 4.3\epsilon_M^3\epsilon_S - \epsilon_M^4\epsilon_S + 0.6\epsilon_S^2 \\ & - 2.3\epsilon_M\epsilon_S^2 + 3.3\epsilon_M^2\epsilon_S^2 - 2.1\epsilon_M^3\epsilon_S^2 + 0.5\epsilon_M^4\epsilon_S^2 \\ & + \epsilon_C(-0.5\epsilon_M^4\epsilon_S^2 + 1.\epsilon_M^4\epsilon_S + 2.1\epsilon_M^3\epsilon_S^2 - 4.3\epsilon_M^3\epsilon_S \\ & - 3.3\epsilon_M^2\epsilon_S^2 + 7.2\epsilon_M^2\epsilon_S + 2.3\epsilon_M\epsilon_S^2 - 5.5\epsilon_M\epsilon_S - 0.5\epsilon_M^4 \\ & + 2.2\epsilon_M^3 - 3.9\epsilon_M^2 + 3.2\epsilon_M - 0.6\epsilon_S^2 + 1.6\epsilon_S - 1). \end{aligned} \quad (61)$$

Fig. 4 shows the resulting bound on the fidelity of the identity channel (left) and the CNOT gate (right) as functions of the channel noise ( $\epsilon_C = \epsilon_C^{\text{Id}}$  and  $\epsilon_C^{\text{CNOT}}$  respectively), assuming that  $\epsilon_S = \epsilon_M = \epsilon_{\text{Setup}}$ .

**Appendix D Systematic approach to robustly certify N-qubit states device-independently**– In this appendix, we describe an approach for the device-independent, robust state certification that allowed us to certify any two-qubit controlled unitary gate. We provide a systematic recipe that can be applied to arbitrary N-qubit states, even though its convergence is not guaranteed. Appendix D.2 and D.3 exploits the Jordan lemma to show how one can obtain device-independent bounds on the fidelity of N-qubit states from a given Bell test. Appendix D.4 shows how to devise Bell tests tailored to the robust certification of N-qubit states. Appendix D.5 finally applies our approach to the maximally entangled two-qubit state and the four-qubit states obtained by applying  $CU_\varphi$  gates on two maximally entangled two-qubit states. Before we start, we discuss a figure of merit, the overlap, to compare two quantum states.

**Appendix D.1 Hilbert-Schmidt inner product and relation to the Uhlmann fidelity**– The Hilbert-Schmidt inner product (aka the overlap) between two states  $\rho$  and  $\bar{\rho}$  defined on the same Hilbert space is given by

$$O(\rho, \bar{\rho}) = \text{Tr}(\rho\bar{\rho}). \quad (62)$$

The overlap does not satisfy some nice properties of the Uhlmann fidelity, but has the advantage to be a linear function of the quantum states and thus allows one to use efficient optimization techniques. There are several known results and methods for robust state certification [2–5], that provide bounds on the overlap between the state to be certified and the target state. As we now show, bounds on the Hilbert-Schmidt product can be directly used to bound the Uhlmann fidelity needed in Eq. (2)-(3) of the main text.

**Lemma .6.** *Given two quantum states  $\rho, \bar{\rho}$ , the following relation holds:*

$$F(\rho, \bar{\rho}) \geq \sqrt{\text{Tr}(\rho\bar{\rho})}, \quad (63)$$

*with equality when one of the two states is pure.*

*Proof.* With  $\bar{\rho} = \sum_i p_i |\bar{\psi}_i\rangle\langle\bar{\psi}_i|$  we have

$$\begin{aligned} F(\rho, \bar{\rho})^2 &= F(\rho, \sum_i p_i |\bar{\psi}_i\rangle\langle\bar{\psi}_i|)^2 \\ &\geq \sum_i p_i F(\rho, |\bar{\psi}_i\rangle\langle\bar{\psi}_i|)^2 \\ &= \sum_i p_i \text{Tr}(\rho |\bar{\psi}_i\rangle\langle\bar{\psi}_i|) \\ &= \text{Tr}(\rho\bar{\rho}). \end{aligned} \quad (64)$$

Here we used the concavity of the square of the Uhlmann fidelity with respect to its second argument [6]. The corresponding inequality is saturated when  $\bar{\rho}$  is pure. By symmetry of the Uhlmann fidelity, equality is also achieved when  $\rho$  is pure.  $\square$

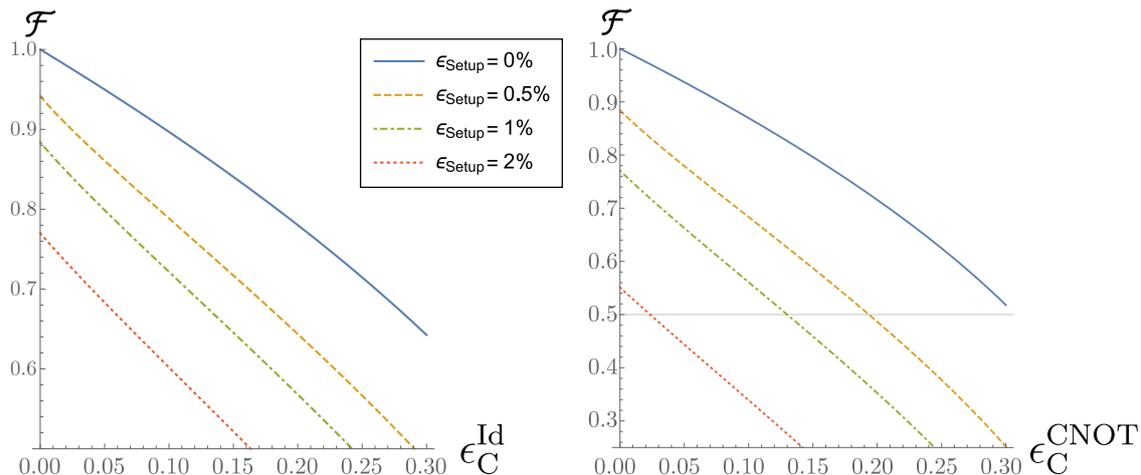


FIG. 4. Device-independently certified fidelity  $\mathcal{F}$  for the identity channel (left) and the CNOT gate (right) in presence of white noise on the certified elements ( $\epsilon_C^{\text{Id(CNOT)}}$ ) and on the other elements of the setup, i.e the sources and measurements ( $\epsilon_{\text{Setup}}$ ).

The definition of the Hilbert-Schmidt product naturally generalizes to the case with a state  $\varrho$  of the global system defined on  $\mathcal{H}_{\text{sys}} \otimes \mathcal{H}_{\text{ext}}$  and a state  $\bar{\rho}$  of a subsystem on  $\mathcal{H}_{\text{sys}}$

$$O(\varrho, \bar{\rho}) = \text{Tr}_{\text{sys}} \left( \text{Tr}_{\text{ext}}(\varrho) \bar{\rho} \right), \quad (65)$$

as it is commonly done. In such a case Lemma.6 ensures that  $F(\text{Tr}_{\text{ext}}(\varrho), \bar{\rho}) \geq O(\varrho, \bar{\rho})$  with equality when  $\bar{\rho}$  is pure.

**Appendix D.2 Device independent certification of states using Jordan's lemma**—In this subsection, we show how the device-independent certification of an  $N$ -qubit state can be reduced to an  $N$ -qubit problem. This reduction is only possible if certain requirements are fulfilled: (i) The Bell test has to involve at most two different measurements per party, and (ii) these measurements must be binary. If these requirements are fulfilled we can use Jordan's lemma, which states the following [7]:

**Lemma .7.** *Let  $X$  and  $Z$  be two Hermitian operators with eigenvalues  $-1$  and  $+1$ . Then there exists a basis in which both operators are block-diagonal, with blocks of dimension  $2 \times 2$  at most.*

This directly implies that if a Bell operator  $\mathcal{B}$  consists of binary measurements then the operator  $\mathcal{B}$  corresponding to the Bell test can be written as

$$\mathcal{B} = \bigoplus_{\alpha_1} \bigoplus_{\alpha_2} \dots \bigoplus_{\alpha_N} \mathcal{B}_{\alpha_1 \dots \alpha_N}. \quad (66)$$

In this expression,  $\mathcal{B}_{\alpha_1 \dots \alpha_N}$  is a  $N$ -qubit Bell operator, and the indices  $\alpha_1, \dots, \alpha_N$  denote the block of parties  $1, \dots, N$  respectively. Therefore, the expectation value of the Bell operator is given by

$$\beta = \sum_{\alpha_1, \dots, \alpha_N} \text{Tr}(\varrho_{\alpha_1 \dots \alpha_N} \mathcal{B}_{\alpha_1 \dots \alpha_N}), \quad (67)$$

where  $\varrho_{\alpha_1 \dots \alpha_N}$  is an unnormalized  $N$ -qubit state corresponding to the projection of the global state  $\rho$  onto the blocks  $(\alpha_1 \dots \alpha_N)$ . We can add the normalization by writing

$$\begin{aligned} \beta &= \sum_{\alpha_1, \dots, \alpha_N} p_{\alpha_1 \dots \alpha_N} \text{Tr}(\rho_{\alpha_1 \dots \alpha_N} \mathcal{B}_{\alpha_1 \dots \alpha_N}) \\ &= \sum_{\alpha_1, \dots, \alpha_N} p_{\alpha_1 \dots \alpha_N} \beta_{\alpha_1 \dots \alpha_N}(\rho_{\alpha_1 \dots \alpha_N}) \end{aligned} \quad (68)$$

where  $\rho_{\alpha_1 \dots \alpha_N} = p_{\alpha_1 \dots \alpha_N}^{-1} \varrho_{\alpha_1 \dots \alpha_N}$ .

This allows to reduce the device independent certification of  $N$ -qubit states to a  $N$ -qubit problem, as we now explain.

**Proposition .8.** *Consider a Bell operator of the block-diagonal form*

$$\mathcal{B} = \bigoplus_{\alpha_1} \bigoplus_{\alpha_2} \dots \bigoplus_{\alpha_N} \mathcal{B}_{\alpha_1 \dots \alpha_N} \quad (69)$$

with  $\alpha_i$  labeling the block of each party  $i = 1 \dots N$ , a global state  $\rho$  yielding the Bell value  $\beta = \text{Tr}(\rho \mathcal{B})$ , and a convex function  $f$ . If for each party and each subspace (block) there exist a fixed local isometry  $\Phi_{\alpha_i}$ , such that the overlap between any possible state  $\tau$  supported on any  $(\alpha_1 \dots \alpha_N)$ -block and the target state  $\bar{\rho}$  is bounded by the Bell value  $\beta_{\alpha_1 \dots \alpha_N}(\tau) = \text{Tr}(\tau \mathcal{B}_{\alpha_1 \dots \alpha_N})$  via

$$O\left(\left(\Phi_{\alpha_1} \otimes \dots \otimes \Phi_{\alpha_N}\right)[\tau], \bar{\rho}\right) \geq f(\beta_{\alpha_1 \dots \alpha_N}(\tau)), \quad (70)$$

then there is an isometry  $\Phi = \Phi_1 \otimes \Phi_2 \otimes \dots \otimes \Phi_N$  for which the global state satisfies

$$O(\Phi[\rho], \bar{\rho}) \geq f(\beta). \quad (71)$$

*Proof.* For each party, let us define the isometry  $\Phi_i = \bigoplus_{\alpha_i} \Phi_{\alpha_i}$ , which independently maps the state in each block  $\alpha_i$  onto the output (qubit) subsystem. Under these isometries the overlap of the global state  $\rho$  with the target state  $\bar{\rho}$  satisfies

$$\begin{aligned}
& O((\Phi_1 \otimes \Phi_2 \otimes \dots \otimes \Phi_N)[\rho], \bar{\rho}) \\
&= O\left(\left(\bigoplus_{\alpha_1, \dots, \alpha_N} \Phi_{\alpha_1} \otimes \dots \otimes \Phi_{\alpha_N}\right)[\rho], \bar{\rho}\right) \\
&= \sum_{\alpha_1, \dots, \alpha_N} p_{\alpha_1 \dots \alpha_N} O\left((\Phi_{\alpha_1} \otimes \dots \otimes \Phi_{\alpha_N})[\rho_{\alpha_1 \dots \alpha_N}], \bar{\rho}\right) \\
&\geq \sum_{\alpha_1, \dots, \alpha_N} p_{\alpha_1 \dots \alpha_N} f\left(\beta_{\alpha_1 \dots \alpha_N}(\rho_{\alpha_1 \dots \alpha_N})\right) \\
&\geq f\left(\sum_{\alpha_1, \dots, \alpha_N} p_{\alpha_1 \dots \alpha_N} \beta_{\alpha_1 \dots \alpha_N}(\rho_{\alpha_1 \dots \alpha_N})\right) = f(\beta),
\end{aligned} \tag{72}$$

where we used the linearity of the Hilbert-Schmidt product and the convexity of  $f$ .  $\square$

Hence, a solution of the  $N$ -qubit problem in the form of a convex bound as in Eq. (70), can be directly applied to device independently certify any state  $\rho$  from the observed Bell value  $\beta$ . This problem can be solved numerically, as we now show.

### Appendix D.3 Convex bound on the $N$ -qubit fidelity

First let us fix the form of the local observable  $X_i$  and  $Z_i$ ,  $i \in \{1, 2, \dots, N\}$ . After applying Jordan's lemma and choosing a suitable coordinate system (via a local basis change for each block that is absorbed into the isometry), we can express the two observables of party  $i$  in the block  $\alpha_i$  as

$$\begin{cases} X_{i, \alpha_i} = \cos(a_{\alpha_i})\sigma_x + \sin(a_{\alpha_i})\sigma_z \\ Z_{i, \alpha_i} = \cos(a_{\alpha_i})\sigma_x - \sin(a_{\alpha_i})\sigma_z \end{cases} \tag{73}$$

where  $\sigma_x$  and  $\sigma_z$  are Pauli matrices,  $r \in \{0, 1\}$  and  $a_{\alpha_i} \in [0, \frac{\pi}{2}]$ . Each Bell operator  $\mathcal{B}_{\alpha_1 \dots \alpha_N} = \mathcal{B}(a_{\alpha_1}, \dots, a_{\alpha_N})$  only involves terms of the form (73) and is uniquely specified by the angles  $a_{\alpha_i}$ .

In order to apply the Proposition.8 we first fix the dependence of the local isometries on the angles  $\Phi_{\alpha_i} = \Phi(a_{\alpha_i})$ . Then we lower bound the overlap  $O\left((\Phi_{\alpha_1} \otimes \dots \otimes \Phi_{\alpha_N})[\tau], \bar{\rho}\right)$  between the target state  $\bar{\rho}$  and all  $N$ -qubit states  $\tau$  whose Bell value  $\text{Tr}(\tau \mathcal{B}(a_{\alpha_1}, \dots, a_{\alpha_N}))$  exceeds a certain value  $\beta'$ . This implies the following optimization:

$$\begin{aligned}
O_{\min}(\beta') &= \min_{\tau, a_1, \dots, a_N} O\left((\Phi(a_1) \otimes \dots \otimes \Phi(a_N))[\tau], \bar{\rho}\right) \\
&\text{s.t. } \text{Tr}(\tau \cdot \mathcal{B}(a_1, \dots, a_N)) \geq \beta' \\
&\tau \geq 0 \\
&\text{Tr}(\tau) = 1 \\
&\tau^\dagger = \tau.
\end{aligned} \tag{74}$$

For simplicity, we dropped the box index  $\alpha_i \rightarrow i$  in these expressions.

If we fix the values of  $\beta'$  and the angles  $a_1, \dots, a_N$ , the problem reduces to a linear optimization, which can be done very efficiently using semidefinite programming for example. So, we use the following strategy: For every  $\beta'$ , we fix  $a_1, \dots, a_N$ , we run the linear optimization and then minimize the overlap over the angles. Finally, the convex function  $f(\beta)$  of Eq. (71) is obtained as the convex roof of  $O_{\min}(\beta')$ .

The optimization (74) is then directly formulated

$$\begin{aligned}
& O\left((\Phi(a_1) \otimes \dots \otimes \Phi(a_N))[\tau], \bar{\rho}\right) = \\
& \text{Tr}\left((\Lambda(a_1) \otimes \dots \otimes \Lambda(a_N))[\tau] \cdot \bar{\rho}\right),
\end{aligned} \tag{75}$$

in terms of the extraction maps  $\Lambda(a_i)[\cdot] = \text{tr}_{\text{ext}} \Phi(a_i)[\cdot]$  induced by the isometries. The resulting  $\Lambda(a_i)$  is some completely positive trace preserving (CPTP) map on a single qubit. Reciprocally any CPTP map can be dilated to an isometry, hence choosing the dependence of the local isometries  $\Phi(a_i)$  on the angle amounts to choose a parametric family of extraction maps  $\Lambda(a_i)$ .

To run the optimization, it remains to choose the dependence of the extraction maps  $\Lambda(a_i)$  on the angles  $a_i$ . To do so, we follow the analytic studies presented in [3], where single-qubit dephasing channels are used:

$$\Lambda(a)[\rho] := \frac{1+g(a)}{2}\rho + \frac{1-g(a)}{2}\Gamma_a \rho \Gamma_a, \tag{76}$$

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

$$\Gamma_a = \begin{cases} \sigma_x & \text{for } a \in [0, \frac{\pi}{4}], \\ \sigma_z & \text{for } a \in [\frac{\pi}{4}, \frac{\pi}{2}]. \end{cases} \tag{77}$$

So far we have assumed that we know the Bell test that is suitable for the robust certification of the target state  $\bar{\rho}$ . We will now give a hint on how, given a target state, the research of such a Bell test can be tackled.

**Appendix D.4 Devising Bell tests tailored to the robust certification of  $N$ -qubit states** – In this appendix, we briefly describe the approach that we used to devise Bell tests tailored to the certification of 4-qubit state family  $|\xi_\varphi\rangle$ . While, in principle, this approach can be used

for any  $N$ -qubit state, its convergence is not guaranteed. In any case, it provides necessary conditions that have to be fulfilled by a Bell test to be suitable for the certification of a given state, drastically limiting the set of potential candidates. The full details will be described elsewhere [8].

Since we want the extracted fidelity to be one in the ideal case, the state  $\bar{\rho}$  has to be the unique state that attains the maximal quantum value of the Bell test for a given realization of the observable. Hence, the simplest necessary condition for a Bell test to be suitable for the certification of a given state, is that this state gives the maximal Bell value for a fixed realization of the observables. As a first step, we thus consider a set of Hermitian operators that have the state  $|\xi_\varphi\rangle$  for unique maximal eigenstate. Such a set of operators can be obtained by applying the gate  $CU_\varphi$  on the convex sums of stabilizers of the initial Bell pairs  $|\phi_2^+\rangle|\phi_2^+\rangle$ . Trivially, any such operator, except those for which the weights of some stabilizers are identically zero, has the state  $|\xi_\varphi\rangle$  as unique maximal eigenstate with eigenvalues 1. Since we want to use the Jordan lemma, we furthermore restrict ourselves to operators that can be expressed as a linear combination of correlators that only contain identity and two other Paulis per party. The Jordan lemma ensures that any such Bell test attains its maximal value for the case of qubits, i.e. when the boxes correspond

to some Pauli matrices and the state is a four-qubit state. Note that different expansions of an operator in correlators (related by local rotations of the Paulis) correspond to different Bell tests. Similarly, a Bell test, seen as a sum of correlators, can correspond to different Bell operators, via different assignments of operators to the measurement boxes. If the aim of this test is to certify the target state, not only the state has to be the maximal eigenstate for a fixed assignment (a fixed Bell operator), but the corresponding eigenvalue, that we conveniently set to  $\lambda = 1$ , has to be globally maximal. Hence, as a second step, we discard the operators among all the candidates to the Bell test whose maximal eigenvalue is not a local maximum with respect to small perturbations of the measurement boxes<sup>4</sup>. This can be done by standard second order eigenvalue perturbation methods: check that  $\lambda$  is not perturbed at first order and all second order terms are negative. Moreover, the sensitivity of maximal eigenvalue to small misalignments of the measurement boxes (the negativity of the second order derivatives of  $\lambda$ ) is a good indication of the fact that the Bell value is the most sensitive to a drop in state's fidelity. This is precisely what is required for a robust certificate. Therefore, as a last step we look for the Bell test for which the maximal eigenvalue is the most sensitive to perturbations in all possible directions.

---

The resulting family of Bell tests for  $|\xi_\varphi\rangle$  reads

$$B_\varphi = \frac{1}{20} \left\langle 2\sqrt{2}B_1^{(1)}(A_0^{(1)} + A_1^{(1)}) + s_\varphi \left( \sqrt{2}B_1^{(1)} + A_0^{(1)} + A_1^{(1)} \right) \left( A_0^{(2)}(B_0^{(2)} + B_1^{(2)}) + A_1^{(2)}(B_0^{(2)} - B_1^{(2)}) \right) \right. \\ \left. + \left( \sqrt{2}c_\varphi + c_\varphi B_1^{(1)}(A_0^{(1)} + A_1^{(1)}) + 2B_0^{(1)}(A_0^{(1)} - A_1^{(1)}) \right) \left( A_0^{(2)}(-B_0^{(2)} + B_1^{(2)}) + A_1^{(2)}(B_0^{(2)} + B_1^{(2)}) \right) \right\rangle, \quad (78)$$

with  $c_\varphi = \cos(\varphi)$ ,  $s_\varphi = \sin(\varphi)$ , and where  $\langle \cdot \rangle$  stands for the expectation value. In practice, one has to determine multiple average values of the form

$$\left\langle B_m^{(1)} A_n^{(1)} A_p^{(2)} B_q^{(2)} \right\rangle = \sum_{b^{(1)}, a^{(1)}, a^{(2)}, b^{(2)}} b^{(1)} a^{(1)} a^{(2)} b^{(2)} P \left( b^{(1)}, a^{(1)}, a^{(2)}, b^{(2)} | B_m^{(1)} A_n^{(1)} A_p^{(2)} B_q^{(2)} \right) \quad (79)$$

Here  $a^{(i)}, b^{(i)} = \pm 1$  label the measurement outcomes for each party, and  $m, n, p, q \in \{0, 1\}$ . If only three or two parties are involved, the definition of the expectation value is adapted accordingly. For the assignments

$$\begin{aligned} B_0^{(1)} &= \sigma_x, & B_1^{(1)} &= \sigma_z, \\ A_0^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), & A_1^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x), \\ A_0^{(2)} &= -c_\varphi \sigma_x + s_\varphi \sigma_z, & A_1^{(2)} &= c_\varphi \sigma_z + s_\varphi \sigma_x, \\ B_0^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), & B_1^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x), \end{aligned}$$

$B_\varphi$  attains the maximal quantum value  $B_\varphi = 1$  for the state  $|\xi_\varphi\rangle = CU_\varphi |\phi_2^+\rangle |\phi_2^+\rangle$ . Recall that the gate  $CU_\varphi = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes e^{-i\varphi\sigma_y}$  is applied on  $A^{(1)}$  holding the control qubit and  $A^{(2)}$  holding the target one. The maximally

---

<sup>4</sup> In our case of complementary qubit measurements  $X$  and  $Z$  the

perturbation is  $X \rightarrow X(1 - \frac{\delta^2}{2}) + \delta Z$  and  $Z \rightarrow Z(1 - \frac{\delta^2}{2}) + \delta X$ .

entangled states  $|\phi_2^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  are shared between  $A^{(1)}$  with  $B^{(1)}$  and  $A^{(2)}$  with  $B^{(2)}$ . A curious reader might also be interested to learn that the local bounds for the Bell tests  $B_\varphi$  are given by

$$B_\varphi \leq \frac{(\sqrt{2}+1)(c_\varphi + s_\varphi) + 2}{5\sqrt{2}} \in [0.62, 0.77]. \quad (80)$$

Let us take a closer look on the  $CNOT = |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes \sigma_x$  gate, which only belongs to the  $CU_\varphi$  family up to local unitaries:

$$CU_{\frac{\pi}{2}} = e^{-i\frac{\pi}{4}} (e^{i\frac{\pi}{4}\sigma_z} \otimes e^{-i\frac{\pi}{4}\sigma_z}) CNOT (\mathbb{1} \otimes e^{i\frac{\pi}{4}\sigma_z}). \quad (81)$$

Using  $\mathbb{1} \otimes e^{i\frac{\pi}{4}\sigma_z} |\phi_2^+\rangle = e^{i\frac{\pi}{4}\sigma_z} \otimes \mathbb{1} |\phi_2^+\rangle$  it follows that the action of the CNOT on the qubits of the two Bell states belonging to Alices yields the state  $|\xi_{CNOT}\rangle = CNOT |\phi_2^+\rangle |\phi_2^+\rangle$  satisfying

$$|\xi_{\frac{\pi}{2}}\rangle = (\mathbb{1}_{B^{(1)}} \otimes (e^{i\frac{\pi}{4}\sigma_z})_{A^{(1)}} \otimes (e^{-i\frac{\pi}{4}\sigma_z})_{A^{(2)}} \otimes (e^{i\frac{\pi}{4}\sigma_z})_{B^{(2)}}) |\xi_{CNOT}\rangle \quad (82)$$

up to an irrelevant global phase. Hence,  $|\xi_{CNOT}\rangle$  can be certified with the Bell test  $B_{\frac{\pi}{2}}$  in Eq. (78) with settings rotated by the local unitaries, i.e. with

$$\begin{aligned} B_0^{(1)} &= \sigma_x, & B_1^{(1)} &= \sigma_z, \\ A_0^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_y + \sigma_z), & A_1^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_y), \\ A_0^{(2)} &= \sigma_z, & A_1^{(2)} &= -\sigma_y, \\ B_0^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_y + \sigma_z), & B_1^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_y). \end{aligned}$$

Finally, recall that to obtain the certificate of the gate the initial state  $|\phi_2^+\rangle |\phi_2^+\rangle$  has to be certified with the same measurement boxes on Bob's side as for the final state. Since the measurement settings for  $B^2$  are different for  $|\xi_{CNOT}\rangle$  compared to  $|\xi_{\frac{\pi}{2}}\rangle$ , we cannot use the settings for  $B_0$  given after Eq. (78). However, this can be easily resolved using the symmetry of the state

$$|\phi_2^+\rangle |\phi_2^+\rangle = (\mathbb{1}_{B^{(1)}} \otimes \mathbb{1}_{A^{(1)}} \otimes (e^{-i\frac{\pi}{4}\sigma_z})_{A^{(2)}} \otimes (e^{i\frac{\pi}{4}\sigma_z})_{B^{(2)}}) |\phi_2^+\rangle |\phi_2^+\rangle. \quad (83)$$

Hence, in the case of the CNOT gate, the initial state can be certified with  $B_0$  in Eq. (78) and the settings

$$\begin{aligned} B_0^{(1)} &= \sigma_x, & B_1^{(1)} &= \sigma_z, \\ A_0^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z), & A_1^{(1)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x), \\ A_0^{(2)} &= \sigma_y, & A_1^{(2)} &= \sigma_z, \\ B_0^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_y + \sigma_z), & B_1^{(2)} &= \frac{1}{\sqrt{2}}(\sigma_z - \sigma_y). \end{aligned}$$

**Appendix D.5 Results**– Below we present the results of the optimization for the states discussed in the main text.

*Maximally entangled two-qubit state*– The maximally entangled two-qubit state  $|\phi_2^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$  has the two following stabilizers  $\sigma_x \otimes \sigma_x$  and  $\sigma_z \otimes \sigma_z$  in the X-Z plane. Any operator

$$\mathcal{B}(p) = p \sigma_x \otimes \sigma_x + (1-p) \sigma_z \otimes \sigma_z, \quad (84)$$

with  $0 < p < 1$ , has  $|\phi_2^+\rangle$  as the unique maximal eigenstate with eigenvalue 1. Applying the approach of the previous section to the operators  $\mathcal{B}(p)$ , we find that the

Bell test

$$\frac{1}{2\sqrt{2}} \langle A_0(B_0 + B_1) + A_1(B_0 - B_1) \rangle, \quad (85)$$

with  $\langle A_m B_n \rangle = \sum_{a,b=\pm 1} a b P(a,b|A_m, B_n)$ , is the most sensitive to the perturbation of the boxes. It corresponds to the operator  $\mathcal{B}(\frac{1}{2})$  for the assignments  $A_0 = \sigma_x$  and  $A_1 = \sigma_z$  for Alice's measurements, and  $B_0 = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$  and  $B_1 = \frac{1}{\sqrt{2}}(\sigma_x - \sigma_z)$  for Bob's measurements. This Bell test is proportional to the well-known CHSH inequality [9].

Let us now illustrate our numerical optimization method on the CHSH inequality. The results we find match the already known analytical result of Ref. [3],

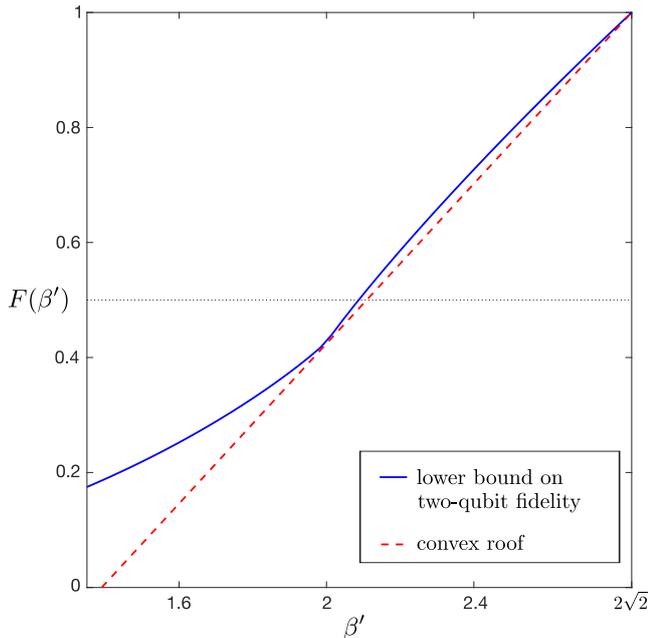


FIG. 5. Application of our numerical approach to CHSH. The blue line is the result of the 2-qubit optimization. The red dashed line is the convex roof and therefore a valid device-independent lower bound on the fidelity of a two-qubit maximally entangled state. The black dashed-dotted line is the non-trivial bound of  $\frac{1}{2}$ .

highlighting the validity and applicability of our approach.

Eq. (73) ensures that Alice's observables  $A_0, A_1$  on her  $k$ -block and Bob's observables  $B_0, B_1$  on his  $l$ -block can be written as

$$\begin{cases} A_{0,k} &= \cos(a_k)\sigma_x + \sin(a_k)\sigma_z \\ A_{1,k} &= \cos(a_k)\sigma_x - \sin(a_k)\sigma_z \end{cases} \quad (86)$$

$$\begin{cases} B_{0,l} &= \cos(b_l)\sigma_x + \sin(b_l)\sigma_z \\ B_{1,l} &= \cos(b_l)\sigma_x - \sin(b_l)\sigma_z. \end{cases} \quad (87)$$

The two-qubit CHSH operator on the  $(k, l)$ -block is thus expressed as

$$\mathcal{B}(a_k, b_l) = 2[(\cos(a_k)\sigma_x + \sin(a_k)\sigma_z) \otimes \cos(b_l)\sigma_x + (\cos(a_k)\sigma_x - \sin(a_k)\sigma_z) \otimes \sin(b_l)\sigma_z]. \quad (88)$$

Up to local unitaries,  $|\phi_2^+\rangle$  is the eigenstate of  $\mathcal{B}(\frac{\pi}{4}, \frac{\pi}{4})$  corresponding to the largest eigenvalue of  $2\sqrt{2}$ .

With the extraction maps of Eq. (76) we do the optimization (74) for  $\beta' \in [1.35, 2\sqrt{2}]$ . We chose this interval to cover the whole range of fidelities  $[0, 1]$ . As depicted in Fig. 5, the obtained bound  $O_{\min}(\beta')$  is not a convex function (solid line), but its convex roof (straight dashed line) is easy to obtain and is given by

$$f(\bar{\beta}) = \frac{1}{2} \left( 1 + \frac{\bar{\beta} - \beta^*}{2\sqrt{2} - \beta^*} \right), \quad (89)$$

with  $\beta^* \approx 2.11$ . Applying the Proposition .8 and the Lemma .6, we conclude that a CHSH value  $\beta$  implies a lower bound on the fidelity

$$F(\text{Tr}_{\text{ext}}(\Phi[\rho]), |\phi_2^+\rangle\langle\phi_2^+|) \geq \sqrt{\frac{1}{2} \left( 1 + \frac{\beta - \beta^*}{2\sqrt{2} - \beta^*} \right)}. \quad (90)$$

*Four-qubit entangled states  $|\xi_\varphi\rangle$* – We follow the same steps for the states  $|\xi_\varphi\rangle$  and corresponding Bell tests  $B_\varphi$ . This allows us to conclude that a Bell value  $\gamma$  leads to the lower bound

$$F(\text{Tr}_{\text{ext}}(\Phi[\rho]), |\xi_\varphi\rangle\langle\xi_\varphi|) \geq \sqrt{\frac{1}{2} \left( 1 + \frac{\gamma - \gamma_\varphi^*}{1 - \gamma_\varphi^*} \right)}, \quad (91)$$

where the constant  $\gamma_\varphi^*$  depends on the parameter  $\varphi$ . For the ten values of  $\varphi$  that we analyzed, we find that  $\gamma_\varphi^* \in [0.795, 0.85]$ , with the best case  $\gamma_{\frac{\pi}{2}}^* \leq 0.795$  that corresponds to the equivalent of a CNOT gate, and the worst case  $\gamma_{\frac{\pi}{4}}^* \leq 0.85$  corresponding to the gate  $CU_{\frac{\pi}{4}}$  which is equivalent to the square root of a CNOT.

---

\* These authors equally contributed

- [1] D. Mayers, A. Yao, QIC, Vol.4 No. 4 July 5, pp273-286 (2004)
- [2] T. H. Yang, T. Vértesi, J-D. Bancal, V. Scarani, M. Navascués, Phys. Rev. Lett. 113, 040401 (2014)
- [3] J. Kaniewski, Phys. Rev. Lett. 117, 070402 (2016)
- [4] A. Coladangelo, K. T. Goh, V. Scarani, Nature Communications 8, 15485 (2017)
- [5] I. Šupić, A. Coladangelo, R. Augusiak, A. Acín, arXiv:1707.06534
- [6] P. E. M. F. Mendonca, R. d. J. Napolitano, M. A. Marchioli, C. J. Foster, Y-C. Liang, Phys. Rev. A 78, 052330 (2008)
- [7] V. Scarani, Acta Physica Slovaca 62, 347 (2012)
- [8] P. Sekatski *et al.* (In preparation)
- [9] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. 23, 880 (1969)



## SELF-TESTING OF QUANTUM MEASUREMENTS

In quantum communication we usually use projective measurements, e.g. when reading out states after their transmission. These measurements are strong measurements in the sense that they extract a maximum amount of information and also disturb the system maximally. The unpleasant consequence is that interesting features such as entanglement are destroyed. While in the previously mentioned scenario of state detection this consequence is irrelevant, there are situations in which one wants to extract information while upholding the entangled structure of the system. An example for such a case is the generation of random numbers in a device-independent way [29]. Whereas with projective measurements one can extract at most one certified random bit per maximally-entangled state, with successive quantum measurements one can in principle extract an arbitrary number of certified bits. We denote measurements that serve this purpose of non-maximal disturbance quantum measurements. They have a classical as well as a quantum output, with the classical output being the measurement outcome (just as in the projective case) while the quantum output is the post-measurement state (for projective measurements the state after the measurement is always of product-form and hence classical).

In this chapter we explain how quantum measurements can be self-tested. The line of argument follows closely the one of chapter 4: Again we use state certificates to lower-bound the fidelity of a channel - in this case the measurement. We also discuss our technique by studying the example of a qubit quantum measurement characterized by two Kraus operators. For this we need the Bell inequalities tailored to partially-entangled two-qubit states that we derived in section 3.2.



Paper No. 3

# Device-independent Characterization of Generalized Measurements

Sebastian Wagner, Jean-Daniel Bancal, Nicolas Sangouard, and Pavel Sekatski

arXiv:1812.02628



# Device-independent characterization of generalized measurements

Sebastian Wagner,<sup>1</sup> Jean-Daniel Bancal,<sup>1</sup> Nicolas Sangouard,<sup>1</sup> and Pavel Sekatski<sup>1</sup>

<sup>1</sup>*Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland*

(Dated: December 7, 2018)

Among certification techniques, those based on the violation of Bell inequalities are appealing because they do not require assumptions on the underlying Hilbert space dimension and on the accuracy of calibration methods. Such device-independent techniques have been proposed to certify the quality of entangled states, unitary operations, projective measurements following von Neumann's model, and rank-one positive-operator-valued measures (POVM). Here, we show that they can be extended to the characterization of generalized measurements with post-measurement states that are not fully determined by the measurement result. We provide concrete recipes that can be used in realistic scenarios where the certification devices are noisy.

*Introduction* — Experiments using either NV centers [1], photon pair sources [2, 3] or neutral atoms [4] have recently been used to test Bell inequalities [5] in a very convincing way. The observed Bell inequality violations have brought new and fascinating insights about nature by showing that some correlations cannot be explained by locally causal models. These experiments also revolutionize branches of applied physics like randomness generation [6–12] by making it device-independent, i.e. the randomness guarantees hold without assumptions on the underlying Hilbert space dimension and on the accuracy of calibration methods.

The possibility of randomness generation from Bell inequalities is clear when one realizes that the only situation allowing for a maximal quantum violation of the simplest Bell inequality [13], within the quantum formalism, consists in using complementary Pauli measurements on a maximally-entangled two-qubit state [14, 15]. This means that the violation of a Bell inequality can certify quantum states and von Neumann measurements directly, without resorting to tomography. Mayers and Yao were among the very first ones to highlight the usefulness of Bell tests as characterization methods, a technique that they called self-testing [16]. Self-testing has been applied to many entangled states [15–19], projective measurements [16, 20–24], and unitary operations [25].

Efforts are being devoted to characterise measurements not captured by the usual von Neumann model. Refs. [26, 27] for example, showed how to characterise rank-one POVMs that are not composed of orthogonal projection. Less is known for measurements whose post-measurement state is not fully determined by the measurement result. In this case, the measurement statistics and the post-measurement states have to be considered together in order to verify that a measurement achieves the ideal trade-off between disturbance and information gain. Such generalized measurements, also sometimes referred to as weak measurements or

quantum instruments, can be more efficient in practice than projective or rank-one measurements e.g. for generating randomness. Whereas randomness generation based on projective measurements requires at least as many maximally-entangled states as the number of certified random bits, an arbitrary number of random bits can in principle be extracted from a single maximally-entangled state by applying successive generalized measurements [28–31]. The certification of such measurements is thus not only of fundamental interest but could be used in practice to characterise the potential of an actual measurement for producing large amounts of device-independent randomness with a single entangled state.

In this manuscript, we provide a recipe to certify generalized measurements by lower-bounding the fidelity of the maps associated with each outcome of the measurement. We also derive a new class of Bell inequalities suitable for the robust self-testing of partially-entangled two-qubit states. Our final recipe is realistically robust to experimental noise.

*Device-independent certification of generalized measurements: Formulation*— Consider an ideal noise-free measurement  $\overline{\mathcal{M}}$  with  $k$  outcomes operating on qubits. It is represented by a collection of  $k$  Kraus operators  $\{\overline{K}_\ell\}_{\ell=0}^{k-1}$  satisfying the completeness relation  $\sum_\ell \overline{K}_\ell^\dagger \overline{K}_\ell = \mathbb{1}$ . Each Kraus operator defines a completely positive map  $\rho \mapsto \overline{K}_\ell \rho \overline{K}_\ell^\dagger$ . Given a state  $\rho_{\mathcal{B}} \in L(\mathbb{C}^2)$ , the probability to observe the outcome  $\ell$  is given by the Born rule  $\overline{p}_\ell = \text{Tr}(\overline{K}_\ell \rho_{\mathcal{B}} \overline{K}_\ell^\dagger)$  and the post-measurement state for this outcome is  $\varrho_\ell = \frac{1}{\overline{p}_\ell} \overline{K}_\ell \rho_{\mathcal{B}} \overline{K}_\ell^\dagger$ .

Von Neumann measurements and rank-one POVMs correspond to the specific case where  $\{\overline{K}_\ell\}$  are proportional to projectors onto pure states  $|\psi_\ell\rangle$ , that is  $\overline{K}_\ell = \eta_\ell |\psi_\ell\rangle\langle\psi_\ell|$ . In this case, the post-measurement state corresponding to the outcome  $\ell$  is  $|\psi_\ell\rangle$  inde-

pendently of the pre-measured state of the system. Performing such a measurement on a physical system extracts full information about its state but also disturbs it maximally, e.g. it breaks all entanglement the system might have with the rest of the world.

Weak measurements on the other hand introduce less disturbance in the system at the price of extracting less information [32]. The corresponding Kraus operators are not represented by rank-one matrices, i.e.  $\overline{K}_\ell^2 \neq \text{Tr}(\overline{K}_\ell)\overline{K}_\ell$  in general. This has the benefit of preserving interesting and useful features such as entanglement while nevertheless revealing information about the system. As an example, the Kraus operators

$$\overline{K}_0(\theta) = \cos(\theta)|0\rangle\langle 0| + \sin(\theta)|1\rangle\langle 1|, \quad (1)$$

$$\overline{K}_1(\theta) = \sin(\theta)|0\rangle\langle 0| + \cos(\theta)|1\rangle\langle 1| \quad (2)$$

are generalized measurements which tend to be projective in the limit  $\theta \rightarrow 0$ , and the identity in the weak limit when  $\theta \rightarrow \pi/4$ . Such a family of measurements is sufficient to implement the scheme proposed in [29–31] to produce more randomness than possible with von Neumann measurements.

Whether the considered measurement is a von Neumann measurement or not, it can be fully characterised by the map

$$\begin{aligned} \overline{\mathcal{M}} : L(\mathbb{C}^2) &\rightarrow L(\mathbb{C}^2 \otimes \mathcal{H}_R), \\ \sigma &\mapsto \sum_\ell \left( \overline{K}_\ell \sigma \overline{K}_\ell^\dagger \right) \otimes |\ell\rangle\langle \ell|_R, \end{aligned} \quad (3)$$

where we have introduced a register  $R$  indicating the outcome. In comparison to the map generated by a single Kraus operator  $\overline{K}_\ell$ , the map  $\overline{\mathcal{M}}$  is trace-preserving by construction, hence defines a quantum channel. Similarly, a possibly noisy measurement acting on a Hilbert space  $\mathcal{H}_\mathcal{B}$  can be described by the following map

$$\begin{aligned} \mathcal{M} : L(\mathcal{H}_\mathcal{B}) &\rightarrow L(\mathcal{H}_\mathcal{B} \otimes \mathcal{H}_R), \\ \sigma &\mapsto \sum_\ell M_\ell[\sigma] \otimes |\ell\rangle\langle \ell|_R, \end{aligned} \quad (4)$$

where  $M_\ell$  are the completely positive maps associated to the outcomes  $\ell$ . In general, these maps may not be expressed in terms of a single Kraus operator, but as a combination of several ones, i.e.  $M_\ell[\sigma] = \sum_m K_{\ell,m} \sigma K_{\ell,m}^\dagger$ , with  $\sum_{\ell,m} K_{\ell,m}^\dagger K_{\ell,m} = \mathbb{1}$ . In order to show that a measurement described by  $\mathcal{M}$  acts like a target measurement  $\overline{\mathcal{M}}$ , it is sufficient to identify a subspace of  $\mathcal{H}_\mathcal{B}$  on which the action of  $\mathcal{M}$  is identical to  $\overline{\mathcal{M}}$ . Moreover, a map is fully described by its action on half of a maximally entangled state. Therefore, one can demonstrate that the considered measurement is identical to the target one

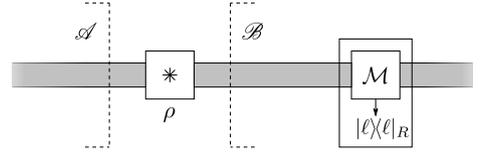


FIG. 1: Scheme of the actual experiment that is used to characterise the measurement box  $\mathcal{M}$  in a device-independent way. The source which is represented by a box with a star, produces an unknown bipartite state  $\rho$  shared between  $\mathcal{A}$  and  $\mathcal{B}$ . Party  $\mathcal{B}$  performs the measurement. The pre- and post-measurement states can be measured with additional measurements named  $A_{0/1}$  for party  $\mathcal{A}$  and  $B_{0/1/2/3}$  for party  $\mathcal{B}$  that are also unknown (not represented).

by showing that there exist completely positive trace-preserving maps

$$\Lambda_{\mathcal{B}}^i : L(\mathbb{C}^2) \rightarrow L(\mathcal{H}_\mathcal{B}), \quad \Lambda_{\mathcal{B}}^o : L(\mathcal{H}_\mathcal{B}) \rightarrow L(\mathbb{C}^2), \quad (5)$$

such that

$$(\mathbb{1} \otimes \Lambda_{\mathcal{B}}^o \circ \mathcal{M} \circ \Lambda_{\mathcal{B}}^i)[|\phi^+\rangle\langle \phi^+|] = (\mathbb{1} \otimes \overline{\mathcal{M}})[|\phi^+\rangle\langle \phi^+|]. \quad (6)$$

The injection and output maps identify subspaces and subsystems in which the measurement  $\mathcal{M}$  acts as the reference measurement  $\overline{\mathcal{M}}$ , see Fig. 2. Note that since all possible outcomes appear in the definition of the maps  $\overline{\mathcal{M}}$  and  $\mathcal{M}$ , equality (6) guarantees at the same time that the outcome states are as expected *and* that each outcome appears with the desired probability.

The previous equality cannot be satisfied in an actual experiment due to unavoidable imperfections. We thus propose an extension for quantifying the distance  $\mathcal{F}(\mathcal{M}, \overline{\mathcal{M}})$  between  $\mathcal{M}$  and  $\overline{\mathcal{M}}$  using

$$\begin{aligned} \mathcal{F}(\mathcal{M}, \overline{\mathcal{M}}) = & \quad (7) \\ \max_{\Lambda_{\mathcal{B}}^i, \Lambda_{\mathcal{B}}^o} & F\left( (\mathbb{1} \otimes \Lambda_{\mathcal{B}}^o \circ \mathcal{M} \circ \Lambda_{\mathcal{B}}^i)[|\phi^+\rangle\langle \phi^+|], (\mathbb{1} \otimes \overline{\mathcal{M}})[|\phi^+\rangle\langle \phi^+|] \right), \end{aligned}$$

where  $F(\rho, \sigma) = \text{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$  is the Uhlmann fidelity between two states  $\rho$  and  $\sigma$ .

*Device-independent certification of generalized measurements: Recipe*– The aim of this section is to show how the quantity (7) can be lower bounded in the setup presented in Fig. 3. In addition to the source producing the bipartite state  $\rho$  and the measurement  $\mathcal{M}$  to be characterised, each party has a measurement box. The box of party  $\mathcal{A}$  has two inputs  $A_0$  and  $A_1$  while the one of party  $\mathcal{B}$  has four inputs  $B_0, B_1, B_2, B_3$ . For each measurement input, a binary outcome is obtained called  $a$  for  $\mathcal{A}$  and  $b$  for  $\mathcal{B}$ , with  $a, b = \pm 1$ . The measurement  $\mathcal{M}$  can be applied by party  $\mathcal{B}$  before the measurement input is chosen. Although there is no assumption about the Hilbert space dimension and on the proper calibration of the measurement devices,  $\mathcal{M}$  can be characterised

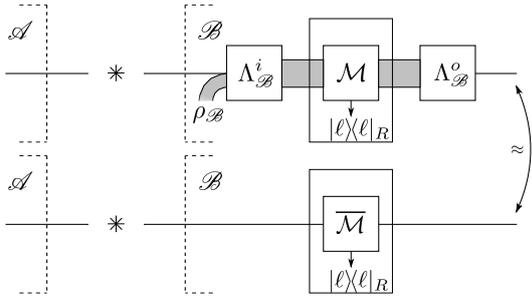


FIG. 2: To characterise an unknown measurement  $\mathcal{M}$ , we compare the action of this black box supplemented with injection maps  $\Lambda_{\mathcal{B}}^i$  and  $\Lambda_{\mathcal{B}}^o$  with the action of a reference measurement  $\bar{\mathcal{M}}$  on one half of a maximally entangled two-qubit state  $|\phi^+\rangle$ .

in two steps: Step I identifies the quality of the state produced by the source while step II is used to characterise the states after each outcome of the measurement to be certified. The certifications associated to steps I and II are then combined to bound  $\mathcal{F}(\mathcal{M}, \bar{\mathcal{M}})$  as defined in Eq. (7)

Let us first focus on step I. In this step, the measurement settings  $A_{0/1}$  and  $B_{0/1}$  are chosen freely and applied directly on the state produced by the source  $\rho$  so that party  $\mathcal{A}$  and  $\mathcal{B}$  can estimate the Clauser, Horne, Shimony, and Holt (CHSH) value [13]

$$\beta = \sum_{k,j=0}^1 (-1)^{k \cdot j} \langle A_k B_j \rangle. \quad (8)$$

Here,  $\langle A_k B_j \rangle = \sum_{a,b} (-1)^{a+b} P(a, b | A_k, B_j)$  is the expectation value of measurements  $A_k$  and  $B_j$ . The CHSH value allows one to bound the fidelity of  $\rho$  with a maximally-entangled two-qubit state. In particular, the results of Ref. [33] show that there exist local extraction maps  $\Lambda_{\mathcal{A}}^i : L(\mathcal{H}_{\mathcal{A}}) \rightarrow \mathbb{C}^2$  and  $\tilde{\Lambda}_{\mathcal{B}}^i : L(\mathcal{H}_{\mathcal{B}}) \rightarrow \mathbb{C}^2$  such that

$$F\left((\Lambda_{\mathcal{A}}^i \otimes \tilde{\Lambda}_{\mathcal{B}}^i)[\rho], |\phi^+\rangle\right) \geq F^i = \sqrt{\frac{1}{2} + \frac{1}{2} \cdot \frac{\beta - \beta^*}{2\sqrt{2} - \beta^*}}, \quad (9)$$

where  $\beta^* = \frac{2(8+7\sqrt{2})}{17} \approx 2.11$ . Whenever  $\beta = 2\sqrt{2}$ , the formula (9) certifies that the source produces  $|\phi^+\rangle$  up to local maps, these maps being explicitly defined from the quantum description of the measurement inputs  $A_{0/1}$  and  $B_{0/1}$ . In this case, one also knows that  $\mathcal{A}$ 's inputs correspond (up to the same maps) to the Pauli measurements  $\bar{A}_0 = \sigma_z$  and  $\bar{A}_1 = \sigma_x$  while  $\mathcal{B}$ 's inputs correspond to  $\bar{B}_{0/1} = \frac{1}{\sqrt{2}}(\sigma_z \pm \sigma_x)$ .

In step II, party  $\mathcal{B}$  applies  $\mathcal{M}$ . Let us first consider the

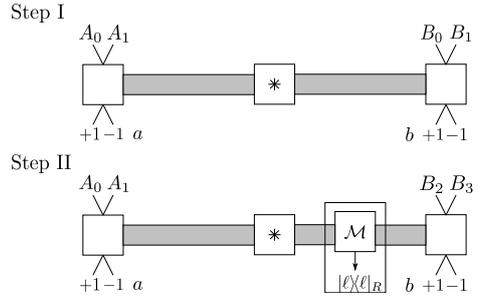


FIG. 3: Recipe for bounding the quality of the measurement box  $\mathcal{M}$  in 2 steps. Step I is used to characterize the state of the source while Step II gives a certificate of the post-measurement states. The statistics recorded in each step is then used to certify the quality of measurement  $\mathcal{M}$  device-independently.

state conditioned on the outcome 0,

$$\varrho_0 = \frac{M_0[\rho]}{\text{Tr}(M_0[\rho])}, \quad (10)$$

which is characterized using  $A_{0/1}$  and  $B_{2/3}$ . In particular, parties  $\mathcal{A}$  and  $\mathcal{B}$  are interested in the Bell inequality

$$\begin{aligned} \mathcal{I}_\theta &= \frac{1}{4} \left[ \frac{\langle A_0(B_2 - B_3) \rangle}{\sin(b_\theta)} + \frac{\sin(2\theta)}{\cos(b_\theta)} \langle A_1(B_2 + B_3) \rangle \right. \\ &\quad \left. + \cos(2\theta) \left( \langle A_0 \rangle + \frac{\langle B_2 - B_3 \rangle}{2 \sin(b_\theta)} \right) \right] \\ &\leq \frac{1}{4} \left[ \cos(2\theta) + (2 + \cos(2\theta)) \sqrt{\frac{7 - \cos(4\theta)}{5 + \cos(4\theta)}} \right] \end{aligned} \quad (11)$$

with  $b_\theta = \arctan \sqrt{(1 + \frac{1}{2} \cos^2(2\theta)) / \sin^2(2\theta)}$ , whose maximal quantum value is one by construction. We derived this Bell inequality using the variational method presented in [25, 34] to self-test partially-entangled two-qubit pure states in a particularly robust manner. Note that for  $\theta = \frac{\pi}{4}$ , Ineq. (11) is the re-normalized CHSH inequality. However, we emphasize that Ineq. (11) is not equivalent to the tilted-CHSH inequality of Refs. [35, 36] but was carefully constructed for the demands of self-testing presented here. The knowledge of  $\mathcal{I}_\theta$  allows one to bound the fidelity of the conditional state  $\varrho_0$ , that is, to guarantee the existence of local maps  $\Lambda_{\mathcal{A}}^i : L(\mathcal{H}_{\mathcal{A}}) \rightarrow \mathbb{C}^2$  and  $\tilde{\Lambda}_{\mathcal{B}}^i : L(\mathcal{H}_{\mathcal{B}}) \rightarrow \mathbb{C}^2$  such that

$$F((\Lambda_{\mathcal{A}}^i \otimes \tilde{\Lambda}_{\mathcal{B}}^i)[\varrho_0], |\phi_\theta^0\rangle) \quad (12)$$

$$\geq F_\theta^0 = \sqrt{\cos^2(\theta) + (1 - \cos^2(\theta)) \frac{\mathcal{I}_\theta - \mathcal{I}_\theta^*}{1 - \mathcal{I}_\theta^*}}. \quad (13)$$

Here,  $\mathcal{I}_\theta^*$  is a cutoff parameter corresponding to the violation for which the fidelity matches the square of the largest Schmidt coefficient of  $|\phi_\theta^0\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle$ . Fig. 4 shows the critical value  $\mathcal{I}_\theta^*$  for both this inequality and the tilted

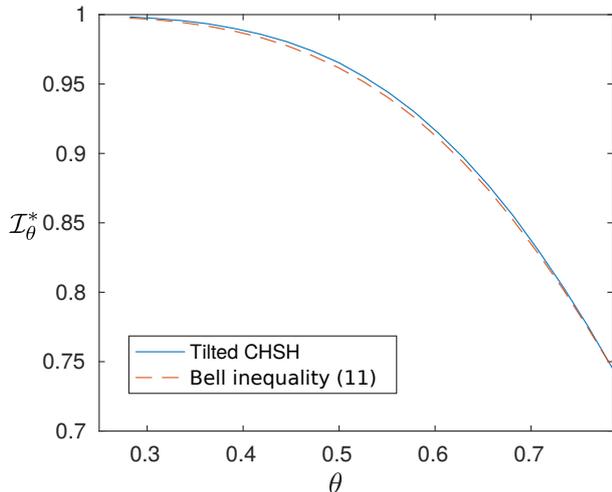


FIG. 4: Plot of the cutoff parameter  $\mathcal{I}_\theta^*$  as a function of  $\theta$ . The newly found Bell inequality provides tighter self-testing bounds for partially entangled states.

CHSH inequality, as calculated in the Appendix. The previous bound shows that whenever  $\mathcal{I}_\theta$  reaches its maximal quantum value  $\mathcal{I}_\theta = 1$ , the state conditioned on the outcome 0 corresponds to the state  $|\phi_\theta^0\rangle$  up to local maps, these maps being explicitly defined from the quantum description of the measurements performed by  $A_{0/1}$  and  $B_{2/3}$  respectively. This also implies in particular that when  $\mathcal{I}_\theta = 1$ ,  $\mathcal{A}$ 's inputs correspond (up to the same maps) to the Pauli measurements  $\bar{A}_0 = \sigma_z$  and  $\bar{A}_1 = \sigma_x$  while  $\mathcal{B}$ 's inputs correspond to  $\bar{B}_{2/3} = \cos(b_\theta)\sigma_z \pm \sin(b_\theta)\sigma_x$ .

Note that the post-measurement state corresponding to the outcome 1 can be characterized with the same measurement boxes  $A_{0/1}$  and  $B_{2/3}$ , as well as the same local extraction maps  $\Lambda_{\mathcal{A}}$  and  $\Lambda_{\mathcal{B}}^{o,\theta}$ , see Appendix. Moreover, by including the classical output of the measurement in a global state including a register, as mentioned before, the overall post-measurement state can be written in a compact form

$$\varrho = \sum_{\ell=0}^1 p_\ell \varrho_\ell \otimes |\ell\rangle\langle\ell|_R, \quad (14)$$

with  $\varrho_\ell = \frac{M_\ell[\rho]}{\text{Tr}(M_\ell[\rho])}$  being the post-measurement state associated to outcome  $\ell$ , and  $p_\ell$  the probability of this outcome. As the certificates for the two branches  $\varrho_0$  and  $\varrho_1$  are obtained with the same isometries, they can be combined into a single certificate for  $\varrho$ . In particular,

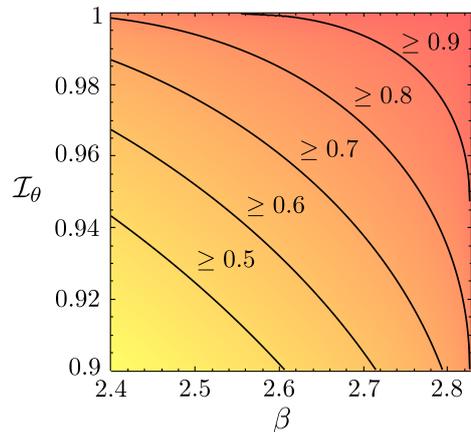


FIG. 5: Lower bounds on the fidelity  $\mathcal{F}(\mathcal{M}, \bar{\mathcal{M}})$  of an unknown measurement  $\mathcal{M}$  defined in Eq. (4) with respect to a reference measurement  $\bar{\mathcal{M}}$  defined in Eq. (3) with the Kraus operators given in Eqs. (1)-(2) for given CHSH violations  $\beta$  and violations  $\mathcal{I}_\theta$  of the Bell inequality (11); we assume that the second output state appears with probability  $p_0 = p_1 = \frac{1}{2}$  achieving the violation  $\mathcal{I}_\theta^0, \mathcal{I}_\theta^1 \geq \mathcal{I}_\theta$ . The plot is for  $\theta = (2\pi + 7)/22 \approx 0.6$ .

using the orthogonality of the register states, we have

$$\begin{aligned} & F((\Lambda_{\mathcal{A}} \otimes \Lambda_{\mathcal{B}}^{o,\theta})[\varrho], \sum_{\ell=0}^1 \frac{1}{2} |\phi_\theta^\ell\rangle\langle\phi_\theta^\ell| \otimes |\ell\rangle\langle\ell|_R) \\ & \geq F_\theta^o = \sum_{\ell} \sqrt{\frac{p_\ell}{2}} F_\theta^\ell. \end{aligned} \quad (15)$$

Taking into account the fact that the fidelity cannot decrease under completely-positive trace-preserving maps and using the triangular inequality, we can prove that the fidelity of the state before and after the measurement can be combined to bound the fidelity of the measurement itself, that is,

$$\mathcal{F}(\mathcal{M}, \bar{\mathcal{M}}) \geq \cos(\arccos(F^i) + \arccos(F_\theta^o)), \quad (16)$$

where we used the Proposition 5 of [25]. Whenever  $\beta = 2\sqrt{2}$  and  $\mathcal{I}_\theta = 1$  for both output states, this bound guarantees that  $\mathcal{F}(\mathcal{M}, \bar{\mathcal{M}}) = 1$ . In noisy scenarios, the fidelity that can be certified is shown in Fig. 5.

Note that in case where the fidelity of the state before the measurement cannot be assessed, the quality of the measurement cannot be certified. Indeed, if the source produces the state  $|\phi_\theta^0\rangle|0\rangle_{B'} + |\phi_\theta^1\rangle|1\rangle_{B'}$ , and the measurement simply reads out the auxiliary  $B'$  system, then all post-measurement statistics are reproduced. Hence, it is necessary to be able to estimate the quality of the pre-measurement state to give a certificate for the proper functioning of the measurement itself.

*Conclusion* — We provided a family of Bell inequalities that can be used to self-test non-maximally-entangled two-qubit states with unprecedented resistance to noise. These results allowed us to derive robust bounds that can be used in practice to certify the quality of measurements beyond the von Neumann model and rank-one POVMs to generalized measurements. The robustness of our certification techniques together with the flexibility of our recipe make us confident that self-testing of generalized measurements could soon be demonstrated experimentally. A natural extension of our result would be to self-test only one Kraus operator within a family. This could be obtained by generalizing the ‘heralded’ fidelity defined in [37] to the current setting.

*Acknowledgements* — This work was supported by the Swiss National Science Foundation (SNSF), through the Grant PP00P2-179109 and 200021-175527. We also acknowledge the Army Research Laboratory Center for Distributed Quantum Information via the project SciNet.

- 
- [1] B. Hensen *et al.* Nature **526**, 682 (2015).  
 [2] L.K. Shalm *et al.* Phys. Rev. Lett. **115**, 250402 (2015).  
 [3] M. Giustina *et al.* Phys. Rev. Lett. **115**, 250401 (2015).  
 [4] W. Rosenfeld *et al.* Phys. Rev. Lett. **119**, 010402 (2017).  
 [5] J.S. Bell, Physics **1**, 195 (1964).  
 [6] R. Colbeck, Quantum And Relativistic Protocols For Secure Multi-Party Computation, Ph.D. thesis, (2009).  
 [7] S. Pironio *et al.* Nature **464**, 1021 (2010).  
 [8] B.G. Christensen *et al.* Phys. Rev. Lett. **111**, 130406 (2013).  
 [9] Yang L. *et al.* Phys. Rev. Lett. **120**, 010503 (2018).  
 [10] Yang L. *et al.* Nature (2018).  
 [11] P. Bierhorst *et al.* Nature **556**, 223 (2018).  
 [12] L. Shen *et al.* arXiv:1805.02828 (2018).  
 [13] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).  
 [14] S. Popescu and D. Rohrlich, Phys. Lett. A **169**, 411 (1992).  
 [15] M. McKague, T. H. Yang, V. Scarani, J. Phys. A: Math. Theor. **45**, 455304 (2012).  
 [16] D. Mayers and A. Yao, Proceedings of the 39th IEEE Conference on Foundations of Computer Science, 1998, page 503; available as arXiv:9809039, see also Quant. Inf. Comput. **4**, 273 (2004).  
 [17] M. McKague, e-print arXiv:1309.5675  
 [18] A. Coladangelo, K.T. Goh, and V. Scarani, Nat. Comm. **8**, 15485 (2017).  
 [19] X. Wu, Y. Cai, T. H. Yang, H. Nguyen Le, J.-D. Bancal, V. Scarani, Phys. Rev. A **90**, 042339 (2014).  
 [20] J.-D. Bancal, M. Navascus, V. Scarani, T. Vrtesi, and T. H. Yang Phys. Rev. A **91**, 022115 (2015).  
 [21] S.-L. Chen, C. Budroni, Y.-C. Liang, and Y.-N. Chen, Phys. Rev. Lett. **116**, 240401 (2016).  
 [22] D. Cavalcanti and P. Skrzypczyk, Phys. Rev. A **93**, 052112 (2016).

- [23] J. Kaniewski, Phys. Rev. A **95**, 062323 (2017).  
 [24] J. Bowles, I. upi, D. Cavalcanti, A. Aco, Phys. Rev. A **98**, 042336 (2018).  
 [25] P. Sekatski, J.-D. Bancal, S. Wagner, and N. Sangouard, Phys. Rev. Lett. **121**, 180505 (2018).  
 [26] E.S. Gomez *et al.* Phys. Rev. Lett. **117**, 260401 (2016).  
 [27] M. Smania *et al.* e-print arXiv: 1811.12851 (2018).  
 [28] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Phys. Rev. Lett. **114**, 250401 (2015).  
 [29] F.J. Curchod *et al.* Phys. Rev. A **95**, 020102 (2017).  
 [30] F.J. Curchod *et al.* arXiv:1802.07962 (2018).  
 [31] B. Coyle, M.J. Hoban, E. Kashefi EPTCS 273, pp. 14-26 (2018).  
 [32] K. Banaszek, and I. Devetak, Phys. Rev. A **64**, 052307 (2001).  
 [33] J. Kaniewski, Phys. Rev. Lett. **117**, 070402 (2016).  
 [34] P. Sekatski *et al.*, in preparation.  
 [35] A. Aco, S. Massar, and S. Pironio, Phys. Rev. Lett. **108**, 100402 (2012).  
 [36] C. Bamps and S. Pironio, Phys. Rev. A **91**, 052111 (2015).  
 [37] J.-D. Bancal, N. Sangouard, and P. Sekatski, e-print arXiv:1807.04941 (2018).

## Appendix

In the supplemental material we will prove the inequality in Eq. (12) of the main text. To do so we have to lower bound the fidelity of an unknown state  $\rho$  with respect to

$$|\phi_\theta^0\rangle = \cos(\theta)|00\rangle + \sin(\theta)|11\rangle \quad (17)$$

as a function of the expected value of the following Bell expression evaluated on the state  $\rho$

$$\mathcal{I}_\theta = \left\langle \frac{1}{4} \left[ \frac{A_0(B_0 - B_1)}{\sin(b_\theta)} + \frac{\sin(2\theta)}{\cos(b_\theta)} A_1(B_0 + B_1) + \cos(2\theta) \left( A_0 + \frac{B_0 - B_1}{2\sin(b_\theta)} \right) \right] \right\rangle, \quad (18)$$

where  $b_\theta = \arctan \sqrt{\frac{1 + \frac{1}{2}c_{2\theta}^2}{s_{2\theta}^2}}$ . Here and in the rest of the appendix we use the short notation  $c_\theta = \cos(\theta)$  and  $s_\theta = \sin(\theta)$ . The Bell expression Eq. (18) has a quantum bound of 1, achieved by measuring precisely the two-qubit state  $|\phi_\theta^0\rangle$  with the observables

$$A_0 = \sigma_z, \quad B_0 = \cos(b_\theta)\sigma_x + \sin(b_\theta)\sigma_z, \quad (19)$$

$$A_1 = \sigma_x, \quad B_1 = \cos(b_\theta)\sigma_x - \sin(b_\theta)\sigma_z. \quad (20)$$

We find that the local bound of Eq. (18) achieved by the deterministic local strategy  $\{A_0 = A_1 = B_0 = 1, B_1 = -1\}$  is given by

$$\mathcal{I}_L = \frac{1}{4} \left[ c_{2\theta} + (2 + c_{2\theta}) \sqrt{\frac{7 - c_{4\theta}}{5 + c_{4\theta}}} \right]. \quad (21)$$

This value is to be compared with the local bound of the well-known (normalized) tilted-CHSH

$$\mathcal{I}_\theta^{(T)} = \left\langle \frac{\alpha_\theta A_0 + A_0(B_0 + B_1) + A_1(B_0 - B_1)}{\sqrt{8 + 2\alpha_\theta}} \right\rangle \quad (22)$$

with  $\alpha_\theta = \frac{2}{\sqrt{1+2\tan^2(2\theta)}}$ , also known to attain the quantum bound of 1 for the partially entangled state of Eq. (17) and well-chosen measurement settings [35, 36]. We find that the local bound of Eq. (21) is always higher than the local bound of the tilted-CHSH  $\mathcal{I}_L \geq \mathcal{I}_L^{(T)} = \frac{2+\alpha_\theta}{\sqrt{8+2\alpha_\theta^2}}$ , meaning that the violation of the tilted-CHSH inequality is more robust to white noise than the violation of our new Bell inequality. Nevertheless, our later studies will show that the new Bell operator allows for a more noise-tolerant state certification. The reason for this counter-intuitive result is the following: comparing the observed violation to the local bound only provides information on the distance between the target state (in our case  $|\phi_\theta^0\rangle$ ) and deterministic strategies, given by parallel measurement setting  $A_0 = \pm A_1$  with  $B_0 = \pm B_1$  and product states  $|\psi_A\rangle|\psi_B\rangle$  with  $A_0|\psi_A\rangle = \pm|\psi_A\rangle$  and  $B_0|\psi_B\rangle = \pm|\psi_B\rangle$ . In self-testing, on the other hand, we need to bound the distance of the target state to arbitrary states for arbitrary measurement settings. It is then crucial that the violation worsens drastically when departing from the perfect settings.

#### Fidelity Bounds

To derive lower bounds on the state fidelity from a Bell violation  $\mathcal{I}_\theta$  we use the tools presented in Ref. [25]. There, it is shown that such a global lower bound can be obtained by solving a two-qubit problem. More precisely, the state fidelity can be bounded by minimizing the quantity

$$O((\Lambda_{\mathcal{A}} \otimes \Lambda_{\mathcal{B}}^{o,\theta})[\rho], |\phi_\theta\rangle\langle\phi_\theta|) \quad (23)$$

with  $O(\rho, \sigma) = \text{Tr}(\sigma\rho)$ , over all possible two-qubit states  $\rho$  and all possible qubit observables  $A_0, A_1, B_0, B_1$  (with eigenvalues  $\pm 1$ ) that are compatible with the value  $\mathcal{I}_\theta$  of the Bell operator. Here,

$$\Lambda_{\mathcal{A}}, \Lambda_{\mathcal{B}}^{o,\theta} : L(\mathbb{C}^2) \rightarrow L(\mathbb{C}^2) \quad (24)$$

are the local extraction channels that can depend on the local observable  $A_0$  with  $A_1$  for  $\Lambda_{\mathcal{A}}$  and  $B_0$  with  $B_1$  for  $\Lambda_{\mathcal{B}}^{o,\theta}$ . The first step, therefore, is to fix these extraction channels.

Before we do so, let us fix some notation for the local observables. Any qubit observable with eigenvalues  $+1$  and  $-1$  can be written as  $\mathbf{n} \cdot \boldsymbol{\sigma}$  with  $|\mathbf{n}| = 1$ . Furthermore, without loss of generality, we can set the local bases such that

$$\begin{aligned} A_0(a) &= \cos(a)H + \sin(a)V \\ A_1(a) &= \cos(a)H - \sin(a)V \end{aligned} \quad (25)$$

$$\begin{aligned} B_0(b) &= \cos(b)\sigma_x + \sin(b)\sigma_z \\ B_1(b) &= \cos(b)\sigma_x - \sin(b)\sigma_z \end{aligned} \quad (26)$$

where  $H = \frac{1}{\sqrt{2}}(\sigma_z + \sigma_x)$ ,  $V = \frac{1}{\sqrt{2}}(\sigma_z - \sigma_x)$ . Hence, in the minimization the pairs of observable  $A_0, A_1$  and  $B_1, B_2$  as well as the extraction channels  $\Lambda_{\mathcal{A}}$  and  $\Lambda_{\mathcal{B}}^{o,\theta}$  only depend on a single parameter  $a$  and  $b$  respectively.

#### Extraction Channels

The extraction channels we will use are adapted versions of the dephasing channels of Ref. [33]. On Alice's side, the observables are maximally dephased if they are parallel or anti-parallel, and unchanged if they are orthogonal. More precisely, for  $a$  being half the angle between Alice's observables in Eq. (25) the dephasing acts according to

$$\Lambda_a[\rho] := \frac{1+g(a)}{2}\rho + \frac{1-g(a)}{2}\Gamma_a\rho\Gamma_a, \quad (27)$$

where  $g(a) = (1 + \sqrt{2})(\cos(a) + \sin(a) - 1)$ , and  $\Gamma_a = H$  if  $a \in [0, \frac{\pi}{4}]$  and  $\Gamma_a = V$  if  $a \in ]\frac{\pi}{4}, \frac{\pi}{2}]$ . Here and from now on,  $\Lambda_a$  is the short notation of  $\Lambda_{\mathcal{A}}(a)$ .

On Bob's side, the observables are also maximally dephased if they are parallel or anti-parallel, but here they are unchanged if half the angle between them equals

$b_\theta$ , where  $b_\theta = \arctan\sqrt{\frac{1+\frac{1}{2}c_\theta^2}{s_{2\theta}^2}}$  for the new inequality and  $b_\theta = \arctan(\sin(2\theta))$  for the tilted-CHSH one. For  $b$  denoting half the angle between Bob's observables in Eq. (26) the dephasing channel on Bob's side is

$$\Lambda_b[\rho] := \frac{1+g(t_\theta(b))}{2}\rho + \frac{1-g(t_\theta(b))}{2}\Omega_b\rho\Omega_b, \quad (28)$$

where

$$t_\theta(b) = \gamma_\theta^{-1} \ln\left(\frac{b - \delta_\theta}{\delta_\theta}\right), \quad (29)$$

$$\gamma_\theta = \frac{4}{\pi} \ln\left(\frac{\frac{\pi}{2} - b_\theta}{b_\theta}\right), \quad (30)$$

$$\delta_\theta = \frac{b_\theta^2}{\pi^2 - 2b_\theta}. \quad (31)$$

For the observables of Bob, the dephasing happens in the direction  $\Omega_b = \sigma_x$  if  $b \in [0, b_\theta]$  and  $\Omega_b = \sigma_z$  if  $b \in ]b_\theta, \frac{\pi}{2}]$ . Again, we introduced the simplified notation  $\Lambda_b = \Lambda_{\mathcal{B}}^{o,\theta}(b)$ .

Using these extraction channels and applying the numerical method presented Ref. [25], we find the values of  $I_\theta^*$  such that the convex bound

$$O((\Lambda_{\mathcal{A}} \otimes \Lambda_{\mathcal{B}}^{o,\theta})[\rho], |\phi_\theta^0\rangle\langle\phi_\theta^0|) \geq (1 - c_\theta^2) \frac{\mathcal{I}_\theta - \mathcal{I}_\theta^*}{1 - \mathcal{I}_\theta^*} + c_\theta^2 \quad (32)$$

holds for our new inequality, as well as for the tilted-CHSH one (using then  $\mathcal{I}_\theta^{(T)}$  instead of  $\mathcal{I}_\theta$ ). Here  $\mathcal{I}_\theta$  is

the observed Bell violation and  $\mathcal{I}_\theta^*$  is the non-trivial cut-off corresponding to the violation for which the fidelity matches the square of the largest Schmidt coefficient of  $|\phi_\theta^0\rangle$ . This means that  $\mathcal{I}_\theta^*$  is the relevant quantity for comparing the self-testing performance of Bell operators in device-independent tasks. Figure 4 in the main text depicts  $\mathcal{I}_\theta^*$  for both the tilted-CHSH and the new inequality.

*Extension to the other state*

We have just shown how the observed violation of a Bell inequalities  $\mathcal{I}_\theta$  gives a bound on the overlap of the state  $\rho$  with the target state

$$|\phi_\theta^0\rangle = c_\theta|00\rangle + s_\theta|11\rangle \quad (33)$$

upon applying the extraction maps  $\Lambda_a$  and  $\Lambda_b$ . We will now show the violation of another inequality  $\mathcal{I}'_\theta$ , related to  $\mathcal{I}_\theta$  by a mere relabeling of some of the inputs and outputs, bounds the overlap of a state  $\rho'$  with the other target state

$$|\phi_\theta^1\rangle = c_\theta|11\rangle + s_\theta|00\rangle. \quad (34)$$

The proof follows the line of [37]. Before we start, note that the two target states are related via  $|\phi_\theta^0\rangle \xrightarrow{\sigma_x \otimes \sigma_x} |\phi_\theta^1\rangle$ . In the ideal case this transformation corresponds to a permutation of the outputs of  $A_1 \leftrightarrow -A_1$  and the exchange of  $B_0$  and  $B_1$ . We will now show that this observation also holds in the non-ideal case.

Let us first have a closer look at Eq. (32). As it holds for any state  $\rho$ , it can be expressed as the expectation value of the operator

$$\Lambda_a \otimes \Lambda_b [|\phi_\theta^0\rangle\langle\phi_\theta^0|] - s\mathcal{B}(a, b) - \mu \geq 0, \quad (35)$$

where  $\mathcal{B}(a, b)$  is the Bell operator obtained by choosing the settings (25), (26) in the Bell expression (18). Here we used the fact that the maps  $\Lambda_{a(b)}$  are self-adjoint and  $s = \frac{1-c_\theta^2}{1-\mathcal{I}_\theta^*}$  and  $\mu = \frac{c_\theta^2 - \mathcal{I}_\theta^*}{1-\mathcal{I}_\theta^*}$ . This operator inequality holds for all measurement angles  $a$  and  $b$ .

Now consider a new Bell expression corresponding to an operator  $\mathcal{B}'$ , obtained by exchanging the outputs of  $A_1$  and exchanging the role of the measurements  $B_0$  and  $B_1$  in the previous expression. The operator  $\mathcal{B}'(a, b)$  can be obtained by applying the rotation  $R := R_{\hat{x}}(\pi) = e^{i\frac{\pi}{2}\sigma_x}$ :

$$\mathcal{B}'(a, b) = (R \otimes R)\mathcal{B}'(\frac{\pi}{2} - a, b)(R^\dagger \otimes R^\dagger) \quad (36)$$

as illustrated in Figure 6.

Since Eq. (35) holds for all angles, it also holds for  $a \rightarrow \frac{\pi}{2} - a$ . Then we act with  $(R \otimes R)$  and  $(R^\dagger \otimes R^\dagger)$  from the left and right respectively. We arrive at

$$(R \otimes R)(\Lambda_{\frac{\pi}{2}-a} \otimes \Lambda_b)[|\phi_\theta^0\rangle\langle\phi_\theta^0|](R^\dagger \otimes R^\dagger) - s\mathcal{B}'(a, b) - \mu \geq 0. \quad (37)$$

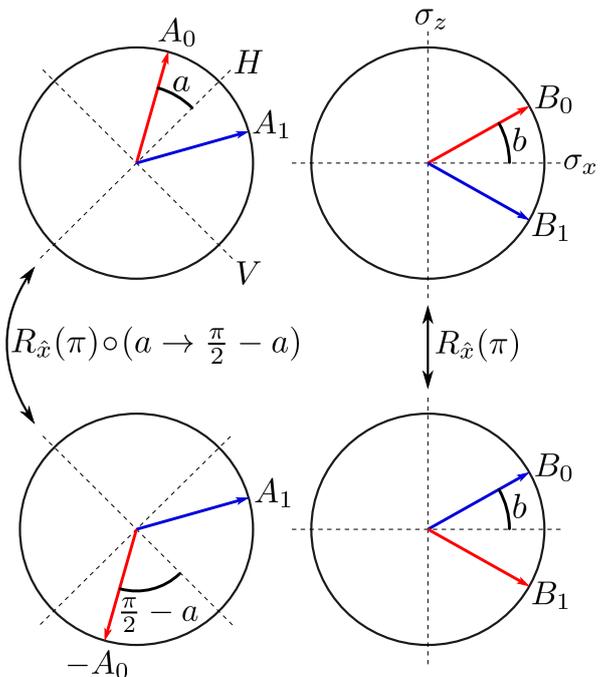


FIG. 6: The first row shows the settings present in the experiment, the left column belongs to Alice, the right one to Bob. In case the outcome of the generalized measurement is 0, they just collect the data. If the outcome is 1, Bob reinterprets his inputs, i.e. if he chose  $B_0$  he saves the result as the outcome of  $B_1$  and vice versa. Alice on the other hand collects her data applying  $A_0 \rightarrow -A_0$ . This transformation is sketched in the second row. The effect of this post-processing on the settings correspond to Alice applying a shift to her measurement angle followed by a rotation around  $\hat{x}$  by  $\pi$ , i.e.  $R_{\hat{x}}(\pi) \circ (a \rightarrow \frac{\pi}{2} - a)$ , where  $R_{\hat{x}}(\pi) = e^{i\frac{\pi}{2}\sigma_x}$ . The observables of Bob are simply rotated by  $R_{\hat{x}}(\pi)$ .

The rotation on Bob's side commutes with the extraction channel  $\Lambda_b$ . This is easily verified by reminding ourselves that  $\Omega_b \in \{\sigma_x, \sigma_z\}$  and therefore  $R\Omega_b\rho\Omega_b R^\dagger = \Omega_b R\rho R^\dagger \Omega_b$ .

On Alice's side, we realize that  $RHR^\dagger = V$ . Therefore  $R\Gamma_{\frac{\pi}{2}-a}R^\dagger = \Gamma_a$ . By also recalling that  $g(\frac{\pi}{2} - a) = g(a)$ , we find that  $R\Lambda_{\frac{\pi}{2}-a}[\rho]R^\dagger = \Lambda_a[R\rho R^\dagger]$ .

Combining everything and using  $R \otimes R|\phi_\theta^0\rangle = |\phi_\theta^1\rangle$ , Eq. (37) is equivalent to

$$\Lambda_a \otimes \Lambda_b [|\phi_\theta^1\rangle\langle\phi_\theta^1|] - s\mathcal{B}'(a, b) - \mu \geq 0, \quad (38)$$

which implies that with the same extraction channels and observables, and a new Bell test  $\mathcal{I}'_\theta$ , one can device-independently self-test the second output state  $|\phi_\theta^1\rangle$ .



## CONCLUSIONS AND PERSPECTIVES

### 6.1 Conclusions

---

In this thesis, we discussed many aspects of device-independent certification. We studied the certification of multipartite systems, the self-testing of quantum channels as well as the certification of quantum measurements.

**Bell-Correlations in Large Systems** The first topic of interest was related to the characterisation of complex systems. We derived a new class of Bell inequalities that applies to systems of arbitrary size and involves an arbitrary number of dichotomic measurements per party. The crucial feature of these inequalities is that they are symmetric under exchange of parties and only consist of one- and two-body correlators. These inequalities are fully device-independent in the sense that they imply possible Bell tests to be conducted on many-body systems. In our work, however, we were interested in detecting non-classical features in such scenarios. Therefore we devised Bell correlation witnesses assuming collective spin measurements are performed on an ensemble of spin- $\frac{1}{2}$  particles. We showed that such a set-up allows one to record violations of said Bell witnesses, highlighting the presence of Bell-correlated states. What is more, we used experimental data of measurements performed on a Bose-Einstein condensate to report on an actual experimental violation that is even more distinct than the violation achieved with a previously-known witness. We analysed our results by studying the effect finite statistics have on the significance of the results. We found that in order to exclude with a high probability that local states were responsible for the violation, a number of experimental runs that scales linearly with the number of spins is both required and sufficient.

**Variational Method for Tailoring Bell Inequalities** In the first chapter related to self-testing, we introduced a new method to tailor Bell inequalities to states. These inequalities are modelled such that they permit one to achieve large self-testing fidelities: The approach guarantees that if the observed violation deviates little from the quantum bound then also the tested state is allowed to deviate only little from the target state. We explained our perturbative approach by constructing a new class of Bell inequalities suitable for the self-testing of partially-entangled two-qubit states. In comparison to the tilted-CHSH one, this new inequality allows for higher self-testing fidelities. The approach is easily generalized to scenarios involving arbitrary numbers of parties and is not restricted to the qubit case. The downside of the approach is that it does not always necessarily provide a conclusive result. Also, since the number of parameters is larger than the number of constraints, using the approach involves some intuition about how to choose some of the parameters. Another possible source for issues is the parametrisation of the observables. One can choose the parametrisation freely but depending on one's choice, further steps such as the diagonalisation of the second-order perturbation matrix (Eq. (3.8)) might be complicated. To give an example, in section 3.2 one could choose a parametrisation that leaves all the observables independent from one another. While this might simplify the notation, it complicates all the succeeding steps making an analytic treatment of the problem infeasible. The parametrisation we chose is more complicated to read but easier to work with. Despite the mentioned (minor) issues, we believe our technique to be very useful for device-independent certification tasks; We used it in chapters 4 and 5 to derive new Bell inequalities for two-qubit and four-qubit target states.

**Self-Testing of Quantum Channels** The second chapter related to self-testing was devoted to the device-independent certification of quantum channels. The overarching objective of our work was providing the tools for the characterisation of building blocks of quantum computers. More specifically, the goal was to find quantitative lower-bounds on the fidelity of an experimental channel with respect to a target channel. We showed that this aim can be achieved by assessing state fidelities, more precisely by determining the fidelity of the state on which the channel acts as well as the fidelity of the output state. We illustrated the power of our technique by lower-bounding the fidelities of a quantum memory and a controlled-unitary gate. In both cases, the bounds are very robust to noise highlighting the applicability of the approach for practical tasks where noise is always present. As a consequence, our work provides the necessary tools to device-independently certify most of the building blocks of a circuit-based quantum computer.

**Self-Testing of Quantum Measurements** After studying the device-independent certification of states and channels, the last work presented in this thesis deals with the certification of measurements. More precisely, in chapter 5 we showed how to self-test quantum measurements, i.e. non-projective measurements. Since these measurements can be interpreted as quantum channels, the work was closely related to the self-testing of channels. A crucial difference is that

the individual post-measurement states (corresponding to different outcomes) have to be certified individually but with common extraction channels. In this way one is able to lower-bound Kraus operator fidelities which one then combines to get a robust lower-bound on the corresponding measurement fidelity. Hence our work established a method for the device-independent characterisation of quantum measurements.

Summing up the previous conclusions, in this thesis we provided new methods for the device-independent certification of entangled quantum states, quantum channels and quantum measurements.

## 6.2 Perspectives

---

Much effort has been spent on certifying quantum states, channels and measurements. While this thesis provides many interesting and relevant results along this line, there are still aspects that can be improved and subjects appealing for further studies.

**Bell-Correlations in Large Systems** The inequalities and witnesses we derived in chapter 2 consist of one- and two-body correlators. This is convenient when performing experiments with indistinguishable particles, just as in the case of a Bose-Einstein condensate. In general, however, it may be advantageous to include higher-order correlators as well. We derived classes of Bell inequalities that, in addition to the correlators already in use, consider higher-order correlators with the crucial feature that the highest correlation order is even. This step has consequences for experimental realisations: In case of the BEC, one must also measure higher-order moments, e.g. the scaled fourth moment  $\zeta_a^4$ , which implies that more measurements are needed. In terms of the finite statistics issue, this means that one has to optimise over more parameters. Since each parameter adds uncertainties to the analysis, this suggests that adding higher-order correlators requires enhancing the number of experimental runs. On the other hand, these correlators also contribute information on the probability distribution of the system which might be beneficial for the statistical analysis. From our intuition and experience, however, we expect adding correlators not to be advantageous for the finite statistics analysis.

An aspect of our work that requires optimisation (or at least a check for optimality) are the witnesses. In their derivation we assumed that the measurements are all situated in one plane and that the angle between successive measurement directions is constant when fixing the number of settings. The first assumption regarding the measurement plane is easily motivated: The squeezed BEC we considered is prepared such that the second moment in one direction is small while the first moment in an orthogonal direction is large. But then the second moment of the third orthogonal direction is very large while the first moment in the same direction is small. Hence it is advisable to construct a witness that does not include all three directions.

The second assumption on the angle distribution can not be easily motivated. Therefore it would be interesting to study the effect other distributions have on the violation spectrum.

Since certifications based on witnesses are device-dependent, a potential task following up on our work could be the realisation of Bell tests based on our inequalities. For very large systems, e.g. BECs of more than 500 atoms, this is a very demanding task since a Bell test requires the individual addressing of the atoms and in order for it to be a true Bell test, the spins must also be spatially separated. We thus propose to start with a comparably small number of spins; already ten would be a great success.

With respect to the finite statistics analysis, we emphasise that our work is not limited to the Bell witnesses presented, but can also be applied to other inequalities and witnesses with second-order correlators - an important example being entanglement witnesses. For Bell inequalities such as the CHSH inequality one could even improve the lower bounds by taking into account the discreteness of the outcomes: In the scenarios considered by us, the first and second moments can take values from a continuous set of outcomes while for the CHSH inequality, the value of a correlator can only take discrete values. This puts restrictions on the probability distributions. Another possible improvement to our bounds could be achieved by taking into account the full statistics, i.e. the outcomes of the individual experimental runs. These outcomes also provide information about the probability distribution.

**Variational Method for Tailoring Bell Inequalities** The method described in chapter 3 is very convenient for tailoring Bell inequalities to quantum states. However, we see a problem regarding the applicability of the approach to multipartite states. The number of parameters in the optimisation grows faster with the number of parties than the number of constraints. In a scenario of  $n$  qubits, for example, the number of variables is  $3n - 1$  ( $n - 1$  for the weightings and  $2n$  for the observables) while the number of non-trivial constraints is  $n$ . The second-order terms in the perturbation method provide optimisations over all parameters, so in principle the problem can still be solved. However, the more variables can be eliminated the easier it is to diagonalise the matrix in the second-order perturbation step. Hence with a growing number of parties, this diagonalisation becomes more difficult.

**Self-Testing of Quantum Channels and Quantum Measurements** Our technique for channel certification is based on the self-testing of states and consists of two almost independent steps - the self-testing of the input and output state. The only link between the two steps is that one party chooses from a fixed set of measurements, irrespective of whether the channel was applied or not. What is not taken into account so far is that the first state certification already provides useful information about correlators that could be used in the second step. By considering this information, one could potentially increase the output fidelity resulting in the improvement of the channel fidelity as well. Our main reason for suggesting studies in this direction, however, concerns the self-testing of states. If one uses the correlation information of the first step, it might be possible to device-independently certify a product output state.

This would be a great achievement since it would mean that self-testing is not restricted to entangled states, but can also be achieved for states not violating a Bell inequality.

Another possible direction is the study of certification in case of composite systems. So far, our method only applies to the individual blocks of a quantum computer. Hopefully we are able to extend our work to self-test systems composed of many building blocks.

**Self-Testing in General** In all our studies on self-testing we always relied on Jordan's lemma. This lemma allowed us to reduce the complexity of the problem significantly by boiling down the scenario to the qubit case. The benefits of the lemma come at a high cost: each party is only allowed to choose from two dichotomic observables. This is of course a major limitation since it prevents us to use our techniques in scenarios involving many settings and/or many outcomes - for example the Bell inequalities of chapter 2. Hence a relevant future project is the study of self-testing in scenarios not allowing the application of Jordan's lemma.

Another point that is worth optimising is related to the extraction channels. In principle in self-testing, we not only have to minimise state fidelities over all possible quantum states, but we also have the freedom to maximise the fidelity over all extraction channels. So far, this last point has not been addressed thoroughly. Instead, we simply fixed the extraction channels and interpreted the observed fidelities as lower-bounds. Parametrising extraction channels and studying their role in the self-testing of simple systems is a promising project on the path to finding tight self-testing bounds. Since the extraction channels of Kaniewski already provide a tight bound for the Mermin inequality [23], it might be that we are already aware of the optimal channels. In any case, the study of optimality is highly interesting.



REFERENCES

- [1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, “Quantum entanglement,” *Reviews of modern physics*, vol. 81, no. 2, p. 865, 2009.
- [2] P. W. Shor, “Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer,” *SIAM review*, vol. 41, no. 2, pp. 303–332, 1999.
- [3] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” *In Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India*, vol. 175, p. 8, 1984.
- [4] A. K. Ekert, “Quantum cryptography based on Bells theorem,” *Physical review letters*, vol. 67, no. 6, p. 661, 1991.
- [5] C. H. Bennett, G. Brassard, and N. D. Mermin, “Quantum cryptography without Bells theorem,” *Physical Review Letters*, vol. 68, no. 5, p. 557, 1992.
- [6] <https://www.idquantique.com/random-number-generation/products/quantis-random-number-generator/>.
- [7] <https://www.picoquant.com/products/category/quantum-random-number-generator/pqrng-150-quantum-random-number-generator>.
- [8] D. Rosset, R. Ferretti-Schöbitz, J.-D. Bancal, N. Gisin, and Y.-C. Liang, “Imperfect measurement settings: Implications for quantum state tomography and entanglement witnesses,” *Physical Review A*, vol. 86, no. 6, p. 062325, 2012.
- [9] A. Acin, N. Gisin, and L. Masanes, “From Bells theorem to secure quantum key distribution,” *Physical review letters*, vol. 97, no. 12, p. 120405, 2006.

- 
- [10] S. Pironio, A. Acin, N. Brunner, N. Gisin, S. Massar, and V. Scarani, “Device-independent quantum key distribution secure against collective attacks,” *New Journal of Physics*, vol. 11, no. 4, p. 045021, 2009.
- [11] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?” *Physical review*, vol. 47, no. 10, p. 777, 1935.
- [12] J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics Physique Fizika*, vol. 1, no. 3, p. 195, 1964.
- [13] S. J. Freedman and J. F. Clauser, “Experimental test of local hidden-variable theories,” *Physical Review Letters*, vol. 28, no. 14, p. 938, 1972.
- [14] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenber, R. F. Vermeulen, R. N. Schouten, C. Abellán *et al.*, “Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres,” *Nature*, vol. 526, no. 7575, p. 682, 2015.
- [15] M. Giustina, M. A. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-Å. Larsson, C. Abellán *et al.*, “Significant-loophole-free test of Bells theorem with entangled photons,” *Physical review letters*, vol. 115, no. 25, p. 250401, 2015.
- [16] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman *et al.*, “Strong loophole-free test of local realism,” *Physical review letters*, vol. 115, no. 25, p. 250402, 2015.
- [17] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, “Proposed experiment to test local hidden-variable theories,” *Physical review letters*, vol. 23, no. 15, p. 880, 1969.
- [18] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, “Bell nonlocality,” *Reviews of Modern Physics*, vol. 86, no. 2, p. 419, 2014.
- [19] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” *Physical Review A*, vol. 40, no. 8, p. 4277, 1989.
- [20] B. S. Cirel’son, “Quantum generalizations of Bell’s inequality,” *Letters in Mathematical Physics*, vol. 4, no. 2, pp. 93–100, 1980.
- [21] D. Mayers and A. Yao, “Self testing quantum apparatus,” *arXiv preprint quant-ph/0307205*, 2003.
- [22] V. Scarani, “The device-independent outlook on quantum physics,” *Acta Physica Slovaca*, vol. 62, no. 4, pp. 347–409, 2012.

- 
- [23] J. Kaniewski, “Analytic and nearly optimal self-testing bounds for the Clauser-Horne-Shimony-Holt and Mermin inequalities,” *Physical Review Letters*, vol. 117, no. 7, p. 070402, 2016.
- [24] N. D. Mermin, “Extreme quantum entanglement in a superposition of macroscopically distinct states,” *Physical Review Letters*, vol. 65, no. 15, p. 1838, 1990.
- [25] G. Svetlichny, “Distinguishing three-body from two-body nonseparability by a Bell-type inequality,” *Physical Review D*, vol. 35, no. 10, p. 3066, 1987.
- [26] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, and A. Acín, “Detecting nonlocality in many-body quantum states,” *Science*, vol. 344, no. 6189, pp. 1256–1258, 2014.
- [27] P. Sekatski, S. Wagner, J.-D. Bancal, and N. Sangouard, *in preparation*.
- [28] A. Acín, S. Massar, and S. Pironio, “Randomness versus nonlocality and entanglement,” *Physical Review Letters*, vol. 108, no. 10, p. 100402, 2012.
- [29] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, “Multiple observers can share the nonlocality of half of an entangled pair by using optimal weak measurements,” *Physical Review Letters*, vol. 114, no. 25, p. 250401, 2015.



## 8.1 Dephasing Channels for "Non-Orthogonal Operators"

---

In this appendix we present how we adapted previously-known extraction channels in order for them to be suitable for "non-orthogonal" operators such as the ones present when self-testing partially-entangled states. Let us first recap the dephasing channels for the CHSH case as used by Kaniewski [23]:

**Standard Dephasing** – There the observables of Bob (Alice accordingly) are maximally dephased if they are parallel or anti-parallel, and unchanged if they are orthogonal. More precisely, if  $b$  denotes half the angle between Bob's observables ( $B_0(b) = \cos(b)\tilde{\sigma}_x + \sin(b)\tilde{\sigma}_z$ ,  $B_1(b) = \cos(b)\tilde{\sigma}_x - \sin(b)\tilde{\sigma}_z$ ), then the dephasing acts according to

$$\Lambda_b[\rho] := \frac{1 + g(b)}{2}\rho + \frac{1 - g(b)}{2}\Gamma_b\rho\Gamma_b, \quad (8.1)$$

where  $g(b) = (1 + \sqrt{2})(\cos(b) + \sin(b) - 1)$ , and  $\Gamma_b = \tilde{\sigma}_x$  if  $b \in [0, \frac{\pi}{4}]$  and  $\Gamma_b = \tilde{\sigma}_z$  if  $b \in [\frac{\pi}{4}, \frac{\pi}{2}]$ . The dephasing function  $g(x)$  is depicted in Fig. 8.1 (right, blue dash-dotted line).

**Change of Dephasing Channels** – In the previous situation, we left the observables unchanged in case they were orthogonal. In what follows, we want the observables to remain unchanged if they differ by an angle  $2\mu < 2 \cdot \frac{\pi}{4}$ .

The property we want to keep from the standard dephasing is the maximal dephasing of parallel and anti-parallel observables, i.e.  $g(0) = g(\frac{\pi}{2}) = 0$ . To achieve this we search for a monotonous transformation  $t_\mu$  which fulfills the following conditions:

$$(i) \ t_\mu(0) = 0, \quad (ii) \ t_\mu\left(\frac{\pi}{2}\right) = \frac{\pi}{2}, \quad (iii) \ t_\mu(\mu) = \frac{\pi}{4}. \quad (8.2)$$

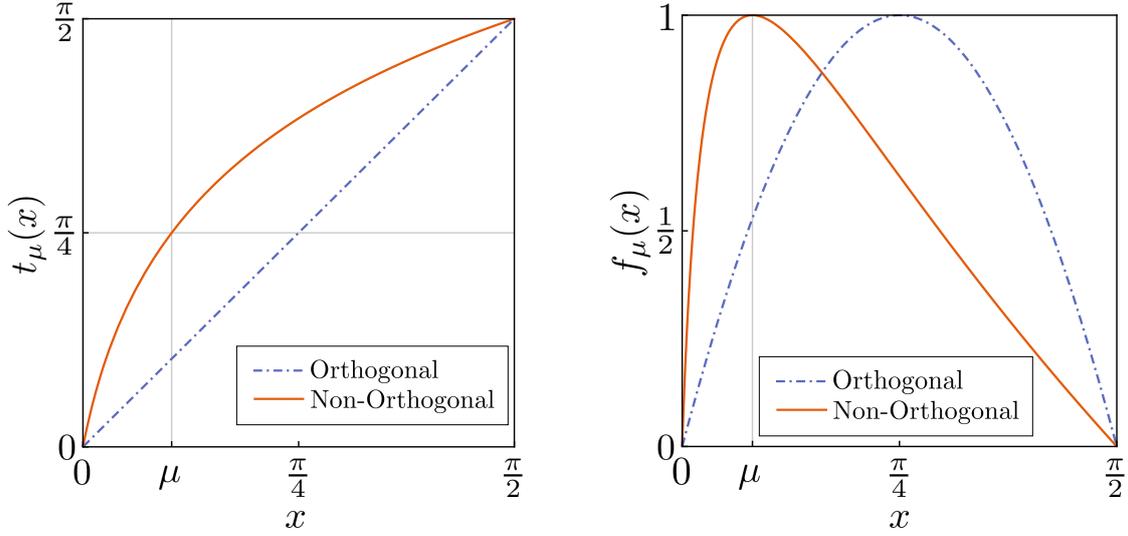


Figure 8.1: Plots of the transformation function (left) and the dephasing function (right). The blue dash-dotted lines are for the standard dephasing, which leaves orthogonal observables unchanged. The orange lines characterize the dephasing for the case where the intended observables differ by an angle  $2\mu$ .

With the help of this transformation we then define the dephasing to be  $f_\mu(b) = g(t_\mu(b))$ . For the transformation, we make the ansatz

$$t_\mu(x) = \gamma_\mu^{-1} \ln \left( \frac{x - \delta_\mu}{\delta_\mu} \right). \quad (8.3)$$

Forcing the conditions Eq. (8.2), we find the parameters

$$\gamma_\mu = \frac{4}{\pi} \ln \left( \frac{\frac{\pi}{2} - \mu}{\mu} \right), \quad (8.4)$$

$$\delta_\mu = \frac{\mu^2}{\pi^2 - 2\mu}. \quad (8.5)$$

The transformation function  $t_\mu(x)$  as well as the dephasing function  $f_\mu(x)$  are sketched in Fig. 8.1.