Bounding sets of sequential quantum correlations and device-independent randomness certification

Joseph Bowles[1], Flavio Baccari[1,2], and Alexia Salavrakos[1]

[1]ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels (Barcelona), Spain

[2]Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, 85748 Garching, Germany

An important problem in quantum information theory is that of bounding sets of correlations that arise from making local measurements on entangled states of arbitrary dimension. Currently, the best-known method to tackle this problem is the NPA hierarchy; an infinite sequence of semidefinite programs that provides increasingly tighter outer approximations to the desired set of correlations. In this work we consider a more general scenario in which one performs sequences of local measurements on an entangled state of arbitrary dimension. We show that a simple adaptation of the original NPA hierarchy provides an analogous hierarchy for this scenario, with comparable resource requirements and convergence properties. We then use the method to tackle some problems in device-independent quantum information. First, we show how one can robustly certify over 2.3 bits of device-independent local randomness from a two-quibt state using a sequence of measurements, going beyond the theoretical maximum of two bits that can be achieved with non-sequential measurements. Finally, we show tight upper bounds to two previously defined tasks in sequential Bell test scenarios.

1 Introduction

The correlations between outcomes of local measurements made on entangled quantum systems are known to exhibit a rich structure. Firstly, they are generally stronger than correlations attainable via classical resources, a phenomenon known as Bell nonlocality [1, 2]. Secondly, sets of quantum correlations are known to contain both smooth and flat boundaries [3, 4], and there exist correlations whose realisation requires infinite-dimensional entangled states [5], even in scenarios involving small and finite alphabet sizes.

All of this makes the problem of characterising, and optimising over, the set of quantum correlations a highly non-trivial and potentially undecidable problem. At the same time, being able to perform an optimisation over the entire set of quantum correlations is crucial for many areas of quantum information theory, principally in the field of device-independent quantum information, where quantum systems are treated as black-boxes and one makes no assumption on the physical dimension of the underlying state. A major breakthrough in this direction came with the discovery of the NPA-hierarchy [6, 7], which provides a characterisation of the set of quantum correlations via a sequence of increasing tighter outer approximations, each expressed in terms of a semi-definite program (SDP). Consequently, the NPA hierarchy has become a vital tool for the study of device-independent protocols in the standard scenario in which they are usually considered, commonly referred to as a Bell test. There, a bipartite state is shared between two parties, each of which makes a number of local measurements in order to generate the data that is used in the protocol.

In recent years a number of works have also considered *sequential Bell test* scenarios, in which the parties make a sequence of local measurements that obey a time-ordered causal structure [8, 9, 10, 11, 12] (see figure 1). Such scenarios have been shown to be relevant for Bell nonlocality via for example the phenomenon of hidden nonlocality [10, 11, 12]. As a result, sequential measurement scenarios are known to provide an advantage in device-independent randomness

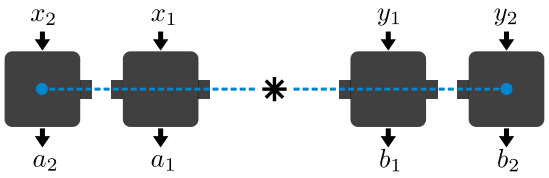

Figure 1: A sequential Bell scenario in which both parties perform a sequence of two measurements on their halves of a bipartite quantum state. In this work we develop methods to characterise the sets of probability distributions that can arise in such scenarios involving arbitrary numbers of parties in each sequence.

certification [13] and, we expect, in many other device-independent protocols. Further to this, sequential measurement scenarios also play a role in demonstrations of contextuality [14] and Leggett-Garg type tests of nonclassicality [15].

It is thus very desirable to develop methods to characterise the correlations arising in sequential Bell test scenarios. In this work we show that such a characterisation is possible by augmenting the original NPA hierarchy with a finite number of additional linear constraints. This provides a sequence of outer approximations to the corresponding set of correlations that can each be defined via a suitable SDP, with analogous resource requirements and convergence properties of the NPA hierarchy. We then apply our hierarchy to several problems in quantum information. First, we investigate device-independent randomness certification. We show how to use the hierarchy to robustly certify over 2.3 bits of local randomness from a two-qubit state via a simple sequential measurement strategy, thus going beyond the theoretical maximum of two bits that is achievable in non-sequential Bell scenarios. We then show that the previously studied strategies for the simultaneous violation of two CHSH inequalities [9] and the violation of the sequential Bell inequality defined in [8] are both optimal for strategies of any dimension, up to numerical precision.

We note that the recent work [16] also describes a sequence of SDP relaxations for generic quantum-causal networks that can be applied to the sequential structures we consider; see the discussion for further information.

2 Preliminaries

2.1 Quantum correlations

In a standard Bell scenario, two spatially-separated players perform measurements on their local share of a bipartite state, chosen according to some random inputs $x, y = 1, \ldots, m$, and then collect the corresponding outputs $a, b = 1, \ldots, d$. The resulting correlations $P(a, b|x, y)$ are called *quantum*, $P(a, b|x, y) \in \mathcal{Q}$, if they can be written $\mathrm{tr}[\rho\, \mathsf{A}_a^x \otimes \mathsf{B}_b^y]$ for some bipartite quantum state ρ and local measurement operators A_a^x and B_b^y. Here one can take ρ pure and the measurements projective without loss of generality, since any measurement on a mixed state can be realised as a projective measurement on a purification of the state [17]. Thus,

$$P(a, b|x, y) \in \mathcal{Q} \iff P(a, b|x, y) = \langle \psi | \mathsf{A}_a^x \otimes \mathsf{B}_b^y | \psi \rangle \tag{1}$$

with

$$\mathsf{A}_a^x \mathsf{A}_{a'}^x = \mathsf{A}_a^x \delta_{a,a'} \ \forall x, a, a' \,, \quad \mathsf{B}_b^y \mathsf{B}_{b'}^y = \mathsf{B}_b^y \delta_{b,b'} \ \forall y, b, b'. \tag{2}$$

Since the state and measurements appearing in (1) are potentially infinite dimensional, the problem of deciding membership in, or optimising over the set $\mathcal{Q}$ is highly non-trivial. Currently, the only general purpose technique to tackle such a problem is the NPA hierarchy [6, 7], which we will recap shortly.

2.2 Sequential quantum correlations

In this work we consider sequential measurement scenarios, where a quantum system is subjected local measurements that obey a time-ordered structure (see Fig. 1). Consider first a single quantum

system $|\psi\rangle$, of potentially uncountable infinite dimension, that undergoes a sequence of n measurements with inputs x_i and outcomes a_i. The first measurement outcome and its corresponding post-measurement state are described by sets of Kraus operators $\{\mathsf{K}^{x_1}_{a_1,\mu_1}\}$. For finite dimensional systems the (sub-normalised) post-measurement state obtained after obtaining outcome a_1 takes the form

$$\rho_{a_1|x_1} = \sum_{\mu_1} \mathsf{K}^{x_1}_{a_1,\mu_1} |\psi\rangle\langle\psi| \mathsf{K}^{x_1\dagger}_{a_1,\mu_1}, \tag{3}$$

with $P(a_1|x_1) = \operatorname{tr} \rho_{a_1|x_1}$, $\sum_{a_1,\mu_1} \mathsf{K}^{\dagger}_{a_1,\mu_1}\mathsf{K}_{a_1,\mu_1} = \mathbb{1}$, and where the sum over μ_1 is needed since we may have multiple Kraus operators associated to a single measurement outcome. Generally, for infinite dimensional systems one replaces the sum with an integral:

$$\rho_{a_1|x_1} = \int_{\mu_1} \mathrm{d}\mu_1 \, \mathsf{K}^{x_1}_{a_1,\mu_1} |\psi\rangle\langle\psi| \mathsf{K}^{x_1\dagger}_{a_1,\mu_1}, \tag{4}$$

where again $P(a_1|x_1) = \operatorname{tr} \rho_{a_1|x_1}$ and $\sum_{a_1} \int \mathrm{d}\mu_1 \mathsf{K}^{\dagger}_{a_1,\mu_1}\mathsf{K}_{a_1,\mu_1} = \mathbb{1}$. Continuing this process for the entire sequence with inputs $\mathbf{x} = (x_1, \cdots, x_n)$ and outputs $\mathbf{a} = (a_1, \cdots, a_n)$, one finds

$$P(\mathbf{a}|\mathbf{x}) = \langle\psi|\mathsf{A}^{\mathbf{x}}_{\mathbf{a}}|\psi\rangle, \qquad \mathsf{A}^{\mathbf{x}}_{\mathbf{a}} = \int \cdots \int \mathrm{d}\mu_1 \cdots \mathrm{d}\mu_n \mathsf{K}^{x_1\dagger}_{a_1,\mu_1} \mathsf{K}^{x_2\dagger}_{a_2,\mu_2} \cdots \mathsf{K}^{x_n\dagger}_{a_n,\mu_n} \mathsf{K}^{x_n}_{a_n,\mu_n} \cdots \mathsf{K}^{x_2}_{a_2,\mu_2} \mathsf{K}^{x_1}_{a_1,\mu_1},$$

$$\sum_{a_i} \int \mathrm{d}\mu_i \mathsf{K}^{x_i\dagger}_{a_i,\mu_i} \mathsf{K}^{x_i}_{a_i,\mu_i} = \mathbb{1}_{\mathrm{A}} \quad \forall x_i. \tag{5}$$

To ease notation we have left the time-step dependence of the Kraus operators implicit. That is, $\{\mathsf{K}^{x_1\dagger}_{a_1,\mu_1}\}$ and $\{\mathsf{K}^{x_2\dagger}_{a_2,\mu_2}\}$ are in general different sets of operators, which is understood from the input/output indices. We define the set of *sequential quantum correlations* $\mathcal{Q}_{\mathrm{SEQ}}$ as those that arise from performing sequential measurements locally on a bipartite quantum state $|\psi\rangle$, i.e.

$$P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \in \mathcal{Q}_{\mathrm{SEQ}} \iff P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \langle\psi|\mathsf{A}^{\mathbf{x}}_{\mathbf{a}} \otimes \mathsf{B}^{\mathbf{y}}_{\mathbf{b}}|\psi\rangle,$$

where the measurement operators $\mathsf{A}^{\mathbf{x}}_{\mathbf{a}}$ and $\mathsf{B}^{\mathbf{y}}_{\mathbf{b}}$ have the sequential structure (5). Can we define a hierarchy, analogous to the NPA hierarchy for $\mathcal{Q}$, to characterise the set $\mathcal{Q}_{\mathrm{SEQ}}$? In this work we show how this can be achieved in an efficient manner, via a simple adaptation of the original NPA hierarchy.

2.3 The NPA hierarchy

Before explaining our method, we review the NPA hierachy [6, 7]. The NPA hierarchy provides a sequence of tests, each of which checks membership in a set $\mathcal{Q}_i \supseteq \mathcal{Q}$ such that $\mathcal{Q}_1 \supseteq \mathcal{Q}_2 \supseteq \cdots \supseteq \mathcal{Q}$. To see how the NPA hierarchy works, consider some state and projective measurements $|\psi\rangle, \mathsf{A}^x_a, \mathsf{B}^y_b$ with corresponding correlations $P(a,b|x,y) = \langle\psi|\mathsf{A}^x_a\mathsf{B}^y_b|\psi\rangle$ (where A^x_a should be understood as $\mathsf{A}^x_a \otimes \mathbb{1}_{\mathrm{B}}$ and B^y_b as $\mathbb{1}_{\mathrm{A}} \otimes \mathsf{B}^y_b$). Define sets S_k, consisting of the identity operator and all products of the operators A^x_a and B^y_b up to degree k;

$$S_1 = \{\mathbb{1}\} \cup_{a,x} \{\mathsf{A}^x_a\} \cup_{b,y} \{\mathsf{B}^y_b\}, \qquad S_{k+1} = S_k \cup_{i,j} \{S^{(i)}_k S^{(j)}_1\} \tag{6}$$

where $S^{(i)}_k$ is the i^{th} element of S_k. Next, define the *moment matrix of order* k, Γ_k, with elements $\Gamma^{(i,j)}_k$

$$\Gamma^{(i,j)}_k = \langle\psi|(S^{(i)}_k)^\dagger S^{(j)}_k|\psi\rangle, \tag{7}$$

By construction, the matrix Γ_k has the following properties:

i. Γ_k satisfies a number of linear constraints stemming from the orthogonality properties (2), the normalisation of the measurement operators, and from the commutation of Alice's and Bob's operators. For example $\langle\psi|\mathsf{A}^x_a\mathsf{A}^x_{a'}|\psi\rangle = 0$ for $a \neq a'$ and $\langle\psi|[\mathsf{A}^x_a, \mathsf{B}^y_b]|\psi\rangle = 0$. We can write these constraints as $\operatorname{tr}[\Gamma_k G_i] = 0$ for some suitable fixed matrices G_i.

ii. Γ_k contains some elements that correspond to observable probabilities. For example $\langle\psi|\mathsf{A}_a^x\mathsf{B}_b^y|\psi\rangle = P(a,b|x,y)$. We write these constraints as $\mathrm{tr}[\Gamma_k F_j] = P_j$, where F_j are fixed matrices and P_j denotes the corresponding observed probability. Similarly, taking $S_k^{(0)} = \mathbb{1}$ we have $\Gamma_k^{(0,0)} = 1$ since $\mathrm{tr}\,|\psi\rangle\langle\psi| = 1$.

iii. $\Gamma_k^\dagger = \Gamma_k$ and Γ_k is positive semi-definite (see [6] for a simple proof).

Imagine that we are given some other correlation $P(a,b,|x,y)$ for which we want to test membership in $\mathcal{Q}$. If $P \in \mathcal{Q}$, there exists a state and measurements leading to P and a corresponding matrix Γ_k satisfying the above conditions. We thus have a necessary condition for $P \in \mathcal{Q}$:

NPA hierarchy (level k):

$$\begin{aligned} \text{Find } \Gamma_k \quad \text{such that} \quad & \Gamma_k \succcurlyeq 0,\ \Gamma_k^\dagger = \Gamma_k,\ \Gamma_k^{(0,0)} = 1, \\ & \mathrm{tr}[\Gamma_k G_i] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k F_i] = P_i\ \forall i. \end{aligned} \tag{8}$$

We denote the set of correlations with a positive solution to the above problem at level k as $\mathcal{Q}_k$. Since the test is a necessary condition for $P \in \mathcal{Q}$ we have $\mathcal{Q}_k \supseteq \mathcal{Q}$. As the test contains only linear and positive-semidefinite constraints, it can be cast as a SDP feasibility problem and solved efficiently (in the size of the matrix Γ_k) by a suitable solver. We thus have a sequence of SDPs, each of which provides a relaxation to the problem of deciding membership in $\mathcal{Q}$. Since Γ_k is a principle sub-matrix of Γ_{k+1}, one has $\Gamma_{k+1} \succcurlyeq 0 \implies \Gamma_k \succcurlyeq 0$ and so $\mathcal{Q}_{k+1} \supseteq \mathcal{Q}_k$. Furthermore, one can perform optimization of linear combinations of the probabilities P_j over $\mathcal{Q}_k$ by removing the final constraint in (8) and defining a linear combination of the elements $\mathrm{tr}[\Gamma_k F_j]$ as an objective function of the SDP. One can then obtain certified upper and lower bounds to the problem via duality theorems of convex optimisation. In practice, relevant problems can be tackled in this way at low levels of the hierarchy that are tractable on a desktop computer.

In principle, one can use other sets than S_k to generate the moment matrix (7), with each choice giving a different relaxation to $\mathcal{Q}$. Often, and in the examples we present later, we will use a level that is mid-way between level 1 and level 2, often called level 1+AB. This level is defined by the set

$$S_{1+AB} = S_1 \cup_{a,x,b,y} \{\mathsf{A}_a^x\mathsf{B}_b^y\}. \tag{9}$$

This set defines the lowest level in the hierarchy of [18], and defines the so-called set of 'almost quantum' correlations [19]. As we will see, this set is often sufficient to get non-trival and even tight bounds to relevant optimisation problems.

3 NPA hierarchy for sequential correlations

First, let us state our main technical result regarding the characterisation of sequential quantum correlations.

Fact 1. *A given set of correlations $P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y})$ belongs to $\mathcal{Q}_{SEQ}$ if and only if it can be realised as $P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \langle\psi|\mathsf{A}_{\mathbf{a}}^{\mathbf{x}} \otimes \mathsf{B}_{\mathbf{b}}^{\mathbf{y}}|\psi\rangle$, with the measurement operators being projective and satisfying one-way 'no-signalling' and orthogonality conditions. That is*

$$\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}\mathsf{A}_{\mathbf{a}'}^{\mathbf{x}} = \delta_{\mathbf{a},\mathbf{a}'}\mathsf{A}_{\mathbf{a}}^{\mathbf{x}} \quad \forall \boldsymbol{x}, \boldsymbol{a}, \boldsymbol{a}' \tag{10}$$

$$\sum_{a_{k+1},\cdots,a_n} \mathsf{A}_{\boldsymbol{a}}^{\boldsymbol{x}} - \mathsf{A}_{\boldsymbol{a}}^{\boldsymbol{x}'} = 0 \quad \begin{aligned} &\forall a_1,\ldots,a_k\,, \\ &\forall \boldsymbol{x},\boldsymbol{x}' s.t\ x_i = x_i'\ (i \le k) \\ &1 \le k \le n-1 \end{aligned} \tag{11}$$

$$\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}\mathsf{A}_{\mathbf{a}'}^{\mathbf{x}'} = 0 \quad \forall \boldsymbol{x}, \boldsymbol{x}', \boldsymbol{a}, \boldsymbol{a}'\ s.t \quad \begin{aligned} &x_i = x_i',\ (i \le k), \\ &(a_1,\ldots,a_k) \ne (a_1',\ldots,a_k'). \\ &1 \le k \le n \end{aligned} \tag{12}$$

and similarly for $\mathsf{B}_{\mathbf{b}}^{\mathbf{y}}$.

Note that the projective condition (10) is in fact implied by the more general condition (12) and so one can equivalently take only (11) and (12) in the above.

Proof. We first prove that any correlations in $\mathcal{Q}_{\text{SEQ}}$ can be realised using measurement operators satisfying (10), (11) and (12). This can be proven by considering Stinespring dilations of the sequential measurements; see appendix A. For example, for a sequence of two measurements (see figure 2), one finds that Alice's full measurement operator can be written

$$\mathsf{A}_{\mathbf{a}}^{\mathbf{x}} = U_1^{x_1\dagger} U_2^{x_2\dagger} (\mathbb{1} \otimes \Pi_{a_1} \otimes \Pi_{a_2}) U_2^{x_2} U_1^{x_1}, \tag{13}$$

which describes a projective measurement and thus satisfies (10). Here, $U_1^{x_1\dagger}$ acts trivially (with the identity) on the third Hilbert space in the product, and $U_2^{x_2\dagger}$ acts trivially on the second Hilbert space, as shown graphically in figure 2.

The constraint (11) is true for any set of measurement operators that are realised sequentially, as can be seen from (5). This is because it reflects the fact that the measurement operators that define the first k measurements (obtained by marginalising $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ over the last $n-k$ outcomes) must be independent of the last $n-k$ inputs, since these occur later in the sequence. The Stinespring dilation of the measurement operators described previously retains the sequential structure of the measurement, and so this constraint holds for the projective measurement operators as well. For example, by summing over a_2 in (13), one finds an operator that is independent of x_2.

Finally, given the Stinespring dilations of the sequential measurements, property (12) follows from the orthogonality conditions of the Π_{a_j}'s. Consider again the sequence of two measurements in (13). One has for $x_1 = x_1'$ and $a_1 \neq a_1'$ (omitting the tensor products and the identity operator)

$$\begin{aligned} \mathsf{A}_{\mathbf{a}}^{\mathbf{x}} \mathsf{A}_{\mathbf{a}'}^{\mathbf{x}'} &= U_1^{x_1\dagger} U_2^{x_2\dagger} \Pi_{a_1} \Pi_{a_2} U_2^{x_2} U_2^{x_2'\dagger} \Pi_{a_1'} \Pi_{a_2'} U_2^{x_2'} U_1^{x_1} \\ &= U_1^{x_1\dagger} U_2^{x_2\dagger} \Pi_{a_2} U_2^{x_2} U_2^{x_2'\dagger} \Pi_{a_1} \Pi_{a_1'} \Pi_{a_2'} U_2^{x_2'} U_1^{x_1} \\ &= 0, \end{aligned} \tag{14}$$

where we have used $[U_2^{x_2}, \Pi_{a_1}] = 0$. Generalising this for a general sequence we find (12)

We now show the opposite direction, i.e. that any measurement operators satisfying (10), (11) and (12) admit a sequential realisation. In fact, to show this we need only conditions (10), (11). Consider a projective measurement with two input labels x_1, x_2 and two output labels a_1, a_2 defined by the measurement operators $\mathsf{A}_{a_1 a_2}^{x_1 x_2}$, and assume that the measurement operators satisfy (11). The measurement can be realised sequentially as follows. The first device performs a measurement with Kraus operators $\mathsf{K}_{a_1}^{x_1} = \sum_{a_2} \mathsf{A}_{a_1 a_2}^{x_1 x_2}$. These operators are projective and independent of x_2 due to (11). The value x_1 is then sent to the second device (using a classical channel) and the second device measures $\mathsf{K}_{a_2}^{x_2} = \sum_{a_1} \mathsf{A}_{a_1 a_2}^{x_1 x_2}$. The measurement operator describing the full sequence is therefore

$$\mathsf{K}_{a_1}^{x_1\dagger} \mathsf{K}_{a_2}^{x_2\dagger} \mathsf{K}_{a_2}^{x_2} \mathsf{K}_{a_1}^{x_1} = \mathsf{K}_{a_1}^{x_1} \mathsf{K}_{a_2}^{x_2} \mathsf{K}_{a_1}^{x_1} = \sum_{a_2'} \mathsf{A}_{a_1 a_2'}^{x_1 x_2} \sum_{a_1'} \mathsf{A}_{a_1' a_2}^{x_1 x_2} \sum_{a_2''} \mathsf{A}_{a_1 a_2''}^{x_1 x_2} = \mathsf{A}_{a_1 a_2}^{x_1 x_2} \tag{15}$$

as required. In the first equality we have used the fact that the Kraus operators are Hermitian and projective by construction. The final equality follows from $\mathsf{A}_{a_1 a_2}^{x_1 x_2}$ being projective. In the above we made use of a communication channel that we have not explicitly modelled but can be realised sequentially ; see appendix B for a proof where this channel is explicit. The general result for sequences of any length can be achieved in the same fashion by applying the same technique inductively on the sequence.

□

Having established Fact 1 we may define a hierarchy of relaxations to $\mathcal{Q}_{\text{SEQ}}$ as follows. Define moment matrices Γ_k as in (7) using the projective measurement operators $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ and $\mathsf{B}_{\mathbf{b}}^{\mathbf{y}}$ (i.e. satisfying (10)), leading to analogous constraints to (8). At this point, the relaxation is equivalent to the standard NPA hierarchy, treating the sequences of measurements as single measurements. The constraints (10), (11) and (12) are linear constraints on the measurement operators and thus imply additional linear constraints on Γ_k. One can therefore add these extra constraints in the form of

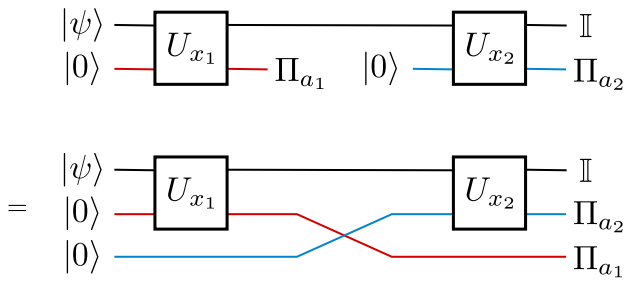

Figure 2: Top: the dilation of a sequence of two measurements. Each measurement in the sequence can be realised by appending an ancila state, performing a joint unitary operation and making a projective measurement on the ancilla space. Bottom: absorbing the ancilla states in the initial state and moving all projective measurements to the end, the scheme is equivalent to a unitary operation followed by a projective measurement. The measurement operators $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ can thus be taken projective WLOG.

extra fixed matrices G_i^{SEQ} to (8). This leads us to the following hierarchy for sequential quantum correlations

Sequential hierarchy (level k):

$$\begin{aligned} \text{Find } \Gamma_k \quad \text{such that} \quad & \Gamma_k \succcurlyeq 0,\ \Gamma_k^\dagger = \Gamma_k,\ \Gamma_k^{(0,0)} = 1, \\ & \mathrm{tr}[\Gamma_k G_i] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k G_i^{\mathrm{SEQ}}] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k F_i] = P_i\ \forall i. \end{aligned} \tag{16}$$

We call $\mathcal{Q}_{\mathrm{SEQ}}^k$ the set defined at level k of this hierarchy. As with the NPA hierachy, the sets $\mathcal{Q}_{\mathrm{SEQ}}^k$ can be optimised over via SDP solvers with a comparable resource overhead. Note that due to the normalisation of measurement operators and (11), some of the measurement operators can be written as linear combinations of others. In practice, this means that such operators can be excluded from the sets S_k (thus increasing efficiency by decreasing the size of Γ_k) since their addition will result in linear dependencies between the rows and columns of Γ_k, which do not affect the constraint $\Gamma_k \succcurlyeq 0$. This process will also introduce further constraints on the now smaller Γ_k. For example, if $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ and $\mathsf{A}_{\mathbf{a}'}^{\mathbf{x}'}$ are two measurement operators that have been removed from S_k through this process, then by expressing them as linear combinations of the remaining elements in S_k, the constraint (12) gives a polynomial operator identity that implies further constraints on the moment matrix.

3.1 Convergence of the hierarchy

Since the conditions (11) characterise precisely the set of sequential measurement operators and are linear constraints, one can use the same methods as in [7] to prove convergence of the hierarchy. In fact, one can extract a quantum state and measurement operators from the moment matrix Γ_∞ corresponding to the asymptotic level of the hierarchy. It is then straightforward to see that the added linear constraints G_i^{SEQ} enforce that the extracted measurement operators satisfy property (11), hence having a sequential realisation. Technically speaking, the convergence is proven to a set $\tilde{\mathcal{Q}}_{\mathrm{SEQ}} \supseteq \mathcal{Q}_{\mathrm{SEQ}}$. Here, $\tilde{\mathcal{Q}}_{\mathrm{SEQ}}$ is the set of sequential quantum correlations where the tensor product structure is replaced by the weaker constraint that Alice and Bob's measurement operators commute, i.e.

$$p(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \in \tilde{\mathcal{Q}}_{\mathrm{SEQ}} \iff p(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \langle\psi|\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}\mathsf{B}_{\mathbf{b}}^{\mathbf{y}}|\psi\rangle.$$

where one has $[\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}, \mathsf{B}_{\mathbf{b}}^{\mathbf{y}}] = 0$ for all $\mathbf{a},\mathbf{x},\mathbf{b},\mathbf{y}$ and the measurement operators have the sequential structure (5). This commuting operator formalism is used in algebraic quantum field theory [20], and it is known that there exist scenarios for which $\mathcal{Q}_{\mathrm{SEQ}} \subset \tilde{\mathcal{Q}}_{\mathrm{SEQ}}$ [21].

3.2 Relaxations of local correlations

The hierarchy can also be used to define semidefinite programming relaxations to the set of 'time ordered local correlations' defined in [8]. Such correlations are those that can be obtained by a local hidden variable model that must respect the sequential causal structure of the scenario. The idea essentially the same as that presented in [22]; as we show in appendix C, any hidden variable model can be seen as a special case of a quantum strategy, where all measurement operators of the same party commute. For the sequential scenario, one therefore just has to add the additional linear constraints to Γ_k implied by the relations $[\mathsf{A}^{\mathbf{x}}_{\mathbf{a}}, \mathsf{A}^{\mathbf{x}'}_{\mathbf{a}'}] = 0$ and $[\mathsf{B}^{\mathbf{y}}_{\mathbf{b}}, \mathsf{B}^{\mathbf{y}'}_{\mathbf{b}'}] = 0$.

4 Applications

In the rest of this article we use our methods to tackle a number of open questions in quantum information theory. Code to implement our method in python can be found in the GitLab repository `https://gitlab.com/josephbowles/sequentialnpa`.

4.1 Robust device-independent certification of more that 2 bits of local randomness

One of the most important applications of the NPA hierarchy is bounding the amount of randomness one can certify from an observed probability distribution in the device-independent setting [23, 24, 25, 26, 27, 28, 29, 30]. A common figure of merit that is used is the *local guessing probability*, defined as the maximum probability with which an adversary—usually called Eve—could guess the value of one of the local outputs for a fixed local input. More precisely, consider the set of tripartite probability distributions $p_{\text{ABE}}(a, b, e|x, y)$ for Alice, Bob and Eve (where Eve has no input and the same output alphabet as Bob) that have a realisation in quantum theory, i.e. $p_{\text{ABE}}(a, b, e|x, y) = \langle\psi|\mathsf{A}^x_a \otimes \mathsf{B}^y_b \otimes \mathsf{E}_e|\psi\rangle \iff p_{ABE} \in \mathcal{Q}$ for some state and measurements. Define $p_{\text{AB}}(a, b|x, y)$ and $p_{\text{BE}}(b, e|y)$ to be the corresponding marginal distributions of $p_{\text{ABE}}(a, b, e|x, y)$. The local guessing probability for Bob's input $y = y^*$ given an observed probability distribution $P_{\text{obs}}(a, b|x, y)$ is the best probability that Eve could guess b given $y = y^*$ while simultaneously reproducing P_{obs} when marginalising over her output. That is,

$$G(y^*) = \max_{p_{\text{ABE}} \in \mathcal{Q}} \sum_{e=1}^{|b|} p_{\text{BE}}(e, e|y^*) \qquad \text{such that} \qquad \begin{aligned} p_{\text{AB}}(a, b|x, y) &= \sum_e p_{\text{ABE}}(a, b, e|x, y) \\ &= P_{\text{obs}}(a, b|x, y) \end{aligned} \tag{17}$$

where $|b|$ is the size of Bob's output alphabet. To define the local guessing probability in the sequential scenario one imposes that the distribution p_{ABE} be realised by a sequential quantum strategy. That is, the local guessing probability for Bob's input $\mathbf{y}^*$ given an observed distribution $P_{\text{obs}}(\mathbf{a}, \mathbf{b}|\mathbf{x}, \mathbf{y})$ becomes

$$G(\mathbf{y}^*) = \max_{p_{\text{ABE}}} \sum_{\mathbf{e}} p_{\text{BE}}(\mathbf{e}, \mathbf{e}|\mathbf{y}^*) \qquad \text{such that} \qquad \begin{aligned} p_{\text{AB}}(\mathbf{a}, \mathbf{b}|\mathbf{x}, \mathbf{y}) &= \sum_{\mathbf{e}} p_{\text{ABE}}(\mathbf{a}, \mathbf{b}, \mathbf{e}|\mathbf{x}, \mathbf{y}) \\ &= P_{\text{obs}}(\mathbf{a}, \mathbf{b}|\mathbf{x}, \mathbf{y}), \end{aligned} \tag{18}$$

where the alphabet of $\mathbf{e}$ is the same as $\mathbf{b}$ and where p_{ABE} has a sequential realisation, i.e.

$$p_{\text{ABE}}(\mathbf{a}, \mathbf{b}, \mathbf{e}|\mathbf{x}, \mathbf{y}) = \langle\psi|\mathsf{A}^{\mathbf{x}}_{\mathbf{a}} \otimes \mathsf{B}^{\mathbf{y}}_{\mathbf{b}} \otimes \mathsf{E}_{\mathbf{e}}|\psi\rangle, \tag{19}$$

where the measurement operators $\mathsf{A}^{\mathbf{x}}_{\mathbf{a}}$ and $\mathsf{B}^{\mathbf{y}}_{\mathbf{b}}$ have the structure (5). In appendix D we show how upper bounds to (18) can be obtained efficiently using our hierarchy.

In the standard Bell scenario, the local guessing probability (17) is always lower bounded by $1/d^2$, where d is the local Hilbert space dimension of the state used to obtain the observed correlations. This follows from the fact that extremal measurements acting on a Hilbert space of dimension d have at most d^2 outcomes [30, 31]. Hence, the amount of randomness, expressed as the min entropy $-\log_2(G)$ is always lower than $2\log_2(d)$ bits. However, if one imposes the sequential structure on the local measurement one can no longer bound the number of outcomes of extremal

measurements. In [13] Curchod et. al. use this to construct a protocol to obtain arbitrarily small local guessing probabilities from any two-qubit entangled pure state using a single Alice and a sequence of Bobs.

The construction in [13] has two disadvantages however. Firstly, the number of measurements that Alice makes grows quickly with the amount of certified randomness. For example, to certify more that two bits of local randomness one needs at least 14 measurements for Alice. Secondly, although the authors prove that the protocol is noise resistant in principle, precise upper bounds on the guessing probability could not be proven for any nonzero level of noise, and the method can therefore not be used in practice. In the following we show that one can use our hierarchy to certify more than two bits of local randomness in a simple sequential scenario using only two measurements for Alice. Moreover, we use our hierarchy to calculate upper bounds to the guessing probabilities in the presence of noise, thus making the scheme experimentally relevant.

To generate the observed correlations P_{obs} we consider a scenario involving one Alice and a sequence of two Bobs (that we call Bob_1 and Bob_2), where Alice and Bob_1 share the two-qubit isotropic state with noise parameter η:

$$\rho(\eta) = (1-\eta)|\phi^+\rangle\langle\phi^+| + \eta\, \mathbb{1}/4 \tag{20}$$

with $|\phi^+\rangle = [|00\rangle + |11\rangle]/\sqrt{2}$. Alice performs one of two measurements given by the observables $\cos\mu\,\sigma_{\text{z}} \pm \sin\mu\,\sigma_{\text{x}}$, where $\tan\mu = \sin 2\epsilon$ and ϵ is a free parameter. Bob_1 performs one of two measurements. For $y_1 = 0$ he performs a projective measurement of σ_{z} with Kraus operators $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. For $y_1 = 1$ he performs the two outcome measurement defined by the Kraus operators

$$\mathsf{K}_+ = \cos\epsilon|+\rangle\langle +| + \sin\epsilon|-\rangle\langle -|, \qquad \mathsf{K}_- = -\cos\epsilon|-\rangle\langle -| + \sin\epsilon|+\rangle\langle +|. \tag{21}$$

The parameter ϵ controls the strength of the measurement: for $\epsilon = 0$, the measurement is a projective measurement in the x direction; for $\epsilon = \pi/4$ the measurement is non-interacting. Bob_2 performs one of three measurements. For $y_2 = 0, 1$ he performs a projective measurement of σ_{z} or σ_{x}. For $y_2 = 2$ he performs the symmetric 3-outcome POVM given by the measurement operators

$$\mathsf{M}_{b_2} = \frac{2}{3}\frac{\mathbb{1} + \mathbf{v}_{b_2}\cdot\vec{\sigma}}{2} \quad b_2 = 0, 1, 2, \tag{22}$$

where $\mathbf{v}_{b_2} = (\sin(\frac{2\pi}{3}b_2), 0, \cos(\frac{2\pi}{3}b_2))$. The inspiration for these measurements is the following. For $y_1 = 1$, the post measurement state shared between Alice and Bob_2 will be one of two partially entangled states, depending on the value of b_1. The correlations obtained by performing the measurements for $x, y_2 = 0, 1$ on these states are known to self-test both of the corresponding state and measurements [32]. We expect (although we have not proven) that this implies that the state shared between Alice and Bob_1 is $|\phi^+\rangle$ and the measurement for Bob_1 (21), which essentially implies that one must have $p(b_2|y_2 = 2) = \frac{1}{3}$, leading to more than two bits of randomness.

In figure 3 we present upper bounds to $G(\mathbf{y}^* = (1, 2))$ obtained in this way as a function of η, with $\epsilon = 7\pi/32$ and calculated using level $1 + AB$ of the hierarchy. For low noise, one can surpass two bits of randomness. Moreover, for close to 4% noise (well within experimental reach) our strategy outperforms the non-sequential strategy where one performs the measurement that maximally violate the CHSH Bell inequality on the same state. We leave a more detailed analysis of noise including detector inefficiencies to future work.

4.2 Monogamy of nonlocality in sequential measurement scenarios

Consider a scenario involving one Alice and two Bobs, where each party has two inputs and two outputs, with inputs and outputs labelled by 0,1. The value of the CHSH Bell functional between Alice and Bob_1 is

$$\text{CHSH}_{\text{AB1}} = \sum_{x,y_1}\sum_{a,b_1}(-1)^{a+b_1+x\cdot y_1} P_{\text{AB1}}(a, b_1|x, y_1) \tag{23}$$

where P_{AB1} is the marginal distribution between Alice and Bob_1. We may define the average CHSH Bell functional between Alice and Bob_2 as

$$\text{CHSH}_{\text{AB2}} = \frac{1}{2}\sum_{b_1,y_1}\sum_{x,y_2}\sum_{a,b_2}(-1)^{a+b_2+x\cdot y_2} P(a, b_1, b_2|x, y_1, y_2),$$

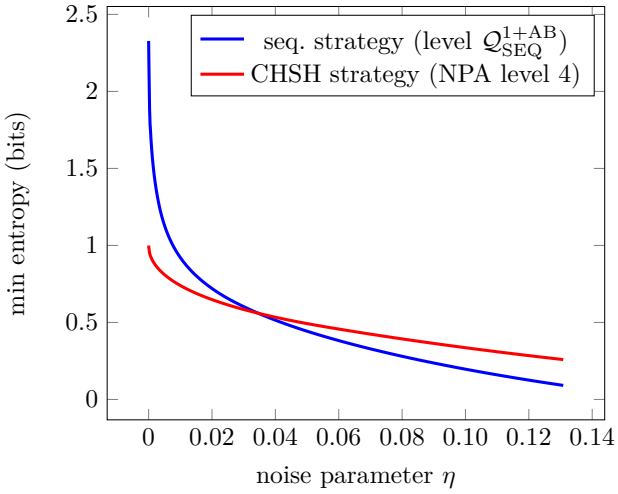

Figure 3: Blue: lower bound to the local randomness as a function of the noise parameter η for our sequential measurement strategy, obtained at level 1+AB of our sequential hierarchy. Red: corresponding local randomness obtainable with the same state in a non-sequential scenario using measurements that lead to the maximal violation of the CHSH Bell inequality, obtained at level 4 of the NPA hierarchy.

i.e. the CHSH Bell functional between Alice and Bob_2, averaged over b_1 and a uniform choice of y_1. The values of CHSH_{AB1} and CHSH_{AB2} are subject to monogamy due to both the monogamy of correlations and the sequential measurement constraints. Silva et. al. investigate this in [9], finding that for two-qubit systems, the optimal trade-off satisfies

$$\text{CHSH}_{\text{AB2}} \leq \sqrt{2}\left(1+\sqrt{1-\frac{(\text{CHSH}_{\text{AB1}})^2}{8}}\right), \tag{24}$$

which can be saturated with an appropriate choice of measurements. We use the sequential NPA hierarchy to investigate this trade-off for systems of general dimension. We numerically maximise the value of CHSH_{AB2} conditioned on values of CHSH_{AB1} at level 1+AB of the hierarchy (see figure 4). We find that the values obtained match those of (24) up to the precision of the SDP solver. Thus, we conjecture that the strategies presented in [9] are optimal for any dimension. This is somewhat surprising since one may expect to gain an advantage from higher dimensional systems. For example, it would allow Bob_1 to communicate perfectly the value of y_1 and b_1 to Bob_2, which in principle could increase the value of $\text{CHSH}_{\text{A,B2}}$.

4.3 Tight bounds on sequential Bell inequalities

In [8] Gallego et. al. present a Bell inequality (see equation 51 therein) that defines a facet of the set of correlations that admit a sequential time-ordered local model. The scenario involves one Alice and two Bobs, with each party performing one of two dichotomic measurements. The Bell inequality is constructed as follows. Define the correlators

$$\begin{aligned}\langle A_x B^2_{y_1 y_2}\rangle &= P(a\cdot b_2=+1|x,y_1,y_2)-P(a\cdot b_2=-1|x,y_1,y_2)\\ \langle A_x B^1_{y_1} B^2_{y_1 y_2}\rangle &= P(a\cdot b_1\cdot b_2=+1|x,y_1,y_2)-P(a\cdot b_1\cdot b_2=-1|x,y_1,y_2).\end{aligned} \tag{25}$$

The inequality is given by

$$\mathcal{I} = \langle A_0(B-B') - A_1(B+B')\rangle \leq 2 \tag{26}$$

where

$$\begin{aligned}B &= \frac{1}{2}[(1+B^1_0)B^2_{01} - (1-B^1_0)B^2_{00}]\\ B' &= \frac{1}{2}[(1-B^1_1)B^2_{11} + (1+B^1_1)B^2_{10}]\end{aligned} \tag{27}$$

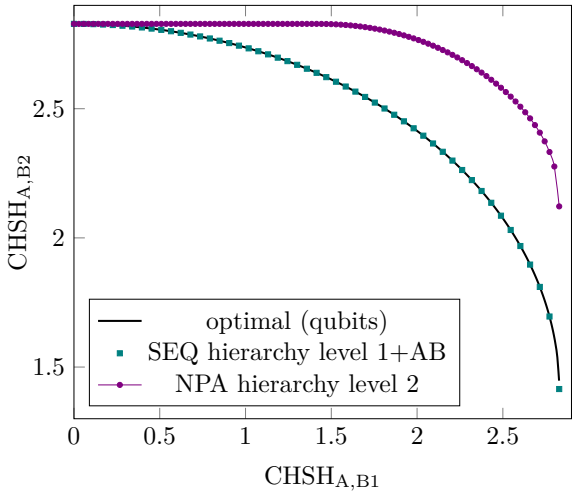

Figure 4: Upper bounds on the maximum value of $\mathrm{CHSH}_{\mathrm{A,B2}}$ as a function of the value of $\mathrm{CHSH}_{\mathrm{A,B1}}$. The values obtained at level 1+AB of the sequential hierarchy match the optimal values for qubit strategies found in [9]. To show the effect of our new constraints, we plot the same bounds obtained via the standard NPA hierarchy at level 2, treating the two Bobs as a single party.

and the bound 2 holds for sequential time-ordered local correlations.

The authors show that it is possible to violate the inequality up to a value of $2\sqrt{2}$ using a sequential quantum strategy, providing a lower bound to the maximum violation using a sequential quantum strategy. Using our hierarchy at level 1+AB, we are able to certify a corresponding upper bound that agrees with the value $2\sqrt{2}$ up to the precision of the SDP solver. We therefore expect that the strategy given in [8] is optimal for this inequality.

5 Discussion

We have presented a general method to bound sets of correlations arising from performing sequential measurements on entangled quantum states. Our techniques can be seen as part of a collection of works that extend the original applicability of the NPA hierarchy to scenarios of restricted dimension [33, 34] and entanglement [18], classicality [22], and modified causality [16**?**].

We note that the techniques described in [16] can in principle deal with the sequential causal structures considered in this work. More specifically, one could use their method to treat 'quantum exogenous' variables by explicitly using the unitaries in (13) as operators in the generating set S_k and defining a resulting relaxation. This method is significantly less efficient however since one needs to go to high levels (with large moment matrices) of the corresponding relaxation, and no convergence properties are proven. Given these points, it would thus be interesting to study whether our method could be extended to other causal scenarios, or be used to improve the efficiency of the method in [16]. For example, can our method be applied to give a convergent hierarchy for a generic causal structure involving latent quantum variables?

The NPA hierarchy is often used as a numerical method to bound fidelities in self-testing protocols [35]. One avenue of research would therefore be to investigate whether sequential measurement scenarios can improve self-testing fidelity bounds, by adapting the current method to our hierarchy, or to investigate the self-testing of quantum channels, to which sequential measurement scenarios are naturally related. Finally, it would also be interesting to use our method to investigate to what extent sequential measurements can improve other device-independent protocols. For example, can our advantages in local guessing probability be translated to practical improvements to rates in randomness extraction or quantum key distribution protocols?

Acknowledgements

We thank Erik Woodhead for pointing out equation (12) and Flavien Hirsch for inspiring preliminary discussions. We also thank Daniel Cavalcanti, Florian Curchod, Dr. Bibounde, Antonio Acin, Remigiusz Augusiak, Marco Tulio Quintino and Peter Wittek for discussions throughout the project.

All authors acknowledge funding from the Spanish MINECO (QIBEQI FIS2016-80773-P, Severo Ochoa SEV-2015-0522, a Severo Ochoa PhD fellowship), Fundacio Cellex, Generalitat de Catalunya (SGR 1381 and CERCA Programme). JB acknowldges funding from the AXA Chair in Quantum Information Science, Juan de la Cierva-formación and the EU Quantum Flagship project QRANGE. FB acknowledges the support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project number 414325145 in the framework of the Austrian Science Fund (FWF): SFB F71.

References

[1] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195, 1964.

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:839–840, 2014. DOI: 10.1103/RevModPhys.86.839.

[3] Koon Tong Goh, J ędrzej Kaniewski, Elie Wolfe, Tamás Vértesi, Xingyao Wu, Yu Cai, Yeong-Cherng Liang, and Valerio Scarani. Geometry of the set of quantum correlations. *Phys. Rev. A*, 97:022104, Feb 2018. DOI: 10.1103/PhysRevA.97.022104.

[4] Mafalda L. Almeida, Jean-Daniel Bancal, Nicolas Brunner, Antonio Acín, Nicolas Gisin, and Stefano Pironio. Guess your neighbor's input: A multipartite nonlocal game with no quantum advantage. *Phys. Rev. Lett.*, 104:230404, Jun 2010. DOI: 10.1103/PhysRevLett.104.230404.

[5] William Slofstra. The set of quantum correlations is not closed, 2017. URL `https://arxiv.org/abs/1703.08618`. arXiv:1703.08618v2.

[6] M. Navascués, S. Pironio, and A. Acín. Bounding the set of quantum correlations. *Phys. Rev. Lett.*, 98:010401, 2007. DOI: 10.1103/PhysRevLett.98.010401.

[7] M. Navascués, S. Pironio, and A. Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New Journal of Physics*, 10(7):073013, 2008. DOI: 10.1088/1367-2630/10/7/073013.

[8] Rodrigo Gallego, Lars Erik Würflinger, Rafael Chaves, Antonio Acín, and Miguel Navascués. Nonlocality in sequential correlation scenarios. *New Journal of Physics*, 16(3):033037, mar 2014. DOI: 10.1088/1367-2630/16/3/033037.

[9] Ralph Silva, Nicolas Gisin, Yelena Guryanova, and Sandu Popescu. Multiple observers can share the nonlocality of half of an entangled pair by using optimal weak measurements. *Phys. Rev. Lett.*, 114:250401, Jun 2015. DOI: 10.1103/PhysRevLett.114.250401.

[10] Flavien Hirsch, Marco Túlio Quintino, Joseph Bowles, and Nicolas Brunner. Genuine hidden quantum nonlocality. *Phys. Rev. Lett.*, 111:160402, Oct 2013. DOI: 10.1103/PhysRevLett.111.160402.

[11] Sandu Popescu. Bell's inequalities and density matrices: Revealing "hidden" nonlocality. *Phys. Rev. Lett.*, 74:2619–2622, Apr 1995. DOI: 10.1103/PhysRevLett.74.2619.

[12] N. Gisin. Hidden quantum nonlocality revealed by local filters. *Physics Letters A*, 210(3):151 – 156, 1996. ISSN 0375-9601. DOI: https://doi.org/10.1016/S0375-9601(96)80001-6.

[13] F. J. Curchod, M. Johansson, R. Augusiak, M. J. Hoban, P. Wittek, and A. Acín. Unbounded randomness certification using sequences of measurements. *Phys. Rev. A*, 95:020102, Feb 2017. DOI: 10.1103/PhysRevA.95.020102.

[14] Matthew F Pusey. Anomalous weak values are proofs of contextuality. *Physical review letters*, 113(20):200401, 2014. DOI: 10.1103/PhysRevLett.113.200401.

[15] Costantino Budroni, Tobias Moroder, Matthias Kleinmann, and Otfried Gühne. Bounding temporal quantum correlations. *Physical review letters*, 111(2):020403, 2013. DOI: 10.1103/PhysRevLett.111.020403.

[16] Elie Wolfe, Alejandro Pozas-Kerstjens, Matan Grinberg, Denis Rosset, Antonio Acín, and Miguel Navascues. Quantum inflation: A general approach to quantum causal compatibility. *arXiv preprint arXiv:1909.10519*, 2019. URL `https://arxiv.org/abs/1909.10519`.

[17] W. Stinespring. Positive functions on C^*-algebras. *Proceedings of the American Mathematical Society*, 6(2):211–211, Jan 1955. DOI: 10.1090/s0002-9939-1955-0069403-4.
[18] Tobias Moroder, Jean-Daniel Bancal, Yeong-Cherng Liang, Martin Hofmann, and Otfried Gühne. Device-independent entanglement quantification and related applications. *Phys. Rev. Lett.*, 111:030501, Jul 2013. DOI: 10.1103/PhysRevLett.111.030501.
[19] Miguel Navascués, Yelena Guryanova, Matty J Hoban, and Antonio Acín. Almost quantum correlations. *Nature communications*, 6(1):1–7, 2015. DOI: 10.1038/ncomms7288.
[20] Rudolf Haag and Daniel Kastler. An algebraic approach to quantum field theory. *Journal of Mathematical Physics*, 5(7):848–861, 1964. DOI: 10.1063/1.1704187.
[21] William Slofstra. Tsirelson's problem and an embedding theorem for groups arising from non-local games. *Journal of the American Mathematical Society 33*, 2020. DOI: 10.1090/jams/929.
[22] F. Baccari, D. Cavalcanti, P. Wittek, and A. Acín. Efficient device-independent entanglement detection for multipartite systems. *Phys. Rev. X*, 7:021042, Jun 2017. DOI: 10.1103/PhysRevX.7.021042.
[23] Roger Colbeck. Quantum and relativistic protocols for secure multi-party computation. *arXiv preprint arXiv:0911.3814*, 2009. URL `https://arxiv.org/abs/0911.3814`.
[24] Roger Colbeck and Adrian Kent. Private randomness expansion with untrusted devices. *Journal of Physics A: Mathematical and Theoretical*, 44(9):095305, 2011. DOI: 10.1088/1751-8113/44/9/095305.
[25] Renato Renner. Security of quantum key distribution. *International Journal of Quantum Information*, 6(01):1–127, 2008. DOI: 10.1142/S0219749908003256.
[26] Olmo Nieto-Silleras, Stefano Pironio, and Jonathan Silman. Using complete measurement statistics for optimal device-independent randomness evaluation. *New Journal of Physics*, 16 (1):013035, 2014. DOI: 10.1088/1367-2630/16/1/013035.
[27] Olmo Nieto-Silleras, Cédric Bamps, Jonathan Silman, and Stefano Pironio. Device-independent randomness generation from several bell estimators. *New journal of physics*, 20(2):023049, 2018. DOI: 10.1088/1367-2630/aaaa06.
[28] Stefano Pironio, Antonio Acín, Serge Massar, A Boyer de La Giroday, Dzmitry N Matsukevich, Peter Maunz, Steven Olmschenk, David Hayes, Le Luo, T Andrew Manning, et al. Random numbers certified by bell's theorem. *Nature*, 464(7291):1021, 2010. DOI: 10.1038/nature09008.
[29] Jean-Daniel Bancal, Lana Sheridan, and Valerio Scarani. More randomness from the same data. *New Journal of Physics*, 16(3):033011, 2014. DOI: 10.1088/1367-2630/16/3/033011.
[30] Antonio Acín, Stefano Pironio, Tamás Vértesi, and Peter Wittek. Optimal randomness certification from one entangled bit. *Physical Review A*, 93(4):040102, 2016. DOI: 10.1103/PhysRevA.93.040102.
[31] Giacomo Mauro D'Ariano, Paoloplacido Lo Presti, and Paolo Perinotti. Classical randomness in quantum measurements. *Journal of Physics A: Mathematical and General*, 38(26):5979, 2005. DOI: 10.1088/0305-4470/38/26/010.
[32] Cédric Bamps and Stefano Pironio. Sum-of-squares decompositions for a family of clauser-horne-shimony-holt-like inequalities and their application to self-testing. *Physical Review A*, 91(5):052111, 2015. DOI: 10.1103/PhysRevA.91.052111.
[33] Miguel Navascués and Tamás Vértesi. Bounding the set of finite dimensional quantum correlations. *Phys. Rev. Lett.*, 115:020501, Jul 2015. DOI: 10.1103/PhysRevLett.115.020501.
[34] Miguel Navascués, Gonzalo de la Torre, and Tamás Vértesi. Characterization of quantum correlations with local dimension constraints and its device-independent applications. *Phys. Rev. X*, 4:011011, Jan 2014. DOI: 10.1103/PhysRevX.4.011011.
[35] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review. *Quantum*, 4:337, September 2020. ISSN 2521-327X. DOI: 10.22331/q-2020-09-30-337.

A Stinespring dilation of sequential measurement

Consider for simplicity a sequence of two measurements with inputs x_1, x_2 and outputs a_1, a_2 on a single quantum system $|\psi\rangle$ that is described by Kraus operators $\mathsf{K}^{x_j}_{a_j,i_j}$. The sequence of

measurement can be realised as follows. Introduce ancilla spaces A_1' and A_1'' and the ancilla state $|0\rangle = |0\rangle_{\mathrm{A}_1'}|0\rangle_{\mathrm{A}_1''}$. Define an operator $U_1^{x_1}$ via its action on the state $|\psi\rangle|0\rangle$ as

$$U_1^{x_1}|\psi\rangle|0\rangle = \sum_{a_1}\int_{\mu_1} \mathrm{d}\mu_1(\mathsf{K}_{a_1,\mu_1}^{x_1}|\psi\rangle)|a_1\rangle|\mu_1\rangle. \tag{28}$$

One has $\langle\phi|\langle 0|\langle 0|U_1^{x_1\dagger}U_1^{x_1}|\psi\rangle|0\rangle|0\rangle = \langle\phi|\psi\rangle$ for all $|\psi\rangle, |\phi\rangle$. It follows that $U_1^{x_1}$ can be extended to a unitary operator acting on $|\psi\rangle|0\rangle$. Measure the A_1' space in the $|a_1\rangle$ basis, obtaining outcome a_1. Conditioning on outcome a_1 and tracing out the A_1' and A_1'' spaces, one finds (4). We have thus reproduced the first measurement in the sequence. Introducing a fresh ancilla and repeating this for the second measurement in the sequence we find

$$\mathsf{A}_{\mathbf{a}}^{\mathbf{x}} = U_1^{x_1\dagger}U_2^{x_2\dagger}(\mathbb{1}\otimes\Pi_{a_1}\otimes\Pi_{a_2})U_2^{x_2}U_1^{x_1}, \tag{29}$$

where the Π_{a_i}'s are projectors onto the corresponding spaces. The full measurement $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ is thus projective. We may repeat this process for a sequence of arbitrary length, and hence $\mathsf{A}_{\mathbf{a}}^{\mathbf{x}}$ can be taken to be projective without loss of generality.

B Detailed proof of fact 1

Here we give a proof of the reverse direction of fact 1, where we explicitly model the communication channel in the Kraus operators. Enlarge the system via an ancilla state so that the full state is $|\psi\rangle\otimes|0\rangle$. This space will be used as a communication channel in the following. The first device performs a measurement with Kraus operators $\mathsf{K}_{a_1}^{x_1} = \sum_{a_2}\mathsf{A}_{a_1a_2}^{x_1x_2}\otimes V_{x_1}$, where V_{x_1} is a unitary operator that maps $|0\rangle$ to $|x_1\rangle$ (for example, if $x_1 = 0, 1$ then $V_{x_1} = \sigma_x^{x_1}$). These operators are independent of x_2 due to (11). The second device measures (projective) Kraus operators $\mathsf{K}_{a_2}^{x_2} = \sum_{x_1,a_1}\mathsf{A}_{a_1a_2}^{x_1x_2}\otimes|x_1\rangle\langle x_1|$. The measurement operator describing the full sequence is therefore

$$\begin{aligned}\mathsf{K}_{a_1}^{x_1\dagger}\mathsf{K}_{a_2}^{x_2\dagger}\mathsf{K}_{a_2}^{x_2}\mathsf{K}_{a_1}^{x_1} =& (\sum_{a_2'}\mathsf{A}_{a_1a_2'}^{x_1x_2}\otimes V_{x_1}^\dagger)(\sum_{x_1',a_1'}\mathsf{A}_{a_1'a_2}^{x_1'x_2}\otimes|x_1'\rangle\langle x_1'|)(\sum_{a_2''}\mathsf{A}_{a_1a_2''}^{x_1x_2}\otimes V_{x_1})\\ =& \sum_{x_1'}\left((\sum_{a_2'}\mathsf{A}_{a_1a_2'}^{x_1x_2}\sum_{a_1'}\mathsf{A}_{a_1'a_2}^{x_1'x_2}\sum_{a_2''}\mathsf{A}_{a_1a_2''}^{x_1x_2})\otimes V_{x_1}^\dagger|x_1'\rangle\langle x_1'|V_{x_1}\right)\end{aligned} \tag{30}$$

The resulting correlations are

$$\begin{aligned}&\langle\psi|\otimes\langle 0|\sum_{x_1'}\left((\sum_{a_2'}\mathsf{A}_{a_1a_2'}^{x_1x_2}\sum_{a_1'}\mathsf{A}_{a_1'a_2}^{x_1'x_2}\sum_{a_2''}\mathsf{A}_{a_1a_2''}^{x_1x_2})\otimes V_{x_1}^\dagger|x_1'\rangle\langle x_1'|V_{x_1}\right)|\psi\rangle\otimes|0\rangle\\ =&\langle\psi|(\sum_{a_2'}\mathsf{A}_{a_1a_2'}^{x_1x_2}\sum_{a_1'}\mathsf{A}_{a_1'a_2}^{x_1x_2}\sum_{a_2''}\mathsf{A}_{a_1a_2''}^{x_1x_2})|\psi\rangle = \langle\psi|\mathsf{A}_{a_1a_2}^{x_1x_2}|\psi\rangle\end{aligned} \tag{31}$$

as desired.

C Hierarchy for time ordered local correlations

Here we show how to modify the our hierarchy for sequential quantum correlations introduced in the main text in order to approximate the set of time ordered local correlations. Following [8], we say that the correlations from a Bell scenario are time ordered local if they can be described by the following model

$$P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \int_\lambda d\lambda\rho(\lambda)p(\mathbf{a}|\mathbf{x},\lambda)p(\mathbf{b}|\mathbf{y},\lambda)\,, \tag{32}$$

where the distribution $p(\mathbf{a}|\mathbf{x},\lambda)$ satisfies the following sequential no-signaling constraint for all values of λ

$$\sum_{a_{k+1},\cdots,a_n} p(\mathbf{a}|\mathbf{x},\lambda) - p(\mathbf{a}|\mathbf{x}',\lambda) = 0 \quad \begin{array}{l} \forall a_1,\ldots,a_k \\ \forall \mathbf{x},\mathbf{x}' \text{ s.t } x_i = x_i', \\ (i \leq k), \end{array} \tag{33}$$

and similarly for $p(\mathbf{b}|\mathbf{y},\lambda)$. Correlations in the above form are the only ones that can be achieve with classical means in a sequential Bell scenario.

It is well know that, by using the constraints in (33), the model (32) can be reduced to a sum over deterministic strategies, namely

$$P(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \sum_{\lambda} p(\lambda) D^{SEQ}(\mathbf{a}|\mathbf{x},\lambda) D^{SEQ}(\mathbf{b}|\mathbf{y},\lambda)\,, \tag{34}$$

where the deterministic probability distributions split into a product

$$D^{SEQ}(\mathbf{a}|\mathbf{x},\lambda) = \prod_{k=1}^{n} D(a_k|x_1,\ldots,x_k,\lambda) \tag{35}$$

and where the expression $D(a_k|x_1,\ldots,x_k,\lambda)$ corresponds to outputting deterministically $a_k = \lambda(x_1,\ldots,x_k)$ depending on the strategy given by $\lambda(.)$ and on all the inputs of previous boxes in the sequence (and similarly for Bob's strategy).

Determining whether a given distribution admits a decomposition in such a form is an instance of linear programming. Indeed, it implies checking if the distribution can be written as a convex combination of a finite amount of extremal points, represented by all the possible choices of deterministic strategies $D^{SEQ}(\mathbf{a}|\mathbf{x},\lambda), D^{SEQ}(\mathbf{b}|\mathbf{y},\lambda)$. This linear program quickly becomes computationally intractable, since the number of extremal points increases exponentially with the number of inputs. Moreover, for each additional box in the sequence, the scaling is even worse than the equivalent multipartite locality scenario, because the possible strategies for each box depend on the inputs of all the previous boxes.

That is why we are interested in relaxing the linear program with an SDP, in a similar spirit as in [22]. In particular, the objective is to have a way of determining whether a distribution is sequentially local that, despite being a relaxation, works in many relevant cases and has a better scaling with the number of inputs/boxes. In the following we show how to do this by adapting our sequential hierarchy. The first step is to find a particular realisation of sequentially local correlations in terms of a quantum measurement on a quantum state; namely we look for realisation of the kind

$$p(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = \mathrm{tr}(\rho_{AB} \mathsf{A}^{\mathbf{x}}_{\mathbf{a}} \otimes \mathsf{B}^{\mathbf{y}}_{\mathbf{b}})\,. \tag{36}$$

Now, it can be easily checked that correlations of the kind (34) can be reproduced by the following choice of state

$$\rho_{AB} = \sum_{\lambda} p(\lambda) |\lambda\rangle\langle\lambda|^{\otimes 2} \tag{37}$$

and measurements for Alice and Bob's side respectively

$$\begin{aligned} \mathsf{A}^{\mathbf{x}}_{\mathbf{a}} &= \sum_{\lambda} D^{SEQ}(\mathbf{a}|\mathbf{x},\lambda)\, |\lambda\rangle\langle\lambda|\,, \\ B^{\mathbf{y}}_{\mathbf{b}} &= \sum_{\lambda} D^{SEQ}(\mathbf{b}|\mathbf{y},\lambda)\, |\lambda\rangle\langle\lambda|\,. \end{aligned} \tag{38}$$

It is also easy to verify that measurements in the above form satisfy the constraints (10) and (11). In particular, the second property follows directly from the fact that the deterministic strategies $D^{SEQ}(\mathbf{a}|\mathbf{x},\lambda)$ and $D^{SEQ}(\mathbf{b}|\mathbf{y},\lambda)$ satisfy the no-signalling condition (33). Moreover, since all measurement operators are diagonal in the $|\lambda\rangle$ basis it follows that $[\mathsf{A}^{\mathbf{x}}_{\mathbf{a}}, \mathsf{A}^{\mathbf{x}'}_{\mathbf{a}'}] = 0$ and $[\mathsf{B}^{\mathbf{y}}_{\mathbf{b}}, \mathsf{B}^{\mathbf{y}'}_{\mathbf{b}'}] = 0$.

In other words, the set of time ordered local correlations can be obtained by means of locally commuting quantum sequential measurements. These commutativity conditions imply additional linear constraints on the moment matrix elements, expressed by some fixed matrices G_i^{LOC}. We can thus define the following hierarchy

Hierarchy for sequential local correlations (level k)

$$\begin{aligned} \text{Find } \Gamma_k \quad \text{such that} \quad & \Gamma_k \succcurlyeq 0,\ \Gamma_k^\dagger = \Gamma_k,\ \Gamma_k^{(0,0)} = 1, \\ & \mathrm{tr}[\Gamma_k G_i] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k G_i^{SEQ}] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k G_i^{LOC}] = 0\ \forall i, \\ & \mathrm{tr}[\Gamma_k F_i] = P_i\ \forall i. \end{aligned} \tag{39}$$

We call $\mathcal{L}^k_{\text{SEQ}}$ the set defined at level k of this hierarchy. By construction, each $\mathcal{L}^k_{\text{SEQ}}$ defines an outer approximation of the set of time ordered local correlations. The computational advantage gained by replacing a linear programming characterisation of the exact set with an SDP relaxation is clear: at each fixed level k, the number of variables involved in the moment matrix Γ_k scales polynomially with the number of input choices for $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$, contrarily to the exponential scaling of the linear programming. This may allow one to probe scenarios which would otherwise be practically impossible using linear programming methods.

D Using the hierarchy to upper bound guessing probabilities

We first review how the standard NPA hierarchy can be used to provide upper bounds to the guessing probability (17) (see also [26, 27, 29]). Define the subnormalised distributions (both normalised to $p_{\text{E}}(e)$)

$$\begin{aligned} \tilde{p}^e_{\text{AB}}(a,b|x,y) &= p_{\text{ABE}}(a,b,e|x,y) = p_{\text{E}}(e)p_{\text{AB}}(a,b|x,y,e), \\ \tilde{p}^e_{\text{B}}(b|y) &= p_{\text{BE}}(b,e|y) = p_{\text{E}}(e)p_{\text{B}}(b|y,e) = \sum_a \tilde{p}^e_{\text{AB}}(a,b|x,y). \end{aligned}$$

Define $\mathcal{Q}^p$ to be the set of quantum correlations, subnormalised to p, that is, $P(a,b|x,y) \in \mathcal{Q}^p$ if $P(a,b|x,y) = p\,P'(a,b|x,y)$ for some $P' \in \mathcal{Q}$. Thus $\tilde{p}^e_{\text{AB}} \in \mathcal{Q}^{p(e)}$ and with this notation (17) reads

$$\begin{aligned} G(y^*) = \max_{\tilde{p}^e_{\text{AB}}} \sum_{e=1}^{|b|} \tilde{p}^e_{\text{B}}(e|y^*), \qquad & \sum_e \tilde{p}^e_{\text{AB}}(a,b|x,y) = P_{\text{obs}}(a,b|x,y) \\ & \tilde{p}^e_{\text{AB}}(a,b|x,y) \in \mathcal{Q}^{p(e)}\ \forall e. \end{aligned} \tag{40}$$

Define the set $\mathcal{Q}_k^{p(e)}$ in the same way as $\mathcal{Q}_k$ but changing the normalisation condition $\Gamma_k^{(0,0)} = 1$ to $\Gamma_k^{(0,0)} = p(e)$ in (8). Thus $\mathcal{Q}_k^{p(e)} \supseteq \mathcal{Q}^{p(e)}$. The problem (40) can therefore be upper bounded by relaxing the condition $\tilde{p}^e_{\text{AB}}(a,b|x,y) \in \mathcal{Q}^{p(e)}$ to $\tilde{p}^e_{\text{AB}}(a,b|x,y) \in \mathcal{Q}_k^{p(e)}$. Practically, this means that one has to consider a set of $|b|$ subnormalised moment matrices in the optimisation.

An analogous procedure can be followed in the sequential scenario by defining the set $\mathcal{Q}^p_{\text{SEQ}}$, the set of subnormalised sequential correlations, and corresponding relaxations $\mathcal{Q}^p_{\text{SEQ},k}$. Following the same logic as above, one arrives at the guessing probability

$$\begin{aligned} G(\mathbf{y}^*) = \max_{\tilde{p}^{\mathbf{e}}_{\text{AB}}} \sum_{\mathbf{e}} \tilde{p}^{\mathbf{e}}_{\text{B}}(\mathbf{e}|\mathbf{y}^*), \qquad & \sum_{\mathbf{e}} \tilde{p}^{\mathbf{e}}_{\text{AB}}(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) = P_{\text{obs}}(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \\ & \tilde{p}^{\mathbf{e}}_{\text{AB}}(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \in \mathcal{Q}^{p(\mathbf{e})}\ \forall \mathbf{e}. \end{aligned} \tag{41}$$

One then relaxes the condition $\tilde{p}^{\mathbf{e}}_{\text{AB}}(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \in \mathcal{Q}^p_{\text{SEQ}}$ to $\tilde{p}^{\mathbf{e}}_{\text{AB}}(\mathbf{a},\mathbf{b}|\mathbf{x},\mathbf{y}) \in \mathcal{Q}^p_{\text{SEQ},k}$, thus needing as many moment matrices as the total alphabet size of $\mathbf{b}$