Naturally restricted subsets of nonsignaling correlations: typicality and convergence

Pei-Sheng Lin[1], Tamás Vértesi[2], and Yeong-Cherng Liang[1,3]

[1]Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan

[2]MTA Atomki Lendület Quantum Correlations Research Group, Institute for Nuclear Research, P.O. Box 51, H-4001 Debrecen, Hungary

[3]Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan

It is well-known that in a Bell experiment, the observed correlation between measurement outcomes—as predicted by quantum theory—can be stronger than that allowed by local causality, yet not fully constrained by the principle of relativistic causality. In practice, the characterization of the set $\mathcal{Q}$ of quantum correlations is carried out, often, through a converging hierarchy of *outer* approximations. On the other hand, some subsets of $\mathcal{Q}$ arising from additional constraints [e.g., originating from quantum states having positive-partial-transposition (PPT) or being finite-dimensional maximally entangled (MES)] turn out to be also amenable to similar numerical characterizations. How, then, at a *quantitative* level, are all these naturally restricted subsets of nonsignaling correlations different? Here, we consider several bipartite Bell scenarios and numerically estimate their volume relative to that of the set of nonsignaling correlations. Within the number of cases investigated, we have observed that (1) for a given number of inputs n_s (outputs n_o), the relative volume of both the Bell-local set and the quantum set increases (decreases) rapidly with increasing n_o (n_s) (2) although the so-called macroscopically local set $\mathcal{Q}_1$ may approximate $\mathcal{Q}$ well in the two-input scenarios, it can be a very poor approximation of the quantum set when $n_s > n_o$ (3) the almost-quantum set $\widetilde{\mathcal{Q}}_1$ is an exceptionally-good approximation to the quantum set (4) the difference between $\mathcal{Q}$ and the set of correlations originating from MES is most significant when $n_o = 2$, whereas (5) the difference between the Bell-local set and the PPT set generally becomes more significant with increasing n_o. This last comparison, in particular, allows us to identify Bell scenarios where there is little hope of realizing the Bell violation by PPT states and those that deserve further exploration.

1 Introduction

The fact that Bell inequalities [11]—constraints derived from the assumption of Bell-locality [17]—can be violated by quantum theory indicates that the set of quantum correlations $\mathcal{Q}$ is intrinsically different from the set of correlations $\mathcal{L}$ allowed by a locally-causal theory [12]. However, it is also known—from the pioneering work of Popescu and Rohrlich [56]—that quantum theory is *not* the most Bell-nonlocal (hereafter abbreviated as nonlocal) among all physical theories that respect the principle of relativistic causality. In bipartite Bell scenarios, this principle gives rise to the so-called nonsignaling (NS) conditions [10], and hence a superset of $\mathcal{Q}$ known as the nonsignaling polytope $\mathcal{NS}$.

Since then, a lot of effort (see, e.g., [16, 32, 34, 42, 46, 49, 50]) has been devoted to understand if *additional* physical, or information-theoretic principles can be supplemented to recover from $\mathcal{NS}$ the set of quantum correlations. In fact, the so-called *almost-quantum* set of correlations [49]—known to be a strict superset to $\mathcal{Q}$—apparently satisfies all information-theoretic principles proposed to

Yeong-Cherng Liang: ycliang@mail.ncku.edu.tw

date. Nonetheless, the extent to which this set, hereafter denoted by $\widetilde{\mathcal{Q}}_1$, differs from $\mathcal{Q}$ itself is not well understood. Indeed, [63] seems to be the only work to date reporting a systematic investigation of the difference in Bell values achievable by these sets.

So far, the only known means that we have in characterizing $\mathcal{Q}$ is via a hierarchy of outer approximations, such as that proposed by Navascués, Pironio, and Acín (NPA) [47]. The NPA hierarchy is known to converge [48] (see also [29])—in the asymptotic limit—to the set of correlations $\mathcal{Q}_{[\cdot,\cdot]}$ achievable assuming quantum theory and with the measurements between spatially separated parties modeled by commuting operators, rather than tensor products. Note that the two different formulations of spatially separated measurements generally lead to different [25, 61, 62] sets of correlations, i.e., $\mathcal{Q} \neq \mathcal{Q}_{[\cdot,\cdot]}$, thereby manifesting the complication of the geometry of these sets [33]. Moreover, very little is known [36] regarding the rate of convergence of this hierarchy towards $\mathcal{Q}_{[\cdot,\cdot]}$.

On the other hand, a few other *subsets* of $\mathcal{Q}$ are naturally also of interest. For example, a somewhat different formulation [45] of the NPA hierarchy has made it possible to characterize— also via a converging hierarchy of outer approximations—the set of correlations $\mathcal{P}$ arising from quantum states having positive partial transposition [52] (PPT). The interest in this stems from a conjecture of Peres [53]—disproved in [65]—concerning the impossibility of bound entangled [37] states violating Bell inequalities. Known counterexamples to Peres' conjecture are, however, too fragile to be demonstrated in any experiment, thus making it desirable to understand how $\mathcal{P}$, being a *restricted* subset of $\mathcal{Q}$, differs from the set of Bell-local correlations $\mathcal{L}$.

Besides, the fact that certain nonlocal features only seem to exist for partially entangled states is also intriguing. One of the first hints along this line is the Hardy paradox [35]. Later, the existence of such correlations was explicitly shown, independently, in [38, 40, 66] (see also [24]), under the name of *more nonlocality with less entanglement*. Interestingly, the set of correlations $\mathcal{M}$ arising from finite-dimensional maximally entangled states, or more precisely its convex hull— as with $\mathcal{Q}$ and $\mathcal{P}$—can also be characterized [39] via a hierarchy of outer approximations, each corresponds to the feasible set of a semidefinite program [15]. To achieve a better understanding of the precise relationship between entanglement and nonlocality, any quantitative estimate of the difference between $\mathcal{M}$ and $\mathcal{Q}$ is surely welcome.

Apart from fundamental interests, nonlocal correlations also play an indispensable role in the context of device-independent quantum information [17, 58]. For instance, from the observation of a Bell violation itself, one can certify the generation of unpredictable random bits [26, 54], guarantee the sharing of truly unconditional secured keys [3], certify various desired features of the underlying systems (see, e.g., [4, 6, 20, 21, 41, 45]), measurements (see, e.g., [7, 20, 23, 57]) or even other more general types of operations [59, 68]. In its strongest form, one could achieve so-called self-testing [44], where the underlying system and measurements are identified uniquely, modulo unimportant local degrees of freedom. For a comprehensive review on this last topic, see [67].

To this end, it is worth noting that the quantitative differences between the "size" of the Bell-local set $\mathcal{L}$, the quantum set $\mathcal{Q}$, and the nonsignaling set $\mathcal{NS}$ has been investigated in [19, 30, 75]. Specifically, in the simplest Bell scenario involving two parties, each performing two binary-outcome measurements, the volume of $\mathcal{L}$ and that of $\mathcal{Q}$, relative to $\mathcal{NS}$, in the *subspace* of "full" correlation functions [70] was first determined in [19]. Then, for the same Bell scenario, the analysis has been generalized [75] to include also the subspace spanned by marginal correlations. Beyond this, numerical estimation of the relative volume of $\mathcal{L}$ to $\mathcal{NS}$ was carried out in [30] for a few Bell scenarios with either *only* two measurement settings or outcomes; some analytic results were also presented therein when restricted to the subspace of full correlations.

Here, we generalize the analysis of [19, 30, 75] in two directions. Firstly, we consider a more extensive list of Bell scenarios, including a few with multiple measurement settings and outcomes, which allows us to make observations that were not possible in prior works. Secondly, we consider not only $\mathcal{L}$, $\mathcal{Q}$ (approximated by relevant outer approximations), and $\mathcal{NS}$, but also the set of correlations $\mathcal{P}$ achievable by PPT quantum states, the (convex hull of the) set of correlations achievable by locally measuring finite-dimensional maximally entangled states with projective measurements $\mathcal{M}^{\mathrm{P}}$, the set $\mathcal{Q}_1$ associated with the principle of *macroscopic locality* [46], the almost-quantum set $\widetilde{\mathcal{Q}}_1$, and more generally the first few levels of the outer approximations of $\mathcal{Q}$ given by the NPA hierarchy [47] as well as the hierarchy of Moroder *et al.* [45]. This last consideration, in particular, allows us to learn the fraction of points in $\mathcal{NS}$ that can be further excluded from $\mathcal{Q}$ as we consider

increasingly tighter outer approximations to $\mathcal{Q}$. Clearly, such information is highly relevant for device-independent analyses that rely on the above hierarchies to outer-approximate $\mathcal{Q}$.

The rest of this paper is structured as follows. In Section 2, we introduce the notations used throughout the paper and the different Bell scenarios considered. Then, we present in Section 3 our main results on the relative volume for the different sets of correlations mentioned above. In Section 4, we give further discussions and comment on some possible directions for future research.

2 Preliminaries

2.1 Notations and Naturally Restricted Subsets of $\mathcal{NS}$

Consider a bipartite Bell experiment where each spatially separated party has a choice over n_s measurement settings where each measurement results in n_o possible outcomes. We shall denote such a Bell scenario by (n_s, n_o). Correlation between the observed outcomes for given measurement settings of the two parties (conventionally called Alice and Bob) may be described by $\vec{P} = \{P(a, b|x, y)\}_{a,b,x,y}$ where we label the measurement settings and outcomes for Alice (Bob), respectively, as x (y) and a (b). Throughout, we consider only Bell scenarios where x, y, a, and b take a finite number of values.

Our starting point is the set of nonsignaling correlations [10, 56], $\mathcal{NS}$, which are all those $\vec{P}$ that satisfy the so-called nonsignaling conditions [10]:

$$\begin{aligned} \sum_a P(a, b|x, y) &\overset{\mathcal{NS}}{=} \sum_a P(a, b|x', y) \ \ \forall\, b, y, x, x', \\ \sum_b P(a, b|x, y) &\overset{\mathcal{NS}}{=} \sum_b P(a, b|x, y') \ \ \forall\, a, x, y, y'. \end{aligned} \tag{1}$$

Physically, these conditions were initially [56] proposed to exclude the possibility of communication by making different local choices of x, y. Since $\mathcal{NS}$ is the intersection of a finite number of hyperplanes defined by Eq. (1) and the direct sum of n_s^2 probability simplices, it is a convex polytope.

An important subset of $\mathcal{NS}$ is the Bell-local [17] set $\mathcal{L}$. For any $\vec{P} \in \mathcal{L}$, it can be shown [31, 55] (see also Ref. [73]) that there exists normalized weights $q(\lambda) \geq 0$ for all λ and $\sum_\lambda q(\lambda) = 1$ such that

$$P(a, b|x, y) \overset{\mathcal{L}}{=} \sum_\lambda q(\lambda) \delta_{a, f_A(x, \lambda)} \delta_{b, f_B(y, \lambda)}, \tag{2}$$

for some choice of local response functions $f_A(x, \lambda)$ and $f_B(y, \lambda)$. Clearly, $\mathcal{L}$ is convex. Since there are only finite possibilities of x, y, a, and b, the set $\mathcal{L}$ forms [55] a convex polytope, with its extreme points correspond to local deterministic strategies given by the Kronecker deltas: $\delta_{a, f_A(x, \lambda)}$ and $\delta_{b, f_B(y, \lambda)}$.

Suppose now that the two parties share a quantum state ρ, then quantum theory dictates that the observed correlation follows Born's rule:

$$P(a, b|x, y) \overset{\mathcal{Q}}{=} \text{tr} \left[\rho\, M_{a|x}^{(A)} \otimes M_{b|y}^{(B)} \right], \tag{3}$$

where $\{M_{a|x}^{(A)}\}_a$ $\left(\{M_{b|y}^{(B)}\}_b \right)$ are positive-operator-valued measures (POVMs) associated with Alice's (Bob's) x-th (y-th) measurement. We denote the set of such correlations by $\mathcal{Q}$. Importantly, in the definition of $\mathcal{Q}$, there is no constraint imposed on the (local) Hilbert space dimension. This, in turn, guarantees the convexity [55] of $\mathcal{Q}$ and the sufficiency of projective measurements (via Naimark's extension [51]) on pure state $\rho = \rho^2$ in the membership test of $\mathcal{Q}$. Even then, the characterization of $\mathcal{Q}$ is by no means computationally easy.

To this end, NPA [47] first pointed out that $\mathcal{Q}$, or more precisely, $\mathcal{Q}_{[\cdot, \cdot]}$ can be characterized [48] *asymptotically* (see also [29]) by solving a hierarchy of semidefinite programs (SDPs). The feasible sets corresponding to these SDPs then define a series of outer approximations $\mathcal{Q}_1 \supseteq \mathcal{Q}_2 \supseteq \cdots \supseteq \mathcal{Q}_\infty = \mathcal{Q}_{[\cdot, \cdot]} \supseteq \mathcal{Q}$, where $\mathcal{Q}_k$ is referred conventionally as NPA level k with k signifying the highest
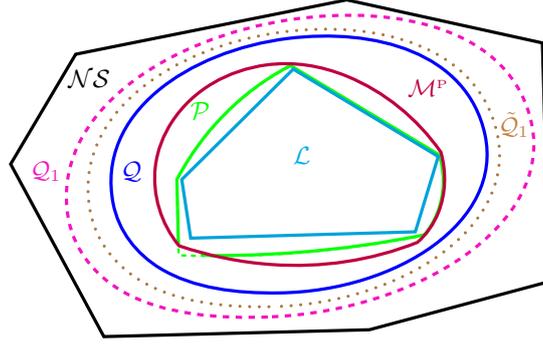
Figure 1: Schematic illustration of the relationships among the various subsets $\mathcal{T} \subset \mathcal{NS}$ considered here. Starting from $\mathcal{NS}$ and moving inwards, we have in solid lines, respectively, the boundary of the quantum set of correlations $\mathcal{Q}$ [Eq. (3), blue], the convex hull $\mathcal{M}^\text{P}$ (red) of the set of correlations attainable by finite-dimensional maximally entangled states in conjunction with projective measurements, the set of correlations attainable by PPT entangled state $\mathcal{P}$ (green), and the Bell-local polytope $\mathcal{L}$ (skyblue). Dashed (pink) and dotted (brown) lines lying between the boundary of $\mathcal{Q}$ and that of $\mathcal{NS}$ mark the boundary of the lowest-level outer approximation of $\mathcal{Q}$, respectively, due to NPA [47] ($\mathcal{Q}_1$) and Moroder *et al.* [45] ($\widetilde{\mathcal{Q}}_1$).

degree of the operator used in defining some matrix of moments (see [47, 48] for details). $\mathcal{Q}_1$, incidentally, is exactly the set of correlations that respects the principle of macroscopic locality [46]. Intermediate levels can also be considered and a prominent example is the so-called NPA level $1 + AB$, which happens to be the lowest level of another converging hierarchy of outer approximations $\widetilde{\mathcal{Q}}_1 \supseteq \widetilde{\mathcal{Q}}_2 \supseteq \cdots \supseteq \widetilde{\mathcal{Q}}_\infty = \mathcal{Q}_{[\cdot,\cdot]} \supset \mathcal{Q}$ due to Moroder *et al.* [45]. The specific outer approximation $\mathcal{Q}_{1+AB} = \widetilde{\mathcal{Q}}_1$ is known in the literature as the *almost quantum* [49] set of correlations, as it seems to satisfy all principles that have been proposed to date to distinguish $\mathcal{Q}$ from $\mathcal{NS}$. In the bipartite scenario, the sets associated with the two hierarchies are known [45] to satisfy the inclusion relations $\mathcal{Q}_1 \supset \widetilde{\mathcal{Q}}_1$ and $\widetilde{\mathcal{Q}}_k \supset \mathcal{Q}_{2k}$ for all integers $k \geq 1$.

As Bell showed in his seminal work [11], there exist quantum correlations arising from entangled quantum states that do not admit a convex decomposition in the form of Eq. (2). At the same time, by considering trivial POVMs consisting only of the identity operator and the null operator, it is straightforward to see that all $\vec{P} \in \mathcal{L}$ can always be cast in the form of Eq. (3). Similarly, all $\vec{P} \in \mathcal{Q}$ are easily seen to satisfy Eq. (1), while Popescu and Rohrlich [56] showed that there exists $\vec{P} \in \mathcal{NS} \setminus \mathcal{Q}$. Together, one arrives at the strict inclusion relations $\mathcal{L} \subsetneq \mathcal{Q} \subsetneq \widetilde{\mathcal{Q}}_1 \subsetneq \mathcal{Q}_1 \subsetneq \mathcal{NS}$.

Even if we restrict our attention to $\mathcal{Q}$ itself, its relationship with the set of quantum states has not been fully understood. For example, although entanglement is necessary [71] for Bell-nonlocality, some entangled states (see, e.g., [9, 71]) are known to produce *only* $\vec{P} \in \mathcal{L}$. In fact, even if Alice and Bob are allowed to share an arbitrary finite-dimensional maximally entangled state $|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle|i\rangle$, it is *impossible* for them to reproduce all $\vec{P} \in \mathcal{Q}$ [24, 38, 40, 66]. To facilitate subsequent discussions, we denote by $\mathcal{M}^\text{P}$ the convex hull of the set of correlations attainable by performing projective measurements on $|\Psi_d\rangle$, with d finite. On the contrary, even the weakest form of entanglement given by PPT entangled states may generate nonlocal correlations [64, 65]. Hereafter, we denote by $\mathcal{P}$ the subset of $\mathcal{Q}$ which arises from ρ being a PPT state, i.e., $\rho^{\text{T}_A} \succeq 0$ where $^{\text{T}_A}$ stands for the partial transposition operation on Alice's Hilbert space.

A schematic diagram explaining the relationships among the various naturally restricted subsets $\mathcal{T}$ of $\mathcal{NS}$ considered in this work is provided in Fig. 1.

2.2 Membership Tests

Our goal is to estimate the relative volume of various subsets $\mathcal{T} \in \{\mathcal{L}, \mathcal{P}, \mathcal{M}^\text{P}, \mathcal{Q}, \widetilde{\mathcal{Q}}_1, \mathcal{Q}_1\}$ of $\mathcal{NS}$. To this end, we perform the membership test $\vec{P} \overset{?}{\in} \mathcal{T}$ for each sampled $\vec{P} \in \mathcal{NS}$ by solving the

following optimization problem:

$$\sup \ v \tag{4a}$$

$$\text{s.t. } v\vec{P} + (1-v)\vec{P}_w \in \mathcal{T}, \tag{4b}$$

where $\vec{P}_w$ is the uniform probability distribution, i.e., $P_w(a,b|x,y) = 1/n_o^2 \ \forall \ x,y$. As $\vec{P}_w$ lies strictly in $\mathcal{L}$, it must also lie in *all* sets $\mathcal{T}$ that are of our interest. Hence, the above optimization problem, which we solve using the optimization software MOSEK implemented in MATLAB, is always feasible by setting $v = 0$. Also, if $\vec{P} \in \mathcal{T}$ then all mixtures with $v \in [0,1]$ are inside $\mathcal{T}$, i.e., the optimum v (denoted by v^*) would be greater than or equal to 1. Hence, $v^* < 1$ indicates that $\vec{P} \notin \mathcal{T}$. Notice that v^*, often called the white-noise visibility, can be understood as the "maximal" weight that can be assigned to $\vec{P}$ when it is admixed with white noise while ensuring that the mixture lies within $\mathcal{T}$. A smaller value of v^*, which corresponds to a larger value of $1 - v^*$, then indicates that the correlation is more robust (in terms of preserving its nonlocal nature) against the mixing with $\vec{P}_w$.

Among the different sets of interest, $\mathcal{L}$ is a convex polytope, and thus its membership test, cf. Eq. (4) is an instance of a linear program [15]. For relatively simple Bell scenarios, this optimization problem can be efficiently solved on a computer. In contrast, for the other sets of interest, including $\mathcal{T} \in \{\mathcal{P}, \mathcal{M}^{\mathrm{P}}, \mathcal{Q}\}$, we rely on a hierarchy of outer approximations, each of which is amenable to semidefinite programming characterizations. In the case of $\mathcal{Q}$, we use both the NPA hierarchy [47, 48] and its variant due to Moroder *et al.* [45] for membership tests. For definiteness, we denote by $\mathcal{Q}_k$ and $\widetilde{\mathcal{Q}}_\ell$, respectively, the level k and the level ℓ outer approximation of $\mathcal{Q}$ based on the NPA hierarchy and the hierarchy of Moroder *et al.* (a summary of both hierarchies can be found in Table V, Appendix B of Ref. [22]). By further requiring the moment matrix $\widetilde{\mathcal{Q}}_\ell$ to be PPT, one immediately obtains a characterization of $\mathcal{P}_\ell$, i.e., the level ℓ outer approximation of $\mathcal{P}$. Notice that all these SDPs can be implemented using the Ncpol2sdpa toolbox developed by Wittek [74]. For the set of correlations $\mathcal{M}^{\mathrm{P}}$ associated with local projective measurements on finite-dimensional maximally entangled states, we make use of a hierarchy adapted from that presented in Ref. [39], the details of which are given in Appendix A. We refer to the level h outer approximation of $\mathcal{M}^{\mathrm{P}}$ obtained thereof as $\mathcal{M}_h^{\mathrm{P}}$.

2.3 Metrics and Relative Volume

The notion of volume for any given region in a space $\mathscr{P}$ is metric-dependent. In our case, $\mathscr{P}$ is the set of conditional probability distributions $\vec{P} = \{P(a,b|x,y)\}_{a,b,x,y}$ where $a,b \in \{1,\dots,n_o\}$ and $x,y \in \{1,2,\dots,n_s\}$. The normalization requirement $\sum_{a,b} P(a,b|x,y) = 1$ for all x and y implies that $\mathscr{P}$ is $(n_s^2)(n_o^2 - 1)$-dimensional. Moreover, we are only interested in $\vec{P}$ that satisfy the nonsignaling constraints of Eq. (1). The nonsignaling polytope $\mathcal{NS}$ and hence the various subsets of interest all lie in a d-dimensional subspace $\mathscr{P}_{\mathcal{NS}}$ of $\mathscr{P}$ where [27] $d = (n_o - 1)^2 n_s^2 + 2n_s(n_o - 1)$.

A convenient, *minimal* parametrization of any $\vec{P} \in \mathcal{NS}$ is given by [27]:

$$\vec{P} = \{P(a|x), P(b|y), P(a,b|x,y)\}_{a,b,x,y} \tag{5}$$

where $P(a|x) = \sum_b P(a,b|x,y)$ and $P(b|y) = \sum_a P(a,b|x,y)$ are, respectively, the marginal conditional probability distributions of Alice and Bob. Note that in this parameterization, the labels a,b in Eq. (5) take only values from $\{1,\dots,n_o - 1\}$. Indeed, the conditional probability distributions for the omitted outcome corresponding to a and/or $b = n_o$ can be determined easily from the components P_i of $\vec{P}$ in Eq. (5) via the normalization of probabilities and the nonsignaling conditions of Eq. (1).

Like in previous works [19, 30, 75], we adopt the Euclidean metric in our computation of the relative volumes (though other options may also be considered, see our remark in Section 4). In this metric $\mathrm{d}s_{\mathrm{E}}^2$, all components of $\vec{P}$ in Eq. (5) are treated on equal footing. Explicitly, $\mathrm{d}s_{\mathrm{E}}^2$ and the corresponding volume element $\mathrm{d}V_{\mathrm{E}}$ are given, respectively, by:

$$\mathrm{d}s_{\mathrm{E}}^2 = \sum_i \mathrm{d}P_i^2 \quad \text{and} \quad \mathrm{d}V_{\mathrm{E}} = \prod_i \mathrm{d}P_i. \tag{6}$$

When there is no risk of confusion, the subscript $_{\mathrm{E}}$ is omitted to simplify the presentation.

We then define the relative volume (RV) for each set $\mathcal{T} \subset \mathcal{NS}$ as:

$$\mathrm{RV}(\mathcal{T}) \equiv \frac{V(\mathcal{T})}{V(\mathcal{NS})} \tag{7}$$

where $V(\mathcal{S})$ is the volume of a set $\mathcal{S}$ in accordance to the (Euclidean) metric. To numerically estimate these RVs, it suffices to sample points $\vec{P} \in \mathcal{NS}$ *uniformly* according to the metric, and determine the fraction of such points that lie in $\mathcal{T}$ via the method explained in Section 2.2. For this purpose, we make use of the MATLAB function cprnd developed by Benham [13], and in particular its *Gibbs sampler* algorithm to perform uniform sampling of $\vec{P}$ in $\mathcal{NS}$. In Appendix B, we give further details on how we generate uniform samples in $\mathcal{NS}$ using the cprnd function. Note that each membership test with respect to $\mathcal{T}$ is exactly a Bernoulli trial with a success probability given by $\mathrm{RV}(\mathcal{T})$.

For each Bell scenario considered, we estimate the RV of $\mathcal{L}$, $\mathcal{Q}_l$, $\widetilde{\mathcal{Q}}_k$, $\mathcal{P}_k$ and $\mathcal{M}_h^{\mathrm{P}}$. In Table 1, we list all the bipartite Bell scenarios considered in this work, the number of samples used in estimating the RV of each target set, the highest level of each type of hierarchies considered, as well as the corresponding RVs. In Appendix C, we further provide the relevant parameters characterizing the size of the optimization problem. In the next section, we present our main results across 23 Bell scenarios.

For all but two of these Bell scenarios, we use $N_{\mathrm{tot}} = 10^6$ sampled correlations in our estimation of the relative volumes. The only exceptions are the (5,4) and the (6,3) Bell scenario where we employ, instead, 3.05×10^5 and 7.85×10^5 samples, respectively. Even in these latter cases, the estimates have converged well with the employed samples (see Fig. 8 in Appendix D for details). Importantly, as we can see in these plots, the number of samples required to reach a good precision does *not* seem to depend on the complexity of the Bell scenario, but rather more on the success probability itself.

3 Numerical Estimates of Relative Volumes

In Table 1, we provide a summary of $\mathrm{RV}(\mathcal{T})$ for $\mathcal{T} \in \{\mathcal{Q}_1, \widetilde{\mathcal{Q}}_1, \mathcal{L}\}$ and certain approximations to $\mathcal{Q}$, $\mathcal{P}$, and $\mathcal{M}^{\mathrm{P}}$ for the various Bell scenarios considered. Note that for 6 of the 23 Bell scenarios, we compute only $\mathrm{RV}(\mathcal{L})$ as it becomes too time consuming to compute the other $\mathrm{RV}(\mathcal{T})$ with a statistically significant number of trials. Throughout, we use $\mathcal{Q}_{**}$, $\mathcal{M}_*^{\mathrm{P}}$, and $\mathcal{P}_*$, respectively, to denote the tightest approximation that we are able to compute for $\mathcal{Q}$, $\mathcal{M}^{\mathrm{P}}$, and $\mathcal{P}$. In the following subsections, we describe in details how the RV of these sets changes in different Bell scenarios. To best illustrate these trends, we make use of line plots showing how each of these relative volumes varies with respect to the relevant parameters.

3.1 $\mathcal{L}$ vs $\mathcal{NS}$

Quantitative estimate of $\mathrm{RV}(\mathcal{L})$ in the 8-dimensional space of $\mathscr{P}_{\mathcal{NS}}$ for the (2,2) Bell scenario was first determined in Ref. [75]. This analysis was then generalized in Ref. [30] to include the (3,2), (4,2), (5,2), (2,3), and the (2,4) Bell scenarios. Among their findings is the observation that for $n_o = 2$, $\mathrm{RV}(\mathcal{L})$ rapidly decreases as n_s increases from 2 to 5. Our findings, as can be seen in Fig. 2, show that this trend holds also for Bell scenarios with $n_o = 3, 4$ (and possibly $n_o = 5$).

Observation 1. *For Bell scenarios with fixed outputs, $\mathrm{RV}(\mathcal{L})$ decreases monotonically with increasing n_s (see Fig. 2).*

Hence, Observation 1 generalizes the observation from [30] for $n_o = 2$. For example, in the $n_o = 3$ case, we observe that $\mathrm{RV}(\mathcal{L})$ decreases from 93.84% (for $n_s = 2$) to 0.01% (for $n_s = 6$), likewise for the $n_o = 4$ case, which decreases from 97.41% (for $n_s = 2$) to 6.31% (for $n_s = 5$), etc.

Before discussing this observed trend, note that our estimate for $\mathrm{RV}(\mathcal{L})$ in the (2,2) and the (3,2) Bell scenario is consistent with that obtained analytically from the software lrs [5], which gives $\mathrm{RV}(\mathcal{L}) = \frac{16}{17} \approx 94.12\%$ and $\mathrm{RV}(\mathcal{L}) = \frac{18\,176}{29\,205} \approx 62.24\%$ respectively. Similarly, our estimate of

(n_s, n_o)	N_{tot} $(\times 10^5)$	$\mathcal{Q}_1$	$\widetilde{\mathcal{Q}}_1$	$\mathcal{Q}_*$	$\widetilde{\mathcal{Q}}_*$	$\mathcal{P}_*$	$\mathcal{M}_*^{\mathsf{P}} \cap \mathcal{Q}_{**}$	$\mathcal{L}$
$(2,2)$	10	99.80%	99.25%	99.17% $(\mathcal{Q}_4)$	99.17% $(\widetilde{\mathcal{Q}}_4)$	94.15% $(\mathcal{P}_4)$	98.68% $(\mathcal{M}_4^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_4)$	94.15%
$(2,3)$	10	99.98%	99.80%	99.74% $(\mathcal{Q}_3)$	99.74% $(\widetilde{\mathcal{Q}}_2)$	93.84% $(\mathcal{P}_2)$	99.50% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	93.84%
$(2,4)$	10	100.00*%	99.99%	99.99% $(\mathcal{Q}_2)$	99.99% $(\widetilde{\mathcal{Q}}_1)$	97.41% $(\mathcal{P}_2)$	99.95% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	97.41%
$(2,5)$	10	100.00**%	100.00*%	100.00*% $(\mathcal{Q}_2)$	100.00*% $(\widetilde{\mathcal{Q}}_1)$	99.41% $(\mathcal{P}_{1.5})$	100.00*% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	99.41%
$(2,6)$	10	100.00**%	100.00**%	100.00** $(\mathcal{Q}_2)$	100.00**% $(\widetilde{\mathcal{Q}}_1)$	99.90% $(\mathcal{P}_{1.5})$	100.00**% $(\mathcal{M}_1^{\mathsf{P}} \cap \mathcal{Q}_2)$	99.90%
$(2,7)$	10	100.00**%	100.00**%	100.00** $(\mathcal{Q}_2)$	100.00**% $(\widetilde{\mathcal{Q}}_1)$	99.98% $(\mathcal{P}_1)$	100.00**% $(\mathcal{M}_1^{\mathsf{P}} \cap \mathcal{Q}_2)$	99.98%
$(2,8)$	10	-	-	-	-	-	-	100.00*%
$(2,9)$	10	-	-	-	-	-	-	100.00*%
$(3,2)$	10	97.23%	92.11%	91.19% $(\mathcal{Q}_3)$	91.19% $(\widetilde{\mathcal{Q}}_2)$	62.18% $(\mathcal{P}_2)$	88.13% $(\mathcal{M}_3^{\mathsf{P}} \cap \mathcal{Q}_2)$	62.18%
$(3,3)$	10	99.78%	97.90%	97.52% $(\mathcal{Q}_2)$	97.90% $(\widetilde{\mathcal{Q}}_1)$	59.88% $(\mathcal{P}_2)$	95.16% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	59.88%
$(3,4)$	10	100.00*%	99.93%	99.91% $(\mathcal{Q}_2)$	99.93% $(\widetilde{\mathcal{Q}}_1)$	81.27% $(\mathcal{P}_{1.25})$	99.65% $(\mathcal{M}_{1.5}^{\mathsf{P}} \cap \mathcal{Q}_2)$	79.79%
$(3,5)$	10	-	-	-	-	-	-	94.93%
$(3,6)$	10	-	-	-	-	-	-	99.16%
$(3,7)$	10	-	-	-	-	-	-	99.89%
$(4,2)$	10	84.62%	69.45%	67.80% $(\mathcal{Q}_2)$	69.45% $(\widetilde{\mathcal{Q}}_1)$	21.14% $(\mathcal{P}_2)$	61.75% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	21.14%
$(4,3)$	10	98.39%	88.69%	98.39% $(\mathcal{Q}_1)$	88.69% $(\widetilde{\mathcal{Q}}_1)$	20.22% $(\mathcal{P}_{1.2})$	88.69% $(\mathcal{M}_1^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_1)$	16.58%
$(4,4)$	10	99.99%	99.59%	99.99% $(\mathcal{Q}_1)$	99.59% $(\widetilde{\mathcal{Q}}_1)$	50.16% $(\mathcal{P}_1)$	99.59% $(\mathcal{M}_1^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_1)$	41.07%
$(4,5)$	10	-	-	-	-	-	-	80.34%
$(5,2)$	10	53.47%	34.49%	32.35% $(\mathcal{Q}_2)$	34.49% $(\widetilde{\mathcal{Q}}_1)$	3.09% $(\mathcal{P}_{1.5})$	27.12% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	2.96%
$(5,3)$	10	89.91%	59.29%	89.91% $(\mathcal{Q}_1)$	59.29% $(\widetilde{\mathcal{Q}}_1)$	3.16% $(\mathcal{P}_1)$	59.29% $(\mathcal{M}_1^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_1)$	1.22%
$(5,4)$	3.05	99.94%	97.03%	99.94% $(\mathcal{Q}_1)$	97.03% $(\widetilde{\mathcal{Q}}_1)$	18.28% $(\mathcal{P}_1)$	97.03% $(\mathcal{M}_1^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_1)$	6.31%
$(6,2)$	10	18.11%	8.80%	7.73% $(\mathcal{Q}_2)$	8.80% $(\widetilde{\mathcal{Q}}_1)$	0.22% $(\mathcal{P}_1)$	5.93% $(\mathcal{M}_2^{\mathsf{P}} \cap \mathcal{Q}_2)$	0.15%
$(6,3)$	7.85	56.67%	17.20%	56.67% $(\mathcal{Q}_1)$	17.20% $(\widetilde{\mathcal{Q}}_1)$	0.12% $(\mathcal{P}_1)$	17.20% $(\mathcal{M}_1^{\mathsf{P}} \cap \widetilde{\mathcal{Q}}_1)$	0.01%

Table 1: Summary of the numerically estimated relative volume $\mathrm{RV}(\mathcal{T})$ for various naturally restricted subsets $\mathcal{T}$ of the set of nonsignaling correlations $\mathcal{NS}$. The second column gives the number of $\vec{P}$ uniformly sampled from $\mathcal{NS}$ using cprnd. From the third column to the rightmost column, we have the estimated relative volume (RV) for, respectively, the macroscopically local set $\mathcal{Q}_1$, the almost quantum set $\widetilde{\mathcal{Q}}_1$, our tightest approximation to $\mathcal{Q}$ based on outer approximations of NPA [47] (denoted by $\mathcal{Q}_*$), our tightest approximation to $\mathcal{Q}$ based on outer approximations of Moroder *et al.* [45] (denoted by $\widetilde{\mathcal{Q}}_*$), our tightest approximation to $\mathcal{P}$ (denoted by $\mathcal{P}_*$), our tightest approximation to $\mathcal{M}^{\mathsf{P}}$ intersecting with $\mathcal{Q}_{**}$ (denoted by $\mathcal{M}_*^{\mathsf{P}} \cap \mathcal{Q}_{**}$), and the Bell-local set $\mathcal{L}$. In the fifth to the eighth column, we include also in bracket the highest level of the SDP hierarchy used in the computation (for an explanation of the various levels and the complexity involved in the computation, see Appendix C). In particular, our best approximation to $\mathrm{RV}(\mathcal{Q})$, given either in the fifth or the sixth column, is highlighted in yellow. For example, in our characterization of $\mathcal{Q}$ in the $(2,3)$ Bell scenario, we are not able to go beyond $\mathcal{Q}_3$ nor $\widetilde{\mathcal{Q}}_2$, neither of which is, *a priori*, a subset of the other. However, since $\widetilde{\mathcal{Q}}_2$ gives a smaller RV (see Table 2), we use it as our $\mathcal{Q}_{**}$ in this Bell scenario. We use 100.00*% and 100.00**% to denote entries where the *estimated* RV$(\mathcal{T})$ satisfies, respectively, RV$(\mathcal{T}) > 99.995\%$ and RV$(\mathcal{T}) > 100\% \left(1 - \frac{1}{N_{\text{tot}}}\right)$. Similarly, all other estimates reported here have a fundamental imprecision of $\frac{100}{N_{\text{tot}}}\%$. Here and below, entries marked with "-" means the corresponding computation has been left out.

RV$(\mathcal{L})$ in the $(2,3)$ Bell scenario (see Table 1) is consistent with that determined from the software vinci [18], giving RV$(\mathcal{L}) \approx 93.82\%$.

How do we understand the observed decreasing trend? Let us remind that for any Bell scenario (n'_s, n_o) with $n'_s > n_s$, *any sub-correlation* $\vec{P}$ extracted from $\vec{P}'$ by considering only n_s out of the n'_s measurement settings (for both Alice and Bob) is a legitimate correlation for the *simpler* Bell scenario (n_s, n_o). Moreover, for $\vec{P}'$ to be in $\mathcal{L}$, all these $\binom{n'_s}{n_s}^2$ sub-correlations $\vec{P}$ extractable from $\vec{P}'$ must also be Bell-local.

Let RV$(\mathcal{L}) = p$ be the success probability of a Bernoulli trial in the (n_s, n_o) Bell scenario. If all such $\vec{P}$ that may be extracted from $\vec{P}'$ could be thought of as being sampled *independently* and *uniformly* from the $\mathcal{NS}$ polytope in the (n_s, n_o) Bell scenario, the success probability of a Bernoulli trial in the (n'_s, n_o) Bell scenario would scale as

$$p^{\binom{n'_s}{n_s}^2} = p^{\frac{(n'_s!)^2}{(n_s!)^2[(n'_s - n_s)!]^2}}. \tag{8}$$
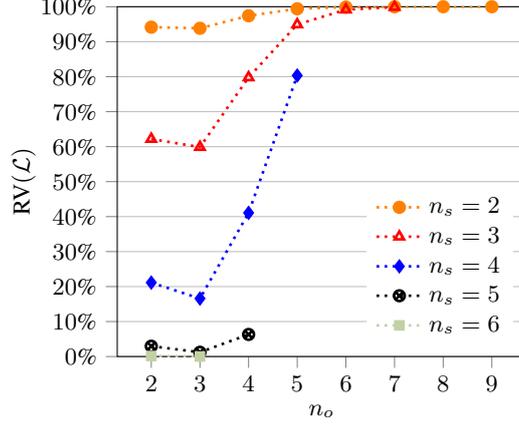
Figure 2: Plots of estimated $\mathrm{RV}(\mathcal{L})$ vs $n_o \in \{2, 3, \ldots, 9\}$ for $n_s \in \{2, 3, \ldots, 6\}$. $\mathrm{RV}(\mathcal{L})$ for scenarios $(2, 5), (2, 6), \ldots, (2, 9), (3, 6)$ and $(3, 7)$ is very close to 100% but not exactly 100%.

Applying this naïve reasoning to the $(2,2)$ and the $(3,2)$ Bell scenario would suggest a decrease of $\mathrm{RV}(\mathcal{L})$ from $\frac{16}{17} \approx 94.12\%$ to 57.95%, which is not too far off from our exact finding that $\mathrm{RV}(\mathcal{L}) = \frac{18}{29} \frac{176}{205} \approx 62.24\%$ in the $(3,2)$ case. Clearly, part of this discrepancy stems from the fact the sub-correlations $\vec{P}$ extractable from $\vec{P}'$ are *not* entirely independent from one another – all these different $\vec{P}$ share a common input with the other $\vec{P}$. Moreover, even if all these sub-correlations $\vec{P}$ are Bell-local, $\vec{P}'$ may still be Bell-nonlocal.

On the other hand, Fig. 2 shows an opposite trend for Bell scenarios with fixed n_s.

Observation 2. *For Bell scenarios with fixed inputs, $RV(\mathcal{L})$ first decreases when n_o varies from 2 to 3, but increases monotonically thereafter with increasing n_o (see Fig. 2).*

This observation generalizes the observation from [30] for $n_s = 2$. Since the way $\mathrm{RV}(\mathcal{L})$ changes with increasing n_o is opposite to that with increasing n_s, it is natural to wonder how $\mathrm{RV}(\mathcal{L})$ changes when $n_s = n_o = k$ increases. To this end, we have the following observation.

Observation 3. *As $n_s = n_o = k$ increases, $RV(\mathcal{L})$ decreases steadily with increasing k. That is, the effect of increasing n_s on $RV(\mathcal{L})$ dominates over that of increasing n_o.*

3.2 $\mathcal{Q}$ vs $\mathcal{NS}$ and $\mathcal{L}$

3.2.1 Convergence of outer approximations towards $\mathcal{Q}_{[\cdot,\cdot]}$

Before discussing how $\mathrm{RV}(\mathcal{Q})$ changes across different Bell scenarios, let us first make a digression to investigate how well the various $\mathcal{Q}_k$ and $\widetilde{\mathcal{Q}}_\ell$ outer-approximate $\mathcal{Q}$ in *each* Bell scenario. Again, their relative volume is a useful figure of merit in this context. From here, we can learn how $\mathrm{RV}(\mathcal{T})$ converges to $\mathcal{Q}$ when we consider approximations $\mathcal{T}$ of $\mathcal{Q}$ with increasing complexity. For definiteness, we make use of the *number of real moment variables* involved in the SDP characterization of $\mathcal{T}$ to serve as our measure of complexity.

Recall from our discussion in Section 3.1 that $\mathcal{L}$ makes up a substantial fraction of $\mathcal{NS}$ for many of the Bell scenarios considered. Thus, to better manifest the convergence graphically, we focus on the nonlocal region of $\mathcal{NS}$, i.e., $\mathcal{NS} \setminus \mathcal{L}$. In other words, for any given outer approximation $\mathcal{T} = \mathcal{Q}_k$ or $\widetilde{\mathcal{Q}}_\ell$, we are interested in the volume of $\mathcal{T} \setminus \mathcal{L}$ *relative* to that of $\mathcal{NS} \setminus \mathcal{L}$, i.e.,

$$f(\mathcal{T}) = \frac{\mathrm{RV}(\mathcal{T} \setminus \mathcal{L})}{\mathrm{RV}(\mathcal{NS} \setminus \mathcal{L})} = \frac{\mathrm{RV}(\mathcal{T} \setminus \mathcal{L})}{1 - \mathrm{RV}(\mathcal{L})}. \tag{9}$$

Evidently, for *each* given Bell scenario, there is some $\mathcal{Q}_k$ or $\widetilde{\mathcal{Q}}_\ell$ considered that gives the smallest $f(\mathcal{T})$. Denoting them, respectively, by $\mathcal{Q}_*$ and $\widetilde{\mathcal{Q}}_*$, then our tightest approximation to $\mathcal{Q}$ is simply $\mathcal{Q}_{**} := \mathrm{argmin}_{\mathcal{T} \in \{\mathcal{Q}_*, \widetilde{\mathcal{Q}}_*\}} \mathrm{RV}(\mathcal{T})$. The actual $\mathcal{Q}_{**}$ in each case can be read off from the corresponding highlighted entry in Table 1. We show in Fig. 3 how $f(\mathcal{T})$ changes with $\mathcal{T}$ for all

those Bell scenarios where we have computed at least two different approximations $\mathcal{T}$'s to $\mathcal{Q}$. For these scenarios, we thus have $\mathcal{Q}_{**} \subsetneq \mathcal{Q}_1$.

(n_s, n_o)	$\mathcal{NS} \setminus \mathcal{Q}_1$	$\mathcal{Q}_1 \setminus \widetilde{\mathcal{Q}}_1$	$\widetilde{\mathcal{Q}}_1 \setminus \mathcal{Q}_2$	$\mathcal{Q}_2 \setminus \mathcal{Q}_3$	$\widetilde{\mathcal{Q}}_1 \setminus \widetilde{\mathcal{Q}}_2$	$\mathcal{Q}_{**} \setminus \mathcal{L}$
(2,2)	2 001	5 458	581	305	886	50 116
(2,3)	187	1781	271	359	631	59 000
(2,4)	1	85	12	-	-	25 808
(2,5)	0	1	0	-	-	5 887
(2,6)	0	0	0	-	-	994
(2,7)	0	0	0	-	-	151
(3,2)	27 711	51 209	5 364	3 818	9 185	290 053
(3,3)	2 190	18 766	3 889	-	-	376 374
(3,4)	10	672	259	-	-	201 161
(4,2)	153 800	151 686	16 481	-	-	466 675
(4,3)	16 131	96 990	-	-	-	721 076
(4,4)	96	3 982	-	-	-	585 222
(5,2)	465 263	189 827	21 394	-	-	293 933
(5,3)	100 943	306 120	-	-	-	580 700
(5,4)	186	8 883	-	-	-	276 676
(6,2)	818 872	93 083	10 792	-	-	75 773
(6,3)	340 105	309 851	-	-	-	134 960

Table 2: Summary of the number of correlations $\vec{P}$ excluded from one approximation of $\mathcal{Q}_{[\cdot,\cdot]}$ to a tighter one. The leftmost column gives the Bell scenario. Except for the Bell scenario (5,4) and (6,3), the total number N_{tot} of correlations sampled from $\mathcal{NS}$ is 10^6 (see Table 1). Further to the right, we have, respectively, the number of $\vec{P} \in \mathcal{NS}$ excluded from $\mathcal{Q}_1$, the number of $\vec{P} \in \mathcal{Q}_1$ excluded from $\widetilde{\mathcal{Q}}_1$, the number of $\vec{P} \in \widetilde{\mathcal{Q}}_1$ excluded from $\mathcal{Q}_2$, the number of $\vec{P} \in \mathcal{Q}_2$ excluded from $\mathcal{Q}_3$, the number of $\vec{P} \in \widetilde{\mathcal{Q}}_1$ excluded from $\widetilde{\mathcal{Q}}_2$, and the number of nonlocal $\vec{P}$ that lie in our tightest approximation $\mathcal{Q}_{**}$ (see Table 1). For the Bell scenario (2,2), we have also performed the membership test for $\widetilde{\mathcal{Q}}_3$, $\mathcal{Q}_4$ and $\widetilde{\mathcal{Q}}_4$. However, no $\vec{P}$ was further excluded from these higher-level relaxations to $\mathcal{Q}_{[\cdot,\cdot]}$. For a graphical representation focussing on the nonlocal region, see Fig. 3.

As is evident from the plots (see also Table 2 and Table 3), the first level of the NPA hierarchy $\mathcal{Q}_1$ (corresponding to the first *filled* symbol on each line) generally does not serve as a very good approximation to $\mathcal{Q}$. In fact, it largely overestimates $RV(\mathcal{Q})$ for several Bell scenarios.

Observation 4. *For two-input Bell scenarios, $\mathcal{Q}_1$ approximates $\mathcal{Q}_{**}$ well but for $n_s > 3 \geq n_o$, $RV(\mathcal{Q}_1)$ overestimates $RV(\mathcal{Q})$ by at least 9.69%. This overestimation even exceeds 30% in both the (5,3) and (6,4) Bell scenario, see Table 3.*

From Fig. 3 as well as Table 2 and Table 3, we also see that the almost quantum set $\widetilde{\mathcal{Q}}_1$ (corresponding to the first *hollow* symbol on each line) generally offers a much better approximation to $\mathcal{Q}$ than $\mathcal{Q}_1$ does. For instance, in all the $n_o = 2$ scenarios considered, $\mathcal{Q}_1$ occupies between 9 to 20% more of the nonlocal region compared to $\widetilde{\mathcal{Q}}_1$, whereas in general, $f(\widetilde{\mathcal{Q}}_1) - f(\mathcal{Q}_{**}) < 5\%$. Even if we measure according to $RV(\mathcal{T})$, the difference between $\widetilde{\mathcal{Q}}_1$ and $\mathcal{Q}_{**}$ remains small for all the 12 Bell scenarios explored beyond $\widetilde{\mathcal{Q}}_1$.

Observation 5. *For all the Bell scenarios investigated, $RV(\widetilde{\mathcal{Q}}_1) - RV(\mathcal{Q}_{**})$ is larger than 1% only for the (4,2), (5,2), and the (6,2) Bell scenario. The largest difference is found in the (5,2) Bell scenario, giving $\approx 2.14\%$, see Table 3.*

3.2.2 $\mathcal{Q}_{**}$ vs $\mathcal{NS}$

Next, we focus on determining how $RV(\mathcal{Q}_{**})$ varies across the different Bell scenarios. The trend of how $RV(\mathcal{Q}_{**})$ changes, with $\mathcal{L}$ included, is similar to that of $\mathcal{L}$ shown in Fig. 2. In a close parallel to Observation 1, we have the following observation for $\mathcal{Q}_{**}$ from Fig. 4.
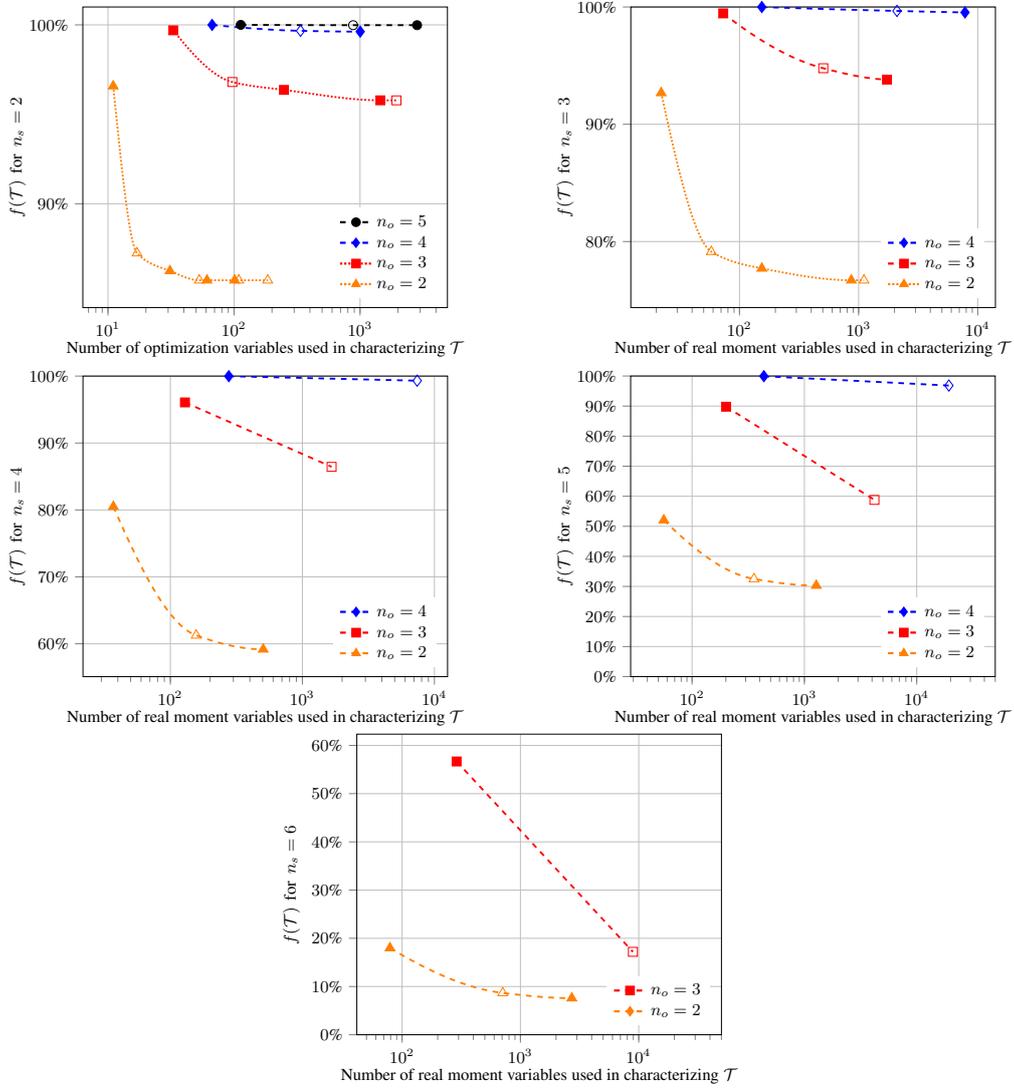
Figure 3: Numerical estimate of $f(\mathcal{T}) = \frac{\mathrm{RV}(\mathcal{T}\setminus\mathcal{L})}{1-\mathrm{RV}(\mathcal{L})}$ for various outer approximations $\mathcal{T}$ of $\mathcal{Q}$, where $\mathcal{T}$ is any member of $\{\mathcal{Q}_k\}_k$ or $\{\widetilde{\mathcal{Q}}_\ell\}_\ell$. In each plot, we use filled (hollow) markers to represent $\mathcal{Q}_k$ ($\widetilde{\mathcal{Q}}_\ell$). The leftmost filled (hollow) marker corresponds to the first level of the NPA [47] (Moroder *et al.* [45]) hierarchy $\mathcal{Q}_1$ ($\widetilde{\mathcal{Q}}_1$). When an increasingly higher level of either hierarchy is considered, the respective SDP characterization involves an increasingly larger number of real (optimization) moment variables (see Appendix C), thus giving rise to markers that are placed more and more to the right of the plot. Plots for the (2,6) and (2,7) Bell scenario have been omitted as $f(\mathcal{T}) = 100\%$ for all approximations considered for these Bell scenarios. As guide for the eye, we have also included a fitting curve (dashed line) in each case. For Bell scenarios with more than three data points, we use the pchip (piecewise cubic Hermite interpolating polynomial) function in MATLAB to generate the respective fitting curve whereas for the other scenarios, we simply use a straight line to join the two data points.

(n_s, n_o)	$\mathcal{Q}_1 \setminus \mathcal{Q}_{**}$	$\widetilde{\mathcal{Q}}_1 \setminus \mathcal{Q}_{**}$	$\mathcal{Q}_{**} \setminus (\mathcal{M}_*^{\mathrm{P}} \cap \mathcal{Q}_{**})$	$\mathcal{P}_* \setminus \mathcal{L}$
(2,2)	0.6344%	0.0886%	0.4817%	0%
(2,3)	0.2412%	0.0631%	0.2429%	0.0012%
(2,4)	0.0097%	0.0012%	0.0364%	0.0012%
(2,5)	0.0001%	0%	0.0016%	0.0017%
(2,6)	0%	0%	0%	0.0003%
(2,7)	0%	0%	0%	0%
(3,2)	6.0394%	0.9185%	3.06%	0%
(3,3)	2.2655%	0.3889%	2.3560%	0.0002%
(3,4)	0.0931%	0.0259%	0.2601%	1.4790%
(4,2)	16.8167%	1.6481%	6.0515%	0.0001%
(4,3)	9.6990%	-	0%	3.6441%
(4,4)	0.3982%	-	0%	9.0878%
(5,2)	21.1221%	2.1394%	5.2313%	0.1267%
(5,3)	30.6120%	-	0%	1.9341%
(5,4)	2.9124%	-	0%	11.968%
(6,2)	10.3875%	1.0792%	1.7932%	0.0679%
(6,3)	39.4714%	-	0%	0.1088%

Table 3: Summary of the difference between the relative volume (RV) of several sets of interest. The leftmost column gives the Bell scenario considered. From the second to the rightmost column, we have, respectively, the difference in the RV between the set associated with the principle of macroscopic locality $\mathcal{Q}_1$ and our tightest approximation to the quantum set $\mathcal{Q}_{**}$, the difference in the RV between the almost quantum set $\widetilde{\mathcal{Q}}_1$ and $\mathcal{Q}_{**}$, the difference in the RV between $\mathcal{Q}_{**}$ and our tightest approximation of the set producible by maximally entangled states with projective measurements $\mathcal{M}_*^{\mathrm{P}}$, and the difference in the RV between our tightest approximation to the set $\mathcal{P}_*$ producible by PPT quantum states and the Bell-local set $\mathcal{L}$.

Observation 6. *For Bell scenarios with the same output n_o, $RV(\mathcal{Q}_{**})$ decreases monotonically with increasing n_s.*

Moreover, since $\mathcal{L}$ is a strict subset of $\mathcal{Q} \subset \mathcal{Q}_{**}$, the following observation may have been anticipated from Observation 2.

Observation 7. *For Bell scenarios with the same input n_s, $RV(\mathcal{Q}_{**})$ increases monotonically with increasing n_o.*
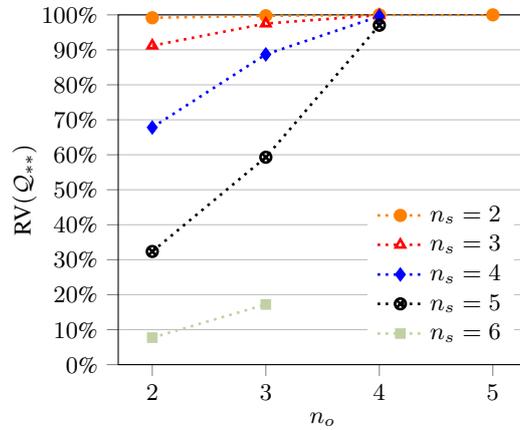
Figure 4: Plots of estimated $RV(\mathcal{Q}_{**})$ vs $n_o \in \{2, 3, \ldots, 5\}$ for $n_s \in \{2, 3, \ldots, 6\}$. The plots for the (2,6) and (2,7) Bell scenarios are omitted as $RV(\mathcal{Q}_{**}) = 100\%$ in both cases.

Despite these similarities, there are also subtle differences. For example, even though for fixed n_s, RV($\mathcal{L}$) generally increases with n_o, it does so after a dip when n_o increases from 2 to 3. More importantly, for Bell scenarios with $n_s = n_o = k$, while RV($\mathcal{L}$) appears to decrease monotonically with increasing k, RV($\mathcal{Q}_{**}$) never seems to get far away from 1. This suggests that for any given n_s, if $n_o \geq n_s$ is large enough, a generic $\vec{P} \in \mathcal{NS}$ is also likely to be a member of $\mathcal{Q}$, i.e., RV($\mathcal{NS} \setminus \mathcal{Q}$) may become vanishingly small.

3.2.3 $\mathcal{Q}_{**}$ vs $\mathcal{L}$

What about the the Bell-*nonlocal* part of the quantum set, i.e., $\mathcal{Q} \setminus \mathcal{L}$? As can be seen from Fig. 5, for all n_s investigated, RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) = RV($\mathcal{Q}_{**}$) − RV($\mathcal{L}$) first increases when n_o increases from 2 to 3. However, for $n_s \leq 4$, this difference in RVs decreases for subsequent values of n_o. Since this is in agreement with the trend of RV($\mathcal{L}$) shown in Fig. 2, the current observation suggests that the trend of RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) for fixed $n_s \leq 4$ is dominated by the trend of RV($\mathcal{L}$). In contrast, the behavior of RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) for varying n_s does not seem to follow immediately from that of RV($\mathcal{Q}_{**}$) nor RV($\mathcal{L}$). In particular, for both the $n_o = 2$ and the $n_o = 3$ case, we see that RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) first increases with n_s (from 2 to 4) but decreases monotonically after that, which differs from the trend found for the $n_o = 4$ scenarios.

Observation 8. *RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) increases monotonically with n_s for the (n_s,4) Bell scenarios and reaches > 90% for the (5,4) Bell scenario.*

A large value of RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) is, *a priori*, unexpected as it requires RV($\mathcal{Q}_{**}$) to be large and RV($\mathcal{L}$) to be small at the same time. However, both requirements happen to hold for the (5,4) Bell scenario.
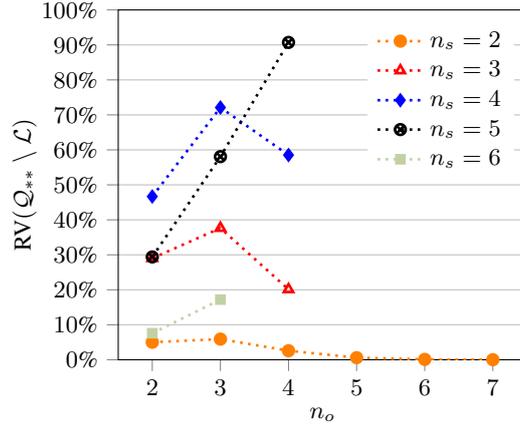
Figure 5: Plots of estimated RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) vs $n_o \in \{2, 3, \ldots, 7\}$ for $n_s \in \{2, 3, \ldots, 6\}$. Here $\mathcal{Q}_{**} = \widetilde{\mathcal{Q}}_1$ for the (4,3), (4,4), (5,3), (5,4), and (6,3) Bell scenario whereas $\mathcal{Q}_{**} \subsetneq \widetilde{\mathcal{Q}}_1$ for all the other Bell scenarios considered.

Let us further remark that when comparing different Bell scenarios, the value of RV($\mathcal{Q}_{**} \setminus \mathcal{L}$) need not correlate with the nonlocality of the correlations contained therein. For example, one might expect that the larger RV($\mathcal{Q}_{**} \setminus \mathcal{L}$), the stronger is the *average* resistance of the associated nonlocal correlations to white noise $\vec{P}_w$. If so, then one might expect the correlations in $\mathcal{Q}_{**} \setminus \mathcal{L}$ for the (5,4) scenario to display the smallest *average* white-noise visibility [cf. Eq. (4)] but the results summarized in Table 4 show otherwise.

An illustration of how various supersets $\mathcal{T}$ of $\mathcal{L}$ contribute towards $\mathcal{NS} \setminus \mathcal{L}$, as measured according to $f(\mathcal{T} \setminus \mathcal{L}) = f(\mathcal{T}) - f(\mathcal{L})$, can be found in the stacked bar chart displayed in Fig. 6 and Fig. 7.

3.3 Other Naturally Restricted Subsets of $\mathcal{Q}$

Next, let us consider $\mathcal{M}^p$ and $\mathcal{P}$, two naturally restricted subsets of $\mathcal{Q}$. Again, our tightest outer approximation to these sets are denoted, respectively, by $\mathcal{M}^p_*$ and $\mathcal{P}_*$.

Scenario	$\mathcal{Q}_{**}$	Max	Min	Mean	σ	Within 1 σ
(2,2)	$\widetilde{\mathcal{Q}}_4$	1.0000	0.7300	0.9265	0.0538	64.75%
(2,3)	$\widetilde{\mathcal{Q}}_2$	1.0000	0.7429	0.9411	0.0461	67.40%
(2,4)	$\mathcal{Q}_2$	1.0000	0.7529	0.9540	0.0393	72.42%
(2,5)	$\mathcal{Q}_2$	1.0000	0.7964	0.9653	0.0308	75.37%
(2,6)	$\mathcal{Q}_2$	1.0000	0.8265	0.9720	0.0261	78.17%
(2,7)	$\mathcal{Q}_2$	0.9999	0.8742	0.9789	0.0226	87.42%
(3,2)	$\widetilde{\mathcal{Q}}_2$	1.0000	0.7214	0.9144	0.0563	62.69%
(3,3)	$\mathcal{Q}_2$	1.0000	0.7303	0.9318	0.0482	64.49%
(3,4)	$\mathcal{Q}_2$	1.0000	0.7286	0.9519	0.0394	70.67%
(4,2)	$\mathcal{Q}_2$	1.0000	0.7179	0.8979	0.0572	62.47%
(4,3)	$\widetilde{\mathcal{Q}}_1$	1.0000	0.7241	0.9136	0.0489	63.33%
(4,4)	$\widetilde{\mathcal{Q}}_1$	1.0000	0.7267	0.9439	0.0411	67.29%
(5,2)	$\mathcal{Q}_2$	1.0000	0.7195	0.8810	0.0538	65.02%
(5,3)	$\widetilde{\mathcal{Q}}_1$	1.0000	0.7336	0.8936	0.0416	67.83%
(5,4)	$\widetilde{\mathcal{Q}}_1$	1.0000	0.7468	0.9265	0.0403	66.17%
(6,2)	$\mathcal{Q}_2$	1.0000	0.7252	0.8672	0.0471	68.18%
(6,3)	$\widetilde{\mathcal{Q}}_1$	0.9998	0.7464	0.8833	0.0301	71.60%

Table 4: Summary of sample statistics associated with the distribution of v^* to $\mathcal{L}$ for all those $\vec{P}$ found to lie in $\mathcal{Q}_{**} \setminus \mathcal{L}$.

3.3.1 $\mathcal{Q}_{**}$ vs $\mathcal{M}_*^{\mathrm{P}}$

Formally, as introduced in Section 2, $\mathcal{M}^{\mathrm{P}}$ is the convex hull of the set of correlations attainable using finite-dimensional maximally entangled states in conjunction with projective measurements. As was first noted in [39], a hierarchy of (increasingly tighter) outer approximations to $\mathcal{M}$ (the analog of $\mathcal{M}^{\mathrm{P}}$ *without* the assumption of measurements being projective) can be obtained via SDPs. In Appendix A, we explain our simplified formulation when the measurement are further *assumed* to be projective.

Note that the hierarchy of SDPs used for the computation of $\mathcal{Q}_{**}$ and that of $\mathcal{M}_*^{\mathrm{P}}$ are independent. Consequently, the two sets $\mathcal{M}_*^{\mathrm{P}}$ and $\mathcal{Q}_{**}$ are generally incomparable, i.e., neither of them is necessarily included in the other, despite the fact that $\mathcal{M}^{\mathrm{P}} \subseteq \mathcal{M} \subsetneq \mathcal{Q}$. To quantitatively understand the difference between $\mathcal{Q}$ and $\mathcal{M}^{\mathrm{P}}$, we thus focus on the difference between $\mathcal{Q}_{**}$ and $\mathcal{M}_*^{\mathrm{P}} \cap \mathcal{Q}_{**}$, i.e., the membership test of any given $\vec{P}$ with respect to $\mathcal{M}_*^{\mathrm{P}}$ is carried out only when it passes the membership test with respect to $\mathcal{Q}_{**}$. For a summary of $\mathcal{Q}_{**}$, $\mathcal{M}_*^{\mathrm{P}}$ involved in the calculation and the sample statistics associated with the distribution of v^* to $\mathcal{M}_*^{\mathrm{P}}$, see Table 5.

Our findings (see Table 1 and Table 3) reveal that the difference between the relative volume of these sets, i.e., $\mathrm{RV}(\mathcal{Q}_*) - \mathrm{RV}(\mathcal{M}_*^{\mathrm{P}} \cap \mathcal{Q}_*)$ is rather small.

Observation 9. *For all the 17 Bell scenarios considered, the largest value of $RV(\mathcal{Q}_*) - RV(\mathcal{M}_*^{P} \cap \mathcal{Q}_*)$ ($\approx$ 6.1%) is found for the Bell scenario (4,2). The Bell scenario (5,2) gives a comparable difference $\approx$ 5.2% whereas for 12 other Bell scenarios, this difference is less than 0.27%. In general, the difference decreases with increasing n_o, see Table 3.*

However, if we restrict our attention to only the nonlocal region of $\mathcal{NS}$, then the difference as quantified by $f(\mathcal{Q}_*) - f(\mathcal{M}_*^{\mathrm{P}} \cap \mathcal{Q}_*)$ is most pronounced ($\approx$8.2%) in the $(2,2)$ Bell scenario, where the original Hardy paradox [35] was proposed, see Fig. 6.

3.4 $\mathcal{P}_*$ vs $\mathcal{L}$

Finally, let us focus on the difference between the set of Bell-local correlations $\mathcal{L}$ and that attainable by locally measuring a PPT (entangled) state $\mathcal{P}$. For a long time, it was believed [53] that no bound entangled state can violate a Bell inequality. Since all PPT entangled states are bound
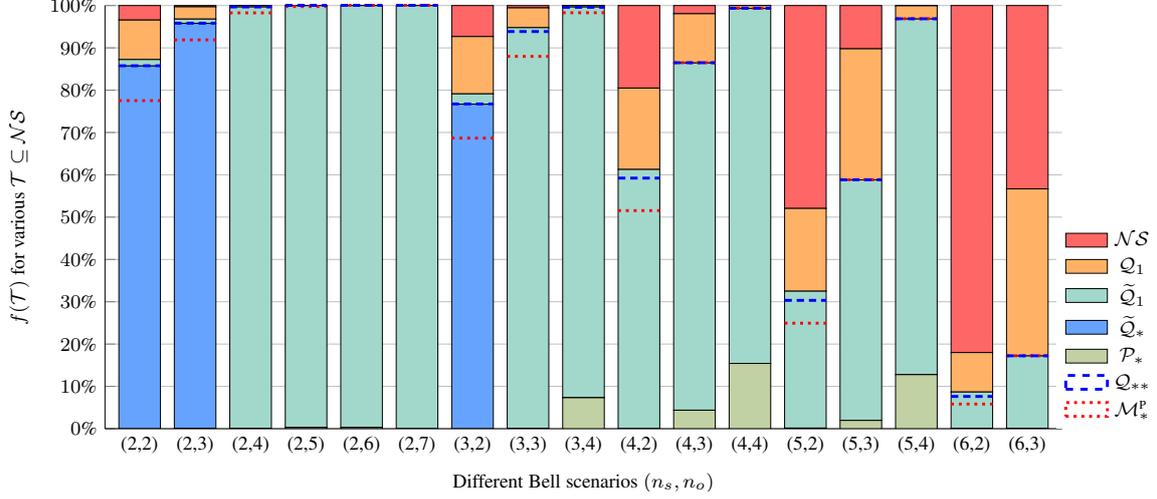
Figure 6: Stacked bar charts showing contributions of various subsets $\mathcal{T} \subseteq \mathcal{NS}$ towards $\mathcal{T} \setminus \mathcal{L}$ where $\mathcal{T}$ is either $\mathcal{M}_*^{\mathrm{P}}$ or any of the subsets listed in the chain of inclusion relations $\mathcal{P}_* \subsetneq \widetilde{\mathcal{Q}}_* \subseteq \widetilde{\mathcal{Q}}_1 \subsetneq \mathcal{Q}_1 \subsetneq \mathcal{NS}$. The RV of a given $\mathcal{T}$ in the nonlocal region, and hence $f(\mathcal{T})$, is the sum over that due to the other sets contained within it. For example, in the (2,2) Bell scenario, $f(\mathcal{Q}_1)$ corresponds to the orange bar as well as all the other bars (turqoise and blue) stacked below it. As $\mathcal{M}_*^{\mathrm{P}}$ and $\mathcal{Q}_{**}$ are not directly comparable to some of the aforementioned sets, we use, respectively, dashed and dotted line to represent them separately. Note also that we have only performed computation for higher levels from the hierarchy of Moroder *et al.* for the (2,2), (2,3), and (3,2) Bell scenarios.
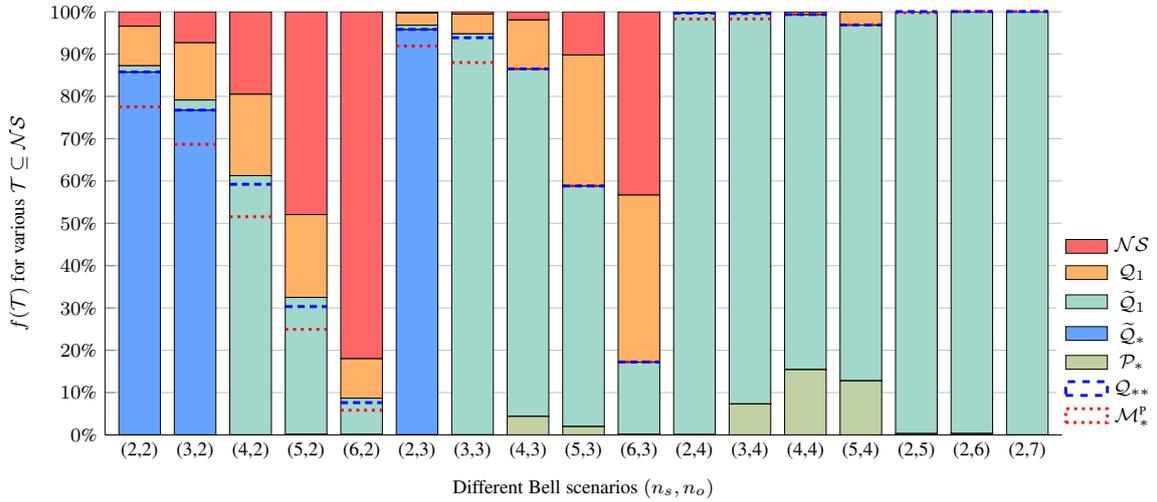
Figure 7: The stacked bar charts of Fig. 6 sorted first in increasing n_s, then followed by increasing n_o.

Scenario	$\widetilde{\mathcal{Q}}_{**}$	$\mathcal{M}_*^{\mathrm{P}}$	Max	Min	Mean	σ	Within $1\,\sigma$
(2,2)	$\widetilde{\mathcal{Q}}_4$	$\mathcal{M}_4^{\mathrm{P}}$	1.3268	0.9405	1.0434	0.0459	77.26%
(2,3)	$\widetilde{\mathcal{Q}}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.2230	0.9419	1.0247	0.0262	82.44%
(2,4)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.1437	0.9470	1.0145	0.0151	85.03%
(2,5)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.0942	0.9574	1.0094	0.0096	86.34%
(2,6)	$\mathcal{Q}_2$	$\mathcal{M}_1^{\mathrm{P}}$	1.0462	1.0000	1.0064	0.0066	85.81%
(2,7)	$\mathcal{Q}_2$	$\mathcal{M}_1^{\mathrm{P}}$	1.0353	1.0000	1.0054	0.0055	87.42%
(3,2)	$\widetilde{\mathcal{Q}}_2$	$\mathcal{M}_3^{\mathrm{P}}$	1.3526	0.9312	1.0364	0.0403	77.87%
(3,3)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.2109	0.9343	1.0155	0.018	81.52%
(3,4)	$\mathcal{Q}_2$	$\mathcal{M}_{1.5}^{\mathrm{P}}$	1.1167	0.9298	1.0084	0.009	85.50%
(4,2)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.3296	0.9348	1.0283	0.0339	78.20%
(4,3)	$\widetilde{\mathcal{Q}}_1$	$\mathcal{M}_1^{\mathrm{P}}$	1.1658	1.0000	1.0125	0.0124	85.33%
(4,4)	$\widetilde{\mathcal{Q}}_1$	$\mathcal{M}_1^{\mathrm{P}}$	1.0741	1.0000	1.0056	0.0056	86.45%
(5,2)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.2661	0.9368	1.0204	0.0275	78.34%
(5,3)	$\widetilde{\mathcal{Q}}_1$	$\mathcal{M}_1^{\mathrm{P}}$	1.1157	1.0000	1.0093	0.0092	84.77%
(5,4)	$\widetilde{\mathcal{Q}}_1$	$\mathcal{M}_1^{\mathrm{P}}$	1.0480	1.0000	1.0039	0.0039	86.18%
(6,2)	$\mathcal{Q}_2$	$\mathcal{M}_2^{\mathrm{P}}$	1.2116	0.9412	1.0141	0.0218	78.53%
(6,3)	$\widetilde{\mathcal{Q}}_1$	$\mathcal{M}_1^{\mathrm{P}}$	1.0815	1.0000	1.0075	0.0073	83.98%

Table 5: Summary of sample statistics associated with the distribution of v^* to $\mathcal{M}_*^{\mathrm{P}}$ of all those $\vec{P}$ found to lie in $\mathcal{Q}_{**}$. The tightest approximations $\mathcal{Q}_{**}$ and $\mathcal{M}_*^{\mathrm{P}}$ are included, respectively, in the second and the third column for ease of reference. The notation for $\mathcal{M}_h^{\mathrm{P}}$ is further explained in Appendix C.

entangled [37], this so-called Peres conjecture would imply that $\mathcal{P} = \mathcal{L}$. Indeed, in the simplest (2,2) Bell scenario, it was shown by Werner and Wolf [69] that these sets do coincide. Even for the (3,2) Bell scenario, numerical results from [45] again indicate that $\mathcal{P} = \mathcal{L}$.

That this conjecture does not hold in full generality was first shown [64] in a tripartite Bell scenario using a three-qubit bound entangled state. Later, the conjecture was also disproved [65] in a *bipartite* setting by considering a two-qutrit PPT entangled state and an asymmetric Bell scenario $\{[2\ 2\ 2]\ [3\ 2]\}$, where the number of entries in each square bracket denotes the number of settings for each party and the actual numbers listed are the number of measurement outcomes for each setting.[1] Consequently, for any symmetric Bell scenario with $n_s, n_o \geq 3$, we must have $\mathcal{L} \subsetneq \mathcal{P}$. Moreover, it can be shown from the results of Ref. [65] and Eq. (4) that the corresponding $\vec{P} \in \mathcal{P}$ has a visibility v^* to $\mathcal{L}$ that is approximately 0.9996.

Here, we make use of the outer approximations proposed in [45] to quantitatively survey the difference between $\mathcal{P}$ and $\mathcal{L}$. Interestingly, even though $\mathcal{P}_*$ only outer approximates $\mathcal{P}$, we see from Table 6 that for the (2,2) and the (3,2) Bell scenario, these approximations work extremely well: among all the 10^6 samples generated for each of these Bell scenarios, there is not even a single $\vec{P} \notin \mathcal{P}_*$ that lies outside $\mathcal{L}$. This is, of course, consistent with the known results given, respectively, in Ref. [69] and Ref. [45]. In fact, for the (2,2) scenario, even the lowest-level approximation given by $\mathcal{P}_1$ does not give rise to any $\vec{P} \notin \mathcal{L}$ from 10^6 samples.

In contrast, except the (2,7) Bell scenario, we do find instances of $\vec{P} \in \mathcal{P}_* \setminus \mathcal{L}$ for *all other* Bell scenarios investigated, see Table 6. For this exceptional case, we note from Table 1 that $\mathcal{L}$ almost spans the entire non-signaling set, thus leaving very little room for $\mathcal{NS} \setminus \mathcal{L}$, let alone $\mathcal{P} \setminus \mathcal{L}$. For most of the other Bell scenarios, even though we know from Table 6 that $\mathcal{P}_* \neq \mathcal{L}$, the difference between $\mathrm{RV}(\mathcal{P}_*)$ and $\mathrm{RV}(\mathcal{L})$ is tiny, if not vanishingly small.

Observation 10. *For all the 17 Bell scenarios considered, the largest value of $\mathrm{RV}(\mathcal{P}_* \setminus \mathcal{L})$ ($\approx 12\%$) is found for the Bell scenario (5,4). The Bell scenario (4,4) shows a comparable difference of $\approx 9.1\%$ whereas for 12 other Bell scenarios, this difference is less than 0.15%. Except the $n_s = 2$ Bell scenarios, $\mathrm{RV}(\mathcal{P}_* \setminus \mathcal{L})$ increases with increasing n_o, see Table 3.*

---

[1]The notation adopted here for this asymmetric Bell scenario follows that introduced in Ref. [8].

Zooming into the behavior of the individual correlation, we see from Table 6 that within these approximations, one can find $\vec{P} \in (\mathcal{P}_* \setminus \mathcal{L})$ that are far more robust with respect to the mixing with white noise in more complex Bell scenarios. For example, in the (4,4) Bell scenario, the most robust $\vec{P} \in \mathcal{P}_*$ was found to give a visibility of 0.9037, as compared with a visibility of 0.9996 found [65] in the (3,3) Bell scenario.

Scenario	$\mathcal{P}_*$	Max	Min	Mean	σ	Within 1 σ
(2,3)	$\mathcal{P}_2$	0.9998	0.9963	0.9983	0.0010	66.67%
(2,4)	$\mathcal{P}_2$	0.9999	0.9972	0.9986	0.0008	66.67%
(2,5)	$\mathcal{P}_{1.5}$	1.0000	0.9908	0.9970	0.0034	76.47%
(2,6)	$\mathcal{P}_{1.5}$	0.9999	0.9985	0.9993	0.0008	66.67%
(3,3)	$\mathcal{P}_2$	1.0000	0.9994	0.9997	0.0004	100%
(3,4)	$\mathcal{P}_{1.25}$	1.0000	0.9326	0.9883	0.0094	70.18%
(4,2)	$\mathcal{P}_2$	1.0000	0.9999	1.0000	0.0001	66.67%
(4,3)	$\mathcal{P}_{1.2}$	1.0000	0.9194	0.9856	0.0109	69.05%
(4,4)	$\mathcal{P}_1$	1.0000	0.9037	0.9819	0.0141	67.90%
(5,2)	$\mathcal{P}_{1.5}$	1.0000	0.9536	0.9918	0.0081	83.50%
(5,3)	$\mathcal{P}_1$	1.0000	0.9088	0.9751	0.0157	64.64%
(5,4)	$\mathcal{P}_1$	1.0000	0.9041	0.9772	0.014	62.32%
(6,2)	$\mathcal{P}_1$	0.9999	0.9087	0.9752	0.0179	65.39%
(6,3)	$\mathcal{P}_1$	0.9996	0.9213	0.9691	0.0139	66.86%

Table 6: Summary of sample statistics associated with the distribution of v^* to $\mathcal{L}$ of all those $\vec{P}$ found to lie in $\mathcal{P}_* \setminus \mathcal{L}$. For the Bell scenarios (2,2), (2,7) and (3,2), we do not find any correlation $\vec{P} \in \mathcal{P}_* \setminus \mathcal{L}$, thus the respective rows are omitted. The tightest approximation $\mathcal{P}_*$ is included in the second column for ease of reference. The notation for $\mathcal{P}_z$ is further explained in Appendix C.

4 Discussions

In the studies of Bell-nonlocality [17] and its applications in device-independent quantum information [58] (DIQI), one often exploits the geometrical features associated with the correlation $\vec{P}$ or the set(s) in which they belong to draw conclusions. For example, the convexity of $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$ and the possibility to describe $\mathcal{L}$ and $\mathcal{NS}$ using a finite number of extreme points are often invoked to simplify the analyses in DIQI. Nonetheless, despite the numerous efforts devoted to these lines of research, much about the geometry of these sets remain to be understood, see, e.g., Ref. [33].

In fact, even though the inclusion relations $\mathcal{L} \subsetneq \mathcal{M}^{\mathrm{P}} \subsetneq \mathcal{Q} \subsetneq \widetilde{\mathcal{Q}}_1 \subsetneq \mathcal{Q}_1 \subsetneq \mathcal{NS}$ and $\mathcal{L} \subseteq \mathcal{P} \subsetneq \mathcal{Q}$ are long known, we do not have any quantitative understanding of their difference beyond the extremely limited exploration carried out in Refs. [19, 30, 75] for $\mathcal{L}$, $\mathcal{Q}$, and $\mathcal{NS}$. In this work, by determining the relative volume (RV) of these sets or their outer approximations, we aim to fill this gap by putting some of the intuitions that the community has developed over the years on more quantitative grounds. As RV$(\mathcal{T})$ can be interpreted as the success probability of a Bernoulli trial, if RV$(\mathcal{T}) \geq 100\% - \epsilon$ for some $\mathcal{T} \subseteq \mathcal{NS}$, the operational task of finding a correlation in $\mathcal{NS} \setminus \mathcal{T}$ by a uniform sampling in $\mathcal{NS}$ can only succeed with probability at most ϵ. En route to the above goal, several intriguing observations are noted.

Firstly, our results suggest that, the fraction of $\mathcal{L}$ in $\mathcal{NS}$ becomes vanishingly small (Observation 1) as we *fix* the number of outputs n_o but *increases* the number of inputs n_s. A similar monotonic behavior is also observed for the macroscopically-local set $\mathcal{Q}_1$, the almost-quantum set $\widetilde{\mathcal{Q}}_1$ (see Table 1), as well as our tightest approximation to the quantum set $\mathcal{Q}_{**}$ (Observation 6). Although the two-output case of Observation 1 was already noted in Ref. [30], our results provide further evidence this trend could well be generic, i.e., independent of the actual value of n_o. Moreover, we demonstrate how a combinatoric argument can be used to understand this decreasing trend.

Interestingly, apart from a dip at the beginning, an essentially opposite trend for $\mathrm{RV}(\mathcal{L})$ is found when we *fix* the number of inputs n_s but *increases* the number of outputs n_o (Observation 2). This increasing trend is even found to be *monotonic* for several supersets of $\mathcal{L}$, including $\mathcal{Q}_1$, $\widetilde{\mathcal{Q}}_1$ (see Table 1), and $\mathcal{Q}_{**}$ (Observation 7). Again, this suggests that the trend—already noted in Ref. [30] for the special case of $n_s = 2$—could well be generic and independent of the specific value of n_s. If so, then for any given n_s, the quantum set $\mathcal{Q}$ and its superset $\mathcal{NS}$ may become essentially indistinguishable when n_o is sufficiently large. While we do not have an intuitive explanation for this observation, it seems plausible that it can lead to interesting consequences on the power of nonlocal quantum resources.

What about the quality of various outer approximations of $\mathcal{Q}$, which is especially relevant for a variety of tasks in DIQI? Our results suggest that for two-input Bell scenarios, $\mathcal{Q}_1$ already provides a superb outer approximation, with $\mathrm{RV}(\mathcal{Q}_1) - \mathrm{RV}(\mathcal{Q}_{**}) < 0.7\%$, but for Bell scenarios with $n_s > 3 \geq n_o$, the reliability of $\mathcal{Q}_1$ as an outer approximation becomes questionable (Observation 4). In these cases, the difference $\mathrm{RV}(\mathcal{Q}_1) - \mathrm{RV}(\mathcal{Q}_{**})$ can even get as large as $\approx 39.5\%$. In contrast, the almost-quantum set $\widetilde{\mathcal{Q}}_1$, as its name suggests, gives consistently a tiny deviations, if at all, from our tightest quantum approximation (Observation 5). Still, our results (see Table 3) suggest that a more noticeable difference may be found in a Bell scenario with two outputs but a larger number of inputs, say, $n_s = 5$.

Moving on to naturally restricted subsets of $\mathcal{Q}$, it is known from the work of Ref. [40] and Ref. [66] that $\mathcal{M}^\mathrm{P} \subseteq \mathcal{M} \subsetneq \mathcal{Q}$, respectively, in the (2,2) and the (3,2) Bell scenario. Our results indicate that for fixed n_s, the difference between $\mathcal{M}^\mathrm{P}$ and $\mathcal{Q}$ may diminish following the increase in n_o (Observation 9). If so, in a Bell scenario where $n_s \ll n_o$, it may be sufficient to consider only finite-dimensional maximally entangled states in conjunction with projective measurements for various tasks in DIQI. However, our observation should not be taken to imply that $\mathcal{M}^\mathrm{P} \to \mathcal{Q}$ for *any* Bell scenario with large enough n_o, see, e.g., Ref. [38].

In contrast, our results do suggest that the difference between $\mathcal{P}$ and $\mathcal{L}$ may become noticeable only when the Bell scenario involved is sufficiently complex (Observation 10). In particular, for the $(n_s, 2)$ and $(2, n_o)$ Bell scenarios that we have investigated, the respective $\mathrm{RV}(\mathcal{P}_*) - \mathrm{RV}(\mathcal{L})$ is always found to be tiny ($< 0.17\%$). For all but two of the remaining Bell scenarios, this difference is also never more than 3.7%. That leaves us only with the (4,4) and the (5,4) Bell scenario—among all those computed—as the most promising candidate for an experimental demonstration of the Peres conjecture violation. Still, even though we get, respectively, $\mathrm{RV}(\mathcal{P}_* \setminus \mathcal{L}) \approx 9\%$ and $\approx 12\%$, and a minimum visibility of ≈ 0.90 for both Bell scenarios, further investigation is clearly needed to confirm its experimental viability.

Apart from this, a few other closely-related research directions may be worth pursuing. For example, even though both $\mathcal{P}$ and $\mathcal{M}^\mathrm{P}$ are known to be subsets of $\mathcal{Q}$, their precise relationship is not known. Intuitively, one would expect $\mathcal{P}$ to be a strict subset of $\mathcal{M}^\mathrm{P}$, which is supported by our observation that $\mathrm{RV}(\mathcal{Q}_{**}) - \mathrm{RV}(\mathcal{M}^\mathrm{P}_*)$ and $\mathrm{RV}(\mathcal{P}_*) - \mathrm{RV}(\mathcal{L})$ are typically small, but proving this does not seem to be trivial. Evidently, a comprehensive estimation of the relative volume of various naturally restricted subsets in the multipartite setting is also desirable. Due to the rich structure in multipartite entanglement [43] and multipartite nonlocality [28], there will be many more natural subsets of $\mathcal{NS}$ to consider in those cases.

Another direction that deserves further investigation is that related to the choice of metric in our sampling. In this work, we have opted for the Euclidean metric defined in the space of $\mathscr{P}_{\mathcal{NS}}$. However, since the space of interest $\mathscr{P}$ is a probability space, an arguably more natural [2] metric is the so-called Fisher (information) metric $\mathrm{d}s_\mathrm{F}^2$. For a set of *unconditional* probability distributions q_i such that $q_i \geq 0$ and $\sum_i q_i = 1$, the Fisher metric and hence the corresponding volume element are given, respectively, by [1]:

$$\mathrm{d}s_\mathrm{F}^2 = \sum_i \frac{\mathrm{d}q_i^2}{q_i} \quad \text{and} \quad \mathrm{d}V_\mathrm{F} = \prod_i \left(\frac{\mathrm{d}q_i}{\sqrt{q_i}} \right). \tag{10}$$

Recall from Section 2 that we are interested in *conditional* distributions lying in the non-signaling subspace, cf. Eq. (1). These requirements, together with a nontrivial dependence of the metric on the coordinate in $\mathscr{P}$, however, make it unclear how we can perform a uniform sampling according to this metric (e.g., using an existing software package like cprnd). Similarly, a careful reader would

have noticed that if we work in the space of $\mathscr{P}$, the uniform distribution $\vec{P}_w$ has the same Euclidean distance to all extreme points of $\mathcal{L}$ (see Appendix E) but if we make use of the parametrization given in Eq. (5), then this invariance is lost. As such, it would be interesting to find a parameterization of $\mathscr{P}_{\mathcal{NS}}$ where this invariance is preserved and repeat the calculation performed here.

Finally, recall that in our studies of convergence of the outer approximations of $\mathcal{Q}$ (see Section 3.2.1), we use the hierarchies of SDPs defined by NPA [47] as well as those given in Moroder *et al.* [45]. These are, however, not the only outer approximations of $\mathcal{Q}$ that discussed in the literature. For example, the SDPs defined by Berta *et al.* [14] are also known to define a converging hierarchy of outer approximations. Moreover, in comparison with the NPA hierarchy, the approximations of Berta *et al.* are known to be (possibly) tighter as they include further non-negativity requirement of certain elements of the moment matrix (see also Ref. [60]). It could thus also be interesting to investigate, how this, and other hierarchies of outer approximations (e.g., those discussed in [22]) appear to converge.

Acknowledgements

A Approximation to the $\mathcal{M}^{\mathbb{P}}$

In this section, we explain how our approximations to the *convex hull* of the set of correlations attainable by finite-dimensional maximally entangled quantum states $\mathcal{M}$ are defined. As was pointed out in Ref. [39], $\mathcal{M}$ can be *outer* approximated by a hierarchy of correlations, each amenable to an SDP characterization. Here, we focus on a subset $\mathcal{M}^{\mathbb{P}}$ of $\mathcal{M}$, where the local POVM elements are further assumed to projectors.

To appreciate how the hierarchy works, let us first remind that for any local POVM element $E_{a|x}$ $E_{b|y}$ acting on a bipartite d-dimensional maximally entangled state $|\Psi_d\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |i\rangle|i\rangle$, we have:

$$P(a,b|x,y) = \langle \Psi_d | E_{a|x} \otimes E_{b|y} | \Psi_d \rangle = \frac{\text{tr}\left(E_{a|x} E_{b|y}^{\text{T}}\right)}{d}, \tag{11}$$

where $(\cdot)^{\text{T}}$ denotes transposition.

The essence of the characterization of $\mathcal{M}$, and hence of our characterization of $\mathcal{M}^{\mathbb{P}}$, is an approximation of the trace function in Eq. (11) by a linear function acting on the POVM elements. To this end, let us define $M_0 = \{\mathbb{I}_d\}$, which is the set consisting of only the d-dimensional identity operator, and its union with a set of *projective* POVM elements

$$M_1 = \{\mathbb{I}_d\} \cup \{\tilde{E}_{a|x}\}_{a,x} \cup \{\tilde{E}_{b|y}\}_{b,y} \tag{12}$$

where $a, b = 1, 2, \ldots, n_o - 1$, $x, y = 1, 2, \ldots, n_s$. Here, the projective nature of the POVM elements implies

$$\tilde{E}_{a|x}\tilde{E}_{a'|x} = \tilde{E}_{a|x}\delta_{a,a'}, \quad \tilde{E}_{b|y}\tilde{E}_{b'|y} = \tilde{E}_{b|y}\delta_{b,b'}. \tag{13}$$

More generally, for any positive integer $k > 1$, let us *define* the set of operators with degree k or less as

$$M_k = \cup_i \left\{\Pi_i(\tilde{E}^1 \ldots \tilde{E}^k)\right\}, \tag{14}$$

where the union is over all possible permutations Π_i of k-fold product of operators chosen from M_1. Notice that as both $\tilde{E}_{a|x}$ and $\tilde{E}_{b|y}$ act on the same Hilbert space, they generally do not commute.

Now, in analogy to the work of Ref. [39], we *define* a bipartite correlation $\vec{P}$ to be a member of $\mathcal{M}_k^{\mathrm{P}}$, $k \geq 1$, if there exists an integer $d \geq 2$ and a linear functional $L : M_k \times M_k \to \mathscr{P}_{\mathcal{NS}}$ such that the following properties hold:

1. $L(\mathbb{I}_d) = 1$.

2. $L(f f^\dagger) \geq 0$ for any $f \in M_k$.

3. $L(f \tilde{E}_{a|x} f^\dagger), L(f \tilde{E}_{b|y} f^\dagger) \geq 0$ for any $f \in M_{k-1}$.

4. $L(f \tilde{E}_{a|x} f^\dagger \tilde{E}_{a'|x'}), L(f \tilde{E}_{b|y} f^\dagger \tilde{E}_{b'|y'}) \geq 0$ and
   $L(f \tilde{E}_{a|x} f^\dagger \tilde{E}_{b|y}), L(f \tilde{E}_{b|y} f^\dagger \tilde{E}_{a|x}) \geq 0$ for any $f \in M_{k-1}$.

5. $L(ST) = L(TS)$ where $ST \in M_{2k}$.

6. $L(\tilde{E}_{a|x} \tilde{E}_{b|y}) = P(a,b|x,y)$ for all a,b,x,y.

Clearly, if $\vec{P} \in \mathcal{M}^{\mathrm{P}}$, then by Eq. (11), a linear functional satisfying all the above properties is guaranteed to exist by taking $L(\cdot) = \frac{1}{d}\mathrm{tr}(\cdot)$ and setting $\tilde{E}_{a|x} = E_{a|x}$ and $\tilde{E}_{b|y} = E_{b|y}^{\mathsf{T}}$. In other words, $\vec{P} \in \mathcal{M}^{\mathrm{P}} \implies \vec{P} \in \mathcal{M}_k^{\mathrm{P}}$ for all $k \geq 1$. Importantly, for any given integer $k \geq 1$, the membership of any given $\vec{P} \overset{?}{\in} \mathcal{M}_k^{\mathrm{P}}$ can be determined by solving an SDP that amounts to requiring the existence of a positive semidefinite moment matrix Γ with its entries given by $\Gamma_{ij} = L(f_i f_j^\dagger)$ where $f_i, f_j \in M_k$ and where all these entries are required to satisfy the linear constraints listed above.

A few remarks are now in order. Firstly, as oppose to the NPA hierarchy where the convexity of $\mathcal{Q}_{[\cdot,\cdot]}$ (and hence of $\mathcal{Q}_k$) is promised by *not* restricting the underlying Hilbert space dimension, convexity has to be assumed in the formulation of $\mathcal{M}$ or $\mathcal{M}^{\mathrm{P}}$ by considering the convex hull of the set of correlations attainable from $|\Psi_d\rangle$. Secondly, in the formulation of $\mathcal{M}$, the projective nature of the POVM elements cannot be taken for granted, since a naïve application of Naimark's extension [51] does not guarantee that the state to which the extended projective POVM elements are applied is maximally entangled.

Consequently, our characterization differs from that given in Ref. [39] in two aspects: (1) our formulation assumes that $\tilde{E}_{a|x}$ and $\tilde{E}_{b|y}$ are projective while that of Ref. [39] does not (2) the formulation given in Ref. [39] actually imposes in property 3. above the more stringent requirement that f can be *any* linear combination of elements in M_{k-1}. Since one of these differences is more constraining while the other is less constraining, our hierarchy $\mathcal{M}_k^{\mathrm{P}}$ is neither a subset nor a superset of the corresponding set $\mathcal{Q}_+^k$ defined in Ref. [39]. Empirically, we have also found that if we keep difference (2) and drop the assumption of $\tilde{E}_{a|x}$, $\tilde{E}_{b|y}$ being projective, then the resulting relaxation of $\mathcal{Q}_+^k$ appears to be hardly constraining. Finally, notice that except for some additional positivity requirement due to property 4 above, the SDP for $\mathcal{M}_1^{\mathrm{P}}$ is the same as that for NPA level 1.

B Sampling methods

To obtain uniformly sampled correlations in $\mathcal{NS}$ using the `cprnd` function, we make use of the minimal parametrization given in Eq. (5). For this purpose, it suffices to input to `cprnd` (1) the total number of samples N_{tot} required (2) the sampling algorithm to be employed,[2] and (3) a complete description of the nonsignaling polytope $\mathcal{NS}$ in terms of its positivity (facet) constraints.

Explicitly, these positivity constraints are

$$P(a|x) \geq 0, \quad P(b|y) \geq 0, \quad P(a,b|x,y) \geq 0, \tag{15}$$

in addition to

$$P(a = n_o|x) = 1 - \sum_{a'=1}^{n_o-1} P(a'|x) \geq 0, \quad P(b = n_o|y) = 1 - \sum_{b'=1}^{n_o-1} P(b'|y) \geq 0, \tag{16}$$

---

[2]In our work, we use "Gibbs" sampler that gives, empirically, better convergence properties than the default "hit-and-run" sampler.

Accepted in ⟨ʠuantum 2022-07-10, click title to verify. Published under CC-BY 4.0.

19

and

$$P(a, b = n_o | x, y) = P(a|x) - \sum_{b'=1}^{n_o-1} P(a, b'|x, y) \geq 0,$$

$$P(a = n_o, b | x, y) = P(b|y) - \sum_{a'=1}^{n_o-1} P(a', b|x, y) \geq 0, \tag{17}$$

$$P(a = n_o, b = n_o | x, y) = 1 - \sum_{a'=1}^{n_o-1} P(a'|x) - \sum_{b'=1}^{n_o-1} P(b'|y) + \sum_{a',b'=1}^{n_o-1} P(a', b'|x, y) \geq 0$$

for all $a, b \in \{1, \dots, n_o - 1\}$ and $x, y \in \{1, 2, \dots, n_s\}$.

C Bell scenarios considered and the complexity involved in the characterization of various sets

We consider bipartite Bell scenarios where each party has n_s measurement settings and where each measurement gives n_o outcomes. Bearing in mind that the generating POVM elements may be chosen to be projectors, the size of the Moroder [45] level ℓ moment matrix, denoted by $D_{\widetilde{\mathcal{Q}}_\ell}$, can be shown to be:

$$D_{\widetilde{\mathcal{Q}}_\ell} = \left[1 + \sum_{j=1}^{\ell} n_s(n_s - 1)^{j-1}(n_o - 1)^j \right]^2. \tag{18}$$

Similarly, the size of the NPA level k moment matrix, denoted by $D_{\mathcal{Q}_k}$ can be shown to be:

$$D_{\mathcal{Q}_k} = 1 + 2n_s \sum_{j=1}^{k} (n_s - 1)^{j-1}(n_o - 1)^j + n_s^2 \sum_{j=2}^{k} (j-1)(n_s - 1)^{j-2}(n_o - 1)^j, \tag{19}$$

where the first sum consists of only contributions of k-fold products of operators from the same party, the last sum consists of k-fold products of operators originating from both parties, and the factor $(j - 1)$ in the last sum accounts for different possibilities in terms of the number of Alice's and Bob's operators.

Finally, the size of the moment matrix corresponding to the characterization of level h of $\mathcal{M}^\text{P}$, denoted by $D_{\mathcal{M}_h^\text{P}}$ is:

$$D_{\mathcal{M}_h^\text{P}} = 1 + 2n_s \sum_{j=1}^{h} (2n_s - 1)^{j-1}(n_o - 1)^j. \tag{20}$$

From the expressions given above, it is clear that for all these hierarchies, the size of the SDP moment matrix increases exponentially with the level of the hierarchy. As such, due to limitation in computational resources, it is also expedient to consider intermediate, non-integer level of these hierarchies in order to obtain a tighter approximation.

For both hierarchies of SDPs due to Moroder *et al.*, let

$$D_{\text{local},\ell} = 1 + \sum_{j=1}^{\ell} n_s(n_s - 1)^{j-1}(n_o - 1)^j. \tag{21}$$

Furthermore, let us denote by $\lfloor q \rfloor$ be the integer part of a positive number $q \geq 1$. Then, we say that an outer approximation is of level q if the considered moment matrix Γ_q contains $\Gamma_{\lfloor q \rfloor}$ as a submatrix while Γ_q itself is a submatrix of $\Gamma_{\lfloor q \rfloor + 1}$. Moreover, the moment matrix of Γ_q is formed by considering only (approximately) the first $(q - \lfloor q \rfloor)(D_{\text{local}, \lfloor q \rfloor + 1} - D_{\text{local}, \lfloor q \rfloor})$ level-$(\lfloor q \rfloor + 1)$ local operators in addition to *all* the level-$\lfloor q \rfloor$ local operators.[3]

In a similar manner, for the hierarchy defined in Appendix A, we say that an outer approximation for $\mathcal{M}^\text{P}$ is of an intermediate level q if the considered moment matrix Γ_q contains $\Gamma_{\lfloor q \rfloor}$ as a

---

[3]Note that our level-ℓ local operators $A_{a_1|x_1} A_{a_2|x_2} \cdots A_{a_\ell|x_\ell}$ are ordered by first increasing the index of a_1, followed by x_1, followed by a_2, etc.

Bell Scenario	k	$(D_{\mathcal{Q}_k}, N_{\mathcal{Q}_k})$	l	$(D_{\widetilde{\mathcal{Q}}_l}, N_{\widetilde{\mathcal{Q}}_l})$	z	$(D_{\mathcal{P}_z}, N_{\mathcal{P}_z})$	h	$(D_{\mathcal{M}_h^{\mathrm{P}}}, N_{\mathcal{M}_h^{\mathrm{P}}}, N_{\mathcal{M}_h^{\mathrm{P}}}^{\geq 0})$	$(D_{\mathcal{L}}, N_{\mathrm{ext}}^{\mathcal{L}})$
(2,2)	4	(41, 101)	4	(81, 185)	4	(81, 185)	4	(161, 806, 239)	(8, 16)
(2,3)	3	(137, 1 449)	2	(169, 1 945)	2	(169, 1 945)	2	(57, 311, 177)	(24, 81)
(2,4)	2	(85, 1 003)	1	(49, 337)	2	(625, 29 353)	2	(121, 1 282, 553)	(48, 256)
(2,5)	2	(145, 2 833)	1	(81, 881)	1.5	(256, 9 106)	2	(209, 3 657, 1 265)	(80, 625)
(2,6)	2	(221, 6 471)	1	(121, 19 21)	1.5	(1 296, 17 946)	1	(21, 171, 171)	(120, 1 296)
(2,7)	2	(313, 12 841)	1	(169, 3 697)	1	(169, 3 697)	1	(25, 241, 241)	(168, 2 401)
(2,8)	1	(29, 323)	1	(225, 6 497)	1	(225, 6 497)	1	(29, 323, 323)	(224, 4 096)
(2,9)	1	(33, 417)	1	(289, 10 657)	1	(289, 10 657)	1	(33, 417, 417)	(288, 6 561)
(3,2)	3	(88, 868)	2	(100, 1 108)	2	(100, 1 108)	3	(187, 1 979, 472)	(15, 64)
(3,3)	2	(97, 1 729)	1	(49, 505)	2	(961, 131 041)	2	(133, 1 808, 673)	(48, 729)
(3,4)	2	(208, 7 822)	1	(100, 2 098)	1.25	(576, 12 456)	1.5	(154, 6 643, 1567)	(99, 4 096)
(3,5)	1	(25, 265)	1	(169, 6 025)	1	(169, 6 025)	1	(25, 265, 265)	(168, 15 625)
(3,6)	1	(31, 406)	1	(256, 13 906)	1	(256, 13 906)	1	(31, 406, 406)	(255, 46 656)
(3,7)	1	(37, 577)	1	(361, 27 793)	1	(361, 27 793)	1	(37, 577, 577)	(360, 117 649)
(4,2)	2	(49, 505)	1	(25, 157)	2	(289, 13 105)	2	(65, 499, 233)	(24, 256))
(4,3)	1	(17, 129)	1	(81, 1 665)	1.2	(361, 7 993)	1	(17, 129, 129)	(80, 6 561)
(4,4)	1	(25, 277)	1	(169, 74 05)	1	(169, 7 405)	1	(25, 277, 277)	(168, 65 536)
(5,2)	2	(76, 1 276)	1	(36, 356)	1.5	(256, 9 106)	2	(101, 1 211, 461)	(35, 1 024)
(5,3)	1	(21, 201)	1	(121, 4 201)	1	(121, 4 201)	1	(21, 201, 201)	(120, 59 049)
(5,4)	1	(31, 436)	1	(256, 19 336)	1	(256, 19 336)	1	(31, 436, 436)	(255, 1 048 576)
(6,2)	2	(109, 2 719)	1	(49, 709)	1	(49, 709)	2	(145, 2 510, 805)	(48, 4 096)
(6,3)	1	(25, 289)	1	(169, 8 929)	1	(169, 8 929)	1	(25, 289, 289)	(168, 531 441)

Table 7: Summary of the Bell scenarios considered in this work (leftmost column), the highest level of the SDP relaxation considered (second, fourth, sixth, and the eighth column), and the corresponding parameters characterizing the complexity of the computation (third, fifth, seventh, and the ninth column). In particular, for the NPA hierarchy (the second and third column), the Moroder hierarchy for $\mathcal{Q}$ (the fourth and fifth column), the Moroder hierarchy for $\mathcal{P}$ (the eighth and ninth column), we provide in bracket the size of the respective moment matrix D and the number of (real) moment variables involved in the corresponding optimization. Similarly, for the hierarchy of SDPs characterizing $\mathcal{M}^{\mathrm{P}}$ (the fifth and sixth column), we provide in bracket the size of the respective moment matrix $D_{\mathcal{T}}$, the number of (real) moment variables $N_{\mathcal{T}}$ involved in the corresponding optimization, and the number of moments that are further required to be non-negative $N_{\mathcal{M}_h^{\mathrm{P}}}^{\geq 0}$. In the last column, we list, accordingly, the key parameters characterizing the complexity of the linear program involved in solving the membership problem $\vec{P} \overset{?}{\in} \mathcal{L}$, i.e., the size $D_{\mathcal{L}}$ of $\vec{P}$, which equals to the dimension d of $\mathcal{NS}$ and the number of extreme points $N_{\mathrm{ext}}^{\mathcal{L}}$ of the respective local polytope $\mathcal{L}$. For completeness, we also include in gray the parameters characterizing the complexity of the various level-1 SDPs for the (2,8), (2,9), (3,5), (3,6), (3,7) Bell scenarios that we did not compute.

submatrix while Γ_q itself is a submatrix of $\Gamma_{\lfloor q \rfloor + 1}$. Moreover, the moment matrix of Γ_q is formed by taking the upper-left submatrix of $\Gamma_{\lfloor q \rfloor + 1}$ with (approximately) $D_{\mathcal{M}_{\lfloor q \rfloor}^{\mathrm{P}}} + (q - \lfloor q \rfloor)(D_{\mathcal{M}_{\lfloor q \rfloor + 1}^{\mathrm{P}}} - D_{\mathcal{M}_{\lfloor q \rfloor}^{\mathrm{P}}})$ rows and columns.[4]

Of course, the size of the moment matrix $D_{\mathcal{T}}$ for a set $\mathcal{T}$ is not only the parameter that determines the computational resource required to solve each of these SDPs. In particular, for the membership test corresponding to $T \in \{\mathcal{Q}_k, \widetilde{\mathcal{Q}}_\ell, \mathcal{M}_h^{\mathrm{P}}, \mathcal{P}_z\}$, the number $N_{\mathcal{T}}$ of real variables (independent moments) involved in the corresponding moment matrix also plays a crucial role. In the case of $\mathcal{M}_h^{\mathrm{P}}$, the number $N_{\mathcal{M}_h^{\mathrm{P}}}^{\geq 0}$ of moments that are further required to be non-negative also play

---

[4]Here, our level-h operators $A_{a_1|x_1} \cdots A_{a_{h-1}|x_{h-1}} A_{a_h|x_h}$ are ordered by first increasing the index of a_h, followed by x_h, followed by a_{h-1}, etc. Moreover, we adopt the convention that $A_{a_i|x_i} = B_{a_i|x_i - n_s}$ for $x_i > n_s$ where $A_{a|x} \leftrightarrow \tilde{E}_{a|x}$, $B_{b|y} \leftrightarrow \tilde{E}_{b|y}$, see Appendix A.
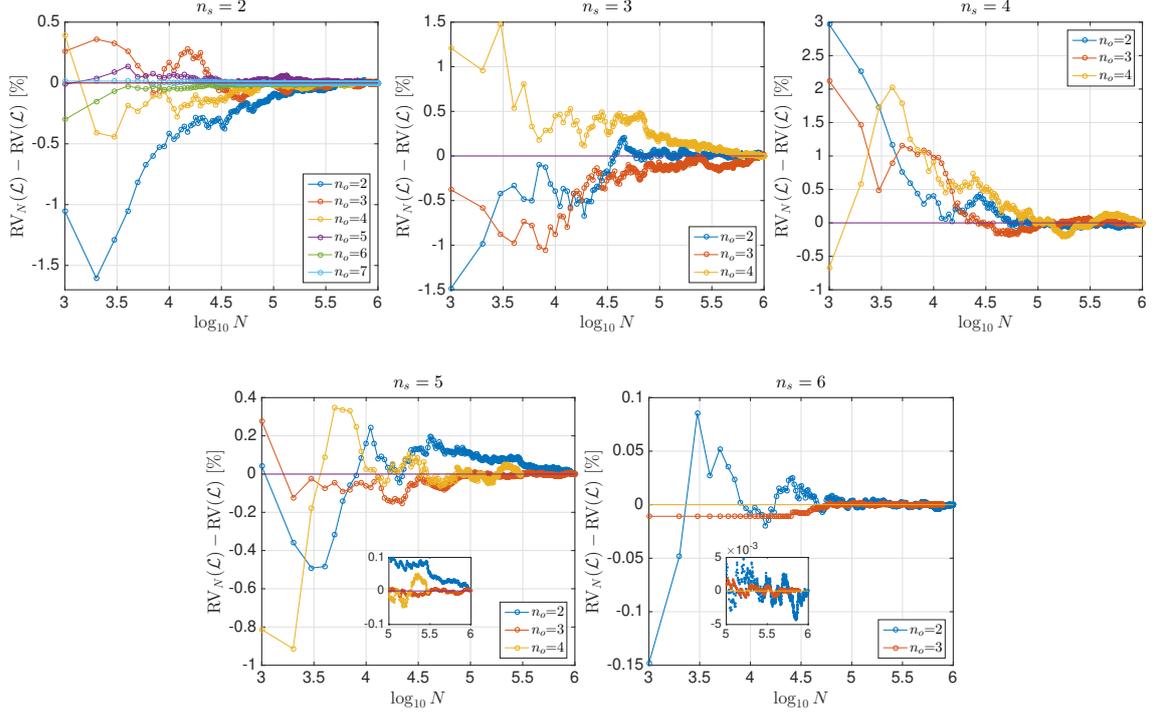
Figure 8: Plots showing how the *estimated* relative volume $\mathrm{RV}(\mathcal{L})$ converges to the reported value as the number of samples N used in the estimation increases towards N_{tot}. More precisely, in each subfigure, we plot $\mathrm{RV}_N(\mathcal{L}) - \mathrm{RV}(\mathcal{L})$ as a function of $\log_{10} N$, where $\mathrm{RV}_N(\mathcal{L})$ is the value of $\mathrm{RV}(\mathcal{L})$ estimated using the first N correlations $\vec{P}$ sampled uniformly from $\mathcal{NS}$. As a reference, note that $\mathrm{RV}(\mathcal{L})$ lies, respectively, within $[93\%,\ 100\%]$, $[59\%,\ 100\%]$, $[16\%,\ 81\%]$, $[1\%,\ 7\%]$, and $[0.01\%,\ 0.15\%]$ for $n_s = 2, 3, 4, 5,$ and 6. For these two latter cases, an inset showing the plots from $\log_{10} N \in [5,\ 6]$ is included to show that $\mathrm{RV}(\mathcal{L})$ has converged to within 10% of its value. As an alternative way to see that our estimated RVs have converged well, one can also employ, e.g., the Wilson score interval [72] to verify that our estimate fits well within the 99% confidence interval of the corresponding estimate.

a part in the complexity of the problem. In Table 7, we provide a summary of the Bell scenarios considered in this work as well as these key parameters relevant to solving the corresponding optimization problems.

D Convergence analysis

In Fig. 8, we provide details showing how our estimates of the relative volume converge as a function of N_{tot}, i.e., the number of sampled correlations.

E Distance from $\vec{P}_w$ to extreme points of $\mathcal{NS}$

In the probability space $\mathscr{P}$ parametrized by all the full conditional distributions $\{P(a,b|x,y)\}_{a,b,x,y}$, the Euclidean distance between two correlations $\vec{P}_1$ and $\vec{P}_2$ is given by:

$$\mathfrak{D}_{\mathrm{E}}(\vec{P}_1|\vec{P}_2) = \sqrt{\sum_{a,b,x,y} \left[P_1(a,b|x,y) - P_2(a,b|x,y)\right]^2}. \tag{22}$$

In any Bell scenario, any Bell-local extreme point of $\mathcal{NS}$, which is also an extreme point of $\mathcal{L}$, can be obtained from

$$P_{\mathcal{L}}(a,b|x,y) = \delta_{a,1}\delta_{b,1} \tag{23}$$

Accepted in 〈ʒuantum 2022-07-10, click title to verify. Published under CC-BY 4.0.

22

via a relabeling of measurement settings, outcomes, and/or parties. Since the uniform distribution $P_w(a,b|x,y) = \frac{1}{n_o^2}$ is always invariant under such a relabeling, we see that their Euclidean distance is invariant under relabeling, and is easily shown to be $\mathfrak{D}_{\mathrm{E}}(\vec{P}_{\mathcal{L}}|\vec{P}_w) = n_s\sqrt{1 - \frac{1}{n_o^2}}$.

Similarly, in the $(2, n_o)$ Bell scenarios, all nonlocal extreme points of $\mathcal{NS}$ can be obtained from one of the followings [10]:

$$P_{\mathcal{NS}}^k(a,b|x,y) = \frac{1}{k}\delta_{(b-a)\bmod k, \, xy}, \quad k \in 2,\ldots,n_o \tag{24}$$

via a relabeling. It then follows from Eq. (22) that $\mathfrak{D}_{\mathrm{E}}(\vec{P}_{\mathcal{NS}}^k|\vec{P}_w) = 2\sqrt{\frac{1}{k} - \frac{1}{n_o^2}}$. Comparing with $\mathfrak{D}_{\mathrm{E}}(\vec{P}_{\mathcal{L}}|\vec{P}_w)$, we thus see that all nonlocal extreme points in the $(2, n_o)$ Bell scenario are actually nearer to $\mathcal{P}_w$ than the local ones.

References

[1] A. Acín. Statistical distinguishability between unitary operations. *Phys. Rev. Lett.*, 87:177901, Oct 2001. DOI: 10.1103/PhysRevLett.87.177901.

[2] Antonio Acín. (private communication).

[3] Antonio Acín, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. Device-independent security of quantum cryptography against collective attacks. *Phys. Rev. Lett.*, 98:230501, Jun 2007. DOI: 10.1103/PhysRevLett.98.230501.

[4] Rotem Arnon-Friedman and Jean-Daniel Bancal. Device-independent certification of one-shot distillable entanglement. *New J. Phys.*, 21(3):033010, 2019. DOI: 10.1088/1367-2630/aafef6.

[5] Davd Avis. *lrs*: A revised implementation of the reverse search vertex enumeration algorithm. (unpublished), 1999. URL http://cgm.cs.mcgill.ca/~avis/doc/avis/Av98a.pdf.

[6] Jean-Daniel Bancal, Nicolas Gisin, Yeong-Cherng Liang, and Stefano Pironio. Device-independent witnesses of genuine multipartite entanglement. *Phys. Rev. Lett.*, 106:250404, Jun 2011. DOI: 10.1103/PhysRevLett.106.250404.

[7] Jean-Daniel Bancal, Nicolas Sangouard, and Pavel Sekatski. Noise-resistant device-independent certification of Bell state measurements. *Phys. Rev. Lett.*, 121:250506, Dec 2018. DOI: 10.1103/PhysRevLett.121.250506.

[8] Tomer Jack Barnea, Jean-Daniel Bancal, Yeong-Cherng Liang, and Nicolas Gisin. Tripartite quantum state violating the hidden-influence constraints. *Phys. Rev. A*, 88:022123, Aug 2013. DOI: 10.1103/PhysRevA.88.022123.

[9] Jonathan Barrett. Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality. *Phys. Rev. A*, 65:042302, Mar 2002. DOI: 10.1103/PhysRevA.65.042302.

[10] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, Feb 2005. DOI: 10.1103/PhysRevA.71.022101.

[11] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, Nov 1964. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[12] J. S. Bell. *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*. Cambridge University Press, 2 edition, 2004. DOI: 10.1017/CBO9780511815676.

[13] Tim Benham. Uniform distribution over a convex polytope. MATLAB central file exchange, 2014. URL https://www.mathworks.com/matlabcentral/fileexchange/34208-uniform-distribution-over-a-convex-polytope.

[14] Mario Berta, Omar Fawzi, and Volkher B. Scholz. Quantum bilinear optimization. *Siam J. Optim.*, 26(3):1529–1564, 2020/04/04 2016. DOI: 10.1137/15M1037731.

[15] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 1 edition, 2004.

[16] Gilles Brassard, Harry Buhrman, Noah Linden, André Allan Méthot, Alain Tapp, and Falk Unger. Limit on nonlocality in any world in which communication complexity is not trivial. *Phys. Rev. Lett.*, 96:250401, Jun 2006. DOI: 10.1103/PhysRevLett.96.250401.

[17] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014. DOI: 10.1103/RevModPhys.86.419.

[18] Benno Büeler, Andreas Enge, and Komei Fukuda. *Exact Volume Computation for Polytopes: A Practical Study*, pages 131–154. Birkhäuser Basel, Basel, 2000. ISBN 978-3-0348-8438-9. DOI: 10.1007/978-3-0348-8438-9_6.

[19] Adán Cabello. How much larger quantum correlations are than classical ones. *Phys. Rev. A*, 72:012113, Jul 2005. DOI: 10.1103/PhysRevA.72.012113.

[20] Shin-Liang Chen, Costantino Budroni, Yeong-Cherng Liang, and Yueh-Nan Chen. Natural framework for device-independent quantification of quantum steerability, measurement incompatibility, and self-testing. *Phys. Rev. Lett.*, 116:240401, Jun 2016. DOI: 10.1103/PhysRevLett.116.240401.

[21] Shin-Liang Chen, Costantino Budroni, Yeong-Cherng Liang, and Yueh-Nan Chen. Exploring the framework of assemblage moment matrices and its applications in device-independent characterizations. *Phys. Rev. A*, 98:042127, Oct 2018. DOI: 10.1103/PhysRevA.98.042127.

[22] Shin-Liang Chen, Costantino Budroni, Yeong-Cherng Liang, and Yueh-Nan Chen. Exploring the framework of assemblage moment matrices and its applications in device-independent characterizations. *Phys. Rev. A*, 98:042127, Oct 2018. DOI: 10.1103/PhysRevA.98.042127.

[23] Shin-Liang Chen, Nikolai Miklin, Costantino Budroni, and Yueh-Nan Chen. Device-independent quantification of measurement incompatibility. *Phys. Rev. Research*, 3:023143, May 2021. DOI: 10.1103/PhysRevResearch.3.023143.

[24] Bradley G. Christensen, Yeong-Cherng Liang, Nicolas Brunner, Nicolas Gisin, and Paul G. Kwiat. Exploring the limits of quantum nonlocality with entangled photons. *Phys. Rev. X*, 5: 041052, Dec 2015. DOI: 10.1103/PhysRevX.5.041052.

[25] Andrea Coladangelo and Jalex Stark. An inherently infinite-dimensional quantum correlation. *Nat. Commun.*, 11(1):3335, 2020. DOI: 10.1038/s41467-020-17077-9.

[26] Roger Colbeck. *Quantum And Relativistic Protocols For Secure Multi-Party Computation*. PhD thesis, University of Cambridge, 2006. URL https://doi.org/10.48550/arXiv.0911.3814.

[27] Daniel Collins and Nicolas Gisin. A relevant two qubit Bell inequality inequivalent to the CHSH inequality. *J. Phys. A: Math. Theo.*, 37(5):1775, 2004. DOI: 10.1088/0305-4470/37/5/021.

[28] Florian John Curchod, Nicolas Gisin, and Yeong-Cherng Liang. Quantifying multipartite nonlocality via the size of the resource. *Phys. Rev. A*, 91:012121, Jan 2015. DOI: 10.1103/PhysRevA.91.012121.

[29] Andrew C. Doherty, Yeong-Cherng Liang, Ben Toner, and Stephanie Wehner. The quantum moment problem and bounds on entangled multi-prover games. In *23rd Annu. IEEE Conf. on Comput. Comp, 2008, CCC'08*, pages 199–210, Los Alamitos, CA, 2008. DOI: 10.1109/CCC.2008.26.

[30] Cristhiano Duarte, Samuraí Brito, Barbara Amaral, and Rafael Chaves. Concentration phenomena in the geometry of Bell correlations. *Phys. Rev. A*, 98:062114, Dec 2018. DOI: 10.1103/PhysRevA.98.062114.

[31] Arthur Fine. Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.*, 48:291–295, Feb 1982. DOI: 10.1103/PhysRevLett.48.291.

[32] T. Fritz, A. B. Sainz, R. Augusiak, J. Bohr Brask, R. Chaves, A. Leverrier, and A. Acín. Local orthogonality as a multipartite principle for quantum correlations. *Nat. Commun.*, 4 (1):2263, 2013. ISSN 2041-1723. DOI: 10.1038/ncomms3263.

[33] Koon Tong Goh, Jędrzej 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.

[34] Tomáš Gonda, Ravi Kunjwal, David Schmid, Elie Wolfe, and Ana Belén Sainz. Almost Quantum Correlations are Inconsistent with Specker's Principle. *Quantum*, 2:87, August 2018. ISSN 2521-327X. DOI: 10.22331/q-2018-08-27-87.

[35] Lucien Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Phys. Rev. Lett.*, 71:1665–1668, Sep 1993. DOI: 10.1103/PhysRevLett.71.1665.

[36] Aram W. Harrow, Anand Natarajan, and Xiaodi Wu. Limitations of semidefinite programs for separable states and entangled games. *Commun. Math. Phys.*, 366(2):423–468, 2019. ISSN 1432-0916. DOI: 10.1007/s00220-019-03382-y.

[37] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80:5239–5242, Jun 1998. DOI: 10.1103/PhysRevLett.80.5239.

[38] M. Junge and C. Palazuelos. Large violation of bell inequalities with low entanglement. *Commun. Math. Phys.*, 306(3):695, 2011. DOI: 10.1007/s00220-011-1296-8.

[39] Ben Lang, Tamás Vértesi, and Miguel Navascués. Closed sets of correlations: answers from the zoo. *J. Phys. A Math. Theor.*, 47(42):424029, oct 2014. DOI: 10.1088/1751-8113/47/42/424029.

[40] Yeong-Cherng Liang, Tamás Vértesi, and Nicolas Brunner. Semi-device-independent bounds on entanglement. *Phys. Rev. A*, 83:022108, Feb 2011. DOI: 10.1103/PhysRevA.83.022108.

[41] Yeong-Cherng Liang, Denis Rosset, Jean-Daniel Bancal, Gilles Pütz, Tomer Jack Barnea, and Nicolas Gisin. Family of Bell-like inequalities as device-independent witnesses for entanglement depth. *Phys. Rev. Lett.*, 114:190401, May 2015. DOI: 10.1103/PhysRevLett.114.190401.

[42] Noah Linden, Sandu Popescu, Anthony J. Short, and Andreas Winter. Quantum nonlocality and beyond: Limits from nonlocal computation. *Phys. Rev. Lett.*, 99:180502, Oct 2007. DOI: 10.1103/PhysRevLett.99.180502.

[43] He Lu, Qi Zhao, Zheng-Da Li, Xu-Fei Yin, Xiao Yuan, Jui-Chen Hung, Luo-Kan Chen, Li Li, Nai-Le Liu, Cheng-Zhi Peng, Yeong-Cherng Liang, Xiongfeng Ma, Yu-Ao Chen, and Jian-Wei Pan. Entanglement structure: Entanglement partitioning in multipartite systems and its experimental detection using optimizable witnesses. *Phys. Rev. X*, 8:021072, Jun 2018. DOI: 10.1103/PhysRevX.8.021072.

[44] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *Quantum Info. Comput.*, 4(4):273–286, July 2004. ISSN 1533-7146. URL http://dl.acm.org/citation.cfm?id=2011827.2011830.

[45] 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.

[46] Miguel Navascués and Harald Wunderlich. A glance beyond the quantum model. *Proc. R. Soc. A*, 466:881, Nov 2009. URL https://doi.org/10.1098/rspa.2009.0453.

[47] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the set of quantum correlations. *Phys. Rev. Lett.*, 98:010401, Jan 2007. DOI: 10.1103/PhysRevLett.98.010401.

[48] Miguel Navascués, Stefano Pironio, and Antonio Acín. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.*, 10(7):073013, 2008. DOI: 10.1088/1367-2630/10/7/073013.

[49] Miguel Navascués, Yelena Guryanova, Matty J. Hoban, and Antonio Acín. Almost quantum correlations. *Nat. Commun.*, 6:6288, 2015. DOI: https://doi.org/10.1038/ncomms7288.

[50] Marcin Pawlowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Zukowski. Information causality as a physical principle. *Nature*, 461(7267):1101–1104, 2009. ISSN 1476-4687. DOI: 10.1038/nature08400.

[51] Asher Peres. Neumark's theorem and quantum inseparability. *Found. Phys.*, 20(12):1441–1453, 1990. DOI: 10.1007/BF01883517.

[52] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, Aug 1996. DOI: 10.1103/PhysRevLett.77.1413.

[53] Asher Peres. All the Bell inequalities. *Found. Phys.*, 29(4):589–614, 1999. DOI: 10.1023/A:1018816310000.

[54] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe. Random numbers certified by Bell's theorems theorem. *Nature (London)*, 464:1021, April 2010. DOI: 10.1038/nature09008.

[55] Itamar Pitowsky. *Quantum Probability - Quantum Logic*. Springer, Berlin, 1989.

[56] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Found. Phys.*, 24(3):379–385, Mar 1994. ISSN 1572-9516. DOI: 10.1007/BF02058098.

[57] Rafael Rabelo, Melvyn Ho, Daniel Cavalcanti, Nicolas Brunner, and Valerio Scarani. Device-independent certification of entangled measurements. *Phys. Rev. Lett.*, 107:050502, Jul 2011. DOI: 10.1103/PhysRevLett.107.050502.

[58] Valerio Scarani. The device-independent outlook on quantum physics. *Acta Physica Slovaca*, 62(4):347, 2012.

[59] Pavel Sekatski, Jean-Daniel Bancal, Sebastian Wagner, and Nicolas Sangouard. Certifying the building blocks of quantum computers from Bell's theorem. *Phys. Rev. Lett.*, 121:180505, Nov 2018. DOI: 10.1103/PhysRevLett.121.180505.

[60] Jamie Sikora and Antonios Varvitsiotis. Linear conic formulations for two-party correlations and values of nonlocal games. *Math. Program., Ser. A*, 162(1):431–463, 2017. DOI: 10.1007/s10107-016-1049-8.

[61] William Slofstra. The set of quantum correlations is not closed. *Forum of Mathematics, Pi*, 7:e1, 2019. DOI: 10.1017/fmp.2018.3.

[62] William Slofstra. Tsirelson's problem and an embedding theorem for groups arising from non-local games. *J. Amer. Math. Soc.*, 33:1–56, 2020. DOI: 10.1090/jams/929.

[63] James Vallins, Ana Belén Sainz, and Yeong-Cherng Liang. Almost-quantum correlations and their refinements in a tripartite Bell scenario. *Phys. Rev. A*, 95:022111, Feb 2017. DOI: 10.1103/PhysRevA.95.022111.

[64] Tamás Vértesi and Nicolas Brunner. Quantum nonlocality does not imply entanglement distillability. *Phys. Rev. Lett.*, 108:030403, Jan 2012. DOI: 10.1103/PhysRevLett.108.030403.

[65] Tamas Vertesi and Nicolas Brunner. Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement. *Nat. Commun.*, 5:5297, 05 2014. DOI: 10.1038/ncomms6297.

[66] Thomas Vidick and Stephanie Wehner. More nonlocality with less entanglement. *Phys. Rev. A*, 83:052310, May 2011. DOI: 10.1103/PhysRevA.83.052310.

[67] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review. *Quantum*, 4:337, Sep 2020. ISSN 2521-327X. DOI: 10.22331/q-2020-09-30-337.

[68] Sebastian Wagner, Jean-Daniel Bancal, Nicolas Sangouard, and Pavel Sekatski. Device-independent characterization of quantum instruments. *Quantum*, 4:243, March 2020. ISSN 2521-327X. DOI: 10.22331/q-2020-03-19-243.

[69] R. F. Werner and M. M. Wolf. Bell's inequalities for states with positive partial transpose. *Phys. Rev. A*, 61:062102, May 2000. DOI: 10.1103/PhysRevA.61.062102.

[70] R. F. Werner and M. M. Wolf. All-multipartite Bell-correlation inequalities for two dichotomic observables per site. *Phys. Rev. A*, 64:032112, Aug 2001. DOI: 10.1103/PhysRevA.64.032112.

[71] Reinhard F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, Oct 1989. DOI: 10.1103/PhysRevA.40.4277.

[72] Edwin B. Wilson. Probable inference, the law of succession, and statistical inference. *J. Amer. Statist. Assoc*, 22(158):209–212, 06 1927. DOI: 10.1080/01621459.1927.10502953.

[73] H M Wiseman. The two Bell's theorems of John Bell. *J. Phys. A Math. Theor.*, 47(42): 424001, 2014. DOI: 10.1088/1751-8113/47/42/424001.

[74] Peter Wittek. Algorithm 950: Ncpol2sdpa—sparse semidefinite programming relaxations for polynomial optimization problems of noncommuting variables. *ACM Trans. Math. Softw.*, 41 (3), jun 2015. ISSN 0098-3500. DOI: 10.1145/2699464.

[75] Elie Wolfe and S. F. Yelin. Quantum bounds for inequalities involving marginal expectation values. *Phys. Rev. A*, 86:012123, Jul 2012. DOI: 10.1103/PhysRevA.86.012123.