

Expanding bipartite Bell inequalities for maximum multipartite randomness

Lewis Woollorton^{1,2,3}, Peter Brown⁴, and Roger Colbeck^{5,1}

¹Department of Mathematics, University of York, Heslington, York, YO10 5DD, United Kingdom

²Quantum Engineering Centre for Doctoral Training, H. H. Wills Physics Laboratory and Department of Electrical & Electronic Engineering, University of Bristol, Bristol BS8 1FD, United Kingdom

³Inria, ENS de Lyon, LIP, 46 Allée d'Italie, 69364 Lyon Cedex 07, France

⁴Télécom Paris, LTCI, Institut Polytechnique de Paris, 19 Place Marguerite Perey, 91120 Palaiseau, France

⁵Department of Mathematics, King's College London, Strand, London, WC2R 2LS, United Kingdom

1st December 2025

Nonlocal tests on multi-partite quantum correlations form the basis of protocols that certify randomness in a device-independent (DI) way. Such correlations admit a rich structure, making the task of choosing an appropriate test difficult. For example, extremal Bell inequalities are tight witnesses of nonlocality, but achieving their maximum violation places constraints on the underlying quantum system, which can reduce the rate of randomness generation. As a result there is often a trade-off between maximum randomness and the amount of violation of a given Bell inequality. Here, we explore this trade-off for more than two parties. More precisely, we study the maximum amount of randomness that can be certified by correlations with a particular violation of the Mermin-Ardehali-Belinskii-Klyshko (MABK) inequality. For any even number of parties, we find that maximum randomness cannot occur beyond a threshold quantum violation, which increases with the number of parties, and we give a conjectured form of the maximum randomness in terms of the MABK value. We also show that maximum randomness can be obtained for any MABK violation for odd numbers of parties. To obtain our results, we derive new families of Bell inequalities certifying maximum randomness from a technique for randomness certification, which we call “expanding Bell inequalities”. Our technique allows a bipartite Bell expression to be used as a seed, and transformed into a multi-partite Bell inequality tailored for randomness certification, showing how intuition learned in the bipartite case can find use in more complex scenarios.

Lewis Woollorton: lewis.woollorton@ens-lyon.fr

Peter Brown: peter.brown@telecom-paris.fr

Roger Colbeck: roger.colbeck@kcl.ac.uk

1 Introduction

Distant measurements on a shared quantum system can display correlations inaccessible to any locally realistic model [1, 2]. Such correlations are termed nonlocal and provide a device-independent (DI) witness of useful quantum characteristics, opening up a new paradigm of information processing [3]. Tasks such as randomness expansion [4, 5, 6, 7, 8], amplification [9] and key distribution [10, 11, 12, 13, 14, 15] can all be achieved in the DI regime, where security is proven based on the observed correlations and without trusting the inner workings of the devices. Moreover, nonlocal correlations can be used to prove the presence of a particular state or sets of measurements, known as self-testing [16, 17, 18, 19, 20, 21].

Certification of DI randomness typically uses a Bell inequality, and different Bell inequalities give different amounts of certified randomness. It is natural to consider extremal Bell inequalities, which form the minimal set separating local and non-local correlations. The set of correlations compatible with high violation is tightly constrained and can negatively affect the certification of randomness. For example, in the simplest bipartite scenario, correlations with higher violation of the only extremal Bell inequality, the Clauser-Horne-Shimony-Holt (CHSH) inequality, cannot certify maximal DI randomness, while those with a lower violation can [22, 23]. Further complications arise when scenarios with more inputs, outputs or parties are considered because the number of different classes of extremal Bell inequalities grows and it becomes more difficult to navigate the trade-offs between extremal Bell violation and maximum certifiable randomness. Alongside foundational interest, this question is also motivated practically, in finding more robust device-independent randomness expansion (DIRE) protocols [23].

In this work we study the family of multi-partite Bell inequalities due to Mermin, Ardehali, Belinskii, and Klyshko (MABK) [24, 25, 26] and their application to DIRE. These all have two inputs and two

outputs per party. The one and two sided randomness given a tripartite MABK violation was bounded in Refs. [27, 28], and compared to other tripartite inequalities in Ref. [29]. Considering global randomness, Ref. [30] gave a bipartite construction for certification of M bits of randomness using M copies of almost unentangled states.

Ref. [31] showed how, for an odd number of parties N , maximum MABK violation can certify N bits of DI randomness, whilst Ref. [32] provided an alternative family of Bell inequalities certifying N bits in the even case. This leaves open the following questions: how suitable is MABK for maximum randomness generation when N is even? Does the MABK violation trade-off with the maximum certifiable randomness, as was shown for the CHSH inequality [23]? Similarly, are there correlations with a non-maximal MABK violation that achieve maximum randomness when N is odd?

To analyse multi-partite nonlocality, we require tools to construct multi-partite Bell tests. To do so, one approach is to generalize intuition from the well understood CHSH scenario; this has been done to extend the number of measurements per party [33, 34], or the number of parties [35], where it was shown how to construct multi-partite Bell inequalities from a bipartite seed. Self-testing results using projections onto bipartite subsystems [36, 37, 38] suggest stronger capabilities of the technique presented in Ref. [35], prompting the question: can it be leveraged for multi-partite randomness certification?

Our work seeks to answer these questions. We first show that maximum MABK violation certifies $N + 1/2 - \log_2(1 + \sqrt{2})/\sqrt{2} \approx N - 0.4$ bits of global randomness when N is even, contrasting the N bits certified when N is odd [31]. We then investigate the limits on certifiable randomness from quantum behaviours that violate an MABK inequality. To do so, we construct new families of N -partite Bell inequalities whose maximum violation certifies N bits for arbitrary N . When N is even, we show that for any MABK violation m up to a non-maximal value m^* , there exists a quantum behaviour achieving m which maximally violates an inequality from our family, and hence certifies maximum global randomness. For violations above m^* , we find (using a second family of inequalities) that randomness monotonically decreases with MABK violation. Our results lower bound this trade-off, and we provide numerical evidence of tightness, indicating an extension of the known tight bound when $N = 2$ [23] to arbitrary N . Additionally, we show that this trade-off diminishes (in the sense that the gap between m^* and maximal MABK violation decreases) as the number of parties grows. Finally, for odd N , we use our family of Bell inequalities to show that N bits of randomness can be certified for a range of MABK violations.

Our constructions are formed from a type of con-

catenation of bipartite Bell inequalities, allowing us to lift our intuition derived from the bipartite case [23] to the multipartite setting. Specifically, we extend the technique presented in Ref. [35], originally introduced to witness genuinely multi-partite nonlocality. Our extension exploits self-testing properties of the seed to certify randomness. Specifically, our main technical contribution is a decoupling lemma (see Lemma 2), which shows that under certain conditions the maximum violation of an expanded Bell expression implies a decoupling between the honest parties and Eve. Such decoupling occurs for extremal quantum correlations, and guarantees security [39]. We envisage that this technique will be a useful tool in multi-partite DI cryptography and of independent interest.

The paper is structured as follows. In Section 2 we provide the necessary background and notation. In Section 3 we detail our enhancement of the technique from [35], along with the decoupling lemma. In Section 4 we describe our constructions, and show that they certify N bits of randomness for a broad range of MABK violations, which we conjecture to be optimal; this result is summarized in Lemmas 5 and 6. We additionally show how the conjectured highest value of MABK at which maximum randomness remains possible tends to the maximum quantum value as the number of parties increases (see Proposition 6). In Section 5 we consider MABK values beyond the conjectured threshold for maximum randomness, up to the maximum quantum value. Here we give a candidate form for the exact trade-off, which is summarized in Lemma 9. We conclude with a discussion in Section 6. All technical proofs can be found in the appendices, and we provide an accompanying python script for the numerical aspects at the GitHub repository [40].

2 Background

2.1 Multiparty DI-scenario

We consider an N party, 2-input 2-output DI scenario, where N isolated devices are given a random input, labelled $x_k \in \{0, 1\}$, and produce an output $a_k \in \{0, 1\}$ stored in a classical register R_k , where $k \in \{1, \dots, N\}$ indexes the party. We use bold to denote tuples; for example, $\mathbf{a} = (a_1, \dots, a_N)$ denotes an N bit string of device outputs. The behaviour of the devices is then described by the joint conditional probability distribution $p(\mathbf{a}|\mathbf{x})$, which must be no-signalling because the devices are isolated.

A behaviour $p(\mathbf{a}|\mathbf{x})$ is quantum if there exists (following Naimark's dilation theorem [41]) a pure state and sets of orthonormal projectors that can reproduce the distribution via the Born rule. Specifically, we consider an adversary Eve, who wishes to guess the device outputs, and let $|\Psi\rangle_{\tilde{Q}E}$ denote the global state in the Hilbert space $\mathcal{H}_{\tilde{Q}} \otimes \mathcal{H}_E$, where $\mathcal{H}_E$ is

held by Eve, $\mathcal{H}_{\tilde{\mathbf{Q}}} = \bigotimes_{k=1}^N \mathcal{H}_{\tilde{\mathbf{Q}}_k}$ and $\mathcal{H}_{\tilde{\mathbf{Q}}_k}$ is held by device k . Throughout the text, tildes will denote elements pertaining to physical Hilbert spaces $\mathcal{H}_{\tilde{\mathbf{Q}}_k}$ (whose dimension is unknown), whilst no tilde will describe qubit Hilbert spaces $\mathcal{H}_{\mathbf{Q}_k}$. We let $\{\tilde{P}_{a_k|x_k}^{(k)}\}_{a_k}$ be binary-outcome projective measurements on $\mathcal{H}_{\tilde{\mathbf{Q}}_k}$, and we denote the corresponding observables of each party by $\tilde{A}_{x_k}^{(k)} = \tilde{P}_{0|x_k}^{(k)} - \tilde{P}_{1|x_k}^{(k)}$; here bracketed superscripts will typically keep track of the party to which the object belongs (i.e., the Hilbert space on which it acts). Following measurements $\mathbf{x}$ by the N honest parties, we obtain the classical quantum state $\rho_{\mathbf{R}E|\mathbf{x}} = \sum_{\mathbf{a}} |\mathbf{a}\rangle\langle\mathbf{a}|_{\mathbf{R}} \otimes \rho_E^{\mathbf{a}|\mathbf{x}}$, where $\rho_E^{\mathbf{a}|\mathbf{x}} = \text{Tr}_{\tilde{\mathbf{Q}}}[(\tilde{P}_{\mathbf{a}|\mathbf{x}} \otimes \mathbb{I}_E)|\Psi\rangle\langle\Psi|_{\tilde{\mathbf{Q}}E}]$ and $\tilde{P}_{\mathbf{a}|\mathbf{x}} = \bigotimes_{k=1}^N \tilde{P}_{a_k|x_k}^{(k)}$. The behaviour (as seen by the N honest parties) is recovered by $p(\mathbf{a}|\mathbf{x}) = \text{Tr}[\rho_E^{\mathbf{a}|\mathbf{x}}]$.

2.2 Multiparty Nonlocality

Given an observed distribution $p(\mathbf{a}|\mathbf{x})$, it will be useful to quantify its distance from the local boundary. When $N = 2$, the CHSH inequality, as the only non-trivial facet, is a natural choice. However, the multiparty scenario is more complex, and the number of classes of Bell inequality increases rapidly with N [42]. Instead, a general way to quantify this distance for an arbitrary scenario can be used, which is related

$$\begin{aligned} \langle M_2 \rangle &= \frac{1}{2} [\langle A_0 B_0 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_1 B_1 \rangle], \\ \langle M_3 \rangle &= \frac{1}{2} [\langle A_0 B_0 C_1 \rangle + \langle A_0 B_1 C_0 \rangle + \langle A_1 B_0 C_0 \rangle - \langle A_1 B_1 C_1 \rangle], \\ \langle M_4 \rangle &= \frac{1}{4} \left[-\langle A_0 B_0 C_0 D_0 \rangle + \langle A_0 B_0 C_0 D_1 \rangle + \langle A_0 B_0 C_1 D_0 \rangle + \langle A_0 B_1 C_0 D_0 \rangle + \langle A_1 B_0 C_0 D_0 \rangle + \langle A_0 B_0 C_1 D_1 \rangle \right. \\ &\quad + \langle A_1 B_1 C_0 D_0 \rangle + \langle A_0 B_1 C_0 D_1 \rangle + \langle A_1 B_0 C_1 D_0 \rangle + \langle A_0 B_1 C_1 D_0 \rangle + \langle A_1 B_0 C_0 D_1 \rangle \\ &\quad \left. - (\langle A_1 B_1 C_1 D_0 \rangle + \langle A_1 B_1 C_0 D_1 \rangle + \langle A_1 B_0 C_1 D_1 \rangle + \langle A_0 B_1 C_1 D_1 \rangle) - \langle A_1 B_1 C_1 D_1 \rangle \right]. \end{aligned} \quad (2)$$

Above, $\langle M_2 \rangle$ is the CHSH functional, whereas $\langle M_3 \rangle$ is the Mermin functional, associated with the GHZ paradox [43]. The property that all input strings of the same Hamming weight have the same coefficient makes the MABK expressions invariant under party relabelings. Additionally, note that for N even, every correlator $\langle A_{\mathbf{x}} \rangle$ for $\mathbf{x} \in \{0,1\}^n$ is included in the expression, whereas for N odd, exactly half of the correlators are included.

2.3 Self-testing

It is known that maximum quantum violation of the MABK family is uniquely achieved, up to local isometries, by the GHZ state and pairs of maximally anticommuting observables [44, 45]. This form of uniqueness arising from a Bell expression is known as self-testing. In this work, we take the choice

to how much the local polytope needs to be diluted to encompass a given non-local correlation. Computing this can be done using linear programming (see Section C.6).

In this work, we choose to study one such CHSH generalization and its relationship to DI randomness; the MABK family [24, 25, 26],

$$\langle M_N \rangle = 2^{\frac{1-N}{2}} \sum_{\mathbf{x}} \cos \left[\frac{\pi}{2} \left(\frac{N-1}{2} - \sum_{k=1}^N x_k \right) \right] \langle \tilde{A}_{\mathbf{x}} \rangle, \quad (1)$$

where $\langle \tilde{A}_{\mathbf{x}} \rangle = \sum_{\mathbf{a}} (-1)^{\sum_{k=1}^N a_k} p(\mathbf{a}|\mathbf{x}) = \langle \Psi | \tilde{A}_{x_1}^{(1)} \otimes \cdots \otimes \tilde{A}_{x_N}^{(N)} \otimes \mathbb{I}_E | \Psi \rangle$ when the behaviour is quantum.

[In effect the prefactor $(-1)^{\sum_{k=1}^N a_k}$ corresponds to relabelling the outcomes $0 \mapsto 1$ and $1 \mapsto -1$ to match the usual formulation of observables with eigenvalues ± 1 .] The local bound is given by $\langle M_N \rangle \leq 1$, and the maximum quantum value is $2^{(N-1)/2}$. Note that the MABK functional is characterized by the fact that the coefficients $c_{\mathbf{x}} = \cos \left[\frac{\pi}{2} \left(\frac{N-1}{2} - \sum_{k=1}^N x_k \right) \right]$ are equal for all strings $\mathbf{x}$ with the same Hamming weight (that is, the number of 1s in the string $\mathbf{x}$). When N is even, $c_{\mathbf{x}} \in \{-1/\sqrt{2}, 1/\sqrt{2}\}$, and when N is odd $c_{\mathbf{x}} \in \{-1, 0, +1\}$. This is illustrated by writing out the first few cases. To simplify the notation, we let $A_x^{(1)} = A_x$, $A_y^{(2)} = B_y$, $A_z^{(3)} = C_z$ and $A_t^{(4)} = D_t$.

$|\psi_{\text{GHZ}}\rangle = (|0\rangle^{\otimes N} + i|1\rangle^{\otimes N})/\sqrt{2}$, $A_0^{(k)} = \cos \theta_N^+ \sigma_X + \sin \theta_N^+ \sigma_Y$, and $A_1^{(k)} = \cos \theta_N^- \sigma_X + \sin \theta_N^- \sigma_Y$, where $\theta_N^{\pm} = (\pi/4)(1/N \pm 1)$.

Whilst self-testing statements can be formally defined between N parties (see, e.g., Refs. [36, 20]), in our formulation it will only be necessary to use a bipartite definition.

Definition 1 (Bipartite self-test). Let $k, l \in \{1, \dots, N\}$ index two distinct parties. Define the sets of target qubit projectors $\{P_{a_k|x_k}^{(k)}\}_{a_k}$, $\{P_{a_l|x_l}^{(l)}\}_{a_l}$, and a target two qubit state $|\Phi\rangle_{Q_k Q_l}$. Let $I^{(k,l)}$ be a Bell operator between parties k and l , with maximum quantum value η^Q . The inequality $\langle I^{(k,l)} \rangle \leq \eta^Q$ self-tests the target state and measurements if, for all physical quantum strategies $(\tilde{\rho}_{\tilde{\mathbf{Q}}_k \tilde{\mathbf{Q}}_l}, \tilde{P}_{a_k|x_k}^{(k)}, \tilde{P}_{a_l|x_l}^{(l)})$ that satisfy $\langle I^{(k,l)} \rangle = \eta^Q$, there exists a local isome-

tery $V : \mathcal{H}_{\tilde{Q}_k} \otimes \mathcal{H}_{\tilde{Q}_l} \otimes \mathcal{H}_E \rightarrow \mathcal{H}_{Q_k} \otimes \mathcal{H}_{Q_l} \otimes \mathcal{H}_{\text{Junk}}$ and ancillary state $|\xi\rangle_{\text{Junk}}$, such that

$$\begin{aligned} V\left(\tilde{P}_{a_k|x_k}^{(k)} \otimes \tilde{P}_{a_l|x_l}^{(l)} \otimes \mathbb{I}_E\right)|\Psi\rangle_{\tilde{Q}_k\tilde{Q}_lE} \\ = \left(P_{a_k|x_k}^{(k)} \otimes P_{a_l|x_l}^{(l)}\right)|\Phi\rangle_{Q_kQ_l} \otimes |\xi\rangle_{\text{Junk}}, \end{aligned} \quad (3)$$

for all a_k, a_l, x_k, x_l , where $|\Psi\rangle$ is a purification of $\tilde{\rho}$.

We define the shifted Bell operator as $\bar{I}^{(k,l)} = \eta^Q \mathbb{I} - I^{(k,l)}$, and we say $\bar{I}^{(k,l)}$ admits a sum-of-squares (SOS) decomposition if there exists a set of polynomials, $\{M_i^{(k,l)}\}_i$, of the operators $\tilde{P}_{a_k|x_k}^{(k)}, \tilde{P}_{a_l|x_l}^{(l)}$, satisfying

$$\bar{I}^{(k,l)} = \sum_i M_i^{(k,l)\dagger} M_i^{(k,l)}. \quad (4)$$

SOS decompositions can be used to enforce algebraic constraints on any quantum state $\tilde{\rho}$, and measurements $\tilde{P}_{a_k|x_k}^{(k)}, \tilde{P}_{a_l|x_l}^{(l)}$, satisfying $\langle \bar{I}^{(k,l)} \rangle = 0$. For example, when $\tilde{\rho} = |\psi\rangle\langle\psi|$ is pure, it must satisfy

$$M_i^{(k,l)}|\psi\rangle = 0, \quad \forall i, \quad (5)$$

(or $\text{Tr}[(M_i^{(k,l)})^\dagger M_i^{(k,l)} \tilde{\rho}] = 0$ more generally). Typically, we find that constraints of the form in Eq. (5) are only satisfied by a unique state (up to local isometries), and when that is the case we call Eq. (5) the self-testing criteria.

2.4 DI randomness certification

We use the conditional von Neumann entropy, $H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*, E)_{\rho_{\mathbf{R}E|\mathbf{x}^*}}$ to measure the asymptotic DI randomness generation rate, where for a bipartite state ρ_{AB} , $H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$ and $H(A)_\rho = -\text{tr}[\rho_A \log \rho_A]$ is the von Neumann entropy of a state ρ_A . $H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*, E)_{\rho_{\mathbf{R}E|\mathbf{x}^*}}$ is the correct quantity for spot-checking DI random number generation, where $\mathbf{x}^*$ is a specified measurement choice used to generate randomness (see [46] for a discussion of this and other possibilities). The asymptotic rate of DI randomness generation is given by a function $R_f(\omega)$ defined as

$$\inf_{\substack{|\Psi\rangle_{\tilde{Q}E}, \{\{\tilde{P}_{a_k|x_k}^{(k)}\}_{a_k}\}_k \\ \text{compatible with } f(P_{\text{obs}})=\omega}} H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*, E)_{\rho_{\mathbf{R}E|\mathbf{x}^*}}, \quad (6)$$

and we require lower bounds on this quantity over states and measurements compatible with a linear constraint on the observed statistics P_{obs} , $f(P_{\text{obs}}) = \omega$, where f defines a Bell expression and ω is its observed value. This definition can be straightforwardly extended to multiple functions. The asymptotic rate can be used as a basis for rates with finite statistics using tools such as the entropy accumulation theorem [47, 48, 49] (an alternative approach for handling finite statistics is the quantum probability estimation

framework [50], but this does not reduce to the single round von Neumann entropy).

Because we study fundamental limits on certifiable DI randomness we work in the noiseless scenario, using self-testing statements to certify DI randomness. Here, $f(P_{\text{obs}})$ is a Bell expression, and we show that all states and measurements that achieve its maximum violation correspond to a post-measurement state in tensor product with Eve, $\rho_{\mathbf{R}E|\mathbf{x}^*} = \rho_{\mathbf{R}|\mathbf{x}^*} \otimes \rho_E$. This allows us to directly evaluate the conditional entropy $H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*, E) = H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*)$, and the infimum becomes trivial since all compatible strategies give rise to the same classical distribution (see Appendix B.3 for details).

3 Expanding Bell inequalities

In this section we discuss and enhance the technique presented in [35], which allows us to derive the main results of this work. Ref. [35] introduced a method for building multi-partite Bell inequalities by expanding a bipartite inequality, called the “seed”, which can be used to witness genuinely multi-partite nonlocality. A new Bell expression is constructed by summing the seed over different subsets of parties, whilst the remaining parties perform some fixed measurement. The more multi-partite nonlocal the correlations, the more bipartite terms are violated. This resembles other recent results that enable multi-partite self-testing by projections onto bipartite subsystems [38, 36, 37].

In this work we extend this technique to make it suitable for DI cryptographic purposes. More precisely, we consider cases where the seed is a bipartite self-test, and use the maximum quantum violation of the expanded Bell expression, constructed in an equivalent manner to [35], to draw conclusions about the post-measurement state held between the honest parties and Eve. These conclusions allow us to derive rates for randomness certification.

3.1 Definition

We begin by introducing a generic formulation for expanding Bell expressions, shortly followed by an example.

Definition 2 (Expanded Bell expressions). Let $\mathbf{C}$ be an $N \times N$ nonzero matrix with entries $c_{k,l}$ satisfying $c_{k,l} \in \{0, 1\}$ if $k < l$, and $c_{k,l} = 0$ otherwise. For each pair k, l for which $c_{k,l} \neq 0$, let $\{I_\mu^{(k,l)}\}_\mu$ denote a set of bipartite Bell expressions between parties k and l , where $\mu \in \{0, 1\}^{N-2}$ is a set of $N - 2$ measurement outcomes of all parties excluding k, l . Further suppose each Bell expression is equivalent up to output relabellings, with a local bound η^L strictly less than the maximum quantum value η^Q . Then the Bell operator

I seeded by the set $\{I_{\mu}^{(k,l)}\}_{\mu}$ is defined as

$$I = \sum_{k,l} c_{k,l} \left(\sum_{\mu} \tilde{P}_{\mu|0}^{(k,l)} I_{\mu}^{(k,l)} \right), \quad (7)$$

where $\tilde{P}_{\mu|0}^{(k,l)} = \prod_{k' \in \{1, \dots, N\} \setminus \{k, l\}} \tilde{P}_{\mu_{k'}|0}^{(k')}$ is the projector for all parties excluding k and l , corresponding to the joint input setting $\mathbf{0}$ and joint outcome μ ¹.

To help understand the definition, consider the following tripartite example. To simplify the notation we label the observables of the three parties $A_x = P_{0|x}^A - P_{1|x}^A$ and similarly for B and C . From Ref. [23], observing saturation of the following Bell expression certifies 2 bits of global randomness,

$$-3\sqrt{3} \leq \langle I^{A,B} \rangle \leq 3\sqrt{3}, \quad (8)$$

where $I^{A,B} = A_0 B_0 + 2(A_0 B_1 + A_1 B_0 - A_1 B_1)$. The bound $\langle I^{A,B} \rangle = \pm 3\sqrt{3}$ is uniquely achieved, up to local isometries, by the bipartite strategy

$$\begin{aligned} \rho_{Q_A Q_B} &= |\Phi_{\pm}\rangle\langle\Phi_{\pm}|, \quad |\Phi_{\pm}\rangle = \frac{|00\rangle \pm i|11\rangle}{\sqrt{2}}, \\ A_0 &= B_0 = \sigma_X, \quad A_1 = B_1 = \frac{-\sigma_X + \sqrt{3}\sigma_Y}{2}. \end{aligned} \quad (9)$$

Now consider a tripartite extension of the above strategy,

$$\begin{aligned} \rho_{Q_A Q_B Q_C} &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0 &= B_0 = C_0 = \sigma_X, \quad A_1 = B_1 = C_1 = \frac{-\sigma_X + \sqrt{3}\sigma_Y}{2}. \end{aligned} \quad (10)$$

One can verify that when all three parties measure $x = y = z = 0$, their outcomes are uniformly distributed, resulting in 3 bits of raw randomness; how can we certify this device-independently? Notice $|\psi_{\text{GHZ}}\rangle$ has the following property: when a single party, say A , measures σ_X , the leftover state held by BC is $|\Phi_{\pm}\rangle$ where the sign depends on the parity of A 's measurement outcome $\mu \in \{0, 1\}$. Parties B and C can now saturate one of the bounds $\langle I^{B,C} \rangle = \pm 3\sqrt{3}$ conditioned on A 's measurement. By using $I_{\mu}^{A,B} = (-1)^{\mu} I^{A,B}$ as the seed, and choosing the matrix $\mathbf{C}$ with party indices $k, l \in \{A, B, C\}$,

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad (11)$$

¹We have written the Bell operator in Eq. (7) in terms of projectors, which is convenient since we consider quantum strategies. We can however rewrite it in a theory independent way, by taking the expectation value and making the substitution $p(\mathbf{a}|\mathbf{x}) = \langle \tilde{P}_{\mathbf{a}|\mathbf{x}} \rangle$.

we can construct the expanded Bell expression according to Definition 2:

$$\begin{aligned} I &= \underbrace{P_{\mu=0|x=0}^A I^{B,C} - P_{\mu=1|x=0}^A I^{B,C}}_{k=A, l=C} \\ &\quad + \underbrace{P_{\mu=0|y=0}^B I^{A,C} - P_{\mu=1|y=0}^B I^{A,C}}_{k=B, l=C} \\ &= A_0 I^{B,C} + B_0 I^{A,C}. \end{aligned} \quad (12)$$

Due to the properties of $|\psi_{\text{GHZ}}\rangle$ discussed above, we find $\langle I \rangle = 2 \cdot 3\sqrt{3}$ is the maximum quantum value of $\langle I \rangle$, and achieved by the strategy in Eq. (10). Moreover, it can be shown that this is strictly greater than the maximum local value of $\langle I \rangle$, implying $\langle I \rangle$ is a non-trivial Bell functional. In later sections, we will prove that $\langle I \rangle = 2 \cdot 3\sqrt{3}$ is a sufficient condition for certifying maximum randomness.

Returning to the general form of I from Eq. (7), we call I an expanded Bell expression, since it is built by combining a bipartite seed, $I_{\mu}^{(k,l)}$, conditioned on fixed measurement settings $\mathbf{0}$ and outcomes μ for the remaining $N-2$ parties. We then sum over all possible outcomes of this $N-2$ party measurement, changing the Bell expression accordingly, and over different combinations of parties (k, l) chosen according to $\mathbf{C}$. There needs to be a gap between the local and quantum bounds of $\langle I \rangle$ for I to define a nontrivial Bell inequality; typically, we find the following upper bound is achievable, which is strictly greater than the local bound:

Lemma 1. *Let I be an expanded Bell expression according to Definition 2. The maximum quantum value of $\langle I \rangle$ is upper bounded by $\eta_N^Q := \sum_{k,l} c_{k,l} \eta^Q$.*

A proof of Lemma 1 is given in Section B.1. If achievable, the only strategy that can give $\langle I \rangle = \eta_N^Q$ is the one satisfying $\langle \sum_{\mu} \tilde{P}_{\mu|0}^{(k,l)} I_{\mu}^{(k,l)} \rangle = \eta^Q$ for all pairing combinations k, l with $c_{k,l} > 0$, i.e., the reduced state held between parties k and l , following the projection $\tilde{P}_{\mu|0}^{(k,l)}$ on the global state, must achieve the maximum quantum value of $\langle I_{\mu}^{(k,l)} \rangle = \eta^Q$. This explains why we must include a dependence on μ for the Bell expression, since it should be tailored to the bipartite state following the outcome μ . For simplicity we assume each bipartite state is equivalent up to local unitaries, meaning we effectively use a single seed $I_{\mu}^{(k,l)}$. In principle however, one could use different seeds depending on the value of μ , which tailors the construction to non-symmetric states [35]. For our work we only consider the symmetric case. Additionally, one could further generalize by choosing specific measurement choices $\mathbf{y}$ dependent on the choice of non-projecting parties k, l , instead of fixing $\mathbf{y} = \mathbf{0}$ as we have done here.

3.2 Expanding Bell expressions and entropy evaluation

Next we present the characteristic of expanded Bell expressions that allows us to certify DI randomness from witnessing its maximum quantum value, forming the main technical contribution of our work.

Lemma 2 (Decoupling lemma). *Let I be an expanded N -party Bell expression defined in Definition 2 with binary inputs and outputs, and $c_{k,l} = 1$ if $l = N$ and $k < N$, and zero otherwise. Suppose for every $I_{\mu}^{(k,N)}$, there exists an SOS decomposition that self-tests the same pure bipartite entangled state $|\Phi\rangle$ between parties k and N , along with some ideal measurements $P_{a_k|x_k}^{(k)}, P_{a_N|x_N}^{(N)}$, according to Definition 1, satisfying $\langle \Phi | P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)} | \Phi \rangle > 0$ for all a_k, a_N . Then for any strategy $|\Psi\rangle_{\tilde{Q}E}, \{\{\tilde{P}_{a_k|x_k}^{(k)}\}_{x_k}\}_k$ that achieves $\langle I \rangle = \eta_N^Q$, the post-measurement state $\rho_{\mathbf{R}E|x}$, for measurement settings $\mathbf{x} = \mathbf{0}$, admits the tensor product decomposition,*

$$\rho_{\mathbf{R}E|\mathbf{0}} = \rho_{\mathbf{R}|\mathbf{0}} \otimes \rho_E. \quad (13)$$

Having established that Eve is decoupled it is then straightforward to evaluate the rate in Eq. (6) conditioned on observing the maximum quantum value of I (the quantity $R_I(\eta_N^Q)$ in the corollary below is an instance of a function of the form $R_f(\omega)$ defined in Eq. (6)).

Corollary 1. *Let I, η_N^Q be defined as in Lemma 2. Then*

$$R_I(\eta_N^Q) = H(\{p(\mathbf{a}|\mathbf{0})\}), \quad (14)$$

where $H(\{p_i\})$ is the Shannon entropy of a distribution $\{p_i\}_i$.

The proof can be found in Section B.3. Lemma 2 allows one to relate observing the maximum quantum value of the expanded Bell expression I , to a condition on the post-measurement state held by the parties and Eve, namely that it must be in tensor product with the purifying system E . This allows one to directly evaluate the conditional entropy according to Corollary 1.

4 Certifying N bits of DI randomness

We now consider constructions for generating N bits of DI randomness. As previously mentioned, when N is odd the MABK family can be used [31]. For an arbitrary number of parties, we apply the previously outlined techniques to derive a suitable Bell expression for this task. This will involve generalizing a one parameter family of quantum strategies, symmetric under permutation of the parties.

4.1 N odd

Using symmetry arguments, the authors of Ref. [31] showed that maximum violation of the MABK inequality, for N odd, implies maximum randomness.

Proposition 1 (Maximum randomness for N odd [31]). *When N is odd, maximum quantum violation of the MABK family of Bell inequalities, given by Eq. (1), certifies N bits of global DI randomness, i.e.*

$$R_{M_N}(2^{(N-1)/2}) = N. \quad (15)$$

From a self-testing point of view, maximum violation of the MABK family can only be achieved with a GHZ state and maximally anticommuting observables [44, 45]. For the form given in Eq. (1), recall the following strategy achieves maximum violation:

$$\begin{aligned} \rho_Q &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0^{(k)} &= \cos\theta_N^+ \sigma_X + \sin\theta_N^+ \sigma_Y, \\ A_1^{(k)} &= \cos\theta_N^- \sigma_X + \sin\theta_N^- \sigma_Y, \\ k &\in \{1, \dots, N\}, \end{aligned} \quad (16)$$

with $\theta_N^\pm = \pi/(4N) \pm \pi/4$, and satisfies $A_0^{(k)} A_1^{(k)} + A_1^{(k)} A_0^{(k)} = 0$ for each k . Following this self-testing property, the infimum in Eq. (15) reduces to a single point up to symmetries, and the fact that the parties hold a pure state implies that Eve must be uncorrelated, i.e., $\rho_{\mathbf{R}E|x^*} = \rho_{\mathbf{R}|x^*} \otimes \rho_E$. By direct calculation, one can verify for $N = 4n - 1$, $n = 1, 2, 3, \dots$, $p(\mathbf{a}|\mathbf{0}) = 1/2^N, \forall \mathbf{a}$, hence $H(\{p(\mathbf{a}|\mathbf{0})\}) = N$. Similarly for $N = 4n + 1$, $n = 1, 2, 3, \dots$, $p(\mathbf{a}|\mathbf{1}) = 1/2^N, \forall \mathbf{a}$, implying $H(\{p(\mathbf{a}|\mathbf{1})\}) = N$.

As an example, taking $N = 3$ in the above construction we have $\langle M_3 \rangle = (\langle \tilde{A}_1^{(0)} \tilde{A}_0^{(1)} \tilde{A}_0^{(2)} \rangle + \langle \tilde{A}_0^{(0)} \tilde{A}_1^{(1)} \tilde{A}_0^{(2)} \rangle + \langle \tilde{A}_0^{(0)} \tilde{A}_0^{(1)} \tilde{A}_1^{(2)} \rangle - \langle \tilde{A}_1^{(0)} \tilde{A}_1^{(1)} \tilde{A}_1^{(2)} \rangle)/2$, and the strategy in Eq. (16) reaches the algebraic bound of 2. We therefore have that $\langle M_3 \rangle$ implies 3 bits of randomness when all three parties use measurement 0. On the other hand, the correlators in the inequality must take the values ± 1 to reach the algebraic bound, so no string of inputs which appears as a correlator in the inequality can generate more than 2 bits of global randomness (in fact they give exactly 2 bits when $\langle M_3 \rangle = 2$).

4.2 N even

When N is even, all combinations $\langle \tilde{A}_{x_1}^{(1)} \tilde{A}_{x_2}^{(2)} \dots \tilde{A}_{x_N}^{(N)} \rangle$ appear in $\langle M_N \rangle$, and all must have non-zero weight to achieve the largest possible quantum value. This is incompatible with maximum randomness from one input combination. Hence, achieving the quantum bound does not certify maximum randomness; by direct calculation, maximum MABK violation certifies

$N + 1/2 - \log_2(1 + \sqrt{2})/\sqrt{2} \approx N - 0.4$ bits of global DI randomness when all parties use measurement 0.

We now proceed to construct new Bell expressions which certify N bits of randomness in the even case. Specifically, we use the techniques introduced in Section 3 to generalize the $N = 2$ case, which was addressed in [23].

4.2.1 Bipartite self-tests

To begin with, we summarize the results obtained for the $N = 2$ case.

Lemma 3 (I_θ family of self-tests). *Define the set $\mathcal{G} = (\pi/4, \pi/2) \cup (\pi/2, 3\pi/4) \cup (5\pi/4, 3\pi/2) \cup (3\pi/2, 7\pi/4)$ ². Let $\theta \in \mathcal{G}$, and define the family of Bell expressions parameterized by θ , for parties $k, l \in \{1, \dots, N\}$,*

$$\begin{aligned} \langle I_\theta^{(k,l)} \rangle &= \cos \theta \cos 2\theta \langle A_0^{(k)} A_0^{(l)} \rangle - \\ &\cos 2\theta (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) + \cos \theta \langle A_1^{(k)} A_1^{(l)} \rangle. \end{aligned} \quad (17)$$

Then we have the following:

(i) *The local bounds are given by $\pm \eta_\theta^L$, where*

$$\eta_\theta^L = \max \left(|\cos \theta (1 - \cos 2\theta)|, |\cos \theta (1 + \cos 2\theta)| + |2 \cos 2\theta| \right). \quad (18)$$

(ii) *The quantum bounds are given by $\pm \eta_\theta^Q$, where $\eta_\theta^Q = 2 \sin^3 \theta$.*

(iii) *Up to local isometries, there exists a unique strategy that achieves $\langle I_\theta \rangle = \eta_\theta^Q$:*

$$\begin{aligned} \rho_{Q_k Q_l} &= |\psi\rangle\langle\psi|, \text{ where } |\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + i|11\rangle), \\ A_0^{(k)} &= A_0^{(l)} = \sigma_X, \\ A_1^{(k)} &= A_1^{(l)} = \cos \theta \sigma_X + \sin \theta \sigma_Y. \end{aligned} \quad (19)$$

The above lemma is a rewriting of Proposition 1 in [23], and can be recovered as a special case of Lemma 7 to follow. Importantly, the self-testing property implies that when their inputs are 0, both parties measure σ_X on $|\psi\rangle$, which results in 2 bits of randomness, $H(\{p(ab|00)\}) = 2$. We also remark that the state $|\psi'\rangle = (|00\rangle - i|11\rangle)/\sqrt{2}$ with the same measurements achieves $\langle I_\theta \rangle = -\eta_\theta^Q$. This is a symmetry of the case above, since relabelling the measurement outcomes of party k yields the new Bell expression $I'_\theta = -I_\theta$. Therefore $\langle I_\theta \rangle = -\eta_\theta^Q$ implies $\langle I'_\theta \rangle = \eta_\theta^Q$, which self-tests $|\psi\rangle$ and the negated measurements, or equivalently $|\psi'\rangle$ and the original measurements.

²Note that $\mathcal{G} = \{\theta \in [0, 2\pi] \mid \cos(2\theta) < 0, \cos(\theta) \neq 0\}$.

4.2.2 Target strategy

For the N partite case, we generalize the above bipartite strategy:

$$\begin{aligned} \rho_Q &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0^{(k)} &= \sigma_X, \end{aligned} \quad (20)$$

$$A_1^{(k)} = \cos \theta \sigma_X + \sin \theta \sigma_Y, \quad k \in \{1, \dots, N\},$$

where $\theta \in \mathcal{G}$ is a free parameter. As before, we find $p(\mathbf{a}|\mathbf{0}) = 1/2^N, \forall \mathbf{a}$, hence this is a target strategy for certifying N bits of global randomness. Our objective now will be to design a tailored Bell inequality that can certify N bits by witnessing its maximum value, attained by the strategy in Eq. (20).

4.2.3 Constructing an N -partite Bell inequality

We use the techniques introduced in Section 3 to construct the desired Bell expression by expanding the seed I_θ . Studying the strategy in Eq. (20), we notice that when $N - 2$ parties perform their first measurement, σ_X , depending on the outcome parity the post-measurement state for the remaining parties is $(|00\rangle \pm i|11\rangle)/\sqrt{2}$. Now, performing the measurements in Eq. (20) on that post-measurement state achieves $\langle I_\theta \rangle = \pm \eta_\theta^Q$, which, according to the self-testing properties of I_θ , will certify 2 bits of randomness. Since the strategy is symmetric under party permutations, we can repeat this argument over different subsets of $N - 2$ parties, and conclude that, when every bipartite self-test is satisfied, the output setting $\mathbf{X} = \mathbf{0}$ must generate uniform DI randomness. The corresponding Bell inequality is constructed according to the following lemma.

Lemma 4 (Bell inequality for maximum randomness). *Let $\boldsymbol{\mu}$ be a tuple of $N - 2$ measurement outcomes for all parties excluding k, l , and $n_\mu \in \{0, 1\}$ be the parity of $\boldsymbol{\mu}$. Let $\theta \in \mathcal{G}$ and $\{I_\mu^{(k,l)}\}_\mu$ be a set of bipartite Bell expressions between parties k, l , where*

$$I_\mu^{(k,l)} = (-1)^{n_\mu} I_\theta^{(k,l)}. \quad (21)$$

Then the expanded Bell expression given by

$$I_\theta = \sum_{k=1}^{N-1} \left(\sum_{\boldsymbol{\mu}} \bar{P}_{\boldsymbol{\mu}|\mathbf{0}} I_\mu^{(k,N)} \right) \quad (22)$$

has quantum bounds $\pm \eta_{N,\theta}^Q$ where $\eta_{N,\theta}^Q = 2(N - 1) \sin^3 \theta$. Moreover, $\langle I_\theta \rangle = \eta_{N,\theta}^Q$ is achieved up to relabellings by the strategy in Eq. (20), and cannot be achieved classically.

The proof can be found in Section C.1.

We can now state our first main result:

Proposition 2 (Maximum randomness certification). *Achieving the maximum quantum value of the Bell inequality in Lemma 4 certifies N bits of global randomness, i.e.,*

$$R_{I_\theta}(\eta_{N,\theta}^Q) = N. \quad (23)$$

The proof is obtained directly by relating maximal Bell violation to decoupling Eve from the post-measurement state, proven in Lemma 2 and Corollary 1. Specifically, using Lemma 4, I_θ is an expanded Bell expression with quantum bounds $\pm \eta_{N,\theta}^Q$. Furthermore, for every sub-expression $I_\mu^{(k,N)}$, there exists an SOS decomposition that self-tests the pure bipartite state and ideal measurements in Eq. (19)³. This meets the conditions required to apply Lemma 2 and Corollary 1. Proposition 2 then follows from the fact that $p(\mathbf{a}|\mathbf{0}) = 1/2^N$ for the strategy in Eq. (20).

4.3 MABK value and maximum randomness

Above we gave a new one parameter construction for certifying N bits of device-independent randomness in any N -party 2-input 2-output scenario. Now we will study some properties of the correlations used to certify maximum randomness. We are specifically interested in how maximal randomness and MABK violation tradeoff against each other. In particular, how large can the MABK violation of a correlation be whilst certifying maximal randomness?

4.3.1 Achievable MABK values with maximum randomness

When N is odd, we have discussed that one can always certify N bits of DI randomness from maximum MABK violation; for the even case, it is unclear how large the MABK value can be whilst certifying maximum randomness. Using the fact that, for the strategy in Eq. (20), $\langle A_x \rangle = \sin n\theta$, where $n = \sum_{k=1}^N x_k$, its MABK value is given by⁴

$$\langle M_N(\theta) \rangle = 2^{\frac{N-1}{2}} \left(\cos^N(\theta/2 + \pi/4) \sin(N\theta/2 + \pi/4) + \cos^N(\theta/2 - \pi/4) \sin(N\theta/2 - \pi/4) \right). \quad (24)$$

Due to Proposition 2, for every $\theta \in \mathcal{G}$, N bits of randomness can be certified by a quantum correlation achieving the MABK value $\langle M_N(\theta) \rangle$. We are hence interested in finding the maximum of this function over $\theta \in \mathcal{G}$. In the following proposition, we introduce θ_N^* , which, for each N , is a candidate value of θ achieving this maximum, and prove that it belongs to the set $\mathcal{G}$.

Proposition 3. *For even N , let*

$$\theta_N^* = \frac{2\pi t_N}{N+1}, \quad (25)$$

³See the SOS decompositions [51, Lemma 9] and the proof of Lemma 7 for their connection to the family I_θ .

⁴This can be established using the identities $\cos(x) = (e^{ix} + e^{-ix})/2$, $\sin(x) = (e^{ix} - e^{-ix})/(2i)$ and $(1 + e^{ix})^N = 2^N e^{ixN/2} \cos^N(x/2)$.

where t_N is the $(N/2)^{\text{th}}$ element of the sequence 1, 1, 5, 7, 3, 3, 11, 13, 5, 5, ... given by

$$t_N = \begin{cases} N/4 + 1/2, & \text{if } N = 8n + 2, \\ N/4, & \text{if } N = 8n + 4, \\ 3N/4 + 1/2, & \text{if } N = 8n + 6, \\ 3N/4 + 1, & \text{if } N = 8n + 8, \quad n \in \mathbb{N}_0. \end{cases} \quad (26)$$

Then $\theta_N^* \in \mathcal{G}$.

Proof can be found in Section C.2. Using this result, we are able to prove the following. It will be convenient to define

$$m_N^* = \langle M_N(\theta_N^*) \rangle. \quad (27)$$

Proposition 4. *Let N be an even integer. For every MABK value s in the range $(1, m_N^*]$, there exists a $\theta_s \in \mathcal{G}$ that satisfies $s = \langle M_N(\theta_s) \rangle$.*

The proof of Proposition 4 can be found in Section C.3. In Fig. 1, we plot the MABK value for a given N as a function of $\theta \in \mathcal{G}$, showing the full range of violations $(1, m_N^*]$ is accessible. Specifically, the dashed lines in Fig. 1 indicate an MABK value of 1, while the peak of each graph is m_N^* . As θ varies, the graph shows a continuous curve between these values. We next state the following technical conjecture.

Conjecture 1. *The maximum MABK value achievable by quantum strategies with uniformly random outputs on input $\mathbf{0}$, i.e., $p(\mathbf{a}|\mathbf{0}) = 1/2^N$, $\forall \mathbf{a}$, is m_N^* .*

Conjecture 1 is known to hold in the case $N = 2$ [23], and we provide numerical evidence that it holds for general N in Appendices C.4 and C.10.

In Fig. 1, we illustrate that m_N^* is the largest MABK value that can be achieved by the family of strategies in Eq. (20), all of which generate maximum randomness. However, we have not ruled out the existence of a strategy outside this family, also generating maximum randomness, achieving an MABK value greater than m_N^* . Conjecture 1 is that this is not the case, and m_N^* is indeed the true maximum. Based on this, we can now state the range of MABK violations for which maximum DI randomness can be certified when N is even.

Lemma 5 (MABK violations achievable with maximum DI randomness, even case). *For even N , we have:*

- (i) *For every $s \in (1, m_N^*]$, there exists a quantum correlation with MABK value s that certifies N bits of device-independent randomness.*

Suppose Conjecture 1 holds. Then we additionally have:

- (ii) *m_N^* is the largest MABK value compatible with strategies generating N bits of device-independent randomness. Or conversely, any*

strategy achieving an MABK value $s > m_N^*$ certifies strictly less than N bits of device-independent randomness.

Part (i) can be established by the following reasoning: by varying $\theta \in \mathcal{G}$, the family of strategies in Eq. (20) achieves every MABK value in the interval $(1, m_N^*]$ by Proposition 4. Since $\theta \in \mathcal{G}$ for all these values, the corresponding bipartite strategy is self-testable according to Lemma 3; we can hence apply Lemma 4 to expand the Bell expression to N parties, and it follows from Corollary 1 that this Bell expression certifies N bits of DI randomness.

An implication of Proposition 3 is that the MABK value m_N^* becomes an achievable lower bound on the maximum MABK value for quantum correlations certifying maximum randomness. This follows from the same reasoning outlined in the previous paragraph, since $\theta_N^* \in \mathcal{G}$ implies maximum DI randomness can be certified from the correlations generated by the associated strategy in Eq. (20). Part (ii) of Lemma 5 is that this is optimal (if Conjecture 1 holds), in the sense that one cannot achieve a larger MABK value whilst simultaneously certifying maximum randomness. As evidence we derive a numerical technique that can generate upper bounds on this MABK value, and show these upper bounds match the lower bounds for some values N . See Appendix C.4 for the details. We also remark that Conjecture 1 is known to hold for the case $N = 2$ [23].

We now consider the case of odd MABK expressions.

Proposition 5. *Let N be an odd integer. For every $s \in (1, 2^{(N-1)/2})$, there exists a $\theta_s \in \mathcal{G}$ that satisfies $s = \langle M_N(\theta_s) \rangle$.*

The proof can be found in Section C.3. Similar to the even case, in Fig. 2 we plot the MABK value as a function of θ for different N , and see that all quantum achievable values can be obtained by Eq. (20) for some $\theta \in \mathcal{G}$. We then have the following consequence:

Lemma 6 (MABK violations achievable with maximum DI randomness, odd case). *Let N be odd. For every $s \in (1, 2^{(N-1)/2}]$ there exists a quantum correlation with MABK value s that certifies N bits of device-independent randomness.*

The above lemma asserts that when N is odd, for all quantum-achievable MABK values maximum randomness can be realized simultaneously. For the case of $\langle M_N \rangle = 2^{(N-1)/2}$, i.e., the maximum quantum value, randomness can be certified according to Proposition 1. For all other MABK values, Proposition 5 implies maximum randomness can be certified by I_θ for some $\theta \in \mathcal{G}$.

We remark on the nontrivial maximal MABK violation compatible with maximum randomness when N is even, contrasting the odd case; as discussed in

Section 4.1, this can be seen as a consequence MABK expressions for even N containing every N party correlator. When maximum randomness is being certified, one of these correlators must be zero, which restricts the maximum MABK violation to be strictly less than the optimal quantum value.

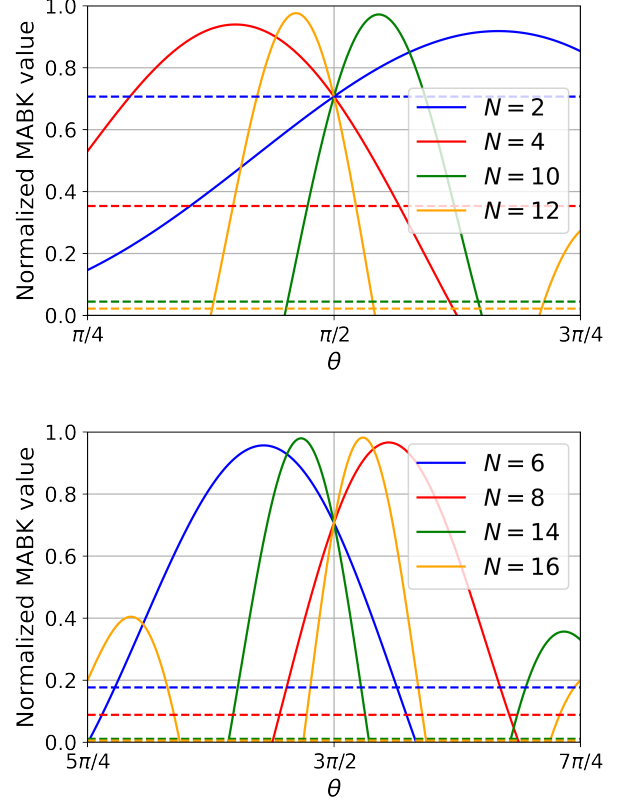


Figure 1: Using the family of strategies in (20) for even N we plot the MABK value renormalized by the maximum quantum value over all strategies (i.e., $2^{(N-1)/2}$) in terms of $\theta \in [\pi/4, 3\pi/4]$, and $\theta \in [5\pi/4, 7\pi/4]$ respectively. The dashed lines indicate where the strategy becomes local. All points in this interval, excluding the boundaries and center point (since $\pi/2$ and $3\pi/2$ are not in $\mathcal{G}$), correspond to strategies that certify N bits of randomness device-independently using our expanded Bell expressions.

4.3.2 Asymptotic behaviour

We now consider the behaviour of the conjectured maximal MABK value achievable with maximum randomness, m_N^* , for increasingly large even N . We show that m_N^* converges to the largest possible quantum value in this limit. Note this is not based on a conjecture; m_N^* is an achievable lower bound, and in the following proposition we show that this lower bound tends towards the global upper bound, namely the maximal quantum MABK value.

Proposition 6 (Maximum randomness in the asymptotic limit). *In the limit of large even N , one can*

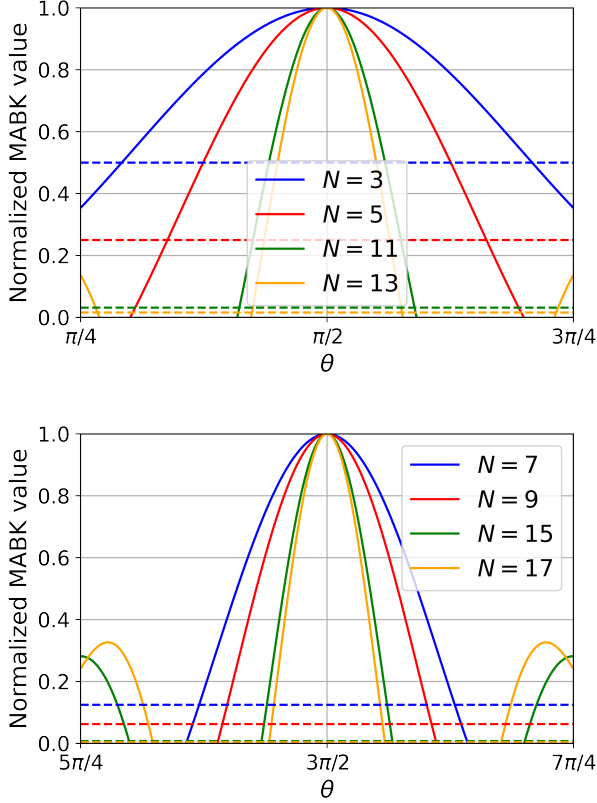


Figure 2: A similar plot to that of Fig. 1 for odd N and second measurement angle of all parties $\theta \in [\pi/4, 3\pi/4]$, and $\theta \in [5\pi/4, 7\pi/4]$ respectively. All points in this interval, excluding the boundaries, correspond to strategies that can certify N bits of randomness device-independently, using our expanded Bell expressions, or using the MABK inequality for the center points. For $N = 3, 7, 11, 15$, we use the MABK expression given by Eq. (1), and for $N = 5, 9, 13, 17$ we use the same expression after relabelling every parties inputs followed by their first measurement's output.

achieve arbitrarily close to the maximum quantum violation of the N party MABK inequality, $2^{(N-1)/2}$, whilst certifying maximum device-independent randomness.

This is proven in Appendix C.5.

4.3.3 Nonlocality and maximum randomness

Whilst we have studied the MABK values achieved by the strategies in Eq. (20), we also consider how nonlocal they are, quantified by how much the local set needs to be “diluted” to contain them. We refer to this measure as the local dilution, which can be computed via a linear program for small N , and we give details in Section C.6. Our findings are presented in Fig. 3.

Interestingly, for $N > 2$ the correlations of the strategy in Eq. (20) are bounded away from the local boundary for any $\theta \in \mathcal{G}$, even for small MABK violation; whilst the violation of a given MABK inequality

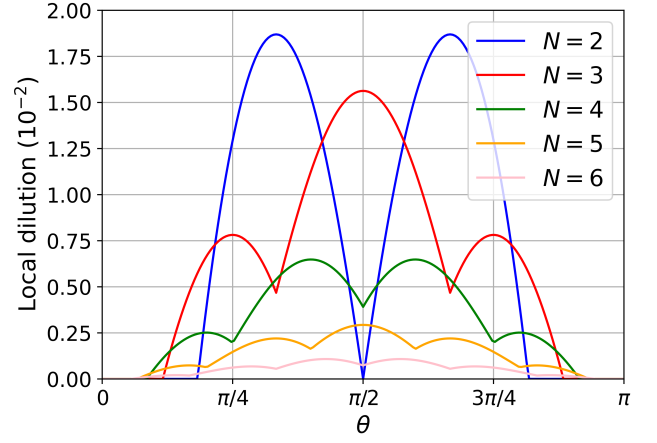


Figure 3: Nonlocality, as measured using local dilution, of the strategies in Eq. (20), for second measurement angle of all parties $\theta \in [0, \pi]$. All values of $\theta \in (\pi/4, \pi/2) \cup (\pi/2, 3\pi/4)$ correspond to strategies which can certify N bits of maximum randomness using our technique for expanding Bell expressions. $\theta = \pi/2$ can also correspond to maximum randomness when N is odd by testing the MABK inequality.

can be made arbitrarily small, the correlations are still distant from the local boundary. In fact, by comparing Fig. 3 with Fig. 1 and Fig. 2, one can see how correlations achieving at or below the maximum local MABK value, $\langle M_N \rangle \leq 1$, can still be nonlocal and certify maximum randomness. To illustrate this point, consider the $N = 3$ (blue) curve in Fig. 2. Provided $\theta \in (\pi/4, \pi/2) \subset \mathcal{G}$, N bits of DI randomness can be certified following previous arguments. However, the blue dashed line in Fig. 2, indicating the local bound ($\langle M_3 \rangle = 1$), intersects the solid curve inside the region $(\pi/4, \pi/2)$, implying that the corresponding strategy in Eq. (20) achieves $\langle M_3 \rangle \leq 1$ while simultaneously certifying maximum randomness. This is explained by the $N = 3$ (red) curve in Fig. 3, where it can be seen that in the closed region $[\pi/4, \pi/2]$ the correlations are nonlocal and hence must violate some other Bell inequality. This illustrates the complexity of the tradeoff between randomness versus nonlocality in the multi-partite scenario.

5 Trade-off between DI randomness and MABK violation

In this section, we present an achievable lower bound on the maximum device-independent randomness that can be generated by correlations achieving any given MABK violation. Moreover, we conjecture this lower bound to be tight based on numerical evidence. This is achieved by introducing a new family of two parameter quantum strategies along with their self-testing Bell expressions. Using this as the seed, we construct a multi-partite Bell expression which certifies their randomness.

5.1 Constructing the Bell expression

5.1.1 Bipartite self-tests

We begin by introducing a versatile family of bipartite self-tests which generalize the results in [23].

Lemma 7 ($J_{\phi,\theta}$ family of self-tests). *Let $(\phi, \theta) \in \mathbb{R}^2$ such that $\cos(2\theta)\cos(2\phi) < 0$ and $\cos(\theta - \phi) \neq 0$ ⁵. Define the family of Bell expressions parameterized by ϕ and θ ,*

$$\begin{aligned} \langle J_{\phi,\theta}^{(k,l)} \rangle &= \cos 2\theta \cos(\theta - \phi) \langle A_0^{(k)} A_0^{(l)} \rangle \\ &\quad - \cos 2\theta \cos 2\phi (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) \\ &\quad + \cos 2\phi \cos(\theta - \phi) \langle A_1^{(k)} A_1^{(l)} \rangle. \end{aligned} \quad (28)$$

Then we have the following:

(i) The local bounds are given by $\pm \eta_{\phi,\theta}^L$, where

$$\begin{aligned} \eta_{\phi,\theta}^L &= \max \{ |\cos(\theta - \phi)(\cos 2\theta - \cos 2\phi)|, \\ &\quad |\cos(\theta - \phi)(\cos 2\theta + \cos 2\phi)| + |2 \cos 2\phi \cos 2\theta| \}. \end{aligned} \quad (29)$$

(ii) The quantum bounds are given by $\pm \eta_{\phi,\theta}^Q$, where $\eta_{\phi,\theta}^Q = 2 \sin^2(\theta + \phi) \sin(\theta - \phi)$.

(iii) $|\eta_{\phi,\theta}^Q| > \eta_{\phi,\theta}^L$.

(iv) Up to local isometries, there exists a unique strategy that achieves $\langle J_{\phi,\theta}^{(k,l)} \rangle = \eta_{\phi,\theta}^Q$:

$$\begin{aligned} \rho_{Q_k Q_l} &= |\psi\rangle\langle\psi|, \text{ where } |\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + i|11\rangle), \\ A_0^{(k)} &= A_0^{(l)} = \cos(\phi) \sigma_X - \sin(\phi) \sigma_Y, \\ A_1^{(k)} &= A_1^{(l)} = \cos(\theta) \sigma_X + \sin(\theta) \sigma_Y. \end{aligned} \quad (30)$$

Note that when $\phi = 0$ Lemma 7 reduces to Lemma 3. Moreover, this Bell expression retains the same symmetry properties of the I_θ family, namely that $\langle J_{\phi,\theta} \rangle = -\eta_{\phi,\theta}^Q$ for the state $|\psi'\rangle$ and the same measurements, and is hence a self-test. Lemma 7 can be obtained as a corollary of the self-testing results from Refs. [52, 53, 51] (see Section C.7 for details).

For future convenience, we define the set $\mathcal{F} = [(-\pi/4, \pi/4) \times \mathcal{G}] \cup [(-\pi/4, \pi/4) \setminus \{0\} \times \{\pi/2, 3\pi/2\}] \subset \mathbb{R}^2$. One can verify that points $(\phi, \theta) \in \mathcal{F}$ satisfy $\cos(2\theta)\cos(2\phi) < 0$ and $\cos(\theta - \phi) \neq 0$, and therefore define a valid self-test according to Lemma 7.

⁵The reason for this condition is explained in Section C.7.

5.1.2 Target strategy

For the N -partite case, we will consider the following strategy:

$$\begin{aligned} \rho_Q &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0^{(k)} &= \cos \phi \sigma_X - \sin \phi \sigma_Y, \\ A_1^{(k)} &= \cos \theta \sigma_X + \sin \theta \sigma_Y, \quad k \in \{1, \dots, N\}. \end{aligned} \quad (31)$$

Using the fact that $\langle A_x \rangle = \sin(n\theta - (N - n)\phi)$, where $n = \sum_k x_k$, the MABK value, defined in Eq. (1), of the above strategy is given by

$$\begin{aligned} \langle M_N(\phi, \theta) \rangle &= 2^{\frac{N-1}{2}} \left(\cos^N [(\theta + \phi)/2 + \pi/4] \right. \\ &\quad \cdot \sin [N(\theta - \phi)/2 + \pi/4] \\ &\quad \left. + \cos^N [(\theta + \phi)/2 - \pi/4] \sin [N(\theta - \phi)/2 - \pi/4] \right). \end{aligned} \quad (32)$$

5.1.3 Constructing an N -partite Bell inequality

Using the previous two building blocks, we construct the following Bell inequality using the $J_{\phi,\theta}$ expressions as the seed.

Lemma 8. *Let μ be a tuple of $N - 2$ measurement outcomes for all parties excluding k , l , and $n_\mu \in \{0, 1\}$ be the parity of μ . Let $(\phi, \theta) \in \mathbb{R}^2$,*

$$\phi' := \frac{\phi N}{2}, \quad \theta' := \theta - \frac{N - 2}{2} \phi, \quad (33)$$

and

$$I_\mu^{(k,l)} = (-1)^{n_\mu} J_{\phi',\theta'}^{(k,l)}. \quad (34)$$

Define the following Bell polynomial

$$I_{\phi',\theta'} := \sum_{k=1}^{N-1} \left(\sum_{\mu} \tilde{P}_{\mu|0}^{(k,N)} I_\mu^{(k,N)} \right). \quad (35)$$

If $(\phi', \theta') \in \mathcal{F}$, $I_{\phi',\theta'}$ is an expanded Bell expression, and has quantum bounds $\pm \eta_{N,\phi',\theta'}^Q$, where $\eta_{N,\phi',\theta'}^Q = 2(N-1) \sin^2(\theta' + \phi') \sin(\theta' - \phi')$, which can be achieved up to relabellings by the strategy

$$\begin{aligned} \rho_Q &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0^{(k)} &= \cos \phi \sigma_X - \sin \phi \sigma_Y, \\ A_1^{(k)} &= \cos \theta \sigma_X + \sin \theta \sigma_Y, \quad k \in \{1, \dots, N\}. \end{aligned} \quad (36)$$

In addition, this quantum bound cannot be achieved classically.

The proof is given in Appendix C.8. Note that the Bell expressions (35) are written in terms of the shifted parameters (ϕ', θ') , instead of the measurement angles (ϕ, θ) . This is because the $N - 2$ parties performing the projection no longer use σ_X , but use $\cos(\phi) \sigma_X - \sin(\phi) \sigma_Y$. This accumulates a phase factor on the state of the remaining parties,

$(|00\rangle + i(-1)^{n_\mu} e^{i(N-2)\phi} |11\rangle)/\sqrt{2} =: |\Phi_{\mu,\phi}\rangle$, which is equivalent to the action of some local unitary U on $(|00\rangle + i(-1)^{n_\mu} |11\rangle)/\sqrt{2}$. To correct for this, we use the Bell expression $J_{\phi',\theta'}^{(k,l)}$ as the seed. See Appendix C.8 for the full details.

5.2 Randomness versus MABK value

Using the previously derived Bell inequality, we consider the trade-off between maximum device-independent randomness and MABK violation. Let us define

$$\theta(\phi) = \frac{N-1}{N+1}\phi + \theta_N^*, \quad (37)$$

where θ_N^* is defined in Eq. (25) and let

$$\phi_N^* = \text{sgn}[\sin(2\theta_N^*)] \frac{\pi}{4N}. \quad (38)$$

For the strategy in Eq. (31), direct calculation yields

$$\langle M_N(\phi_N^*, \theta(\phi_N^*)) \rangle = 2^{(N-1)/2}, \quad (39)$$

and $\langle M_N(0, \theta(0)) \rangle = m_N^*$. Hence, by varying $\phi \in [0, \phi_N^*]$ ⁶ and choosing $\theta = \theta(\phi)$, one can obtain the desired range of MABK values $[m_N^*, 2^{(N-1)/2}]$, from the conjectured maximum MABK value with maximum randomness to the maximum quantum value. Choosing this parameterization, the raw randomness, $H(\mathbf{R}|\mathbf{X} = \mathbf{x}^*) = H(\{p_\phi(\mathbf{a}|\mathbf{0})\})$, of the strategy in Eq. (31) (with $\theta = \theta(\phi)$), as a function of ϕ , is given by

$$H(\{p_\phi(\mathbf{a}|\mathbf{0})\}) \equiv r(\phi) = N - 1 + H_{\text{bin}}\left[\frac{1 - \sin N\phi}{2}\right], \quad (40)$$

where H_{bin} is the binary entropy function and $r(\phi)$ is a smooth, monotonically decreasing function of ϕ in the range $[0, \phi_N^*]$.

As shown in Section 4 (by combining Lemma 4 with Proposition 3), when $\phi = 0$ we can certify maximum DI randomness. It also follows from the self-testing properties of the MABK family that when $\phi = \phi_N^*$ we obtain $N - 1 + 3/2 - \log_2(1 + \sqrt{2})/\sqrt{2} \approx N - 0.4$ bits of global DI randomness when all parties use measurement 0. What remains then, is to apply the new Bell expression constructed in Lemma 8 to certify $r(\phi)$ bits of DI randomness for $\phi \in (0, \phi_N^*)$. To do so, the following proposition shows that for every $\phi \in [0, \phi_N^*]$, there exists a valid Bell expression given by Lemma 8.

Proposition 7. *Let $\phi \in [0, \phi_N^*]$, and $\theta = \theta(\phi)$ as defined in Eqs. (37) and (38). Then $(\phi', \theta') \in \mathcal{F}$, where ϕ' and θ' are defined in Lemma 8.*

This is proven in Section C.9. Proposition 7 implies that, for every ϕ in the range we are interested in, the

⁶To ease notation, we write $[0, \phi_N^*]$ when $\phi_N^* > 0$ and $\phi_N^* < 0$, interpreting the latter case as $[\phi_N^*, 0]$.

bipartite expression $(-1)^{n_\mu} J_{\phi',\theta'}$ is a valid self-test of the strategy

$$\begin{aligned} \rho_{Q_k Q_l} &= |\Phi_{\mu,\phi}\rangle\langle\Phi_{\mu,\phi}|, \\ A_0^{(k)} &= A_0^{(l)} = \cos \phi' \sigma_X - \sin \phi' \sigma_Y, \\ A_1^{(k)} &= A_1^{(l)} = \cos \theta' \sigma_X + \sin \theta' \sigma_Y. \end{aligned} \quad (41)$$

After applying the local unitary $U_\phi \otimes U_\phi$, where $U_\phi = |0\rangle\langle 0| + e^{i(N-2)\phi/2} |1\rangle\langle 1|$, this is equivalent to

$$\begin{aligned} \rho_{Q_k Q_l} &= |\Phi_{\mu,\phi}\rangle\langle\Phi_{\mu,\phi}|, \\ A_0^{(k)} &= A_0^{(l)} = \cos \phi \sigma_X - \sin \phi \sigma_Y, \\ A_1^{(k)} &= A_1^{(l)} = \cos \theta \sigma_X + \sin \theta \sigma_Y, \end{aligned} \quad (42)$$

which is exactly the bipartite strategy we wanted to self-test, since it is the one held by parties k and N after the projector $P_{\mu|\mathbf{0}}^{(k,N)}$ is applied to the global state $|\psi_{\text{GHZ}}\rangle$. As a result, the correlations generated by the strategy in Eq. (31), by choosing $\theta = \theta(\phi)$ and varying $\phi \in [0, \phi_N^*]$, maximally violate the Bell inequality $J_{\phi',\theta'}$ with ϕ' and θ' given in Eq. (33). We can therefore employ the decoupling lemma to make the rate unconditioned on Eve, $r(\phi)$, device-independent.

Proposition 8. *Achieving the maximum quantum value of the Bell inequality in Lemma 8 certifies $r(\phi)$ bits of randomness, i.e.,*

$$R_{I_{\phi',\theta'}}(\eta_{N,\phi',\theta'}^Q) = r(\phi). \quad (43)$$

We have established that, for every MABK value $s \in [m_N^*, 2^{(N-1)/2}]$, there exists a $\phi_s \in [0, \phi_N^*]$ which defines a quantum strategy achieving s , for which its generated randomness $r(\phi_s)$ is device-independent. We have therefore related every MABK value with a DI rate. We now conjecture that the curve $(s, r(\phi_s))$ is optimal in terms of raw randomness, which can then be made device-independent following the above discussion — see Lemma 9.

Conjecture 2. *For N even, the maximum randomness unconditioned on Eve, r , that can be generated by quantum strategies achieving an MABK value, s , is given by*

$$r(s) = \begin{cases} N, & s \in (1, m_N^*], \\ r(\phi_s), & s \in (m_N^*, 2^{(N-1)/2}], \end{cases} \quad (44)$$

where $r(\phi)$ is defined in Eq. (40), and

$$\phi_s = \arg \min \{|\phi| : \phi \in [0, \phi_N^*], \langle M_N(\phi, \theta(\phi)) \rangle = s\}. \quad (45)$$

For the range $s \in (1, m_N^*]$, and $s \in (m_N^*, 2^{(N-1)/2}]$, the rate $r(s)$ and MABK value s is achieved by the family of quantum strategies in Eq. (20) and Eq. (31), and certified device-independently by the Bell expressions in Lemma 4 and Lemma 8, respectively. Similarly to Conjecture 1, Conjecture 2 is known to hold for the case $N = 2$ [23].

The minimization in the definition of ϕ_s is included since we have not shown the set $\{\phi \in [0, \phi_N^*] : \langle M_N(\phi, \theta(\phi)) \rangle = s\}$ is unique. Intuitively, the closer ϕ_s is to zero the closer we are to maximum randomness; hence we expect $r(\phi)$ to be monotonically decreasing with $|\phi|$ for $\phi \in [0, \phi_N^*]$, and minimization guarantees the best rate at a given s . For the examples we have computed, the minimization turns out to be trivial, i.e., ϕ_s is the unique solution to $\langle M_N(\phi_s, \theta(\phi_s)) \rangle = s$.

Lemma 9. *Suppose Conjecture 2 holds. Then the maximum amount of device-independent randomness that can be certified for the range of MABK values $(1, 2^{(N-1)/2}]$ is given by Eq. (44). Moreover, this rate is achievable.*

The above lemma tells us that if the rate in Eq. (44) is optimal without conditioning on Eve, it will also be optimal conditioned on Eve, i.e., the rate can be made device-independent. This is because Propositions 3 and 7 ensure that, for every MABK value, the corresponding quantum strategy achieving it, given in Eq. (20) or Eq. (31), can generate certifiable randomness using the corresponding Bell expression I_θ or $I_{\phi', \theta'}$. A rewriting of Lemma 9 that does not rely on Conjecture 2, would say that (44) corresponds to an achievable lower bound on the maximum DI randomness as a function of MABK value. Our results are summarized in Fig. 4.

We provide evidence for Conjectures 1 and 2 in Appendix C.10, where we derive a numerical technique for upper bounding the maximum randomness that can be generated by quantum strategies achieving a given MABK value s . An implementation of this technique in python can be found in our GitHub repository [40]. By studying the numerical results, we find our analytical lower bound agrees with the numerical upper bound to at least 16 digits.

5.3 Other extremal Bell inequalities

In the N -partite 2-input 2-output scenario, the MABK family is just one class of extremal Bell inequalities, and the techniques developed in this section can be readily applied to others. For example, when $N = 3$ and restricting to Bell inequalities containing only three party correlators, there are a total of 3 non-trivial, new classes, one being the MABK inequality M_3 , and two being of the form [42]

$$S_1 = \frac{1}{4} \sum_{x,y,z} A_x B_y C_z - A_1 B_1 C_1, \quad (46)$$

$$S_2 = A_0 B_0 (C_0 + C_1) - A_1 B_1 (C_0 - C_1), \quad (47)$$

where we used A_x for $A_{x_1}^{(1)}$, B_y for $A_{x_2}^{(2)}$ etc. for legibility. S_1 and S_2 have local bounds of 1 and 2, and maximum quantum values $5/3$ and $2\sqrt{2}$, respectively. We applied our numerical technique (see Appendix C.10)

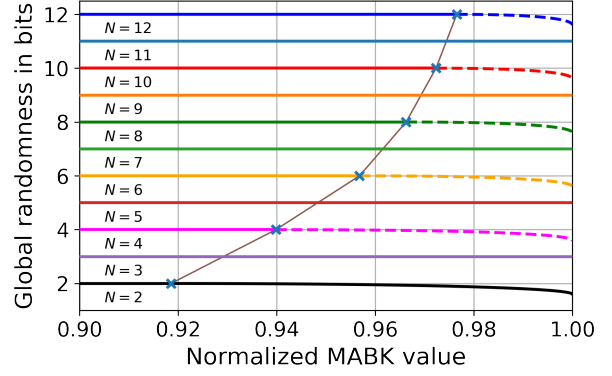


Figure 4: The conjectured curves of maximum device-independent randomness versus MABK value, where the MABK value is normalized by its maximum quantum value $2^{(N-1)/2}$. When N is odd, maximum global randomness can be achieved for any MABK value between the local and quantum bound, indicated by solid lines. The blue crosses indicate the conjectured maximum MABK value for which maximum randomness can be achieved when N is even, m_N^* , which tends to the maximum quantum value as $N \rightarrow \infty$. To the left of the blue crosses, N bits of randomness can be achieved for MABK values between the local bound (not shown on this plot) and m_N^* ; since this is the global maximum, it is optimal, indicated by solid lines. To the right of the blue crosses, dashed lines indicate lower bounds on the trade-off between maximum randomness and MABK value from m_N^* to the maximum quantum value, which are conjectured to be tight. The case of $N = 2$ was proven in Ref [23], which we reproduce with the results of the present paper.

to find upper bounds on the maximum amount of randomness certifiable whilst achieving a given violation of S_1 and S_2 . A comparison can be found in Fig. 5, and we provide additional figures and details in Section C.11. We find that both S_1 and S_2 exhibit a trade-off for violations close to the quantum maximum. Due to its structure, S_2 exhibits identical trade-off characteristics to the CHSH inequality found in [23], with the maximum CHSH value achievable with maximum randomness being numerically close to $3\sqrt{3}/2$. On the other hand, we find that more randomness can be generated from maximum violation of S_1 , which is still strictly less than the maximum certified by MABK. Exact values are given in Section C.11.

6 Discussion

In this work, we studied how MABK violation constrains the generation of global DI randomness for an arbitrary number of parties. Whilst there is no trade-off in the odd case, we conjectured the precise trade-off for the even case. This conjecture is supported by an analytical lower bound and a numerical upper bound that agree up to at least the first 16 digits

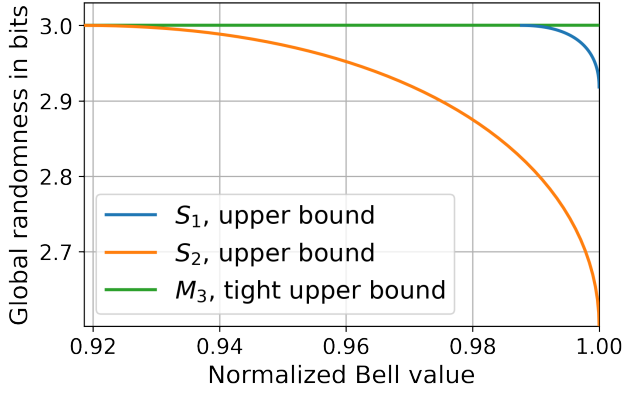


Figure 5: Upper bounds on the trade-off between asymptotic global DI randomness and violation of non-trivial extremal Bell inequalities, for the tripartite scenario with two binary measurements per party. The violation has been normalized by the maximum quantum value. S_1 and S_2 are given by Eq. (46) and Eq. (47) respectively, and M_3 is the MABK expression.

and have been checked for up to $N = 12$. We additionally showed that, in the asymptotic limit, this trade-off vanishes. Our main technical contribution was the extension of a recent tool, originally introduced to witness genuinely multi-partite nonlocality, to randomness certification.

The difference between the odd and even cases is explained by the odd MABK expressions containing half the full party correlators, allowing for input combinations not included in the Bell expression to be uniform. For even N , all correlators are included and must be non-zero to obtain maximum violation, resulting in a trade-off between randomness and MABK value. The number of correlators is 2^N , whereas the quantum bound is $2^{(N-1)/2}$. Maximum violation is achieved when the product of all correlators and their MABK coefficients have the same value, which in this case will be $2^{-(N+1)/2}$. Hence we find the penalty for fixing one correlator to zero (which is necessary for maximum randomness) diminishes as N grows, and vanishes in the asymptotic limit.

We also quantified the nonlocality (measured using the dilution) of our constructions for low values of N , calculated via a linear program. We found our correlations are always separated from the local boundary (except for the case $N = 2$). It would be interesting to find correlations that lie arbitrarily close to the local boundary whilst still certifying maximum randomness, as was found in the $N = 2$ case [22, 23]. This could entail breaking the symmetry between the second measurements of each party, and the techniques presented here can be applied. Moreover, one could hope to go further and bound the trade-off between randomness and nonlocality in this setting; however, the challenge becomes finding a suitable measure of nonlocality that is efficient to work with. Specifically, for the case $N = 2$, constraining the amount

of nonlocality of a quantum correlation is equivalent to constraining its CHSH value, which is linear in the correlators. When used as a constraint for maximizing randomness, this can then be handled using the techniques presented in Section C.10.1, which rely on the Navascués-Pironio-Acín hierarchy [54, 55]. For $N > 2$, the dilution measure is calculated using a linear program, therefore requiring a different approach.

We hope that the results found here will inform future experiments in DIRE; initial works have already made progress in this direction [56]. In such experiments, one should consider the optimal protocol for DIRE. In this work, we have proposed protocols that rely on spot-checking [7, 8, 46], where randomness is generated by a single input combination. An alternative approach has been explored in Ref. [46], where it was found that generating randomness from all inputs and averaging the result boosts the rate when the CHSH inequality is used. In the case of maximally violating a self-testing Bell inequality, spot-checking will always be optimal when a specific input combination leads to maximum randomness; having one maximally random setting necessarily implies that at least one other setting combination will not be maximally random for the correlations to be nonlocal, hence averaging will only decrease the overall rate. Whether this continues to hold when the presence of noise reduces the Bell value remains an open question, and the trade-off between using spot-checking and weighted averaging for our constructions deserves future investigation, with the aim of further improving practical rates.

Another interesting direction for future research is to connect expanded Bell inequalities to multi-partite self-testing. Our decoupling result allows the structure of the post-measurement state to be certified following the maximum violation of an expanded Bell expression; the open question we pose is whether an expanded Bell inequality can constitute a self-test of the N -party state, opening up a new range of useful applications beyond randomness certification. Intuitively, if each pair of parties can self-test all of their measurements along with a projected state, the marginal information could be sufficient to make a statement about the global state. Such a strategy has already been adopted for self-testing certain multi-partite strategies [36]. However, there are subtleties to consider. For example, expanded Bell expressions tailored to some quantum strategies can result in trivial inequalities (i.e., ones that exhibit no classical-quantum gap), despite the underlying strategy being nonlocal, and furthermore self-testable. We therefore believe that an additional ingredient is needed to self-test generic multi-partite states using the techniques in this work.

Finally, whilst we studied the MABK family of inequalities, one could consider a similar analysis for other families of extremal multi-partite Bell expres-

sions [42]. For example, we explored upper bounds on the trade-offs for all extremal inequalities when $N = 3$, and found the MABK inequality to be optimal for global randomness; it would be interesting to find out if the upper bounds presented for the other inequalities are achievable. One might try to find explicit constructions that match these bounds, and use our techniques to make them device-independent. Analysing other extremal inequalities will build a more complete picture of how DI randomness can be generated in the multi-partite scenario, better informing the way forward for future multi-partite experiments.

Acknowledgements

This work was supported by the UK’s Engineering and Physical Sciences Research Council (EP-SRC) via the Quantum Communications Hub (Grant No. EP/T001011/1), the Integrated Quantum Networks Hub (Grant No. EP/Z533208/1), Grant No. EP/SO23607/1 and the European Union’s Horizon Europe research and innovation programme under the project “Quantum Secure Networks Partnership” (QSNP, grant agreement No. 101114043).

References

- [1] J. S. Bell. “Speakable and unspeakable in quantum mechanics”. [Cambridge University Press](#). (1987).
- [2] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. “Bell nonlocality”. [Reviews of Modern Physics](#) **86**, 419–478 (2014).
- [3] Jonathan Barrett, Noah Linden, Serge Massar, Stefano Pironio, Sandu Popescu, and David Roberts. “Nonlocal correlations as an information-theoretic resource”. [Physical Review A](#) **71**, 022101 (2005).
- [4] Roger Colbeck. “Quantum and relativistic protocols for secure multi-party computation”. PhD thesis. University of Cambridge. (2007). [arXiv:0911.3814](#).
- [5] S. Pironio, A. Acin, 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 theorem”. [Nature](#) **464**, 1021–1024 (2010).
- [6] Roger Colbeck and Adrian Kent. “Private randomness expansion with untrusted devices”. [Journal of Physics A](#) **44**, 095305 (2011).
- [7] Carl A. Miller and Yaoyun Shi. “Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices”. [Proceedings of the 46th Annual ACM Symposium on Theory of Computing \(STOC ’14\) \(New York, USA\)](#) pages 417–426 (2014).
- [8] Carl A. Miller and Yaoyun Shi. “Universal security for randomness expansion from the spot-checking protocol”. [SIAM Journal of Computing](#) **46**, 1304–1335 (2017).
- [9] Roger Colbeck and Renato Renner. “Free randomness can be amplified”. [Nature Physics](#) **8**, 450–454 (2012).
- [10] Artur K. Ekert. “Quantum cryptography based on Bell’s theorem”. [Physical Review Letters](#) **67**, 661–663 (1991).
- [11] John Barrett, Lucien Hardy, and Adrian Kent. “No signalling and quantum key distribution”. [Physical Review Letters](#) **95**, 010503 (2005).
- [12] Antonio Acin, Nicolas Brunner, Nicolas Gisin, Serge Massar, Stefano Pironio, and Valerio Scarani. “Device-independent security of quantum cryptography against collective attacks”. [Physical Review Letters](#) **98**, 230501 (2007).
- [13] Stefano Pironio, Antonio Acin, Nicolas Brunner, Nicolas Gisin, Serge Massar, and Valerio Scarani. “Device-independent quantum key distribution secure against collective attacks”. [New Journal of Physics](#) **11**, 045021 (2009).
- [14] Umesh Vazirani and Thomas Vidick. “Fully device-independent quantum key distribution”. [Physical Review Letters](#) **113**, 140501 (2014).
- [15] Rotem Arnon-Friedman, Frédéric Dupuis, Omar Fawzi, Renato Renner, and Thomas Vidick. “Practical device-independent quantum cryptography via entropy accumulation”. [Nature Communications](#) **9**, 459 (2018).
- [16] Dominic Mayers and Andrew Yao. “Self testing quantum apparatus” (2004). [arXiv:quant-ph/0307205](#).
- [17] M. McKague, T. H. Yang, and V. Scarani. “Robust self-testing of the singlet”. [Journal of Physics A: Mathematical and Theoretical](#) **45**, 455304 (2012).
- [18] Tzyh Haur Yang and Miguel Navascués. “Robust self-testing of unknown quantum systems into any entangled two-qubit states”. [Physical Review A](#) **87**, 050102 (2013).
- [19] Andrea Coladangelo, Koon Tong Goh, and Valerio Scarani. “All pure bipartite entangled states can be self-tested”. [Nature Communications](#) **8**, 15485 (2017).
- [20] Ivan Šupić, Joseph Bowles, Marc-Olivier Renou, Antonio Acin, and Matty J. Hoban. “Quantum networks self-test all entangled states”. [Nature Physics](#) **19**, 670–675 (2023).
- [21] Ivan Šupić and Joseph Bowles. “Self-testing of quantum systems: a review”. [Quantum](#) **4**, 337 (2020).
- [22] Antonio Acin, Serge Massar, and Stefano Pironio. “Randomness versus nonlocality and en-

- tanglement”. *Physical Review Letters* **108**, 100402 (2012).
- [23] Lewis Woollorton, Peter Brown, and Roger Colbeck. “Tight analytic bound on the trade-off between device-independent randomness and nonlocality”. *Physical Review Letters* **129**, 150403 (2022).
- [24] N. David Mermin. “Extreme quantum entanglement in a superposition of macroscopically distinct states”. *Physical Review Letters* **65**, 1838–1840 (1990).
- [25] M. Ardehali. “Bell inequalities with a magnitude of violation that grows exponentially with the number of particles”. *Physical Review A* **46**, 5375–5378 (1992).
- [26] A. V. Belinskii and D. N. Klyshko. “Interference of light and Bell’s theorem”. *Physics-Uspekhi* **36**, 653 (1993).
- [27] Erik Woodhead, Boris Bourdoncle, and Antonio Acín. “Randomness versus nonlocality in the Mermin-Bell experiment with three parties”. *Quantum* **2**, 82 (2018).
- [28] Federico Grasselli, Gláucia Murta, Hermann Kampermann, and Dagmar Bruß. “Entropy bounds for multiparty device-independent cryptography”. *PRX Quantum* **2**, 010308 (2021).
- [29] Federico Grasselli, Gláucia Murta, Hermann Kampermann, and Dagmar Bruß. “Boosting device-independent cryptography with tripartite nonlocality”. *Quantum* **7**, 980 (2023).
- [30] Cédric Bamps, Serge Massar, and Stefano Pironio. “Device-independent randomness generation with sublinear shared quantum resources”. *Quantum* **2**, 86 (2018).
- [31] Chirag Dhara, Giuseppe Prettico, and Antonio Acín. “Maximal quantum randomness in Bell tests”. *Physical Review A* **88**, 052116 (2013).
- [32] Gonzalo de la Torre, Matty J. Hoban, Chirag Dhara, Giuseppe Prettico, and Antonio Acín. “Maximally nonlocal theories cannot be maximally random”. *Physical Review Letters* **114**, 160502 (2015).
- [33] Samuel L. Braunstein and Carlton M. Caves. “Wringing out better Bell inequalities”. *Annals of Physics* **202**, 22–56 (1990).
- [34] I. Šupić, R. Augusiak, A. Salavrakos, and A. Acín. “Self-testing protocols based on the chained Bell inequalities”. *New Journal of Physics* **18**, 035013 (2016).
- [35] Florian J. Curchod, Mafalda L. Almeida, and Antonio Acín. “A versatile construction of Bell inequalities for the multipartite scenario”. *New Journal of Physics* **21**, 023016 (2019).
- [36] Ivan Šupić, A Coladangelo, R Augusiak, and A Acín. “Self-testing multipartite entangled states through projections onto two systems”. *New Journal of Physics* **20**, 083041 (2018).
- [37] M. Zwerger, W. Dür, J.-D. Bancal, and P. Sekatski. “Device-independent detection of genuine multipartite entanglement for all pure states”. *Physical Review Letters* **122**, 060502 (2019).
- [38] Xinhui Li, Yu Cai, Yunguang Han, Qiaoyan Wen, and Valerio Scarani. “Self-testing using only marginal information”. *Physical Review A* **98**, 052331 (2018).
- [39] T. Franz, F. Furrer, and R. F. Werner. “Extremal quantum correlations and cryptographic security”. *Physical Review Letters* **106**, 250502 (2011).
- [40] Lewis Woollorton. “Expanding bipartite Bell inequalities for maximum multi-partite randomness: accompanying python script”. <https://github.com/LewisWoollorton/expanding-BI.git> (2025).
- [41] Vern Paulsen. “Completely bounded maps and operator algebras”. *Cambridge Studies in Advanced Mathematics*. Cambridge University Press. (2003).
- [42] R. F. Werner and M. M. Wolf. “All-multipartite Bell-correlation inequalities for two dichotomic observables per site”. *Physical Review A* **64**, 032112 (2001).
- [43] D. M. Greenberger, M. Horne, and A. Zeilinger. “Going beyond Bell’s theorem”. In M. Kafatos, editor, *Bell’s Theorem, Quantum Mechanics and Conceptions of the Universe*. Pages 69–72. Kluwer Academic, Dordrecht, The Netherlands (1989).
- [44] Jędrzej Kaniewski. “Analytic and nearly optimal self-testing bounds for the Clauser-Horne-Shimony-Holt and Mermin inequalities”. *Physical Review Letters* **117**, 070402 (2016).
- [45] Jędrzej Kaniewski. “Self-testing of binary observables based on commutation”. *Physical Review A* **95**, 062323 (2017).
- [46] Rutvij Bhavsar, Sammy Ragy, and Roger Colbeck. “Improved device-independent randomness expansion rates using two sided randomness”. *New Journal of Physics* **25**, 093035 (2023).
- [47] Frédéric Dupuis, Omar Fawzi, and Renato Renner. “Entropy accumulation”. *Communications in Mathematical Physics* **379**, 867–913 (2020).
- [48] Frederic Dupuis and Omar Fawzi. “Entropy accumulation with improved second-order term”. *IEEE Transactions on Information Theory* **65**, 7596–7612 (2019).
- [49] Tony Metger, Omar Fawzi, David Sutter, and Renato Renner. “Generalised entropy accumulation”. *Communications in Mathematical Physics* **405**, 261 (2024).
- [50] Yanbao Zhang, Honghao Fu, and Emanuel Knill. “Efficient randomness certification by quantum probability estimation”. *Phys. Rev. Res.* **2**, 013016 (2020).

- [51] Lewis Wooltorton, Peter Brown, and Roger Colbeck. “Device-independent quantum key distribution with arbitrarily small nonlocality”. *Physical Review Letters* **132**, 210802 (2024).
- [52] Thinh P. Le, Chiara Meroni, Bernd Sturmfels, Reinhard F. Werner, and Timo Ziegler. “Quantum Correlations in the Minimal Scenario”. *Quantum* **7**, 947 (2023).
- [53] Victor Barizien, Pavel Sekatski, and Jean-Daniel Bancal. “Custom Bell inequalities from formal sums of squares”. *Quantum* **8**, 1333 (2024).
- [54] Miguel Navascués, Stefano Pironio, and Antonio Acín. “Bounding the set of quantum correlations”. *Physical Review Letters* **98**, 010401 (2007).
- [55] Miguel Navascués, Stefano Pironio, and Antonio Acín. “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations”. *New Journal of Physics* **10**, 073013 (2008).
- [56] Alban Jean-Marie Seguinard, Amélie Piveteau, Piotr Mironowicz, and Mohamed Bourennane. “Experimental certification of more than one bit of quantum randomness in the two inputs and two outputs scenario”. *New Journal of Physics* **25**, 113022 (2023).
- [57] Marco Tomamichel. “Quantum information processing with finite resources”. *Springer International Publishing*. (2016).
- [58] C. Jordan. “Essai sur la géométrie à n dimensions”. *Bulletin de la S. M. F.* **3**, 103–174 (1875).
- [59] Armin Tavakoli, Alejandro Pozas-Kerstjens, Peter Brown, and Mateus Araújo. “Semidefinite programming relaxations for quantum correlations”. *Reviews of Modern Physics* **96**, 045006 (2024).
- [60] 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”. *Physical Review A* **97**, 022104 (2018).
- [61] Cédric Bamps and Stefano Pironio. “Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing”. *Physical Review A* **91**, 052111 (2015).
- [62] Cezary Śliwa. “Symmetries of the Bell correlation inequalities”. *Physics Letters A* **317**, 165–168 (2003).
- [63] Peter Brown, Hamza Fawzi, and Omar Fawzi. “Computing conditional entropies for quantum correlations”. *Nature Communications* **12**, 575 (2021).

A Notation

Symbol	Meaning
a_k	Single bit associated to party k
$\mathbf{a}$	N bit string $(a_1, \dots, a_N)$
$\boldsymbol{\mu}$	$N - 2$ bit string
$\overline{\boldsymbol{\mu}_{k_1 k_2 \dots k_m}}$	$N - m$ bit string of measurement outcomes for parties $\{1, \dots, N\} \setminus \{k_1, k_2, \dots, k_m\}$
A_k	System associated to party k
$\mathbf{A}$	N partite system $A_1 \dots A_N$
$\overline{\mathbf{A}_{k_1 k_2 \dots k_m}}$	$N - m$ partite system $A_1 \dots A_N$ excluding $A_{k_1} A_{k_2} \dots A_{k_m}$
$\tilde{Q}_k$	Quantum system of party k
Q_k	Qubit system of party k
$\tilde{P}_{a_k x_k}^{(k)}$	Projector of party k associated with input $x_k \in \{0, 1\}$ and output $a_k \in \{0, 1\}$
$\tilde{A}_{x_k}^{(k)}$	Observable of party k associated with input x_k
$\tilde{P}_{\mathbf{a} \mathbf{x}}$	$\bigotimes_{k=1}^N \tilde{P}_{a_k x_k}^{(k)}$ where $\mathbf{a} = (a_1, \dots, a_k)$ and $\mathbf{x} = (x_1, \dots, x_k)$
$\tilde{P}_{\boldsymbol{\mu} \mathbf{0}}^{(\overline{k,l})}$	$\bigotimes_{k' \in \{1, \dots, N\} \setminus \{k, l\}} \tilde{P}_{a_{k'} x_{k'}}^{(k')}$ where $\boldsymbol{\mu} = (a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_{l-1}, a_{l+1}, \dots, a_N)$
$R_f(\omega)$	Asymptotic rate defined in Eq. (6) with the constraint $f(P_{\text{obs}}) = \omega$
$\mathcal{G}$	The set $(\pi/4, \pi/2) \cup (\pi/2, 3\pi/4) \cup (5\pi/4, 3\pi/2) \cup (3\pi/2, 7\pi/4)$
$\mathcal{F}$	The set $\left[(-\pi/4, \pi/4) \times \mathcal{G}\right] \cup \left[(-\pi/4, \pi/4) \setminus \{0\} \times \{\pi/2, 3\pi/2\}\right]$
$\langle M_N \rangle$	N partite MABK functional (see Eq. (1))

Table 1: Summary of notation.

B Proofs for expanded Bell expressions

We then have

B.1 Proof of Lemma 1

Lemma 1. *Let I be an expanded Bell expression according to Definition 2. The maximum quantum value of $\langle I \rangle$ is upper bounded by $\eta_N^Q := \sum_{k,l} c_{k,l} \eta^Q$.*

Proof. To prove the above, we show that the operator expression $\bar{I} = \eta_N^Q \mathbb{I} - I$ is non-negative. Since each bipartite expression satisfies $\langle I_{\boldsymbol{\mu}}^{(k,l)} \rangle \leq \eta^Q$ for all $\boldsymbol{\mu}, k, l$, we have positive expressions

$$\bar{I}_{\boldsymbol{\mu}}^{(k,l)} = \eta^Q \mathbb{I} - I_{\boldsymbol{\mu}}^{(k,l)} \succeq 0, \quad \forall \boldsymbol{\mu}, k, l. \quad (48)$$

$$\begin{aligned}
\bar{I}_{k,l} &:= \eta^Q \mathbb{I} - \sum_{\boldsymbol{\mu}} \tilde{P}_{\boldsymbol{\mu}|\mathbf{0}}^{(\overline{k,l})} I_{\boldsymbol{\mu}}^{(k,l)} \\
&= \sum_{\boldsymbol{\mu}} \tilde{P}_{\boldsymbol{\mu}|\mathbf{0}}^{(\overline{k,l})} \left(\eta^Q \mathbb{I} - I_{\boldsymbol{\mu}}^{(k,l)} \right) \\
&= \sum_{\boldsymbol{\mu}} \tilde{P}_{\boldsymbol{\mu}|\mathbf{0}}^{(\overline{k,l})} \bar{I}_{\boldsymbol{\mu}}^{(k,l)} \succeq 0.
\end{aligned} \quad (49)$$

Finally, it follows from the above that $\bar{I} \succeq 0$, which proves the upper bound on the maximum quantum value. We can also upper bound the maximum local value by $\sum_{k < l} c_{k,l} \eta^L < \sum_{k < l} c_{k,l} \eta^Q = \eta_N^Q$ since $\eta^L < \eta^Q$. $\square$

B.2 Uniqueness of binary distributions with fixed marginals

Before proving Lemma 2, we establish the following fact about classical distributions of bit strings when their conditional distributions are fixed. We will later use this result to justify that the distribution used for randomness achieving the quantum bound of an expanded Bell inequality is unique when the seed is a self-test.

Lemma 10. *Let $\mathbf{A} = A_1 \dots A_N$, $N \geq 3$, be a random N bit string, which takes values $\mathbf{a} \in \{0, 1\}^N$ according to the distribution $p_{\mathbf{A}}(\mathbf{a})$. Let $\mathbf{A}_{\bar{k}}$ be an $N - 2$ bit partition of the string $\mathbf{A}$, excluding bits N and $k \in \{1, \dots, N - 1\}$, taking values $\mathbf{a}_{\bar{k}} \in \{0, 1\}^{N-2}$. Then the distribution $p_{\mathbf{A}}(\mathbf{a})$ is entirely determined by the set of conditional distributions $\{p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k a_N | \mathbf{a}_{\bar{k}})\}_{k \in \{1, \dots, N-1\}}$, provided $p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k a_N | \mathbf{a}_{\bar{k}}) > 0$ for all $k, a_k, a_N, \mathbf{a}_{\bar{k}}$.*

Proof. The proof can be established using a recursive argument to solve for a set of marginal terms $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}})$, for a fixed $l \in \{1, \dots, N - 1\}$, from which the full distribution can be recovered. Let us choose another partition $k \in \{1, \dots, N - 1\}$ with $k \neq l$. Then we have

$$\begin{aligned} p_{\mathbf{A}}(\mathbf{a}) &= p_{\mathbf{A}_{\bar{k}}}(\mathbf{a}_{\bar{k}}) p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k a_N | \mathbf{a}_{\bar{k}}) \\ &= p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}}) p_{A_l A_N | \mathbf{A}_{\bar{l}}}(a_l a_N | \mathbf{a}_{\bar{l}}). \end{aligned} \quad (50)$$

By choosing different values of a_k , this implies the two sets of equations

$$\begin{aligned} p_{\mathbf{A}_{\bar{k}}}(\mathbf{a}_{\bar{k}}) p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k = 0, a_N | \mathbf{a}_{\bar{k}}) &= p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}}) p_{A_l A_N | \mathbf{A}_{\bar{l}}}(a_l a_N | a_k = 0, \mathbf{a}_{\bar{k}l}), \\ p_{\mathbf{A}_{\bar{k}}}(\mathbf{a}_{\bar{k}}) p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k = 1, a_N | \mathbf{a}_{\bar{k}}) &= p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}}) p_{A_l A_N | \mathbf{A}_{\bar{l}}}(a_l a_N | a_k = 1, \mathbf{a}_{\bar{k}l}). \end{aligned} \quad (51)$$

Since $p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k a_N | \mathbf{a}_{\bar{k}}) \neq 0$, we can rearrange both for $p_{\mathbf{A}_{\bar{k}}}$ and equate:

$$\begin{aligned} \frac{p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}}) p_{A_l A_N | \mathbf{A}_{\bar{l}}}(a_l a_N | a_k = 0, \mathbf{a}_{\bar{k}l})}{p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k = 0, a_N | \mathbf{a}_{\bar{k}})} &= \frac{p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}}) p_{A_l A_N | \mathbf{A}_{\bar{l}}}(a_l a_N | a_k = 1, \mathbf{a}_{\bar{k}l})}{p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k = 1, a_N | \mathbf{a}_{\bar{k}})}, \end{aligned} \quad (52)$$

where we have explicitly written in the a_k variable, and $\bar{k}l$ denotes a tuple excluding parties k and l . The above equation tells us, for every choice of $\mathbf{a}_{\bar{k}l} \in \{0, 1\}^{N-3}$, the unknowns $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, \mathbf{a}_{\bar{k}l})$ and $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 1, \mathbf{a}_{\bar{k}l})$ are linearly dependent. Hence we consider the 2^{N-3} unknowns $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, \mathbf{a}_{\bar{k}l})$, since the $a_k = 1$ terms can be computed via this linear dependence.

Consider choosing $k' \in \{1, \dots, N - 1\} \setminus \{k, l\}$. We can find a new set of equations by repeating the above

process with k' instead of k , to find:

$$\begin{aligned} \frac{p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 0, \mathbf{a}_{\bar{k}k'l})}{p_{A_{k'} A_N | \mathbf{A}_{\bar{k}'}}(a_{k'} = 0, a_N | a_k = 0, \mathbf{a}_{\bar{k}k'l})} &\cdot p_{A_l A_N | \mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 0, \mathbf{a}_{\bar{k}k'l}) \\ &= \frac{p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 1, \mathbf{a}_{\bar{k}k'l})}{p_{A_{k'} A_N | \mathbf{A}_{\bar{k}'}}(a_{k'} = 1, a_N | a_k = 0, \mathbf{a}_{\bar{k}k'l})} \\ &\cdot p_{A_l A_N | \mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 1, \mathbf{a}_{\bar{k}k'l}), \end{aligned} \quad (53)$$

where we have only included the $a_k = 0$ case, and written $\bar{k}k'l$ to denote a tuple excluding parties k, k', l . Notice that we have now identified linear dependence between $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 0, \mathbf{a}_{\bar{k}k'l})$ and $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 1, \mathbf{a}_{\bar{k}k'l})$. Neither of these equations contain a marginal with $a_k = 1$, so they relate pairs of unknowns distinctly to the parings from the previous equations, and are hence linearly independent. As before, we can proceed with the 2^{N-4} unknowns $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}} | a_k = 0, a_{k'} = 0, \mathbf{a}_{\bar{k}k'l})$ and recover the others by linear dependence.

We can apply the above procedure iteratively a total of $N - 2$ times (for every $k \in \{1, \dots, N - 1\} \setminus \{l\}$), halving the number of unknowns on every iteration by establishing linear dependence. Starting with 2^{N-2} marginals $p_{\mathbf{A}_{\bar{l}}}(\mathbf{a}_{\bar{l}})$, this leaves us with 1 unknown, which can be found from normalization, completing the proof. $\square$

Now we can present a corollary which will be needed for the proof of Lemma 2, showing the conditional distribution used for randomness, $p(\mathbf{a} | \mathbf{0})$, achieving the quantum bound of an expanded Bell expression is unique.

Corollary 2. *Let I be an expanded Bell inequality according to Definition 2 with $c_{k,N} = 1$ if $k < N$ and zero otherwise, and η_N^Q be as defined in Lemma 1. Let I be constructed using a seed with the self-testing properties described in Lemma 2, and $p(\mathbf{a} | \mathbf{x})$ denote a quantum behaviour that achieves $\langle I \rangle = \eta_N^Q$. Then provided $p(a_k a_N | \mathbf{b}_{\bar{k}}, \mathbf{x} = \mathbf{0}) > 0$ for all $k, a_k, a_N, \mathbf{b}_{\bar{k}}$, the conditional the distribution $p(\mathbf{a} | \mathbf{0})$ is unique.*

Proof. If the quantum value $\langle I \rangle = \eta_N^Q$ is achieved, we must have

$$\sum_{k=1}^{N-1} \langle \eta^Q \mathbb{I} - \sum_{\mu} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle = 0, \quad (54)$$

which implies $\langle \eta^Q \mathbb{I} - \sum_{\mu} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle = 0$ from Eq. (49). Next, we observe

$$\begin{aligned} \langle \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle &= \text{Tr} \left[\tilde{P}_{\mu|0}^{(k,N)} \rho_{\tilde{Q}E} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \right] \\ &\leq |\text{Tr} \left[\tilde{P}_{\mu|0}^{(k,N)} \rho_{\tilde{Q}E} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \right]| \\ &\leq \|\tilde{P}_{\mu|0}^{(k,N)} \rho_{\tilde{Q}E} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)}\|_1, \end{aligned} \quad (55)$$

where the second inequality follows from the fact $|\text{Tr}[A]| \leq \|A\|_1$ for any operator A (see e.g. [57, Lemma 3.3]). Above, $\rho_{\tilde{Q}E} = |\Psi\rangle\langle\Psi|$, where $|\Psi\rangle$ is the quantum state that realizes the correlations, and identities on E are implicit. Next we combine Eq. (55) with Hölder's inequality to obtain

$$\begin{aligned} \langle \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle &\leq \|\tilde{P}_{\mu|0}^{(k,N)} \rho_{\tilde{Q}E} \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)}\|_1 \\ &\leq \|\tilde{P}_{\mu|0}^{(k,N)} \rho_{\tilde{Q}E} \tilde{P}_{\mu|0}^{(k,N)}\|_1 \|I_{\mu}^{(k,N)}\|_{\infty} = \langle \tilde{P}_{\mu|0}^{(k,N)} \rangle \eta^Q. \end{aligned} \quad (56)$$

Since

$$\sum_{\mu} \left(\langle \tilde{P}_{\mu|0}^{(k,N)} \rangle \eta^Q - \langle \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle \right) = 0 \quad (57)$$

we conclude $\langle \tilde{P}_{\mu|0}^{(k,N)} I_{\mu}^{(k,N)} \rangle = \langle \tilde{P}_{\mu|0}^{(k,N)} \rangle \eta^Q$. Now we can write

$$\langle \Psi' | I_{\mu}^{(k,N)} | \Psi' \rangle = \eta^Q, \quad (58)$$

where we define the normalized state $|\Psi'\rangle = \tilde{P}_{\mu|0}^{(k,N)} |\Psi\rangle / \sqrt{\langle \tilde{P}_{\mu|0}^{(k,N)} \rangle}$. The self-testing properties of $I_{\mu}^{(k,N)}$ (defined in Lemma 2) then determine the unique correlations $\langle \Psi' | \tilde{P}_{a_k|x}^{(k)} \tilde{P}_{a_N|y}^{(N)} | \Psi' \rangle = p_{\mu}^{k,N}(a_k a_N | xy)$. Using the definition of $|\Psi'\rangle$, we find

$$\langle \tilde{P}_{\mu|0}^{(k,N)} \tilde{P}_{a_k|0}^{(k)} \tilde{P}_{a_N|0}^{(N)} \rangle = \langle \tilde{P}_{\mu|0}^{(k,N)} \rangle p_{\mu}^{k,N}(a_k a_N | 00). \quad (59)$$

We can now directly apply Lemma 10. Explicitly, we have $p_A(\mathbf{a}) = \langle \tilde{P}_{\mu|0}^{(k,N)} \tilde{P}_{a_k|0}^{(k)} \tilde{P}_{a_N|0}^{(N)} \rangle$, and the $N-1$ fixed conditional distributions $p_{A_k A_N | \mathbf{A}_{\bar{k}}}(a_k a_N | \mathbf{a}_{\bar{k}}) = p(a_k a_N | \mathbf{a}_{\bar{k}}, \mathbf{x} = \mathbf{0}) = p_{\mu}^{k,N}(a_k a_N | 00)$, where we have identified $\mu = \mathbf{a}_{\bar{k}}$. $\square$

B.3 Proof of Lemma 2

We can now proceed to prove Lemma 2 in the main text, which is restated below:

Lemma 2. *Let I be an expanded N -party Bell expression defined in Definition 2 with binary inputs and outputs, and $c_{k,l} = 1$ if $l = N$ and $k < N$, and zero otherwise. Suppose that for every $I_{\mu}^{(k,N)}$, there exists an SOS decomposition that self-tests the same pure bipartite entangled state $|\Phi\rangle$ between parties k and N , along with some ideal measurements $P_{a_k|x_k}^{(k)}, P_{a_N|x_N}^{(N)}$, according to Definition 1, satisfying $\langle \Phi | P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)} | \Phi \rangle > 0$ for all a_k, a_N . Then for any strategy $|\Psi\rangle_{\tilde{Q}E}, \{\{\tilde{P}_{a_k|x_k}^{(k)}\}_{x_k}\}_k$ that achieves $\langle I \rangle = \eta_N^Q$, the post-measurement state $\rho_{RE|x}$, for measurement settings $\mathbf{x} = \mathbf{0}$, admits the tensor product decomposition,*

$$\rho_{RE|0} = \rho_{R|0} \otimes \rho_E. \quad (60)$$

Before proceeding, we summarize the proof. First, we provide an SOS decomposition for the shifted Bell operator $\bar{I}$. This is constructed using the SOS of every bipartite Bell expression $I_{\mu}^{(k,N)}$, which exist by assumption. As a result, any state and set of measurements achieving $\bar{I} = 0$ (equivalently, the optimal quantum bound of I) must satisfy the same algebraic relations imposed by the SOS polynomials of $I_{\mu}^{(k,N)}$ after an appropriate projection is performed by the other parties. In the next step, we employ Jordan's lemma, reducing the problem to qubits and simplifying the following analysis. In the third step we show that, for each N qubit system in the Jordan decomposition, the SOS constraints self-test the bipartite state shared by parties k and N after the other $N-2$ parties have performed a measurement. Finally, we write down the global post-measurement state and show how the aforementioned self-testing properties imply that it must be in a tensor product with E .

Proof. We proceed in the four steps described above.

Step 1: SOS decompositions. By assumption, there exist SOS decompositions for the bipartite Bell operators $\bar{I}_{\mu}^{(k,N)}$, i.e., we can write

$$\bar{I}_{\mu}^{(k,N)} = \sum_i M_{\mu,i}^{(k,N)\dagger} M_{\mu,i}^{(k,N)}, \quad (61)$$

where $M_{\mu,i}^{(k,N)}$ is a polynomial of the operators $A_{x_k}^{(k)}, A_{x_N}^{(N)}$. We now use this to build an SOS decomposition for $\bar{I}$,

$$\begin{aligned} \bar{I} &= \sum_{k=1}^{N-1} \bar{I}_{k,N} \\ &= \sum_{k=1}^{N-1} \sum_{\mu} \tilde{P}_{\mu|0}^{(k,N)} \bar{I}_{\mu}^{(k,N)} \\ &= \sum_{k=1}^{N-1} \sum_{\mu} \sum_i \left(\tilde{P}_{\mu|0}^{(k,N)} M_{\mu,i}^{(k,N)} \right)^{\dagger} \left(\tilde{P}_{\mu|0}^{(k,N)} M_{\mu,i}^{(k,N)} \right). \end{aligned} \quad (62)$$

Now, for some physical state $|\Psi\rangle$ and observables, observation of the maximum quantum value of the Bell expression $\langle I \rangle$, or equivalently $\langle \bar{I} \rangle = 0$, implies the algebraic constraints

$$\tilde{P}_{\mu|0}^{(k,N)} M_{\mu,i}^{(k,N)} |\Psi\rangle = 0, \quad \forall k \in \{1, \dots, N-1\}, \mu, i. \quad (63)$$

Since $\tilde{P}_{\mu|0}^{(k,N)}$ and $M_{\mu,i}^{(k,N)}$ commute, we define the post-measurement states

$$|\Psi_{\mu,k}\rangle = \frac{\tilde{P}_{\mu|0}^{(k,N)} |\Psi\rangle}{\sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{(k,N)} | \Psi \rangle}}, \quad (64)$$

and arrive at the set of algebraic constraints, for a fixed k, μ ,

$$M_{\mu,i}^{(k,N)} |\Psi_{\mu,k}\rangle = 0, \quad \forall i. \quad (65)$$

Step 2: Jordan's lemma. Next, we employ Jordan's lemma [58], which allows us to reduce our analysis to a convex combination of N -qubit systems. Specifically, we will use the version presented in [46, Lemma 4], and apply [46, Lemma 6] to consider (without loss of generality) the set of post-measurement states ρ_{RE} arising from block diagonal measurement operators with block size 2×2 , and a state which has support only on each 2×2 block. We write $\mathcal{H}_{Q_k}$ to denote the qubit Hilbert space of party k , and introduce a flag system $\mathcal{H}_{F_k}$ which indicates the 2×2 block. Then for every party $\mathcal{H}_{\bar{Q}_k} = \mathcal{H}_{Q_k} \otimes \mathcal{H}_{F_k}$, and the purified state takes the form

$$|\Psi\rangle_{\bar{Q}E} = \sum_{\mathbf{f}} \sqrt{p_{\mathbf{f}}} |\mathbf{f}\rangle_F \otimes |\varphi^{\mathbf{f}}\rangle_{QE} \otimes |\mathbf{f}\rangle_{E'}, \quad (66)$$

where $\rho^{\mathbf{f}} = \text{Tr}_E[|\varphi^{\mathbf{f}}\rangle\langle\varphi^{\mathbf{f}}|]$ is an N -qubit state corresponding to block $\mathbf{f} = (f_1, \dots, f_N)$, where f_k indexes the block for party k and $p_{\mathbf{f}} > 0$ form a probability distribution. Tracing out EE' , the total state received by the parties is then a classical quantum state, where the classical register indexes which N -qubit state is in the quantum system. Eve then holds a purification of the state in register E along with a label in register E' telling her which block it pertains to. Similarly, the projectors decompose

$$\tilde{P}_{a_k|x_k}^{(k)} = \sum_{f_k} |f_k\rangle\langle f_k|_{F_k} \otimes P_{a_k|x_k}^{(f_k)}, \quad (67)$$

where $P_{a_k|x_k}^{(f_k)}$ is a rank one single qubit projector on $\mathcal{H}_{Q_k}$ corresponding to block f_k [58, 13, 46]. The projector performed on all systems except k and N for inputs 0 is

$$\tilde{P}_{\mu|0}^{(\bar{k}, \bar{N})} = \sum_{\mathbf{f}_{\bar{k}\bar{N}}} |\mathbf{f}_{\bar{k}\bar{N}}\rangle\langle\mathbf{f}_{\bar{k}\bar{N}}|_{F_{\bar{k}\bar{N}}} \otimes P_{\mu|0}^{(\mathbf{f}_{\bar{k}\bar{N}})}, \quad (68)$$

where the notation $\bar{k}\bar{N}$ is understood to denote a tuple excluding entries k, N , e.g., $\mathbf{f}_{\bar{k}\bar{N}}$ is a tuple of indices $f_{k'}$ for $k' \neq k, N$. $P_{\mu|0}^{(\mathbf{f}_{\bar{k}\bar{N}})} = \bigotimes_{k' \neq k, N} P_{\mu_{k'}|0}^{(f_{k'})}$ is then a rank one projector on the $N - 2$ qubit Hilbert space shared by all parties excluding k and N , corresponding to the block $\mathbf{f}_{\bar{k}\bar{N}}$.

We can now apply Jordan's lemma to the projected states $|\Psi_{\mu,k}\rangle := \tilde{P}_{\mu|0}^{(\bar{k}, \bar{N})} |\Psi\rangle / \sqrt{\langle \tilde{P}_{\mu|0}^{(\bar{k}, \bar{N})} \rangle}$, by computing

$$\begin{aligned} |\Psi_{\mu,k}\rangle &= \sum_{\mathbf{f}} \sqrt{p_{\mathbf{f}}} |\mathbf{f}\rangle_F \\ &\quad \otimes \frac{\left(\mathbb{I}_{Q_k Q_N E} \otimes P_{\mu|0}^{(\mathbf{f}_{\bar{k}\bar{N}})} \right) |\varphi^{\mathbf{f}}\rangle_{QE}}{\sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{(\bar{k}, \bar{N})} | \Psi \rangle}} \otimes |\mathbf{f}\rangle_{E'} \\ &= \sum_{\mathbf{f} \in S_{\mu}^k} \sqrt{p_{\mathbf{f}}} |\mathbf{f}\rangle_F \otimes |\phi_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle_{Q_{\bar{k}\bar{N}}} \\ &\quad \otimes |\psi_{\mu,k}^{\mathbf{f}}\rangle_{Q_k Q_N E} \otimes |\mathbf{f}\rangle_{E'} \end{aligned} \quad (69)$$

where S_{μ}^k is the set of blocks for which the projection is non-zero. Here $\mathbf{f}$ is understood to be the concatenation of $f_k, f_N, \mathbf{f}_{\bar{k}\bar{N}}$, and we used the fact that the qubit projectors are rank one to write the state as a tensor product,

$$\begin{aligned} &\frac{\left(\mathbb{I}_{Q_k Q_N E} \otimes P_{\mu|0}^{(\mathbf{f}_{\bar{k}\bar{N}})} \right) |\varphi^{\mathbf{f}}\rangle_{QE}}{\sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{(\bar{k}, \bar{N})} | \Psi \rangle}} \\ &= \begin{cases} |\phi_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle_{Q_{\bar{k}\bar{N}}} \otimes |\psi_{\mu,k}^{\mathbf{f}}\rangle_{Q_k Q_N E} & \text{if } \mathbf{f} \in S_{\mu}^k, \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (70)$$

where $|\phi_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle = \bigotimes_{k' \neq k, N} |\phi_{\mu_{k'}}^{f_{k'}}\rangle$, and $P_{\mu_{k'}|0}^{(f_{k'})} = |\phi_{\mu_{k'}}^{f_{k'}}\rangle\langle\phi_{\mu_{k'}}^{f_{k'}}|$. $|\psi_{\mu,k}^{\mathbf{f}}\rangle_{Q_k Q_N E}$ is then the state held by parties k, N and Eve following the projection. By writing $|\tilde{\phi}_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle_{\bar{Q}_{\bar{k}\bar{N}}} = |\mathbf{f}_{\bar{k}\bar{N}}\rangle_{F_{\bar{k}\bar{N}}} \otimes |\phi_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle_{Q_{\bar{k}\bar{N}}}$, and $|\tilde{\psi}_{\mu,k}^{\mathbf{f}}\rangle_{\bar{Q}_k \bar{Q}_N E} = |f_k f_N\rangle_{F_k F_N} \otimes |\psi_{\mu,k}^{\mathbf{f}}\rangle_{Q_k Q_N E}$, we can re-write the constraints in Eq. (65) as (suppressing the identity operator on Eve's systems)

$$\begin{aligned} &\sum_{\mathbf{f} \in S_{\mu}^k} \sqrt{p_{\mathbf{f}}} |\tilde{\phi}_{\mu}^{\mathbf{f}_{\bar{k}\bar{N}}}\rangle_{\bar{Q}_{\bar{k}\bar{N}}} \\ &\quad \otimes M_{\mu,i}^{(k,N)} |\tilde{\psi}_{\mu,k}^{\mathbf{f}}\rangle_{\bar{Q}_k \bar{Q}_N E} \otimes |\mathbf{f}\rangle_{E'} = 0, \end{aligned} \quad (71)$$

implying, for a fixed k, μ ,

$$M_{\mu,i}^{(k,N)} |\tilde{\psi}_{\mu,k}^{\mathbf{f}}\rangle_{\bar{Q}_k \bar{Q}_N E} = 0, \quad \forall i, \forall \mathbf{f} \in S_{\mu}^k. \quad (72)$$

This subsequently implies

$$M_{\mu,i}^{(k,N)} |f_k, f_N\rangle |\psi_{\mu,k}^{\mathbf{f}}\rangle_{Q_k Q_N E} = 0, \quad \forall i, \forall \mathbf{f} \in S_{\mu}^k, \quad (73)$$

where we wrote

$$M_{\mu,i}^{(k,N)} = \sum_{f_k, f_N} |f_k, f_N\rangle\langle f_k, f_N|_{F_k F_N} \otimes M_{\mu,i}^{(k,N)} |f_k, f_N\rangle \quad (74)$$

using the decomposition in Eq. (67).

Step 3: Bipartite self-tests. For the next part of the proof, we use the self-testing properties of the bipartite Bell expression $I_{\mu}^{(k,l)}$, which are in turn derived from the SOS polynomials, and the constraints they impose on any state that achieves the maximum quantum value. Explicitly, for any state which satisfies the polynomial constraints, there exists a local isometry which transforms that state into an ideal state in tensor product with some junk. Due to Jordan's lemma, the constraints in Eq. (73) certify the existence of a pair of local unitaries on the qubit registers of parties k and N , such that, for every measurement outcome μ and block combination $\mathbf{f} \in S_{\mu}^k$, $U_{\mu,k}^{\mathbf{f}}$, there exists a fixed two qubit state $|\Phi\rangle$ and

⁷Note that the notation $U_{\mu,k}^{\mathbf{f}}$ suppresses the tensor product of local unitaries to ease notation, that is, $U_{\mu,k}^{\mathbf{f}} = V \otimes W$ where V is a unitary on system Q_k and W is a unitary on system Q_N .

projectors $P_{a_k|x_k}^{(k)}$, $P_{a_N|x_N}^{(N)}$ satisfying

$$\begin{aligned} & \left(\mathbb{I}_{F_k F_N} \otimes U_{\mu,k}^f \otimes \mathbb{I}_E \right) \left(\tilde{P}_{a_k|x_k}^{(k)} \otimes \tilde{P}_{a_N|x_N}^{(N)} \right) |\tilde{\psi}_{\mu,k}^f\rangle_{\tilde{Q}_k \tilde{Q}_N E} \\ &= |f_k f_N\rangle_{F_k F_N} \otimes \left(P_{a_k|x_k}^{(k)} \otimes P_{a_N|x_N}^{(N)} \right) |\Phi\rangle_{Q_k Q_N} \otimes |\lambda_{\mu,k}^f\rangle_E, \end{aligned} \quad (75)$$

where $|\lambda_{\mu,k}^f\rangle_E$ is the junk system left over in the Eve register, which a priori may depend on μ, k and f . Since we required that all bipartite expressions self-test the same state up to local unitaries, the above equation holds for all μ .

Step 4: Post-measurement state. Finally, we use the existence of the local unitaries $U_{\mu,k}^f$ to analyze the global post-measurement state held by all parties following their zero measurement, and storing their outcomes in the register R , which takes the form

$$\begin{aligned} \rho_{REE'|0} &= \sum_{a_k, a_N, \mu} |a_k a_N \mu\rangle \langle a_k a_N \mu|_R \otimes \\ &\text{Tr}_{\tilde{Q}} \left[\left(\tilde{P}_{a_k|0}^{(k)} \otimes \tilde{P}_{a_N|0}^{(N)} \otimes \tilde{P}_{\mu|0}^{\overline{(k,N)}} \otimes \mathbb{I}_{EE'} \right) |\Psi\rangle \langle \Psi| \right]. \end{aligned} \quad (76)$$

We can rewrite the term inside the partial trace

$$\begin{aligned} & \left(\tilde{P}_{a_k|0}^{(k)} \otimes \tilde{P}_{a_N|0}^{(N)} \otimes \tilde{P}_{\mu|0}^{\overline{(k,N)}} \otimes \mathbb{I}_{EE'} \right) |\Psi\rangle \\ &= \sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle} \left(\tilde{P}_{a_k|0}^{(k)} \otimes \tilde{P}_{a_N|0}^{(N)} \otimes \mathbb{I}_{\tilde{Q}_{(k,N)}} \otimes \mathbb{I}_{EE'} \right) |\Psi_{\mu,k}\rangle \\ &= \sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle} \sum_{f \in S_{\mu}^k} \sqrt{p_f} |\tilde{\phi}_{\mu}^{f_{kN}}\rangle_{\tilde{Q}_{kN}} \\ &\quad \otimes \left(\tilde{P}_{a_k|0}^{(k)} \otimes \tilde{P}_{a_N|0}^{(N)} \right) |\tilde{\psi}_{\mu,k}^f\rangle_{\tilde{Q}_k \tilde{Q}_N E} \otimes |f\rangle_{E'} \\ &= \sqrt{\langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle} \sum_{f \in S_{\mu}^k} \sqrt{p_f} |\tilde{\phi}_{\mu}^{f_{kN}}\rangle_{\tilde{Q}_{kN}} \\ &\quad \otimes |f_k f_N\rangle_{F_k F_N} \otimes U_{\mu,k}^{f\dagger} \left(P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)} \right) |\Phi\rangle_{Q_k Q_N} \\ &\quad \otimes |\lambda_{\mu,k}^f\rangle_E \otimes |f\rangle_{E'}. \end{aligned} \quad (77)$$

Now, tracing out $\tilde{Q}$ equates to tracing out systems FQ , which yields

$$\begin{aligned} & \text{Tr}_{\tilde{Q}} \left[\left(\tilde{P}_{a_k|0}^{(k)} \otimes \tilde{P}_{a_N|0}^{(N)} \otimes \tilde{P}_{\mu|0}^{\overline{(k,N)}} \otimes \mathbb{I}_{EE'} \right) |\Psi\rangle \langle \Psi| \right] \\ &= \langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle \langle \Phi | \left(P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)} \right) | \Phi \rangle \\ &\quad \cdot \sum_{f \in S_{\mu}^k} p_f |\lambda_{\mu,k}^f\rangle \langle \lambda_{\mu,k}^f|_E \otimes |f\rangle \langle f|_{E'}. \end{aligned} \quad (78)$$

By employing Corollary 2, we know that, following observation of $\langle I \rangle = \eta^Q$, the marginals $\langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle$ are unique; hence the following equality holds:

$$\langle \Psi | \tilde{P}_{\mu|0}^{\overline{(k,N)}} | \Psi \rangle \langle \Phi | \left(P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)} \right) | \Phi \rangle = p(a_k, a_N, \mu|0), \quad (79)$$

where $\{p(a_k, a_N, \mu|0)\}_{a_k, a_N, \mu}$ are those unique correlations needed to achieve $\langle I \rangle = \eta^Q$. We then find

$$\begin{aligned} \rho_{REE'|0} &= \sum_{a_k, a_N, \mu} p(a_k, a_N, \mu|0) |a_k a_N \mu\rangle \langle a_k a_N \mu|_R \\ &\quad \otimes \sum_{f \in S_{\mu}^k} p_f |\lambda_{\mu,k}^f\rangle \langle \lambda_{\mu,k}^f|_E \otimes |f\rangle \langle f|_{E'}. \end{aligned} \quad (80)$$

For a fixed $k \in \{1, \dots, N-1\}$, Eve can still establish correlations with the N parties excluding k and N , by choosing the sets $\{S_{\mu}^k\}_{\mu}$ to be disjoint. She could then distinguish the different outcomes μ by a projective measurement on E' . She can also establish correlations by making the set $\{|\lambda_{\mu,k}^f\rangle\}_{\mu}$ distinguishable and measuring E . To show that Eve is in fact uncorrelated with all parties, we use the fact that there are $N-1$ choices of k , i.e., one can choose different combinations of parties to maximally violate the bipartite self-tests.

We now include extra notation $\overline{kN}$, $\mu \rightarrow \mu_{\overline{kN}}$, which denotes a tuple of $N-2$ outcomes for all parties excluding k and N , and define the following state, which is just a rewriting of Eq. (80) making the choice k explicit:

$$\begin{aligned} \rho_{REE'|0}^k &:= \sum_{a_k, a_N, \mu_{\overline{kN}}} p(a_k, a_N, \mu_{\overline{kN}}|0) \\ &\quad \cdot |a_k a_N \mu_{\overline{kN}}\rangle \langle a_k a_N \mu_{\overline{kN}}|_R \\ &\quad \otimes \sum_{f \in S_{\mu_{\overline{kN}}}^k} p_f |\lambda_{\mu_{\overline{kN}},k}^f\rangle \langle \lambda_{\mu_{\overline{kN}},k}^f|_E \otimes |f\rangle \langle f|_{E'}. \end{aligned} \quad (81)$$

Now, we could equally define a second state $\rho_{REE'|0}^{k'}$ where $k' \neq k, N$ in the exact same way, except we choose k' instead of k as the self-testing party. Since both states are a rewriting of the same post measurement state $\rho_{REE|0}$ given by Eq. (80), they must be equal, and we establish the equalities

$$\rho_{REE'|0}^k = \rho_{REE'|0}^{k'}, \quad \forall k, k' \in \{1, \dots, N-1\}, \quad k' \neq k. \quad (82)$$

These equalities will allow us to place constraints on the sets $S_{\mu_{\overline{kN}}}^k$, and ultimately show that they are all equal. Specifically, Eq. (82) implies⁸

$$\begin{aligned} & \sum_{f \in S_{\mu_{\overline{kN}}}^k} p_f |\lambda_{\mu_{\overline{kN}},k}^f\rangle \langle \lambda_{\mu_{\overline{kN}},k}^f|_E \otimes |f\rangle \langle f|_{E'} \\ &= \sum_{f \in S_{\mu_{\overline{k'N}}}^{k'}} p_f |\lambda_{\mu_{\overline{k'N}},k'}^f\rangle \langle \lambda_{\mu_{\overline{k'N}},k'}^f|_E \otimes |f\rangle \langle f|_{E'}. \end{aligned} \quad (83)$$

By tracing out system E , and using the fact that $p_f \neq 0$, Eq. (83) can only hold if

$$S_{\mu_{\overline{kN}}}^k = S_{\mu_{\overline{k'N}}}^{k'}, \quad \forall a_1, \dots, a_{N-1}. \quad (84)$$

⁸Note that we have implicitly equated the rewriting of the same outcome string $\mathbf{a} = a_k a_N \mu_{\overline{kN}} = a_{k'} a_N \mu_{\overline{k'N}}$, and only considered cases that appear in the summation, i.e., strings with nonzero probability, $p(\mathbf{a}|0) > 0$.

Consider the case $k = 1$ and $k' = 2$. Then we have for a fixed $a_3, \dots, a_{N-1}$,

$$S_{(a_2, a_3, \dots, a_{N-1})}^1 = S_{(a_1, a_3, a_4, \dots, a_{N-1})}^2, \quad \forall a_2. \quad (85)$$

Since the RHS is independent of a_2 , we must have that $S_{(a_2, a_3, \dots, a_{N-1})}^1 \equiv S_{(a_3, \dots, a_{N-1})}^1$, i.e., the LHS is also independent of a_2 . We can apply the same argument for $k' = 3, 4, \dots, N-1$, deducing that

$$S_{(a_2, a_3, \dots, a_{N-1})}^1 \equiv S, \quad (86)$$

which is independent of the outcome string $\mu_{1N} = (a_2, a_3, \dots, a_{N-1})$. There is nothing unique about choosing $k = 1$, allowing us to write

$$\begin{aligned} \rho_{REE'|0}^k &= \sum_{a_k, a_N, \mu_{kN}} p(a_k, a_N, \mu_{kN}|0) \\ &\quad \cdot |a_k a_N \mu_{kN}\rangle \langle a_k a_N \mu_{kN}|_R \\ &\quad \otimes \sum_{f \in S} p_f |\lambda_{\mu_{kN}, k}^f\rangle \langle \lambda_{\mu_{kN}, k}^f|_E \otimes |f\rangle \langle f|_{E'}, \end{aligned} \quad (87)$$

Hence Eve can learn nothing about any of the outcomes by measuring her register E' . What remains is to deal with the vectors $|\lambda_{\mu_{kN}, k}^f\rangle$. We will show that the set $\{|\lambda_{\mu_{kN}, k}^f\rangle\}_{\mu_{kN}}$ contains one linearly independent vector, hence Eve can learn nothing about the outcome μ_{kN} by measuring her register E .

Using the same approach as before, Eq. (82) now implies, for every $f \in S$,

$$\begin{aligned} &\sum_{a_k a_N \mu_{kN}} \langle \Psi | \tilde{P}_{\mu_{kN}}^{(k, N)} | \Psi \rangle \langle \Phi | (P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)}) | \Phi \rangle \\ &\quad \cdot |a_k a_N \mu_{kN}\rangle \langle a_k a_N \mu_{kN}| \otimes |\lambda_{\mu_{kN}, k}^f\rangle \langle \lambda_{\mu_{kN}, k}^f| \\ &= \sum_{a_{k'} a_N \mu_{k'N}} \langle \Psi | \tilde{P}_{\mu_{k'N}}^{(k', N)} | \Psi \rangle \langle \Phi | (P_{a_{k'}|0}^{(k')} \otimes P_{a_N|0}^{(N)}) | \Phi \rangle \\ &\quad \cdot |a_{k'} a_N \mu_{k'N}\rangle \langle a_{k'} a_N \mu_{k'N}| \otimes |\lambda_{\mu_{k'N}, k'}^f\rangle \langle \lambda_{\mu_{k'N}, k'}^f|. \end{aligned} \quad (88)$$

Now equating for a given value of $\mathbf{a} = (a_1, \dots, a_k, \dots, a_{k'}, \dots, a_N)$, we obtain

$$\begin{aligned} &\langle \Psi | \tilde{P}_{\mu_{kN}}^{(k, N)} | \Psi \rangle \langle \Phi | (P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)}) | \Phi \rangle \\ &\quad \cdot |\lambda_{\mu_{kN}, k}^f\rangle \langle \lambda_{\mu_{kN}, k}^f| \\ &= \langle \Psi | \tilde{P}_{\mu_{k'N}}^{(k', N)} | \Psi \rangle \langle \Phi | (P_{a_{k'}|0}^{(k')} \otimes P_{a_N|0}^{(N)}) | \Phi \rangle \\ &\quad \cdot |\lambda_{\mu_{k'N}, k'}^f\rangle \langle \lambda_{\mu_{k'N}, k'}^f|. \end{aligned} \quad (89)$$

We can hence conclude the two vectors are linearly dependent,

$$|\lambda_{\mu_{kN}, k}^f\rangle \sim |\lambda_{\mu_{k'N}, k'}^f\rangle, \quad (90)$$

where $\sim$ denotes equality up to a constant⁹. Notice

⁹Note it is sufficient to just consider the cases for which the probabilities $\langle \Psi | \tilde{P}_{\mu_{kN}}^{(k, N)} | \Psi \rangle \langle \Phi | (P_{a_k|0}^{(k)} \otimes P_{a_N|0}^{(N)}) | \Phi \rangle$ are nonzero.

If they are zero, then the vector $|\lambda_{\mu_{kN}, k}^f\rangle$ will not appear in the summation, and Eve will never observe it.

that on the left hand side of the above equation, one has the freedom to choose any value of $a_{k'}$ without changing the right hand side up to a constant. Therefore

$$|\lambda_{\mu_{kN}}^f|_{a_{k'}=0, k}\rangle \sim |\lambda_{\mu_{kN}}^f|_{a_{k'}=1, k}\rangle. \quad (91)$$

There are 2^{N-3} such equations written above, corresponding to the different choices of μ_{kN} once $a_{k'}$ is fixed. We have 2^{N-2} unknowns that we want to fix, hence this is sufficient to show linear dependence for the $N = 3$ case. If $N > 3$, we can select another party k'' not equal to k, k', N and add another 2^{N-3} equations,

$$|\lambda_{\mu_{kN}, k}^f\rangle \sim |\lambda_{\mu_{k''N}, k''}^f\rangle, \quad (92)$$

allowing us to write

$$\begin{aligned} |\lambda_{\mu_{kN}}^f|_{a_{k'}=0, k}\rangle &\sim |\lambda_{\mu_{kN}}^f|_{a_{k'}=1, k}\rangle \\ |\lambda_{\mu_{kN}}^f|_{a_{k''}=0, k}\rangle &\sim |\lambda_{\mu_{kN}}^f|_{a_{k''}=1, k}\rangle. \end{aligned} \quad (93)$$

The combination of the two will solve the $N = 4$ case. We can see that for the first choice of k' , we get 2^{N-3} equations, which halves the number of unknowns to 2^{N-3} . Each new choice of k' , of which there are $N-2$ in total, halves the number of unknowns, until we are left with one, which shows that every vector can be written as a scalar multiple of a fixed vector, i.e.,

$$|\lambda_{\mu_{kN}, k}^f\rangle = \beta_{\mu_{kN}}^k |\lambda_k^f\rangle, \quad (94)$$

for some constants $\beta_{\mu_{kN}}^k$, which, due to normalization, must be equal to a phase. Substituting the above into (88) (noting that any phase factors will cancel when taking the projector), we find $|\lambda_k^f\rangle = |\lambda_{k'}^f\rangle$, hence we can drop the k label.

Now we finally conclude

$$\begin{aligned} \rho_{REE'|0} &\equiv \rho_{REE'|0}^k = \left(\sum_{a_k, a_N, \mu_{kN}} p(a_k, a_N, \mu_{kN}|0) \right. \\ &\quad \cdot |a_k a_N \mu_{kN}\rangle \langle a_k a_N \mu_{kN}|_R \\ &\quad \left. \otimes \left(\sum_{f \in S} p_f |\lambda_k^f\rangle \langle \lambda_k^f|_E \otimes |f\rangle \langle f|_{E'} \right) \right), \end{aligned} \quad (95)$$

which is of the desired tensor product form, and completes the proof. $\square$

B.4 Proof of Corollary 1

Corollary 1. Let I, η_N^Q be defined as in Lemma 2. Then

$$R_I(\eta_N^Q) = H(\{p(\mathbf{a}|\mathbf{0})\}), \quad (96)$$

where $H(\{p_i\})$ is the Shannon entropy of a distribution $\{p_i\}_i$.

Proof. Corollary 1 directly follows from Lemma 2 since the state on which the entropy is evaluated has Eve's part in tensor product. This implies, for all compatible post measurement states, $H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_{\rho_{\mathbf{RE}|\mathbf{0}}} = H(\mathbf{R}|\mathbf{X} = \mathbf{0})_{\rho_{\mathbf{R}|\mathbf{0}}}$. Since $\rho_{\mathbf{R}|\mathbf{0}}$, derived explicitly in the proof of Lemma 2, is the same for all compatible states and measurements, we obtain the claim:

$$\begin{aligned} & \inf_{\substack{|\Psi\rangle_{\tilde{Q}E}, \{\{\tilde{P}_{a_k|x_k}^{(k)}\}_{a_k}\}_k \\ \text{compatible with } \langle I \rangle}} H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_{\rho_{\mathbf{RE}|\mathbf{0}}} \\ &= \eta_N^Q \\ &= H(\{p(a_k, a_l, \boldsymbol{\mu}|\mathbf{0})\}). \end{aligned} \quad (97)$$

□

C Proofs and details for DI randomness certification

Recall the bipartite Bell-inequalities I_θ which are defined as

$$\begin{aligned} \langle I_\theta^{(k,l)} \rangle &= \cos \theta \cos 2\theta \langle A_0^{(k)} A_0^{(l)} \rangle - \\ &\cos 2\theta (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) + \cos \theta \langle A_1^{(k)} A_1^{(l)} \rangle \end{aligned} \quad (98)$$

for $\theta \in \mathcal{G}$ where

$$\mathcal{G} = (\pi/4, \pi/2) \cup (\pi/2, 3\pi/4) \cup (5\pi/4, 3\pi/2) \cup (3\pi/2, 7\pi/4). \quad (99)$$

We now prove a few results about these Bell-inequalities that were stated in the main text.

C.1 Proof of Lemma 4

Lemma 4. Let $\boldsymbol{\mu}$ be a tuple of $N - 2$ measurement outcomes for all parties excluding k, l , and $n_\mu \in \{0, 1\}$ be the parity of $\boldsymbol{\mu}$. Let $\theta \in \mathcal{G}$ and $\{I_\mu^{(k,l)}\}_\mu$ be a set of bipartite Bell expressions between parties k, l , where

$$I_\mu^{(k,l)} = (-1)^{n_\mu} I_\theta^{(k,l)}. \quad (100)$$

Then the expanded Bell expression given by

$$I_\theta = \sum_{k=1}^{N-1} \left(\sum_{\boldsymbol{\mu}} \tilde{P}_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)} I_\mu^{(k,N)} \right) \quad (101)$$

has quantum bounds $\pm \eta_{N,\theta}^Q$ where $\eta_{N,\theta}^Q = 2(N - 1)\sin^3 \theta$. Moreover, $\langle I_\theta \rangle = \eta_{N,\theta}^Q$ is achieved up to relabelings by the strategy in Eq. (20), and cannot be achieved classically.

Proof. The above is an example of an expanded Bell expression, constructed according to Definition 2, where we used the bipartite expression in Lemma 3 as a seed, and chose $c_{k,l} = 1$ if $l = N$ and $k \in \{1, \dots, N - 1\}$, and 0 otherwise. First we notice that if the bound

$\eta_{N,\theta}^Q$ is achievable, then I_θ must define a nontrivial Bell inequality. This is because the maximum local value is upper bounded by $(N - 1)\eta_\theta^L < \eta_{N,\theta}^Q$, since $\eta_\theta^L < \eta_\theta^Q$. A similar argument holds for the minimum local value.

We now proceed to show $\langle I \rangle = \eta_{N,\theta}^Q$ is achieved by the strategy in Eq. (20). Notice that, for the ideal operators $P_{a_{k'}|0}^{(k')} = |+\rangle\langle+|$ ($|-\rangle\langle-|$) if $a_{k'} = 0$ ($a_{k'} = 1$),

$$P_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)} |\psi_{\text{GHZ}}\rangle = \sqrt{\langle \psi_{\text{GHZ}} | P_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)} | \psi_{\text{GHZ}} \rangle} |\phi_\mu\rangle \otimes |\Phi_\mu\rangle, \quad (102)$$

where $|\phi_\mu\rangle$ spans the support of $P_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)}$, and $|\Phi_\mu\rangle = (|00\rangle + i(-1)^{n_\mu}|11\rangle)/\sqrt{2}$. We then have

$$\begin{aligned} & \sum_{\boldsymbol{\mu}} \langle \psi_{\text{GHZ}} | P_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)} \otimes I_\mu^{(k,N)} | \psi_{\text{GHZ}} \rangle \\ &= \sum_{\boldsymbol{\mu}} \langle \psi_{\text{GHZ}} | P_{\boldsymbol{\mu}|\mathbf{0}}^{(k,N)} | \psi_{\text{GHZ}} \rangle \langle \Phi_\mu | I_\mu^{(k,N)} | \Phi_\mu \rangle \\ &= 2 \sin^3 \theta, \quad \forall k \in \{1, \dots, N - 1\}, \end{aligned} \quad (103)$$

where the final equality comes from the fact that when $n_\mu = 0$, we recover the state and measurements in Eq. (19) which yields the maximum quantum value of $I_\theta^{(k,N)}$, and when $n_\mu = 1$, we obtain a strategy equivalent to Eq. (19) up to a relabelling, which yields the maximum quantum value of $-I_\theta^{(k,N)}$.

Since the maximum quantum bound is achievable, a nontrivial Bell inequality follows. The minimum quantum value $-\eta_{N,\theta}^Q$ is simply achieved by relabelling the outcomes. □

C.2 Proof of Proposition 3

Proposition 3. For even N , let

$$\theta_N^* = \frac{2\pi t_N}{N + 1}, \quad (104)$$

where t_N is the $(N/2)^{\text{th}}$ element of the sequence 1, 1, 5, 7, 3, 3, 11, 13, 5, 5, ... given by

$$t_N = \begin{cases} N/4 + 1/2, & \text{if } N = 8n + 2, \\ N/4, & \text{if } N = 8n + 4, \\ 3N/4 + 1/2, & \text{if } N = 8n + 6, \\ 3N/4 + 1, & \text{if } N = 8n + 8, \quad n \in \mathbb{N}_0. \end{cases} \quad (105)$$

Then $\theta_N^* \in \mathcal{G}$.

Proof. The sequence θ_N^* splits naturally into four subsequences for $n \in \mathbb{N}_0$

$$\begin{aligned} \theta_{8n+2}^* &= \frac{2\pi(2n+1)}{8n+3} & \theta_{8n+4}^* &= \frac{2\pi(2n+1)}{8n+5} \\ \theta_{8n+6}^* &= \frac{2\pi(6n+5)}{8n+7} & \theta_{8n+8}^* &= \frac{2\pi(6n+7)}{8n+9}. \end{aligned} \quad (106)$$

First notice that the first terms in each of these subsequences are contained in one of the subintervals of $\mathcal{G}$

$$\begin{aligned} \theta_2^* &\in (\pi/2, 3\pi/4) & \theta_4^* &\in (\pi/4, \pi/2) \\ \theta_6^* &\in (5\pi/4, 3\pi/2) & \theta_8^* &\in (3\pi/2, 7\pi/4). \end{aligned} \quad (107)$$

Looking at the ratio of subsequent terms in each subsequence

$$\begin{aligned} \frac{\theta_{8n+2}^*}{\theta_{8n+10}^*} &= \frac{16n^2 + 30n + 11}{16n^2 + 30n + 9} & \frac{\theta_{8n+4}^*}{\theta_{8n+12}^*} &= \frac{16n^2 + 34n + 13}{16n^2 + 34n + 15} \\ \frac{\theta_{8n+6}^*}{\theta_{8n+14}^*} &= \frac{48n^2 + 130n + 75}{48n^2 + 130n + 77} & \frac{\theta_{8n+8}^*}{\theta_{8n+16}^*} &= \frac{48n^2 + 158n + 119}{48n^2 + 158n + 117} \end{aligned}$$

we see that the subsequences θ_{8n+2}^* and θ_{8n+8}^* are strictly monotonically decreasing sequences whereas θ_{8n+4}^* and θ_{8n+6}^* are strictly monotonically increasing sequences. Furthermore, their limits are

$$\begin{aligned} \lim_{n \rightarrow \infty} \theta_{8n+2}^* &= \lim_{n \rightarrow \infty} \theta_{8n+4}^* = \pi/2 \\ \lim_{n \rightarrow \infty} \theta_{8n+6}^* &= \lim_{n \rightarrow \infty} \theta_{8n+8}^* = 3\pi/2. \end{aligned} \quad (108)$$

The limits correspond to the boundaries of the respective subinterval of $\mathcal{G}$ that contain the first terms of the subsequences. As they converge strictly monotonically to these limits, every term in the subsequences is contained within their respective subintervals and hence $\theta_N^* \in \mathcal{G}$ for all even N . $\square$

C.3 Proof of Propositions 4 and 5

Proposition 4. *Let N be an even integer. For every MABK value s in the range $(1, m_N^*]$, there exists a $\theta_s \in \mathcal{G}$ that satisfies $s = \langle M_N(\theta_s) \rangle$.*

Proof. We examine each subsequence independently.

Case 1. For $N = 8n+2$, recall that $\theta_N^* \in (\pi/2, 3\pi/4)$. When $n = 0$, the claim can be seen from Fig. 1, which indicates MABK values below 1 for $\theta \in (\pi/2, 3\pi/4)$, and the fact that $M_N(\theta)$ is continuous. We hence consider $n \geq 1$ in the following. First note

$$\langle M_{8n+2}(\pi/2) \rangle = 2^{4n}. \quad (109)$$

Consider the point $\tilde{\theta}_n = \pi/2 - \pi/(4n+1) > \pi/4$. We have

$$\begin{aligned} \langle M_{8n+2}(\tilde{\theta}_n) \rangle &= -2^{4n} \left(\sin^{8n+2} \left(\frac{\pi}{8n+2} \right) \right. \\ &\quad \left. + \cos^{8n+2} \left(\frac{\pi}{8n+2} \right) \right) < 0. \end{aligned} \quad (110)$$

Therefore, by continuity of the function $\langle M_N(\theta) \rangle$ in θ , the range of achievable MABK values corresponding to $\theta \in (\tilde{\theta}_n, \pi/2) \subset \mathcal{G}$ contains the sub-interval $(1, 2^{4n})$. By a similar argument, the sub-interval $(2^{4n}, m_{8n+2}^*)$ must also be achievable for $\theta \in (\pi/2, \theta_{8n+2}^*) \subset \mathcal{G}$. For the value 2^{4n} , notice

$$\langle M_{8n+2}(3\pi/4) \rangle = 2^{(8n+1)/2} (-1)^n \sin^{8n+2}(3\pi/8), \quad (111)$$

and $\sin^{8n+2}(3\pi/8) < 1/\sqrt{2}$. Hence there must exist a $\theta \in (\theta_{8n+2}^*, 3\pi/4) \subset \mathcal{G}$ such that $\langle M_{8n+2}(\theta) \rangle = 2^{4n}$.

Case 2. For $N = 8n+4$, recall $\theta_N^* \in (\pi/4, \pi/2)$. As before, we only need to consider $n \geq 1$ by analyzing Fig. 1. Let $\tilde{\theta}_n = \pi/2 + \pi/(4n+2) < 3\pi/4$. Then

$$\begin{aligned} \langle M_{8n+4}(\tilde{\theta}_n) \rangle &= 2^{4n+1} \left(\sin^{8n+4} \left(\frac{\pi}{8n+4} \right) \right. \\ &\quad \left. - \cos^{8n+4} \left(\frac{\pi}{8n+4} \right) \right) < 0. \end{aligned} \quad (112)$$

We also note

$$\langle M_{8n+4}(\pi/2) \rangle = 2^{4n+1}, \quad (113)$$

and

$$\begin{aligned} \langle M_{8n+4}(\pi/4) \rangle &= 2^{4n+1} (-1)^n \left(\cos^{8n+4}(3\pi/8) \right. \\ &\quad \left. + \sin^{8n+4}(3\pi/8) \right) < 2^{4n+1}. \end{aligned} \quad (114)$$

We can therefore apply the same arguments as in the previous case.

Case 3. For $N = 8n+6$, recall $\theta_N^* \in (5\pi/4, 3\pi/2)$. We observe that the claim holds for $n = 0$ by examining Fig. 1, and consider $n \geq 1$. Let $\tilde{\theta}_n = 3\pi/2 + \pi/(4n+3)$, and note

$$\begin{aligned} \langle M_{8n+6}(\tilde{\theta}_n) \rangle &= -2^{4n+2} \left(\sin^{8n+6} \left(\frac{\pi}{8n+6} \right) \right. \\ &\quad \left. + \cos^{8n+6} \left(\frac{\pi}{8n+6} \right) \right) < 0. \end{aligned} \quad (115)$$

By checking

$$\langle M_{8n+6}(3\pi/2) \rangle = 2^{4n+2}, \quad (116)$$

and

$$\begin{aligned} \langle M_{8n+6}(5\pi/4) \rangle &= -2^{(8n+5)/2} (-1)^n \sin^{8n+6}(\pi/8) < 2^{4n+2}, \end{aligned} \quad (117)$$

we can apply the same arguments as case 1.

Case 4. Finally, for $N = 8n+8$, we have $\theta_N^* \in (3\pi/2, 7\pi/4)$. The case $n = 0$ can be seen by examining Fig. 1, and for $n \geq 1$ we define $\tilde{\theta}_n = 3\pi/2 - \pi/(4n+4) > 5\pi/4$. Then

$$\begin{aligned} \langle M_{8n+8}(\tilde{\theta}_n) \rangle &= 2^{4n+3} \left(\sin^{8n+8} \left(\frac{\pi}{8n+8} \right) \right. \\ &\quad \left. - \cos^{8n+8} \left(\frac{\pi}{8n+8} \right) \right) < 0. \end{aligned} \quad (118)$$

We also have

$$\langle M_{8n+8}(3\pi/2) \rangle = 2^{4n+3}, \quad (119)$$

and

$$\begin{aligned} \langle M_{8n+8}(7\pi/4) \rangle &= 2^{4n+3} (-1)^{n+1} \left(\cos^{8n+8}(\pi/8) \right. \\ &\quad \left. - \sin^{8n+8}(\pi/8) \right) < 2^{4n+3}. \end{aligned} \quad (120)$$

By the arguments in case 1, this completes the proof. $\square$

Proposition 5. Let N be an odd integer. For every MABK value s in the range $(1, 2^{(N-1)/2})$, there exists a $\theta_s \in \mathcal{G}$ that satisfies $s = \langle M_N(\theta_s) \rangle$.

Proof. As was done for the even case, we consider four subsequences.

Case 1. For $N = 8n + 3$, we note $\langle M_{8n+3}(\pi/2) \rangle = 2^{4n+1}$. When $n = 0$, the claim can be verified from Fig. 2, hence we consider $n \geq 1$. Consider the point $\tilde{\theta}_n = \pi/2 + \pi/(4n+2) \in (\pi/2, 3\pi/4) \subset \mathcal{G}$. Then

$$\langle M_N(\tilde{\theta}_n) \rangle = -2^{4n+1} \left(\sin^{8n+3} \left(\frac{\pi}{8n+4} \right) \sin(\pi f(n)) + \cos^{8n+3} \left(\frac{\pi}{8n+4} \right) \cos(\pi f(n)) \right), \quad (121)$$

where $f(n) = \frac{16n^2+24n+7}{8n+4}$. Notice that

$$\begin{aligned} & \sin(\pi f(n)) + \sin\left(\frac{\pi}{8n+4}\right) \\ &= 2 \sin(\pi(n+1)) \cos\left(\frac{\pi}{2} \frac{8n^2+12n+3}{4n+2}\right) = 0. \end{aligned} \quad (122)$$

Similarly

$$\begin{aligned} & \cos(\pi f(n)) - \cos\left(\frac{\pi}{8n+4}\right) \\ &= -2 \sin(\pi(n+1)) \sin\left(\frac{\pi}{2} \frac{8n^2+12n+3}{4n+2}\right) = 0. \end{aligned} \quad (123)$$

Hence

$$\langle M_N(\tilde{\theta}_n) \rangle = 2^{4n+1} \left(\sin^{8n+4} \left(\frac{\pi}{8n+4} \right) - \cos^{8n+4} \left(\frac{\pi}{8n+4} \right) \right) < 0. \quad (124)$$

Therefore, the achievable MABK values from the interval $(\pi/2, \tilde{\theta}_n) \subset \mathcal{G}$ include the sub-interval $(2^{4n+1}, 1)$ by continuity of the function $\langle M_N(\theta) \rangle$ in θ .

Case 2. For $N = 8n + 7$, note $\langle M_{8n+7}(3\pi/2) \rangle = 2^{4n+3}$. The claim for $n = 0$ can be verified using Fig. 2, so we consider $n \geq 1$. Let $\tilde{\theta}_n = 3\pi/2 + \pi/(4n+4) < 7\pi/4$, and note

$$\langle M_N(\tilde{\theta}_n) \rangle = 2^{4n+3} \left(\sin^{8n+7} \left(\frac{\pi}{8n+8} \right) \cos(\pi f(n)) - \cos^{8n+7} \left(\frac{\pi}{8n+8} \right) \sin(\pi f(n)) \right), \quad (125)$$

where $f(n) = \frac{48n^2+100n+51}{8n+8}$. Following the same steps as before, we have

$$\sin(\pi f(n)) - \cos\left(\frac{\pi}{8n+8}\right) = 0, \quad (126)$$

and

$$\cos(\pi f(n)) - \sin\left(\frac{\pi}{8n+8}\right) = 0, \quad (127)$$

implying

$$\langle M_N(\tilde{\theta}_n) \rangle = 2^{4n+3} \left(\sin^{8n+8} \left(\frac{\pi}{8n+8} \right) - \cos^{8n+8} \left(\frac{\pi}{8n+8} \right) \right) < 0. \quad (128)$$

The same arguments used in case 1 prove the claim.

Case 3. For $N = 8n + 5$, we consider the MABK expression obtained by relabelling the inputs of every party, followed by the output of every party's first measurement, i.e., $\tilde{A}_0^{(k)} \mapsto \tilde{A}_1^{(k)}$ and $\tilde{A}_1^{(k)} \mapsto -\tilde{A}_0^{(k)}$. The resulting MABK value of the strategy in Eq. (20) is given by¹⁰

$$\langle \tilde{M}_N(\theta) \rangle = 2^{\frac{N-1}{2}} \left(\cos^N(\theta/2 + \pi/4) \sin(-N\theta/2 + \pi/4) - \cos^N(\theta/2 - \pi/4) \sin(N\theta/2 + \pi/4) \right). \quad (129)$$

We then have $\langle \tilde{M}_{8n+5}(\pi/2) \rangle = 2^{4n+2}$. We can verify the claim for $n = 0$ from Fig. 2, hence we consider $n \geq 1$. Let $\tilde{\theta} = \pi/2 + \pi/(4n+3) < 3\pi/4$, and note

$$\langle \tilde{M}_N(\tilde{\theta}_n) \rangle = -2^{4n+2} \left(\sin^{8n+5} \left(\frac{\pi}{8n+6} \right) \cos(\pi f(n)) + 8 \cos^{8n+5} \left(\frac{\pi}{8n+6} \right) \sin(\pi f(n)) \right), \quad (130)$$

where $f(n) = \frac{8n^2+16n+7}{4n+3}$. We have

$$\sin\left(\frac{\pi}{8n+6}\right) - \cos(\pi f(n)) = 0, \quad (131)$$

and

$$\cos\left(\frac{\pi}{8n+6}\right) - \sin(\pi f(n)) = 0, \quad (132)$$

which implies

$$\langle \tilde{M}_N(\tilde{\theta}_n) \rangle = -2^{4n+2} \left(\sin^{8n+6} \left(\frac{\pi}{8n+6} \right) + \cos^{8n+6} \left(\frac{\pi}{8n+6} \right) \right) < 0. \quad (133)$$

The claim then follows from the same arguments as the previous cases.

Case 4. For $N = 8n + 9$, we also consider the relabelled MABK expression in Eq. (129), and note $\langle \tilde{M}_{8n+5}(3\pi/2) \rangle = 2^{4n+4}$. Let $\tilde{\theta}_n = 3\pi/2 + \pi/(4n+5) < 7\pi/4$, and consider

$$\langle \tilde{M}_N(\tilde{\theta}_n) \rangle = 2^{4n+4} \left(\sin^{8n+9} \left(\frac{\pi}{8n+10} \right) \sin(\pi f(n)) - \cos^{8n+9} \left(\frac{\pi}{8n+10} \right) \cos(\pi f(n)) \right), \quad (134)$$

¹⁰Rather than relabelling the MABK expression, we could equivalently consider the strategy in Eq. (31), which achieves the MABK value in Eq. (32) of the original expression (Eq. (1)), with the substitution $\phi \mapsto \theta$, $\theta \mapsto 0$. This also results in Eq. (129).

where $f(n) = \frac{48n^2+124n+79}{8n+10}$. We have

$$\cos\left(\frac{\pi}{8n+10}\right) - \cos(\pi f(n)) = 0, \quad (135)$$

and

$$\sin\left(\frac{\pi}{8n+10}\right) + \sin(\pi f(n)) = 0, \quad (136)$$

implying

$$\begin{aligned} \langle \tilde{M}_N(\tilde{\theta}_n) \rangle &= -2^{4n+4} \left(\sin^{8n+10}\left(\frac{\pi}{8n+10}\right) \right. \\ &\quad \left. + \cos^{8n+10}\left(\frac{\pi}{8n+10}\right) \right) < 0. \end{aligned} \quad (137)$$

Following the same steps as before establishes the claim. $\square$

C.4 Evidence for Conjecture 1

Conjecture 1. *The maximum MABK value achievable by quantum strategies that generate maximum randomness, i.e., strategies with $p(\mathbf{a}|\mathbf{0}) = 1/2^N$, $\forall \mathbf{a}$, is m_N^* .*

Conjecture 1 is that the analytical lower bound, m_N^* , on the maximum MABK value achievable whilst generating maximum randomness, is optimal, in the sense there no other set of quantum correlations exists with a higher MABK value while achieving maximum randomness. To justify this claim, we developed a numerical technique to upper bound the maximum MABK value of quantum correlations achieving maximum randomness, and checked this against the achievable lower bounds m_N^* . We found close agreement between our numerical calculations and the analytic lower bounds, modulo a small discrepancy which pushed the numerical values slightly below the analytical ones. Since the SDP being solved is a minimization, we judge this to be caused by numerical precision, and an accumulation of errors pushing the global optimum below its true value. A comparison between our analytical lower bound and the numerics can be found in Table 2, and details of the numerical method can be found in Section C.10.1. We also remark that this conjecture was proven analytically in the $N = 2$ case in Ref. [23].

C.5 Proof of Proposition 6

Proposition 6. *In the limit of large even N , one can achieve arbitrarily close to the maximum quantum violation of the N party MABK inequality, $2^{(N-1)/2}$, whilst certifying maximum device-independent randomness.*

Proof. We consider the asymptotic behaviour of m_N^* , which gives an achievable lower bound on the maximum MABK value compatible with maximum randomness. Our aim is to show that as N becomes very

large, whilst remaining an even integer, the MABK violation tends to the maximum quantum violation $2^{(N-1)/2}$, i.e.,

$$\lim_{\substack{N \rightarrow \infty \\ N \text{ even}}} \frac{m_N^*}{2^{\frac{N-1}{2}}} = 1. \quad (138)$$

This can be established directly from the asymptotic behaviour of θ_N^* given in Eq. (108), as follows.

We first insert the expression for $m_N^* = \langle M_N(\theta_N^*) \rangle$ in Eq. (24) into the left hand side of (138) to give

$$\begin{aligned} \lim_{N \rightarrow \infty} \left(\cos^N(\theta_N^*/2 + \pi/4) \sin(N\theta_N^*/2 + \pi/4) \right. \\ \left. + \cos^N(\theta_N^*/2 - \pi/4) \sin(N\theta_N^*/2 - \pi/4) \right). \end{aligned} \quad (139)$$

We then consider two cases:

Case 1: $N = 8n + 2$ or $N = 8n + 4$. For these values of N , we know $\lim_{N \rightarrow \infty} \theta_N^* = \pi/2$. One can then verify

$$\begin{aligned} \lim_{N \rightarrow \infty} \cos^N(\theta_N^*/2 + \pi/4) &= 0, \\ \lim_{N \rightarrow \infty} \cos^N(\theta_N^*/2 - \pi/4) &= 1. \end{aligned} \quad (140)$$

Now we split into two sub-cases to evaluate $N\theta_N^*$ in the limit (139).

1a: $N = 8n + 2$. Here we find

$$\frac{N\theta_N^*}{2} = \frac{\pi(8n+2)(2n+1)}{8n+3}, \quad (141)$$

which has the oblique asymptote $2\pi n + 3\pi/4$. We hence find

$$\lim_{N \rightarrow \infty} \sin(N\theta_N^*/2 - \pi/4) = \sin(\pi/2) = 1. \quad (142)$$

1b: $N = 8n + 4$. Here we find

$$\frac{N\theta_N^*}{2} = \frac{4\pi(2n+1)^2}{8n+5}, \quad (143)$$

which also has the oblique asymptote $2\pi n + 3\pi/4$, implying Eq. (142).

Case 2: $N = 8n + 6$ or $N = 8n + 8$. For these values of N , we know $\lim_{N \rightarrow \infty} \theta_N^* = 3\pi/2$. One can then verify

$$\begin{aligned} \lim_{N \rightarrow \infty} \cos^N(\theta_N^*/2 + \pi/4) &= 1, \\ \lim_{N \rightarrow \infty} \cos^N(\theta_N^*/2 - \pi/4) &= 0. \end{aligned} \quad (144)$$

Splitting into two sub-cases:

2a, $N = 8n + 6$: Here

$$\frac{N\theta_N^*}{2} = \frac{\pi(8n+6)(6n+5)}{8n+7}, \quad (145)$$

which has the oblique asymptote $6\pi n + 17\pi/4$. Inserting into the limit we find

$$\lim_{N \rightarrow \infty} \sin(N\theta_N^*/2 + \pi/4) = \sin(9\pi/2) = 1. \quad (146)$$

N	Analytical lower bound, m_N^*	Numerical upper bound
2	$3\sqrt{3}/4 \approx 1.299038105676658$	1.299038105676658
4	$(5/8)\sqrt{(5/2)(5+\sqrt{5})} \approx 2.658283776100125$	2.658283776100125
6	$(7/8)(\cos(\pi/14) + 3\sqrt{2}\cos(3\pi/28) - 3\cos(\pi/7) + 5\cos(3\pi/14)) \approx 5.412519474140316$	5.412519474140316
8	$45/16 + 33\sqrt{3}/4 + (45\sqrt{2}/16)(3 + \sqrt{3})\cos(\pi/36) - (1/16)(180\sqrt{3} + 90)\cos^2(\pi/36) \approx 10.93208547970192$	10.93208547970192
10	$(11/32)(\cos(\pi/22) + 30\sqrt{2}\cos(3\pi/44) - 5\cos(\pi/11) + 15\cos(3\pi/22) + 5\sqrt{2}\cos(7\pi/44) - 30\cos(2\pi/11) + 42\cos(5\pi/22)) \approx 22.00126183851497$	22.00126183851497
12	$(429\sqrt{2}/16)\cos(\pi/52) + (39\sqrt{2}/32)\cos(9\pi/52) + (715\sqrt{2}/64)\cos(5\pi/52) + (13/64)\cos(\pi/26) + (1287/64)\cos(5\pi/26) - (715/64)\cos(2\pi/13) + (143/32)\cos(3\pi/26) - (39/32)\cos(\pi/13) - (429/16)\cos(3\pi/13) \approx 44.19316043036512$	44.19316043036512

Table 2: Comparison between upper and lower bounds on the maximum MABK value which can be achieved by quantum correlations certifying maximum randomness, for different numbers of parties N , rounded to 16 digits. The analytical lower bound is provided by the family of strategies Eq. (20) in the main text, and the numerical upper bound is calculated with an SOS approach using the SDPA-GMP solver. The numerical upper bounds agree exactly with the analytical lower bounds up to 16 digits of precision, supporting the conjecture that our analytical lower bound is tight.

2b, $N = 8n + 8$: For the final case

$$\frac{N\theta_N^*}{2} = \frac{\pi(8n+8)(6n+7)}{8n+9}, \quad (147)$$

which has the oblique asymptote $6\pi n + 25\pi/4$. Inserting into the limit we find

$$\lim_{N \rightarrow \infty} \sin(N\theta_N^*/2 + \pi/4) = \sin(13\pi/2) = 1. \quad (148)$$

This shows that the MABK violation m_N^* tends to the maximum quantum value in the regime of large N . Note however, in the limit θ tends to $\pi/2$ or $3\pi/2$; these values are not self-testable, since θ does not lie in a valid interval for there to be a gap between the local and quantum bound for the two party Bell inequality. However, the points either side of both angles are self-testable. We hence obtain the result that one can achieve arbitrarily close to maximum MABK violation whilst certifying maximum DI randomness, for an arbitrarily large finite N . $\square$

C.6 Local dilution measure of nonlocality

In this section we briefly outline the definition of local dilution (see also [59, Section I.B.1] for a description of this technique) which was plotted in Fig. 3. Given a distribution $\{p(\mathbf{a}|\mathbf{x})\}$, we can represent it as a vector

in a high dimensional space, $p \in \mathbb{R}^{2^{2N}}$. We denote the set of 2^{2N} local deterministic distributions as $\{d_\lambda\}_\lambda$, and the local polytope can be constructed by taking its convex hull,

$$\mathcal{L} = \left\{ \sum_\lambda q_\lambda d_\lambda \mid \sum_\lambda q_\lambda = 1, q_\lambda \geq 0 \right\}. \quad (149)$$

To check whether p admits a local description, we can search for a set $\{q_\lambda\}_\lambda$, where $\sum_\lambda q_\lambda = 1$, $q_\lambda \geq 0$, using a standard linear program.

To quantify the distance of a given p from $\mathcal{L}$, we define the diluted sets, for $\epsilon \geq 0$,

$$\mathcal{L}_\epsilon = \left\{ \sum_\lambda q_\lambda d_\lambda \mid \sum_\lambda q_\lambda = 1, q_\lambda \geq -\epsilon \right\}. \quad (150)$$

As $\epsilon \rightarrow \infty$, $\mathcal{L}_\epsilon$ tends to the affine hull of the points $\{d_\lambda\}_\lambda$, and we have the inclusion: $\mathcal{L} \subset \mathcal{L}_{\epsilon_1} \subset \mathcal{L}_{\epsilon_2} \subset \mathcal{L}_\infty$, where $0 < \epsilon_1 < \epsilon_2$. We then define the following measure, which we call the local dilution:

$$\mathcal{D}[p] := \min\{\epsilon \mid p \in \mathcal{L}_\epsilon\}. \quad (151)$$

$\mathcal{D}[p]$ can be interpreted as the smallest dilution of the local polytope that contains p . If $\mathcal{D}[p_1] < \mathcal{D}[p_2]$, one has to enlarge the local set more to include p_2 than to include p_1 . We can hence understand p_2 to be more nonlocal than p_1 .

C.7 Proof of Lemma 7

In the main text, we detailed a two parameter family of bipartite Bell expressions which are used to build our N -partite expressions. These are generalizations of the two families introduced in [23], and can be recovered as sub-families of the self-tests studied in [53, 51, 52].

Lemma 7. *Let $(\phi, \theta) \in \mathbb{R}^2$ such that $\cos(2\theta)\cos(2\phi) < 0$ and $\cos(\theta - \phi) \neq 0$. Define the family of Bell expressions parameterized by ϕ and θ ,*

$$\begin{aligned} \langle J_{\phi, \theta}^{(k, l)} \rangle &= \cos 2\theta \cos(\theta - \phi) \langle A_0^{(k)} A_0^{(l)} \rangle \\ &\quad - \cos 2\theta \cos 2\phi (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) \\ &\quad + \cos 2\phi \cos(\theta - \phi) \langle A_1^{(k)} A_1^{(l)} \rangle. \end{aligned} \quad (152)$$

Then we have the following:

(i) *The local bounds are given by $\pm \eta_{\phi, \theta}^L$, where*

$$\begin{aligned} \eta_{\phi, \theta}^L &= \max \{ |\cos(\theta - \phi)(\cos 2\theta - \cos 2\phi)|, \\ &\quad |\cos(\theta - \phi)(\cos 2\theta + \cos 2\phi)| + |2 \cos 2\phi \cos 2\theta| \}. \end{aligned} \quad (153)$$

(ii) *The quantum bounds are given by $\pm \eta_{\phi, \theta}^Q$, where*

$$\eta_{\phi, \theta}^Q = 2 \sin^2(\theta + \phi) \sin(\theta - \phi).$$

(iii) *$|\eta_{\phi, \theta}^Q| > \eta_{\phi, \theta}^L$.*

(iv) *Up to local isometries, there exists a unique strategy that achieves $\langle J_{\phi, \theta}^{(k, l)} \rangle = \eta_{\phi, \theta}^Q$:*

$$\begin{aligned} \rho_{Q_k Q_l} &= |\psi\rangle\langle\psi|, \text{ where } |\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + i|11\rangle), \\ A_0^{(k)} &= A_0^{(l)} = \cos \phi \sigma_X - \sin \phi \sigma_Y, \\ A_1^{(k)} &= A_1^{(l)} = \cos \theta \sigma_X + \sin \theta \sigma_Y. \end{aligned} \quad (154)$$

Proof. The above is a direct corollary of [51, Proposition 1]. Specifically, let $\beta = -2\phi$, $\gamma = \theta - \phi$ and $\alpha = \theta + \phi$. Then the Bell expression $B_{\alpha, \beta, \gamma}$ of [51, Proposition 1] is equal to the expression $J_{\phi, \theta}$ up to a factor of $\cos(\theta - \phi)$. The condition for $\eta_{\phi, \theta}^L < |\eta_{\phi, \theta}^Q|$ is given by [51, Equation 4], which reads

$$\begin{aligned} \cos(\alpha + \beta) \cos \beta \cos(\alpha + \gamma) \cos \gamma &= \\ \cos^2(\theta - \phi) \cos 2\theta \cos 2\phi &< 0, \end{aligned} \quad (155)$$

and is satisfied if and only if $\cos(2\theta)\cos(2\phi) < 0$ and $\cos(\theta - \phi) \neq 0$. $\square$

Lemma 3 can be proven as a special case of the above, by setting $\phi = 0$. This follows from $\mathcal{G} \subset \mathcal{F}$, where $\mathcal{F}$ is a set of valid self-testing points defined below Lemma 7.

One can also recover the I_δ and J_γ family of Bell expressions from [23], as special cases of the $J_{\phi, \theta}$ expressions. By choosing $\phi = 0$, and $\theta = \delta + \pi/2$, $\delta \in (0, \pi/6]$, we obtain

$$\begin{aligned} \langle J_{0, \delta + \pi/2} \rangle &= \sin(\delta) \cos(2\delta) \langle A_0^{(k)} A_0^{(l)} \rangle \\ &\quad + \cos(2\delta) (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) \\ &\quad - \sin(\delta) \langle A_1^{(k)} A_1^{(l)} \rangle \leq 2 \cos^3(\delta), \end{aligned} \quad (156)$$

which is identical to the expression found in [23]. One can obtain the J_γ family by choosing $\phi = -3\gamma/2$ and $\theta = \phi/3 + 2\pi/3$ using Eq. (37), and multiplying by $4 \cos^2(\gamma + \pi/6) / \sin^2(2\gamma + \pi/3)$:

$$\begin{aligned} \frac{4 \cos^2(\gamma + \pi/6)}{\sin^2(2\gamma + \pi/3)} \langle J_{-3\gamma/2, -\gamma/2 + 2\pi/3} \rangle &= \langle A_0^{(k)} A_0^{(l)} \rangle \\ &\quad + \left(4 \cos^2(\gamma + \pi/6) - 1 \right) (\langle A_0^{(k)} A_1^{(l)} \rangle + \langle A_1^{(k)} A_0^{(l)} \rangle) \\ &\quad - \langle A_1^{(k)} A_1^{(l)} \rangle \leq 8 \cos^3(\gamma + \pi/6). \end{aligned} \quad (157)$$

C.8 Proof of Lemma 8

Lemma 8. *Let μ be a tuple of $N - 2$ measurement outcomes for all parties excluding k, l , and $n_\mu \in \{0, 1\}$ be the parity of μ . Let $(\phi, \theta) \in \mathbb{R}^2$,*

$$\phi' := \frac{\phi N}{2}, \quad \theta' := \theta - \frac{N-2}{2} \phi, \quad (158)$$

and

$$I_\mu^{(k, l)} = (-1)^{n_\mu} J_{\phi', \theta'}^{(k, l)}. \quad (159)$$

Define the following Bell polynomial

$$I_{\phi', \theta'} := \sum_{k=1}^{N-1} \left(\sum_{\mu} \tilde{P}_{\mu|0}^{(k, N)} I_\mu^{(k, N)} \right). \quad (160)$$

If $(\phi', \theta') \in \mathcal{F}$, $I_{\phi', \theta'}$ is an expanded Bell expression, and has quantum bounds $\pm \eta_{N, \phi', \theta'}^Q$ where $\eta_{N, \phi', \theta'}^Q = 2(N-1) \sin^2(\theta' + \phi') \sin(\theta' - \phi')$, which can be achieved up to relabellings by the strategy

$$\begin{aligned} \rho_Q &= |\psi_{\text{GHZ}}\rangle\langle\psi_{\text{GHZ}}|, \\ A_0^{(k)} &= \cos \phi \sigma_X - \sin \phi \sigma_Y, \\ A_1^{(k)} &= \cos \theta \sigma_X + \sin \theta \sigma_Y, \quad k \in \{1, \dots, N\}. \end{aligned} \quad (161)$$

In addition, this quantum bound cannot be achieved classically.

Proof. Suppose $(\phi', \theta') \in \mathcal{F}$, and consider the ideal projection operators $P_{a_{k'}}^{(k')} = |\varphi_{a_{k'}}\rangle\langle\varphi_{a_{k'}}|$ derived from the observables in Eq. (31), i.e., $\cos \phi \sigma_X - \sin \phi \sigma_Y = |\varphi_0\rangle\langle\varphi_0| - |\varphi_1\rangle\langle\varphi_1|$ where

$$|\varphi_{a_{k'}}\rangle = \frac{|0\rangle + (-1)^{a_{k'}} e^{-i\phi} |1\rangle}{\sqrt{2}}. \quad (162)$$

We then have

$$P_{\mu|0}^{(k,N)}|\psi_{\text{GHZ}}\rangle = \sqrt{\langle\psi_{\text{GHZ}}|P_{\mu|0}^{(k,N)}|\psi_{\text{GHZ}}\rangle}|\varphi_{\mu}\rangle \otimes |\Phi_{\mu,\phi}\rangle_{Q_k Q_N}, \quad (163)$$

where $|\varphi_{\mu}\rangle$ spans the support of $P_{\mu|0}^{(k,N)}$ and

$$|\Phi_{\mu,\phi}\rangle = \frac{|00\rangle + i(-1)^{n_{\mu}}e^{i(N-2)\phi}|11\rangle}{\sqrt{2}}. \quad (164)$$

Notice $|\Phi_{\mu,\phi}\rangle = (U_{\phi} \otimes U_{\phi})|\Phi_{\mu}\rangle$ where $U_{\phi} = |0\rangle\langle 0| + e^{i(N-2)\phi/2}|1\rangle\langle 1|$. We now calculate

$$\begin{aligned} \langle\Phi_{\mu,\phi}|I_{\mu}^{(k,l)}|\Phi_{\mu,\phi}\rangle \\ = (-1)^{n_{\mu}}\langle\Phi_{\mu}|(U_{\phi} \otimes U_{\phi})^{\dagger}J_{\phi',\theta'}^{(k,N)}(U_{\phi} \otimes U_{\phi})|\Phi_{\mu}\rangle. \end{aligned} \quad (165)$$

We can now calculate how $I_{\mu}^{(k,N)}$ transforms under $(U_{\phi} \otimes U_{\phi})^{\dagger}$. Let $O = \cos(t)\sigma_X + \sin(t)\sigma_Y$ for $t \in \mathbb{R}$ be an arbitrary observable. Then

$$U_{\phi}^{\dagger}OU_{\phi} = \cos\left(-\frac{N-2}{2}\phi+t\right)\sigma_X + \sin\left(-\frac{N-2}{2}\phi+t\right)\sigma_Y. \quad (166)$$

Substituting $t = -\phi$ and $t = \theta$ into the above, we find

$$\begin{aligned} U_{\phi}^{\dagger}A_0^{(k)}U_{\phi} &= \cos\phi'\sigma_X - \sin\phi'\sigma_Y =: A_0'^{(k)} \\ U_{\phi}^{\dagger}A_1^{(k)}U_{\phi} &= \cos\theta'\sigma_X + \sin\theta'\sigma_Y =: A_1'^{(k)}, \end{aligned} \quad (167)$$

and similarly for party N , where we used the definitions of $A_x^{(k)}$ in Eq. (154). We therefore find

$$\begin{aligned} J' &:= (U_{\phi} \otimes U_{\phi})^{\dagger}J_{\phi',\theta'}^{(k,N)}(U_{\phi} \otimes U_{\phi}) \\ &= \cos 2\theta' \cos(\theta' - \phi')A_0'^{(k)} \otimes A_0'^{(N)} \\ &\quad - \cos 2\theta' \cos 2\phi' (A_0'^{(k)} \otimes A_1'^{(N)} + A_1'^{(k)} \otimes A_0'^{(N)}) \\ &\quad + \cos 2\phi' \cos(\theta' - \phi')A_1'^{(k)} \otimes A_1'^{(N)}, \end{aligned} \quad (168)$$

The largest and smallest eigenvalues of J' are $\pm\eta_{\theta',\phi'}^Q$, and are associated with the eigenvectors $(|00\rangle \pm i|11\rangle)/\sqrt{2}$. We therefore have

$$\langle\Phi_{\mu,\phi}|I_{\mu}^{(k,N)}|\Phi_{\mu,\phi}\rangle = (-1)^{n_{\mu}}\langle\Phi_{\mu}|J'|\Phi_{\mu}\rangle = \eta_{\phi',\theta'}^Q. \quad (169)$$

Putting everything together, we obtain

$$\begin{aligned} \sum_{\mu} \langle\psi_{\text{GHZ}}|P_{\mu|0}^{(k,N)} \otimes I_{\mu}^{(k,N)}|\psi_{\text{GHZ}}\rangle \\ = \sum_{\mu} \langle\psi_{\text{GHZ}}|P_{\mu|0}^{(k,N)}|\psi_{\text{GHZ}}\rangle \langle\Phi_{\mu,\phi}|I_{\mu}^{(k,N)}|\Phi_{\mu,\phi}\rangle \\ = \eta_{\phi',\theta'}^Q, \quad \forall k \in \{1, \dots, N-1\}. \end{aligned} \quad (170)$$

The quantum bound $\eta_{N,\phi',\theta'}^Q$ then follows by summing over k , which shows achievability. The minimum quantum value $-\eta_{N,\phi',\theta'}^Q$ is achieved by relabelling the outcomes. Moreover, the quantum bound cannot be achieved classically since $I_{\phi',\theta'}$ corresponds to an expanded Bell inequality when $(\phi', \theta') \in \mathcal{F}$ (cf. the discussion below Eq. (49)). $\square$

C.9 Proof of Proposition 7

Proposition 7. Let $\phi \in [0, \phi_N^*]$, and $\theta = \theta(\phi)$ as defined in Eqs. (37) and (38). Then $(\phi', \theta') \in \mathcal{F}$, where ϕ' and θ' are defined in Lemma 8.

Proof. For convenience, we restate the set $\mathcal{F}$ below,

$$\mathcal{F} = \left[(-\pi/4, \pi/4) \times \mathcal{G}\right] \cup \left[(-\pi/4, \pi/4) \setminus \{0\} \times \{\pi/2, 3\pi/2\}\right]. \quad (171)$$

First we will consider the range of ϕ' as we vary ϕ . By definition, $\phi' = \phi N/2$, which takes values over the interval $[0, \phi_N^* N/2]$. Then

$$\frac{\phi_N^* N}{2} = \text{sgn}[\sin(2\theta_N^*)] \frac{\pi}{8}, \quad (172)$$

which implies $-\pi/4 < \phi' < \pi/4$.

Next we consider θ' as a function of ϕ ,

$$\theta' = \theta(\phi) - \frac{N-2}{2}\phi = -\frac{N(N-3)}{2(N+1)}\phi + \theta_N^*. \quad (173)$$

When $\phi = 0$, $\theta' = \theta_N^*$, and when $\phi = \phi_N^*$,

$$\theta' = -\text{sgn}[\sin(2\theta_N^*)] \frac{\pi}{8} \frac{N-3}{N+1} + \theta_N^* =: \tilde{\theta}'. \quad (174)$$

We first consider $N = 8n + 2$, which satisfies $\text{sgn}[\sin(2\theta_N^*)] = -1$ and $\pi/2 < \theta_N^* < 3\pi/4$. By inserting the definition of θ_N^* from Proposition 3,

$$\tilde{\theta}' = \frac{\pi}{8} \frac{8n-1}{8n+3} + \frac{2\pi}{8n+3} \left(\frac{8n+2}{4} + \frac{1}{2}\right) = \frac{5\pi}{8} \in (\pi/2, 3\pi/4). \quad (175)$$

We therefore see $\theta' \in (\pi/2, 3\pi/4)$ (since θ' is a linear function of ϕ and lies in that range for the extremal values of ϕ).

For $N = 8n + 8$, $\text{sgn}[\sin(2\theta_N^*)] = -1$ and $3\pi/2 < \theta_N^* < 7\pi/4$. We have

$$\begin{aligned} \tilde{\theta}' &= \frac{\pi}{8} \frac{8n+5}{8n+9} + \frac{2\pi}{8n+9} \left(\frac{3(8n+8)}{4} + 1\right) \\ &= \frac{13\pi}{8} \in (3\pi/2, 7\pi/4), \end{aligned} \quad (176)$$

implying $3\pi/2 < \theta' < 7\pi/4$. For $N = 8n + 4$, we have $\text{sgn}[\sin(2\theta_N^*)] = +1$, $\pi/4 < \theta_N^* < \pi/2$ and

$$\tilde{\theta}' = -\frac{\pi}{8} \frac{8n+1}{8n+5} + \frac{2\pi}{8n+5} \frac{8n+4}{4} = \frac{3\pi}{8} \in (\pi/4, \pi/2), \quad (177)$$

implying $\pi/4 < \theta' < \pi/2$. Finally, for $N = 8n + 6$, $\text{sgn}[\sin(2\theta_N^*)] = +1$, $5\pi/4 < \theta_N^* < 3\pi/2$ and

$$\begin{aligned} \tilde{\theta}' &= -\frac{\pi}{8} \frac{8n+3}{8n+7} + \frac{2\pi}{8n+7} \left(\frac{3(8n+6)}{4} + \frac{1}{2}\right) \\ &= \frac{11\pi}{8} \in (5\pi/4, 3\pi/2), \end{aligned} \quad (178)$$

implying $\pi/4 < \theta' < \pi/2$. For every N , we have shown $\theta' \in \mathcal{G}$ and $\phi' \in (-\pi/4, \pi/4)$ for $\phi \in [0, \phi_N^*]$, hence $(\phi', \theta') \in \mathcal{F}$. $\square$

C.10 Evidence for Conjecture 2

Conjecture 2. For N even, the maximum randomness unconditioned on Eve, r , that can be generated by quantum strategies achieving an MABK value, s , is given by

$$r(s) = \begin{cases} N, & s \in (1, m_N^*], \\ r(\phi_s), & s \in (m_N^*, 2^{(N-1)/2}], \end{cases} \quad (179)$$

where $r(\phi)$ is defined in Eq. (40), and

$$\phi_s = \arg \min \{ |\phi| \mid \phi \in [0, \phi_N^*], \langle M_N(\phi, \theta(\phi)) \rangle = s \}. \quad (180)$$

C.10.1 Numerical approach for upper bounding maximum randomness versus MABK value

The objective of this section is to generalize the technique introduced in [23], which found upper bounds on the maximum randomness versus CHSH value using SOS decompositions. A python implementation of this technique is provided in our GitHub repository [40]. We begin by defining the problem in detail.

Let $\mathcal{Q}$ denote the set of quantum distributions, and $\mathcal{M}(P)$ denote the MABK value of a distribution $P \in \mathcal{Q}$. Moreover, let $H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_P$ be the conditional von Neumann entropy of the outputs $\mathbf{R}$ for inputs $\mathbf{X} = \mathbf{0}$ given the observed distribution P , taking the infimum over all quantum strategies that could give rise to P , i.e.,

$$\begin{aligned} H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_P &:= \inf_{\substack{|\Psi\rangle_{\tilde{\mathcal{Q}}E}, \\ \{\{\tilde{P}_{a_k|x_k}^{(k)}\}_{a_k}\}_k, \\ \text{compatible with } P}} H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_{\rho_{RE|\mathbf{0}}}. \end{aligned} \quad (181)$$

Similarly, let $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P$ be the Shannon entropy of the distribution on $\mathbf{R}$ for inputs $\mathbf{X} = \mathbf{0}$. Then the curve $R : [m_N^*, 2^{(N-1)/2}] \rightarrow [0, N]$, $s \mapsto R(s)$ we want to find is defined by the optimization

$$\begin{aligned} R(s) &= \max_P H(\mathbf{R}|\mathbf{X} = \mathbf{0}, E)_P \\ \text{s.t. } &\mathcal{M}(P) = s, \\ &P \in \mathcal{Q}. \end{aligned} \quad (182)$$

Our technique proceeds by defining a sequence of upper bounds on $\text{Graph}[R(s)] = \{(s, r) \mid r = R(s)\}$. [We speak of upper bounds on $\text{Graph}[R(s)]$ because we either upper bound $R(s)$ or $R^{-1}(r)$, and, because the functions turn out to be monotonic, either gives an upper bound.]

We begin by removing the dependence on Eve, which results in an upper bound $R(s) \leq \bar{R}(s)$ on the previous problem via strong subadditivity of the von Neumann entropy. Explicitly,

$$\begin{aligned} R(s) &\leq \bar{R}(s) = \max_P H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P \\ \text{s.t. } &\mathcal{M}(P) = s, \\ &P \in \mathcal{Q}. \end{aligned} \quad (183)$$

Next we state the following two lemmas, adapted from Ref. [23].

Lemma 11. [Monotonicity of $\bar{R}(s)$] Suppose m_N^* is the largest MABK value compatible with $p(\mathbf{a}|\mathbf{0}) = 2^{-N}$ (cf. Conjecture 1). Then the function $\bar{R}(s)$ is strictly decreasing on $[m_N^*, 2^{(N-1)/2}]$.

We will prove Lemma 11 as a corollary of the following.

Lemma 12. Let $\mathcal{Q}'$ be a convex set, $H : \mathcal{Q}' \rightarrow \mathbb{R}$ be concave, $\mathcal{M}' : \mathcal{Q}' \rightarrow \mathbb{R}$ be linear and $\bar{R}'(s) = \max_P H(P)$ subject to $\mathcal{M}'(P) = s$, $P \in \mathcal{Q}'$. Then $\bar{R}'(s)$ is convex.

Proof. We directly calculate

$$\begin{aligned} \bar{R}'[\lambda s_1 + (1 - \lambda)s_2] &= \max_P H(P) \\ \text{s.t. } &\mathcal{M}'(P) = \lambda s_1 + (1 - \lambda)s_2, \\ &P \in \mathcal{Q}', \\ &\geq \max_{P_1, P_2} H(\lambda P_1 + (1 - \lambda)P_2) \\ \text{s.t. } &\mathcal{M}'(P_1) = s_1, \mathcal{M}'(P_2) = s_2, \\ &P_1, P_2 \in \mathcal{Q}', \\ &\geq \max_{P_1, P_2} \lambda H(P_1) + (1 - \lambda)H(P_2) \\ \text{s.t. } &\mathcal{M}'(P_1) = s_1, \mathcal{M}'(P_2) = s_2, \\ &P_1, P_2 \in \mathcal{Q}', \\ &= \lambda \bar{R}'(s_1) + (1 - \lambda)\bar{R}'(s_2), \end{aligned} \quad (184)$$

where the first inequality follows due to the linearity of $\mathcal{M}'$ and convexity of $\mathcal{Q}'$, and the second inequality due to the concavity of H . $\square$

Proof of Lemma 11. Note that $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P$ is concave in P , the quantum set, $\mathcal{Q}$, is convex and the MABK value, $\mathcal{M}$, is linear. Lemma 12 hence implies that $\bar{R}(s)$ is concave in s .

Then note that $\bar{R}(m_N^*) > \bar{R}(s) \forall s \in (m_N^*, 2^{(N-1)/2}]$. This follows from the assumption that the largest MABK value achievable when $p(\mathbf{a}|\mathbf{0}) = 2^{-N}$ is m_N^* , and hence $\bar{R}(s)$ is initially decreasing, which implies the claim. $\square$

In the case $N = 2$, it is known that m_2^* is the largest MABK value compatible with $p(a, b|0, 0) = 1/4$; see [23, Corollary 1]. For $N = 4, 6, 8, 10, 12$, we have verified tightness numerically, as shown in Table 2¹¹

Lemma 13 (Inverse function of $\bar{R}(s)$). Suppose $r = \bar{R}(s)$. The function $\bar{R}(s)$ has the following inverse, denoted $\bar{R}^{-1}$, that satisfies $s = \bar{R}^{-1}(r)$, given by

$$\begin{aligned} \bar{R}^{-1}(r) &= \max_P \mathcal{M}(P) \\ \text{s.t. } &H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P = r, \\ &P \in \mathcal{Q}. \end{aligned} \quad (185)$$

¹¹Monotonicity of $\bar{R}(s)$ is not needed to establish [23, Corollary 1] or the upper bounds presented in Table 2.

Proof. We prove Lemma 13 by showing $\bar{R}^{-1}(\bar{R}(s)) = s$, and $\bar{R}(\bar{R}^{-1}(r)) = r$, and using Lemma 11. First consider $\bar{R}^{-1}(\bar{R}(s)) = s$,

$$\begin{aligned} \bar{R}^{-1}(\bar{R}(s)) &= \max_P \mathcal{M}(P) \\ \text{s.t. } H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P &= \bar{R}(s), \\ P &\in \mathcal{Q}. \end{aligned} \quad (186)$$

The constraint $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P = \bar{R}(s)$ implies that the achievable MABK values for the distribution P lies to the left of s , i.e., $\mathcal{M}(P) \leq s$, since the curve $\bar{R}(s)$ is strictly decreasing (cf. Lemma 11). To see this, suppose $s' = \mathcal{M}(P) > s$. Then $\bar{R}(s') < \bar{R}(s)$ by Lemma 11. But, $\bar{R}(s') \geq H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P$, implying $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P < \bar{R}(s)$, which by assumption cannot hold, giving a contradiction. Therefore $\bar{R}^{-1}(\bar{R}(s)) \leq \max_{\{P \in \mathcal{Q} \text{ s.t. } \mathcal{M}(P) \leq s\}} \mathcal{M}(P) = s$. Moreover, this bound is saturated by the behaviour P^* , which achieves the maximum defining $\bar{R}(s)$ ¹², i.e., satisfies $\bar{R}(s) = H(AB|X = 0, Y = 0)_{P^*}$ and $\mathcal{M}(P^*) = s$. This establishes $\bar{R}^{-1}(\bar{R}(s)) = s$.

For the other direction $\bar{R}(\bar{R}^{-1}(r))$, the same reasoning holds. The constraint $\mathcal{M}(P) = \bar{R}^{-1}(r)$ implies that $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P \leq r$ since any distribution that achieves a CHSH value of $\bar{R}^{-1}(r)$ can generate no more than r bits of randomness. Hence $\bar{R}(\bar{R}^{-1}(r)) \leq r$. This bound is saturated by the behaviour P^* that achieves the maximum defining $\bar{R}^{-1}(s)$, completing the proof. $\square$

From the above lemma, we can solve for upper bounds on the points $(s, R(s)) \in \text{Graph}[R(s)]$ using the inverse function, i.e., $(s, R(s)) \leq (s, \bar{R}(s)) = (\bar{R}^{-1}(r), r)$ where $\bar{R}(s) = r$. What remains is to upper bound $\bar{R}^{-1}(r)$, which will correspond to an upper bound on $R(s)$ due to the monotonicity argument. Next we obtain an upper bound by applying a symmetrization step to the distribution $\{p(\mathbf{a}|\mathbf{0})\}$:

Lemma 14. *Let $\mathcal{E}$ be the local channel that maps all even (odd) parity outcome strings to a uniform mixture of all even (odd) parity outcome strings,*

$$\mathcal{E}[p(\mathbf{a}|\mathbf{0})] = \begin{cases} \frac{1}{2^{N-1}} \sum_{\mathbf{b} \text{ even}} p(\mathbf{b}|\mathbf{0}), & \text{if } \mathbf{a} \text{ has even parity,} \\ \frac{1}{2^{N-1}} \sum_{\mathbf{b} \text{ odd}} p(\mathbf{b}|\mathbf{0}), & \text{if } \mathbf{a} \text{ has odd parity.} \end{cases} \quad (187)$$

The entropy after applying $\mathcal{E}$ is non-decreasing, and the MABK value is invariant under $\mathcal{E}$.

Proof. The first claim follows from the data processing inequality, $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P \leq H(\mathbf{R}|\mathbf{X} = \mathbf{0})_{\mathcal{E}(P)}$. The second claim comes from the fact that the correlators $\langle A_{\mathbf{x}} \rangle$ are invariant under $\mathcal{E}$. $\square$

¹²Note that the function $P \mapsto H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P$ is continuous, and the quantum set $\mathcal{Q}$ is compact (see the discussion in [60, Appendix B]), from which it follows that $\mathcal{Q}_s = \{P \in \mathcal{Q} : \mathcal{M}(P) = s\}$ is also compact. This implies that the objective function in (183) attains its maximum over its domain.

After the map $\mathcal{E}$ is applied, the probabilities are symmetrized, i.e., $p(\mathbf{a}|\mathbf{0}) = \epsilon$ if $\mathbf{a}$ is even and $p(\mathbf{a}|\mathbf{0}) = (1 - 2^{N-1}\epsilon)/2^{N-1} = (2 - 2^N\epsilon)/2^N$ if $\mathbf{a}$ is odd, for $\epsilon \in [0, 2^{1-N}]$. We can hence define the following upper bound on $\bar{R}(s) \leq \bar{\bar{R}}(s)$ where

$$\begin{aligned} \bar{\bar{R}}(s) &= \max_{\mathcal{E}(P)} H(\mathbf{R}|\mathbf{X} = \mathbf{0})_{\mathcal{E}(P)} \\ \text{s.t. } \mathcal{M}(\mathcal{E}(P)) &= s, \\ P &\in \mathcal{Q} \\ &= \max_P H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P \\ \text{s.t. } \mathcal{M}(P) &= s, \\ p(\mathbf{a}|\mathbf{0}) &= \begin{cases} \epsilon & \text{if } \mathbf{a} \text{ is even,} \\ (2 - 2^N\epsilon)/2^N & \text{if } \mathbf{a} \text{ is odd,} \end{cases} \\ P &\in \mathcal{Q}. \end{aligned} \quad (188)$$

We can now apply Lemma 13 (which still holds after applying $\mathcal{E}$, since it preserves the convexity of $\mathcal{Q}$), and define the function

$$\begin{aligned} \bar{\bar{R}}^{-1}(r) &= \max_P \mathcal{M}(P) \\ \text{s.t. } H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P &= r, \\ p(\mathbf{a}|\mathbf{0}) &= \begin{cases} \epsilon & \text{if } \mathbf{a} \text{ is even,} \\ (2 - 2^N\epsilon)/2^N & \text{if } \mathbf{a} \text{ is odd,} \end{cases} \\ P &\in \mathcal{Q}. \end{aligned} \quad (189)$$

We can further relax this by considering the correlator after symmetrization, $\langle A_0 \rangle = 2^{N-1}\epsilon - 2^{N-1}(2 - 2^N\epsilon)/2^N = 2^N\epsilon - 1$. Then we have $\bar{\bar{R}}^{-1}(r) \leq \bar{\bar{\bar{R}}}^{-1}(r)$ where

$$\begin{aligned} \bar{\bar{\bar{R}}}^{-1}(r) &= \max_P \mathcal{M}(P) \\ \text{s.t. } H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P &= r, \\ \langle A_0 \rangle &= 2^N\epsilon - 1, \\ P &\in \mathcal{Q}. \end{aligned} \quad (190)$$

This follows from the fact that the feasible region of the former optimization is a subset of the latter.

Next, consider the expression for the entropy $H(\mathbf{R}|\mathbf{X} = \mathbf{0})_P$ of a symmetrized distribution P . By direct computation we find

$$\begin{aligned} H(\mathbf{R}|\mathbf{X} = \mathbf{0})_{\mathcal{E}(P)} &= - \sum_{\mathbf{a} \text{ even}} \epsilon \log \epsilon \\ &\quad - \sum_{\mathbf{a} \text{ odd}} \frac{2 - 2^N\epsilon}{2^N} \log \frac{2 - 2^N\epsilon}{2^N} \\ &= -2^{N-1} \left(\epsilon \log \epsilon + \frac{2 - 2^N\epsilon}{2^N} \log \frac{2 - 2^N\epsilon}{2^N} \right) \\ &= -2^{N-1} \left(\epsilon \log 2^{N-1}\epsilon - \epsilon(N-1) \right) \\ &\quad + \frac{1}{2^{N-1}} (1 - 2^{N-1}\epsilon) \\ &\quad \cdot (\log(1 - 2^{N-1}\epsilon) - (N-1)) \\ &= N - 1 + H_{\text{bin}}(2^{N-1}\epsilon). \end{aligned} \quad (191)$$

Since $2^{N-1}\epsilon \in [0, 1/2]$ for $\epsilon \in [0, 1/2^N]$, there is a one-to-one correspondence between $H(\mathbf{R}|\mathbf{X}=\mathbf{0})_{\mathcal{E}(P)}$ and ϵ . Hence if $H(\mathbf{R}|\mathbf{X}=\mathbf{0})_{\mathcal{E}(P)} = r$, there exists a unique $\epsilon_r \in [0, 1/2^N]$ that satisfies $r = N - 1 + H_{\text{bin}}(2^{N-1}\epsilon_r)$. We can therefore write

$$\begin{aligned} \bar{\bar{R}}^{-1}(r) &= \max_P \mathcal{M}(P) \\ \text{s.t. } \langle \mathbf{A}_0 \rangle &= 2^N \epsilon_r - 1, \\ P &\in \mathcal{Q}. \end{aligned} \quad (192)$$

Next we introduce the function

$$\begin{aligned} \tilde{R}^{-1}(r) &= \min_{t,z,S} t \\ \text{s.t. } t\mathbb{I} - M_N &= S + z \left(\mathbf{A}_0 - (2^N \epsilon_r - 1)\mathbb{I} \right) \\ S &\succeq 0. \end{aligned} \quad (193)$$

Here S is an SOS decomposition for the operator expression

$$S = \left(t + z(2^N \epsilon_r - 1) \right) \mathbb{I} - \left(M_N + z \mathbf{A}_0 \right) \succeq 0, \quad (194)$$

hence, for any feasible point (t, z, S) , all quantum correlations satisfying $\langle \mathbf{A}_0 \rangle = 2^N \epsilon_r - 1$ also satisfy $t \geq \langle M_N \rangle$. This implies $\bar{\bar{R}}^{-1}(r) \leq \tilde{R}^{-1}(r)$. Now, by choosing a basis of monomials for S , one can evaluate the function $\tilde{R}^{-1}(r)$ numerically using a standard SDP solver. We refer the reader to Ref. [61] for details of how to find such formulation, and provide a self-contained description below.

If the monomial basis is chosen to be $\{R_\mu\}_\mu$, where each R_μ is a product of operators from the set $\{A_{x_k}^{(k)}\}_{x_k, k}$, then one can construct polynomials of the form $M_i = \sum_\mu q_i^\mu R_\mu$, $q_i^\mu \in \mathbb{C}$. S then admits an SOS decomposition if q_i^μ, R_μ are chosen such that

$$S = \sum_i M_i^\dagger M_i = \sum_{\mu\nu} M^{\mu\nu} R_\mu^\dagger R_\nu, \quad (195)$$

where $M^{\mu\nu} = \sum_i (q_i^\mu)^* q_i^\nu$ are matrix elements $\mathbf{M}$, the Gram matrix of the set of vectors $\{\mathbf{q}^\mu\}_\mu$, where $\mathbf{q}^\mu = [q_1^\mu, \dots, q_i^\mu, \dots]^T$. One now arrives at the following optimization problem

$$\begin{aligned} \min_{t,z,\mathbf{M}} t \\ \text{s.t. } \sum_{\mu\nu} M^{\mu\nu} R_\mu^\dagger R_\nu &= \left(t + z(2^N \epsilon_r - 1) \right) \mathbb{I} \\ &\quad - \left(M_N + z \mathbf{A}_0 \right), \\ \mathbf{M} &\succeq 0. \end{aligned} \quad (196)$$

Now one can write $R_\mu^\dagger R_\nu = \sum_i F_{\mu\nu}^i E_i$, where $\mathbf{F}^i$ is some complex matrix with elements $F_{\mu\nu}^i$, for $E_i \in \text{Canonical}[\{R_\mu^\dagger R_\nu\}_{\mu\nu}]$, which denotes the canonical reduction of the set $\{R_\mu^\dagger R_\nu\}_{\mu\nu}$ using the commutation and projective relations (see Ref. [61] for details).

Similarly, one can express the operator expression in this basis, $\left(t + z(2^N \epsilon_r - 1) \right) \mathbb{I} - \left(M_N + z \mathbf{A}_0 \right) = \sum_i s^i E_i$, and arrive at an SDP of the form

$$\begin{aligned} \min_{t,z,\mathbf{M}} t \\ \text{s.t. } \sum_{\mu\nu} F_{\mu\nu}^i M^{\mu\nu} &= s^i \quad \forall i, \\ \mathbf{M} &\succeq 0. \end{aligned} \quad (197)$$

For our numerical calculations, we choose the monomial basis $\left\{ \prod_{k=1}^{N/2} A_{x_k}^{(k)}, \prod_{k=N/2+1}^N A_{x_k}^{(k)} \right\}_{x_k}$, consisting of $2^{N/2+1}$ elements.

C.10.2 Numerical results

We present the numerical results which support our conjecture in Fig. 6, for $N = 4, 6, 8, 10, 12$. We find the two curves coincide very well, suggesting the lower bounds are indeed tight. The case of $N = 2$ was proven in Ref. [23]. The technique also generates the numerical results in Table 2.

C.11 Other Bell inequalities for $N = 3$

The above technique can be directly applied to find bounds on other Bell inequalities. We applied this to the tripartite expressions in Eqs. (46) and (47), which, along with the MABK expression, trivial extensions of CHSH and positivity cover all inequivalent classes of facets for the local polytope with uniform marginals (see Ref. [62] for the general case). The exact trade-offs we found are detailed in Fig. 7, and we summarize various bounds of each class in Table 3. We restate the Bell expressions studied below [42]:

$$\begin{aligned} M_3 &= A_0(B_0C_1 + B_1C_0)/2 + A_1(B_0C_0 - B_1C_1)/2, \\ S_1 &= \frac{1}{4} \sum_{x,y,z} A_x B_y C_z - A_1 B_1 C_1, \\ S_2 &= A_0 B_0 (C_0 + C_1) - A_1 B_1 (C_0 - C_1), \\ S_3 &= A_0 (B_0 C_0 + B_0 C_1 + B_1 C_0 - B_1 C_1), \\ S_4 &= A_0 B_0 C_0, \end{aligned} \quad (198)$$

where we have made the substitution $A_{x_0}^{(0)} \rightarrow A_x$, $A_{x_1}^{(1)} \rightarrow B_y$ etc. for readability. For S_1 and S_2 , we applied our numerical technique by substituting them in for the MABK expression. When $\epsilon = 1/8$, the SDP returns an upper bound on the maximum violation with maximum randomness, which we label s^* (the facet class in question will be clear from context). For S_1 , we numerically identified that $\epsilon = 1/6$ achieves the maximum quantum value of $5/3$, for which the corresponding entropy is given by Eq. (191), which evaluates to $2 + H_{\text{bin}}(2/3)$; by varying $\epsilon \in [1/8, 1/6]$ we achieve the trade off in Fig. 7, which corresponds to an upper bound. Similarly, for S_2 we

found $\epsilon = (2 + \sqrt{2})/16$ achieves the maximum quantum value $2\sqrt{2}$, so by varying $\epsilon \in [1/8, (2 + \sqrt{2})/16]$ we find the curve in Fig. 7. The corresponding entropy is $2 + H_{\text{bin}}((2 + \sqrt{2})/4)$ which is exactly 1 plus the amount certified by maximum CHSH violation [46, 63, 23].

We checked the value of our constructions in Eq. (20), with $N = 3$, for the expressions $\langle S_1 \rangle$ and

$\langle S_2 \rangle$. We find

$$\begin{aligned}\langle S_1 \rangle &= 2 \sin(3\theta/2) \cos^3(\theta/2), \\ \langle S_2 \rangle &= \sin \theta - \sin 2\theta + \sin 3\theta.\end{aligned}\quad (199)$$

By varying θ , we find one can achieve a wide range of values from the local bound, however the values of s^* are inaccessible using this parameterization; one will need to consider a larger family of strategies to further investigate if s^* is achievable with maximum randomness.

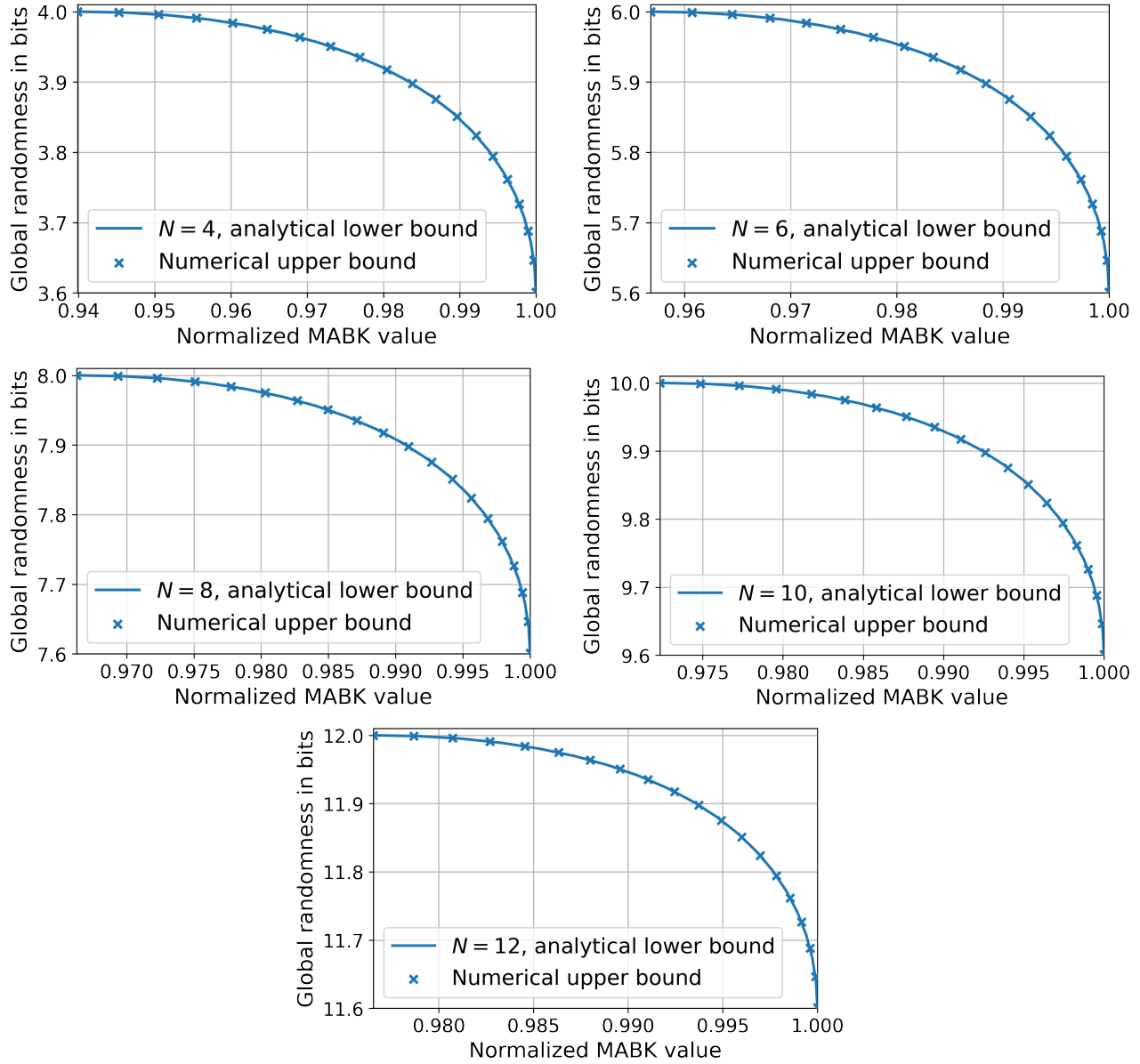


Figure 6: Lower bounds on the maximum DI randomness versus MABK value, normalized by the maximum quantum value $2^{(N-1)/2}$, for even N . These are compared to numerical upper bounds computed via the SOS technique. Each plot starts at the conjectured maximum MABK value achievable with maximum randomness, m_N^* , and ends with the maximum quantum MABK value. The points coincide closely with each curve, supporting our conjecture that the analytical lower bound provided in the text is tight.

Facet	η^L	s^*	η^Q	$r(\langle S_3 \rangle = \eta^Q)$
M_3	1	2	2	3
S_1	1	1.64621108	$5/3 \approx 1.66666667$	$2 + H_{\text{bin}}(2/3) \approx 2.91829583$
S_2	2	2.59807617	$2\sqrt{2} \approx 2.82842712$	$7/2 - \log_2(1 + \sqrt{2})/\sqrt{2} \approx 2.60087604$
S_3	2	$3\sqrt{3}/2 \approx 2.59807621$	$2\sqrt{2} \approx 2.82842712$	$5/2 - \log_2(1 + \sqrt{2})/\sqrt{2} \approx 1.60087604$
S_4	1	0	1	0

Table 3: Relevant bounds on violations of all inequivalent facet classes of the local polytope in the $N = 3$ scenario, with two binary measurement per party, given in Eq. (198). M_3 is the MABK expression, S_1, S_2 are the other non trivial facets for this scenario, S_3 is the trivial extension of CHSH and S_4 is a positivity facet. η^L and η^Q are the local and quantum bounds respectively, and s^* is an upper bound on the maximum value achievable with maximum randomness, computed numerically for S_1 and S_2 while all other cases are analytic. $r(\langle S_3 \rangle = \eta^Q)$ is an upper bound on the asymptotic global DI randomness when the maximum quantum violation is achieved. The cases M_3 and S_3 are known to be tight [31, 23]. Due to its structure, S_2 retains the same characteristics as CHSH [23], whereas S_1 exhibits an unexplored trade-off. Since S_4 cannot be violated, it cannot be used for DI randomness certification.

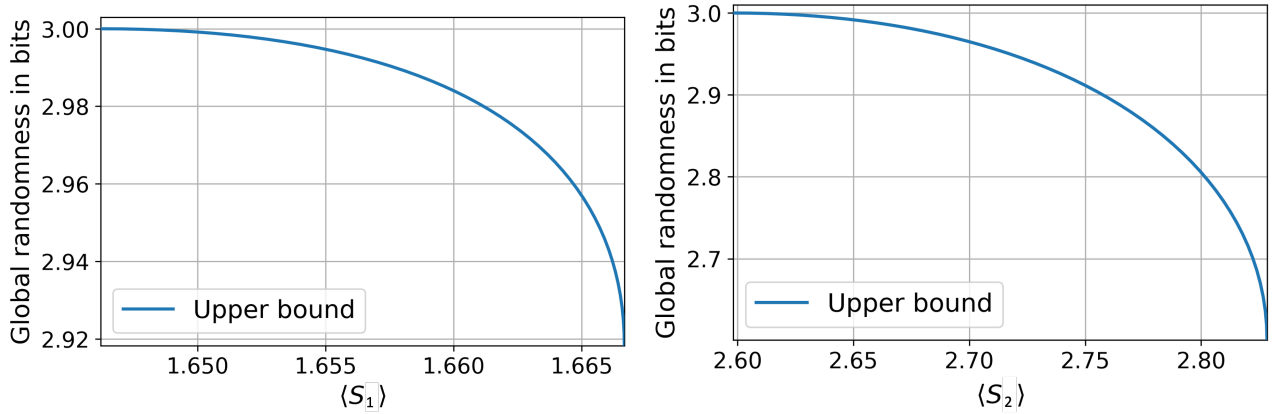


Figure 7: Upper bounds on the trade-off between global DI randomness generation and violation of non-trivial facets in the tripartite scenario with two binary measurements per party. The bounds are numerical and generated by our SOS technique. S_1 and S_2 are two non-trivial extremal Bell expressions given in Eq. (198). Details of the exact numerical values are given in Table 3.