

## Entropy Bounds for Multiparty Device-Independent Cryptography

Federico Grasselli<sup>Ⓞ,\*</sup>, Gláucia Murta<sup>Ⓞ,†</sup>, Hermann Kampermann, and Dagmar Bruß

*Institut für Theoretische Physik III, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, Düsseldorf D-40225, Germany*



(Received 24 July 2020; accepted 7 December 2020; published 14 January 2021)

Multiparty quantum cryptography based on distributed entanglement will find its natural application in the upcoming quantum networks. The security of many multipartite device-independent (DI) protocols, such as DI conference-key agreement, relies on bounding the von Neumann entropy of the parties' outcomes conditioned on the eavesdropper's information, given the violation of a multipartite Bell inequality. We consider three parties testing the Mermin-Ardehali-Belinskii-Klyshko (MABK) inequality and certify the privacy of their outcomes by bounding the conditional entropy of a single party's outcome and the joint conditional entropy of two parties' outcomes. From the former bound, we show that genuine multipartite entanglement is necessary to certify the privacy of a party's outcome, while the latter significantly improves previous results. We obtain the entropy bounds thanks to two general results of independent interest. The first one drastically simplifies the quantum setup of an  $N$ -partite Bell scenario. The second one provides an upper bound on the violation of the MABK inequality by an arbitrary  $N$ -qubit state, as a function of the state's parameters.

DOI: [10.1103/PRXQuantum.2.010308](https://doi.org/10.1103/PRXQuantum.2.010308)

### I. INTRODUCTION

Stimulated by data-security concerns and by commercial opportunities, several companies and governments are increasingly investing resources in quantum-cryptography technologies [1,2]. Those include, most prominently, quantum key distribution (QKD) [3–10] and quantum random-number generation [11,12]. The former enables two parties to establish a secret key (shared random bitstring), while the latter is considered the only source of genuine randomness. In the context of emerging quantum networks [13–20], the task of QKD can be generalized to quantum conference-key agreement (CKA) [21–27]. Here,  $N$  parties establish a common secret key to securely broadcast messages within their network, as proved by recent CKA experiments [28,29]. However, it is challenging to ensure that the assumptions on the implementation of these cryptographic tasks are met in practice, hence jeopardizing their security.

This led to the development of device-independent (DI) cryptographic protocols, whose security holds independently of the actual functioning of the quantum devices

and is based on the observation of a Bell inequality violation [30]. Such protocols include DIQKD [31–39] and DICKA [40–43] schemes, where a secret key is shared by two or more parties, respectively. Otherwise, with DI randomness generation (DIRG) protocols [44–50] one can generate random bitstrings that are guaranteed to be private thanks to a Bell violation.

A crucial aspect of any DI protocol is the ability to carefully estimate, from the observed Bell violation, the minimum amount of uncertainty that a potential eavesdropper, Eve, could have about the protocol's outputs. Indeed, this quantity determines the length of the secret random bitstring that can be distilled from the protocol's outputs. Eve's uncertainty is quantified by appropriate conditional von Neumann entropies [6,33,34,38] of the effective quantum state shared by the parties in a generic round of the protocol. The goal is to minimize the entropy over all the possible states yielding the observed Bell inequality violation.

This task can be carried out numerically, however the available techniques [51–54] focus on minimizing a lower bound on the von Neumann entropy, namely the min-entropy [6], thus producing suboptimal results. Here we follow an analytical approach that reduces the degrees of freedom of the generic state shared by the parties without loss of generality, thereby allowing a direct minimization of the conditional von Neumann entropy. This can result in a tight bound of Eve's uncertainty, hence in longer secret bitstrings and higher noise tolerance for the DI protocol. Such an analytical procedure has only been developed by

\*federico.grasselli@hhu.de

†glaucia.murta@hhu.de

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

Pironio *et al.* [34] for the case of two parties testing the Clauser-Horne-Shimony-Holt (CHSH) inequality [55].

In this work we develop an analytical procedure applicable to multipartite DI scenarios. Specifically, we consider  $N$  parties, each equipped with two measurement settings with binary outcomes, testing a generic full-correlator Bell inequality—i.e., an inequality where each correlator involves every party [56]. Remarkably and without loss of generality, we reduce the generic state shared by the  $N$  parties (in one protocol round) to a mere  $N$ -qubit state almost diagonal in the GHZ basis. Notably, we recover the result of Pironio *et al.* when  $N = 2$ .

We then focus on the Mermin-Ardehali-Belinskii-Klyshko (MABK) inequality [57–59] and derive an analytical bound on the maximal violation of the MABK inequality yielded by rank-one projective measurements on an given  $N$ -qubit state. This is a result of independent interest, which generalizes the bound for the bipartite case of Ref. [60] and constitutes the first of its kind valid for an arbitrary  $N$ -qubit state.

Our results on the state reduction in a multipartite DI scenario and on the MABK violation upper bound, allow us to quantify Eve’s uncertainty about the parties’ outcomes when three parties, Alice, Bob, and Charlie, test the MABK inequality (see Fig. 1). Specifically, we obtain an analytical lower bound on the von Neumann entropy of Alice’s outcome conditioned on Eve’s information. We compare our bound with a numerical estimation of

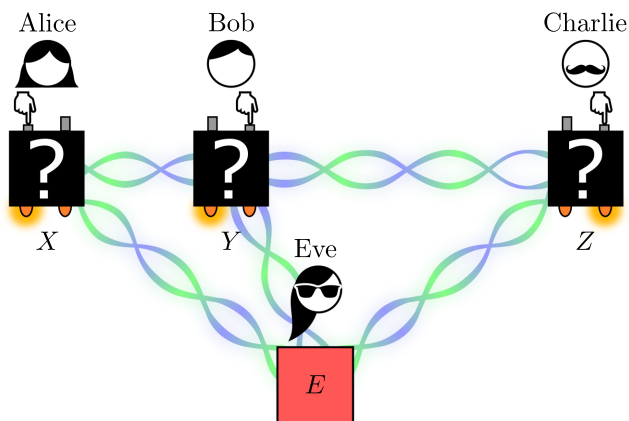


FIG. 1. Alice, Bob, and Charlie generate device-independent randomness from the input-output correlations of their unknown quantum devices by testing the MABK inequality. Each device is equipped with two inputs and two outputs. The eavesdropper Eve might hold a quantum memory  $E$  entangled with the parties’ devices and use it to guess their outcomes  $X$ ,  $Y$ , or  $Z$ . We quantify Eve’s uncertainty on Alice’s outcome  $X$  by computing the conditional von Neumann entropy  $H(X|E)$ . Additionally, we assume that Alice and Bob are colocated and collaborate to generate global randomness from their outcomes  $X$  and  $Y$ . We quantify Eve’s joint uncertainty on their outcomes by computing  $H(XY|E)$ .

the corresponding tight entropy bound and with previous results. We additionally prove that genuine multipartite entanglement is a necessary resource to certify the privacy of a party’s outcome in any  $N$ -party MABK scenario. The bound can find potential application in the security of DIRG based on multipartite nonlocality. We also provide a heuristic argument for which full-correlator Bell inequalities, such as the MABK inequality, are unlikely to be employed in any DICKA protocol.

In the same tripartite scenario of Fig. 1, we derive a lower bound on the joint conditional von Neumann entropy of Alice and Bob’s outcomes, which substantially improves the result derived in Ref. [50], where the authors bound the corresponding min-entropy. The derived bound can be employed in proving the security of DI global randomness generation schemes.

## II. REDUCTION OF THE $N$ -PARTY QUANTUM STATE

Let  $Alice_1, Alice_2, \dots, Alice_N$  be  $N$  parties who want to generate private randomness (random bitstrings) from the outcomes of their uncharacterized quantum devices. In order to certify device independently that the generated randomness is private, they test a generic full-correlator Bell inequality [56], where each party can choose among two measurement settings with binary outcomes on her respective device. We identify this as the  $(N, 2, 2)$  DI scenario.

An eavesdropper, Eve, wants to learn the randomness generated by some of the parties. We consider the most adversarial scenario where Eve herself may distribute arbitrary quantum states to the parties’ devices, which could be forged by Eve. The device of each party measures the binary observable  $A_x^{(i)}$  ( $i = 1, \dots, N$ ) on the received quantum state, according to  $Alice_i$ ’s measurement input ( $x = 0, 1$ ). Note that the observables measured by each device may be pre-established by Eve.

The tested Bell inequality is a linear combination of full correlators of the form:

$$\langle A_{x_1}^{(1)} \dots A_{x_N}^{(N)} \rangle. \quad (1)$$

From the observed Bell violation, the parties can quantify Eve’s uncertainty on their random bitstrings by computing an appropriate conditional von Neumann entropy. With this result, a party could enhance the privacy of her bitstring (with privacy amplification [6]) and use it for various cryptographic tasks (e.g., DICKA or DIRG).

Here we present a fundamental result that enables a direct computation of the conditional von Neumann entropy of interest. Indeed, our result drastically simplifies, without loss of generality, the general quantum setup described above. For instance, the generic quantum state distributed to the parties is reduced to an  $N$ -qubit state (almost) diagonal in the GHZ basis.

The GHZ basis [21] for the Hilbert space of  $N$  qubits is defined as follows.

**Definition 1.** *The GHZ basis for the set of  $N$ -qubit states is composed of the following  $2^N$  states:*

$$|\psi_{\sigma, \vec{u}}\rangle = \frac{1}{\sqrt{2}} \left[ |0\rangle |\vec{u}\rangle + (-1)^\sigma |1\rangle |\vec{u}\rangle \right], \quad (2)$$

where  $\sigma \in \{0, 1\}$  while  $\vec{u} \in \{0, 1\}^{N-1}$  and  $\vec{u} = \vec{1} \oplus \vec{u}$  are  $(N-1)$ -bit strings. In particular, for a three-qubit state, the GHZ basis reads

$$|\psi_{i,j,k}\rangle = \frac{1}{\sqrt{2}} \left[ |0, j, k\rangle + (-1)^i |1, \bar{j}, \bar{k}\rangle \right] \quad i, j, k \in \{0, 1\}, \quad (3)$$

where the bar over a bit indicates its negation.

We now formally state the first major result of this work, the proof of which is reported in Sec. VIII A.

**Theorem 1.** *Let  $N$  parties test an  $(N, 2, 2)$  full-correlator Bell inequality in order to certify the privacy of their outcomes. It is not restrictive to assume that, in each round, Eve distributes a mixture  $\sum_\alpha p_\alpha \rho_\alpha$  of  $N$ -qubit states  $\rho_\alpha$ , together with a flag  $|\alpha\rangle$  (known to her), which determines the measurements performed on  $\rho_\alpha$  given the parties' inputs. Without loss of generality, the measurements performed by each device on  $\rho_\alpha$  are rank-one binary projective measurements in the  $(x, y)$  plane of the Bloch sphere. Moreover, each state  $\rho_\alpha$  is diagonal in the GHZ basis, except for some purely imaginary off-diagonal terms:*

$$\rho_\alpha = \sum_{\vec{u} \in \{0, 1\}^{N-1}} \left[ \lambda_{0\vec{u}}^\alpha |\psi_{0, \vec{u}}\rangle \langle \psi_{0, \vec{u}}| + \lambda_{1\vec{u}}^\alpha |\psi_{1, \vec{u}}\rangle \langle \psi_{1, \vec{u}}| \right. \\ \left. + \text{i} s_{\vec{u}}^\alpha (|\psi_{0, \vec{u}}\rangle \langle \psi_{1, \vec{u}}| - |\psi_{1, \vec{u}}\rangle \langle \psi_{0, \vec{u}}|) \right], \quad (4)$$

Finally,  $N$  arbitrary off-diagonal terms  $s_{\vec{u}}^\alpha$  can be assumed to be zero and the corresponding diagonal elements  $(\lambda_{0\vec{u}}^\alpha, \lambda_{1\vec{u}}^\alpha)$  can be arbitrarily ordered (e.g.,  $\lambda_{0\vec{u}}^\alpha \geq \lambda_{1\vec{u}}^\alpha$ ).

In the following we focus our analysis on a given state  $\rho_\alpha$ . Hence, for ease of notation we drop the symbol  $\alpha$  in the parameters related to the state  $\rho_\alpha$  (e.g.,  $\lambda_{0\vec{u}}^\alpha$  and  $s_{\vec{u}}^\alpha$ ) when there is no ambiguity.

Note that, for  $N = 2$ , we recover the result of Ref. [34]. By applying Theorem 1 to the case of  $N = 3$  parties, it is not restrictive to assume that they share a mixture of states  $\rho_\alpha$ , with the following matrix representation in the GHZ

basis:

$$\rho_\alpha = \begin{bmatrix} \lambda_{000} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_{100} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{001} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{101} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{010} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{110} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_{011} & \text{i} s \\ 0 & 0 & 0 & 0 & 0 & 0 & -\text{i} s & \lambda_{111} \end{bmatrix}. \quad (5)$$

The eigenvalues of Eq. (5) are given by

$$\rho_{ijk} = \lambda_{ijk} \quad (j, k) \neq (1, 1) \\ \rho_{i11} = \frac{\lambda_{011} + \lambda_{111} + (-1)^i \sqrt{(\lambda_{011} - \lambda_{111})^2 + 4s^2}}{2}. \quad (6)$$

In order to accurately quantify Eve's uncertainty on the parties' outcomes via conditional von Neumann entropies, one also needs an analytical expression for the maximal violation of the tested Bell inequality. In Sec. III we establish such a result for the MABK inequality.

### III. UPPER BOUND ON MABK VIOLATION

The MABK inequality [57–59] is one possible generalization of the CHSH inequality [55] and is derived on the following MABK operator.

**Definition 2.** *The MABK operator  $M_N$  is defined by recursion [61, 62]*

$$M_2 = G_{\text{CHSH}}[A_0^{(1)}, A_1^{(1)}, A_0^{(2)}, A_1^{(2)}] \\ \equiv A_0^{(1)} \otimes A_0^{(2)} + A_0^{(1)} \otimes A_1^{(2)} + A_1^{(1)} \otimes A_0^{(2)} \\ - A_1^{(1)} \otimes A_1^{(2)} \quad (7) \\ M_N = \frac{1}{2} G_{\text{CHSH}}[M_{N-1}, \overline{M_{N-1}}, A_0^{(N)}, A_1^{(N)}],$$

where  $A_{x_i}^{(i)}$  for  $x_i = 0, 1$  are the binary observables of Alice <sub>$i$</sub>  (each observable satisfies  $[A_{x_i}^{(i)}]^\dagger = A_{x_i}^{(i)}$  and  $[A_{x_i}^{(i)}]^2 \leq \text{id}$ , where “id” is the identity operator) and where  $\overline{M}_i$  is the operator obtained from  $M_i$  by replacing every observable  $A_{x_i}^{(i)}$  with  $A_{1-x_i}^{(i)}$ . For  $N = 3$ , the MABK operator reads

$$M_3 = A_0 \otimes B_0 \otimes C_1 + A_0 \otimes B_1 \otimes C_0 \\ + A_1 \otimes B_0 \otimes C_0 - A_1 \otimes B_1 \otimes C_1, \quad (8)$$

where  $A_x$ ,  $B_y$ , and  $C_z$  are Alice's, Bob's, and Charlie's observables, respectively.

The  $N$ -partite MABK inequality reads [61,62]

$$\langle M_N \rangle = \text{Tr}[M_N \rho] \leq \begin{cases} 2, & \text{classical bound} \\ 2^{N/2}, & \text{GME threshold} \\ 2^{(N+1)/2} & \text{quantum bound} \end{cases}, \quad (9)$$

where  $M_N$  is the MABK operator and a violation of the GME threshold implies that the parties share a genuine multipartite entangled (GME) state.

The second major result of this work is an upper bound on the maximal MABK violation obtained when  $N$  parties share an  $N$ -qubit state and perform rank-one projective measurements on the respective qubits. The bound is state dependent and tight on certain classes of states (proof and tightness conditions in Sec. VIII B). This is the first bound of such a kind for an  $N$ -partite Bell inequality. Recently, the authors in Ref. [63] derived a similar bound in the  $N = 3$  case. Our bound is tight on a larger set of states (discussion in Sec. VIII B) and is valid for general  $N$ .

**Theorem 2.** *The maximum violation  $\mathcal{M}_\rho$  of the  $N$ -partite MABK inequality, Eq. (9), attained with rank-one projective measurements on an  $N$ -qubit state  $\rho$ , satisfies*

$$\mathcal{M}_\rho \leq 2\sqrt{t_0 + t_1}, \quad (10)$$

where  $t_0$  and  $t_1$  are the largest and second-to-the-largest eigenvalues of the matrix  $T_\rho T_\rho^T$ , where  $T_\rho$  is the correlation matrix of  $\rho$ .

We define the correlation matrix of an  $N$ -qubit state as follows.

**Definition 3.** *The correlation matrix of an  $N$ -qubit state  $\rho$ ,  $T_\rho$ , is a square or rectangular matrix defined by the elements  $[T_\rho]_{ij} = \text{Tr}[\rho \sigma_{v_1} \otimes \dots \otimes \sigma_{v_N}]$  such that*

$$\begin{aligned} i &= 1 + \sum_{k=1}^{\lceil N/2 \rceil} 3^{\lceil N/2 \rceil - k} (v_k - 1), \\ j &= 1 + \sum_{k=\lceil N/2 \rceil + 1}^N 3^{N-k} (v_k - 1), \end{aligned} \quad (11)$$

where  $v_1, \dots, v_N \in \{1, 2, 3\}$ ,  $\sigma_{v_i}$  are the Pauli operators, and  $\lceil x \rceil$  returns the smallest integer greater or equal to  $x$ .

**Remark.** *We remark that the most general measurements to be considered in computing the maximal MABK violation are projective measurements defined by observables  $[A_{x_i}^{(i)}]^2 = \text{id}$  [56], since positive-operator valued measures (POVMs) never provide higher violations [64,65]. Such measurements on qubits reduce to either (i) rank-one*

*projective measurements given by  $A_{x_i}^{(i)} = \vec{a}_{x_i}^i \cdot \vec{\sigma}$  with unit vectors  $\vec{a}_{x_i}^i \in \mathbb{R}^3$  and where  $\vec{\sigma} = (X, Y, Z)^T$  is the vector of Pauli operators, or (ii) rank-two projective measurements given by the identity  $A_{x_i}^{(i)} = \pm \text{id}$ , i.e., measurements with a fixed outcome. While for  $N = 2$  parties the identity does not lead to any violation [60] and the optimal measurements are described by case (i), in a multipartite scenario case (ii) cannot be ignored.*

For instance, if  $N = 3$  parties share the state  $\text{id}/2 \otimes |\psi_{00}\rangle\langle\psi_{00}|$  (with  $|\psi_{00}\rangle$  given in Definition 1), an MABK violation of  $2\sqrt{2}$  is achieved if the first party measures  $A_0^{(1)} = A_1^{(1)} = \text{id}$ , whereas no violation is obtained if her measurements are restricted to  $A_{x_i}^{(i)} = \vec{a}_{x_i}^i \cdot \vec{\sigma}$ .

We point out that previous works [63,66–68] addressing the violation of multipartite Bell inequalities achieved by a given multiqubit state have neglected case (ii) and only considered case (i). By applying the above example, we stress that the results of Refs. [63,66–68] characterizing Bell violations yielded by multiqubit states are, in fact, less general than claimed.

Nevertheless, for states whose maximal violation is above the GME threshold, the bound we provide in Theorem 2 is general and holds independently of the parties' measurements. Indeed, measuring the identity cannot lead to violations above the GME threshold and thus case (i) is already the most general.

By applying Theorem 2 to the state  $\rho_\alpha$  in Eq. (5), we obtain an upper bound on the maximal MABK violation  $\mathcal{M}_\alpha$  achievable on  $\rho_\alpha$  with rank-one projective measurements.

**Corollary 1.** *For a tripartite state  $\rho_\alpha$  of the form given in Eq. (5), the maximal violation  $\mathcal{M}_\alpha$  of the MABK inequality achieved with rank-one projective measurements satisfies*

$$\mathcal{M}_\alpha \leq \mathcal{M}_\alpha^\dagger = 4 \sqrt{\sum_{j,k=0}^1 (\rho_{0jk} - \rho_{1jk})^2}, \quad (12)$$

where  $\{\rho_{ijk}\}$  are the eigenvalues of the state  $\rho_\alpha$ , as specified in Eq. (6).

In Sec. VIII B, we provide the tightness conditions, Eq. (78), for which the upper bound in Eq. (12) is achieved.

#### IV. ONE-OUTCOME CONDITIONAL ENTROPY BOUND

Consider the (3, 2, 2) DI scenario of Fig. 1. Alice, Bob, and Charlie test the tripartite MABK inequality in order to quantify Eve's uncertainty on the generic outcome  $X$  of one of Alice's observables, by computing the conditional von Neumann entropy  $H(X|E)$ . We emphasize that, in a DIRG protocol, the entropy  $H(X|E)$  determines the



asymptotic rate of secret random bits extracted by applying privacy amplification [6] on Alice's  $X$  outcomes [69,70]. Similarly, in DICKA the secret key rate is determined by  $H(X|E)$  decreased by the amount of classical information disclosed by the parties in the other steps of the protocol [42,62,69].

We derive an analytical lower bound on  $H(X|E)$  as a function of the observed MABK violation. Theorem 1 guarantees that we can restrict the computation of the conditional entropy  $H(X|E_{\text{tot}})$  over a mixture of states  $\rho_\alpha$  of the form Eq. (5) and to rank-one projective measurements performed by the parties. We emphasize that the total information  $E_{\text{tot}} = E \Xi$  available to Eve includes the knowledge of the flag  $\Xi$ , which carries the value of  $\alpha$  (see Sec. VIII A). The goal is to lower bound the conditional entropy  $H(X|E_{\text{tot}})$  with a function  $F$  of the observed MABK violation  $m$ .

Thanks to Theorem 1, we can express the conditional entropy  $H(X|E_{\text{tot}})$  as follows:

$$\begin{aligned} H(X|E_{\text{tot}}) &= \sum_{\alpha} p_{\alpha} H(X|E \Xi = \alpha) \\ &= \sum_{\alpha} p_{\alpha} H(X|E)_{\rho_{\alpha}}, \end{aligned} \quad (13)$$

as a matter of fact the state on which  $H(X|E_{\text{tot}})$  is computed is a classical-quantum state [see Eq. (41)]. At the same time, the observed violation  $m$  can be expressed as

$$m = \sum_{\alpha} p_{\alpha} m_{\alpha}. \quad (14)$$

In Eq. (13), the entropy  $H(X|E)_{\rho_{\alpha}}$  is the conditional entropy of Alice's outcome given that Eve distributed the state  $\rho_{\alpha}$ , while  $p_{\alpha}$  is the probability distribution of the mixture prepared by Eve. In Eq. (14),  $m_{\alpha}$  is the violation that the parties would observe had they shared the state  $\rho_{\alpha}$  in every round of the protocol and performed the corresponding rank-one projective measurements.

We then aim at lower bounding  $H(X|E)_{\rho_{\alpha}}$  with a convex function  $F$  of the MABK violation  $m_{\alpha}$ :

$$H(X|E)_{\rho_{\alpha}} \geq F(m_{\alpha}). \quad (15)$$

Indeed, by combining Eqs. (13)–(15) and the convexity of  $F$ , one can obtain the desired lower bound on  $H(X|E_{\text{tot}})$  as a function of the observed violation  $m$ :

$$H(X|E_{\text{tot}}) \geq F(m). \quad (16)$$

The bound is tight if, for any given MABK violation  $m$ , there exist a quantum state and a set of measurements that achieve violation  $m$  and whose outcome's conditional entropy is exactly given by  $F(m)$ . We now obtain the function  $F$  by minimizing  $H(X|E)_{\rho_{\alpha}}$  over all the states  $\rho_{\alpha}$  yielding a violation  $m_{\alpha}$ .

The eigenvectors of the state  $\rho_{\alpha}$ , corresponding to the eigenvalues in Eq. (6), read

$$\begin{aligned} |\rho_{ijk}\rangle &= |\psi_{i,j,k}\rangle \quad (j,k) \neq (1,1), \\ |\rho_{011}\rangle &= \cos(t) |\psi_{0,1,1}\rangle - \text{i} \sin(t) |\psi_{1,1,1}\rangle, \\ |\rho_{111}\rangle &= \cos(t) |\psi_{1,1,1}\rangle - \text{i} \sin(t) |\psi_{0,1,1}\rangle, \end{aligned} \quad (17)$$

where the parameter  $t$  is defined as

$$t = \arctan \frac{2s}{\lambda_{011} - \lambda_{111} + \sqrt{(\lambda_{011} - \lambda_{111})^2 + 4s^2}}. \quad (18)$$

By combining the freedom in ordering the diagonal elements  $\lambda_{ijk}$  of  $\rho_{\alpha}$  (cf. Theorem 1) with the definition of the eigenvalues  $\rho_{ijk}$  in Eq. (6), one can impose the following constraints on the eigenvalues:

$$\rho_{0jk} \geq \rho_{1jk} \quad \forall j,k. \quad (19)$$

The entropy  $H(X|E)_{\rho_{\alpha}}$  is computed on the classical-quantum state:

$$\rho_{XE}^{\alpha} = (\mathcal{E}_X \otimes \text{id}_E) \text{Tr}_{BC} [|\phi_{ABCE}^{\alpha}\rangle \langle \phi_{ABCE}^{\alpha}|], \quad (20)$$

where  $|\phi_{ABCE}^{\alpha}\rangle$  is a purification of  $\rho_{\alpha}$  (Eve holds the purifying system  $E$ ), while  $\mathcal{E}_X$  represents one of the two projective measurements of Alice, defined by the eigenvectors:

$$|a\rangle_X = \frac{1}{\sqrt{2}} [ |0\rangle + (-1)^a e^{\text{i}\varphi_X} |1\rangle ] \quad a \in \{0,1\}, \quad (21)$$

where  $\varphi_X \in [0, 2\pi)$  identifies the measurement direction in the  $(x,y)$  plane of the Bloch sphere. For definiteness, we choose  $\varphi_X$  to be the measurement direction of Alice's observable  $A_0$ :  $\varphi_X = \varphi_{A_0}$ . Hence, we are deriving a lower bound on  $H(X_{A_0}|E_{\text{tot}})$ , where  $X_{A_0}$  is the outcome of Alice's observable  $A_0$ .

The entropy minimization can be simplified if, instead of minimizing over the matrix elements  $\{\lambda_{ijk}\}$  and  $s$  of  $\rho_{\alpha}$ , one minimizes over its eigenvalues  $\{\rho_{ijk}\}$  and over  $t$ . This change of variables is legitimized by the bijective map linking the two sets of parameters, defined by Eqs. (6) and (18).

The solution of the following optimization problem yields a tight lower bound on  $H(X_{A_0}|E)_{\rho_{\alpha}}$ :

$$\begin{aligned} \min_{\{\rho_{ijk}, t, \vec{\varphi}\}} & H(X_{A_0}|E)_{\rho_{\alpha}}(\rho_{ijk}, t, \varphi_{A_0}) \\ \text{subject to} & \langle M_3 \rangle_{\rho_{\alpha}}(\rho_{ijk}, t, \vec{\varphi}) \geq m_{\alpha}; \quad \rho_{0jk} \geq \rho_{1jk}; \\ & \sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned} \quad (22)$$

where  $\vec{\varphi} := (\varphi_{A_0}, \varphi_{A_1}, \varphi_{B_0}, \varphi_{B_1}, \varphi_{C_0}, \varphi_{C_1})$  contains the measurement directions identified by the observables

$A_0, A_1, B_0, B_1, C_0$ , and  $C_1$ . Notably, due to the symmetries of the MABK inequality, all the tight lower bounds on  $H(X_{A_i}|E)_{\rho_\alpha}$ ,  $H(Y_{B_i}|E)_{\rho_\alpha}$ , and  $H(Z_{C_i}|E)_{\rho_\alpha}$  (for  $i = 0, 1$ ) coincide. Thus, solving Eq. (22) yields a tight lower bound on the conditional entropy  $H(X|E)_{\rho_\alpha}$  of any single party's outcome  $X$ .

We drastically simplify the optimization problem in Eq. (22), by replacing the MABK expectation value  $\langle M_3 \rangle_{\rho_\alpha}$  with its upper bound  $\mathcal{M}_\alpha^\dagger$  derived in Eq. (12). Indeed, this allows us to independently minimize  $H(X|E)_{\rho_\alpha}$  over  $t$  and  $\varphi_{A_0}$  without affecting the MABK violation. The resulting conditional entropy is minimized by  $t = \varphi_{A_0} = 0$  and reads

$$\begin{aligned} H(X|E)_{\rho_\alpha}(\rho_{ijk}, t = 0, \varphi_{A_0} = 0) \\ = 1 - H(\{\rho_{ijk}\}) + H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}), \end{aligned} \quad (23)$$

where the Shannon entropy of a probability distribution  $\{p_i\}_i$  is defined as  $H(\{p_i\}) = \sum_i -p_i \log_2 p_i$ .

We are thus left to solve the following optimization problem:

$$\begin{aligned} \min_{\{\rho_{ijk}\}} \quad & 1 - H(\{\rho_{ijk}\}) + H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}) \\ \text{subject to} \quad & \mathcal{M}_\alpha^\dagger(\rho_{ijk}) \geq m_\alpha; \quad \rho_{0jk} \geq \rho_{1jk}; \\ & \sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned} \quad (24)$$

whose solution is a lower bound on the solution of the original optimization problem, Eq. (22). We analytically solve Eq. (24) and provide the complete proof in Appendix E.

Importantly, the following family of states solves Eq. (24) for every value of the violation  $m_\alpha$ :

$$\tau(v_m) = v_m |\psi_{0,0,0}\rangle \langle \psi_{0,0,0}| + (1 - v_m) |\psi_{0,1,1}\rangle \langle \psi_{0,1,1}|, \quad (25)$$

where the parameter  $v_m$  is fixed by the violation  $m_\alpha$  by

$$m_\alpha = \mathcal{M}_\tau^\dagger(v_m) = 4\sqrt{2v_m^2 - 2v_m + 1}, \quad (26)$$

where we use Eq. (12) in the second equality. The lower bound on the conditional entropy  $H(X|E)_{\rho_\alpha}$  is thus given by the entropy of the states in Eq. (25):

$$H(X|E)_{\rho_\alpha} \geq F(m_\alpha) := H(X|E)_{\tau(v_m)}. \quad (27)$$

The entropy of the states in Eq. (25) is easily computed from Eq. (23) and can be expressed in terms of the violation  $m_\alpha$  by reverting Eq. (26). We obtain

$$F(m_\alpha) = 1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{m_\alpha^2}{8} - 1}\right), \quad (28)$$

where  $h(p) = -p \log_2 p - (1 - p) \log_2(1 - p)$  is the binary entropy. Finally, the lower bound, Eq. (28), is a convex

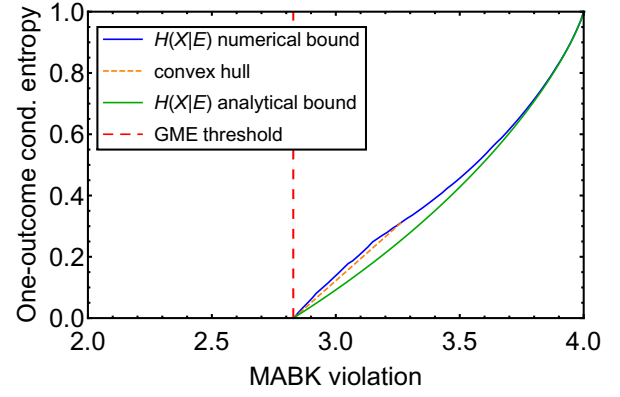


FIG. 2. Analytical lower bound on the conditional von Neumann entropy  $H(X|E_{\text{tot}})$  as a function of the MABK inequality violation [green line, Eq. (29)] observed by three parties. We compare it to the numerical optimization of Eq. (22) (blue line), whose convex hull (dashed orange) yields an upper limit on the lowest value of  $H(X|E_{\text{tot}})$ . We notice that Eve has no uncertainty on Alice's outcome  $X$  for violations below the genuine multipartite entanglement (GME) threshold.

function, hence we can employ it in Eq. (16) and obtain the desired lower bound on  $H(X|E_{\text{tot}})$  as a function of the observed MABK violation:

$$H(X|E_{\text{tot}}) \geq 1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{m^2}{8} - 1}\right). \quad (29)$$

In Fig. 2 we plot the lower bound on the conditional entropy derived in Eq. (29), as well as a numerical optimization of Eq. (22), which yields an upper bound on the minimal value of  $H(X|E)_{\rho_\alpha}$ . We can conclude that the tight lower bound on  $H(X|E_{\text{tot}})$  lies in the plot region delimited by the convex hull of the numerical curve [the bound in Eq. (15) must be convex] and our analytical lower bound.

From Fig. 2, we deduce that our analytical lower bound on  $H(X|E_{\text{tot}})$  leaves little room for improvement (compared to the ideal tight bound) and that it is actually tight up to the GME threshold of  $m = 2\sqrt{2}$ . We prove this by showing that the state  $\tau(1/2)$ , which yields the analytical bound at  $m = 2\sqrt{2}$ , is also an optimal solution of the original optimization problem, Eq. (22). Indeed, when  $m = 2\sqrt{2}$ , the tightness conditions of the MABK upper bound, Eq. (78), applied to  $\tau(1/2)$  are verified for  $\varphi_{A_0} = \varphi_{A_1} = 0$ . In other words, there exist observables that Alice, Bob, and Charlie can measure on  $\tau(1/2)$  such that  $\langle M_3 \rangle_{\tau(1/2)} = 2\sqrt{2}$ . In particular, Alice's optimal observables are the Pauli  $X$ . Under these conditions, the entropy in Eq. (22) reads  $H(X|E)_{\tau(1/2)}(\varphi_{A_0} = 0) = 0$ , which must be the solution of Eq. (22) since in general it holds  $H(X|E)_{\rho_\alpha} \geq 0$ . Thus the lower bound, Eq. (29), is tight for  $m = 2\sqrt{2}$  and is equal to zero.

By combining this with the fact that the tight lower bound on  $H(X|E_{\text{tot}})$  is monotonically nondecreasing in  $m$  by construction [see Eq. (22)], we deduce that the conditional entropy of a party's outcome is zero for every violation below the GME threshold of  $2\sqrt{2}$ . Hence, GME states are a necessary resource to guarantee private randomness of a party's outcome in a tripartite MABK scenario.

Notably, the claim on the necessity of GME can be generalized to an  $N$ -party MABK scenario. Consider the following family of states that generalizes Eq. (25) to  $N$  parties:

$$\tau(\nu) = \nu|\psi_{0,\bar{0}}\rangle\langle\psi_{0,\bar{0}}| + (1-\nu)|\psi_{0,\bar{1}}\rangle\langle\psi_{0,\bar{1}}|. \quad (30)$$

For  $\nu = 1/2$ , we have that the  $N$ -party MABK violation upper bound, Eq. (10), yields  $2\sqrt{2}$  and its tightness conditions (see Appendix D) are satisfied when Alice measures  $\pm X$  for both of her observables. With these settings, the  $N$ -party conditional entropy reads  $H(X|E)_{\tau(1/2)}(\varphi_{A_0} = 0, \pi) = 0$ . By repeating the argument on the monotonicity of the entropy, we deduce that GME is necessary to certify the privacy of a party's outcome in any MABK scenario.

Since private randomness of a party's outcome is a prerequisite of any DICKA protocol, it is an open question whether GME is a necessary ingredient for DICKA. Note, instead, that GME has been shown not to be necessary for device-dependent CKA [71]. Besides, in Sec. V we argue on the apparent incompatibility of full-correlator Bell inequalities and DICKA protocols.

Finally, we mention that a lower bound on  $H(X|E_{\text{tot}})$  as a function of the MABK inequality violation is also derived in Ref. [62], for the general  $N$ -party scenario. The conditional entropy bound obtained in Ref. [62] reads

$$H(X|E_{\text{tot}}) \geq 1 - h\left(\frac{1}{2} + \frac{1}{2}\sqrt{\frac{m^2}{2^N} - 1}\right), \quad (31)$$

where  $m$  is the observed violation of the  $N$ -partite MABK inequality, Eq. (9). Surprisingly, despite the fact that the bound in Ref. [62] is derived with a completely different approach without aiming at optimality, the lower bound, Eq. (31), for  $N = 3$  coincides with the bound, Eq. (29), obtained in this work.

## V. FULL-CORRELATOR BELL INEQUALITIES AND DICKA

We provide an heuristic argument on why full-correlator Bell inequalities with two dichotomic observables per party, such as the MABK inequality, seem to be useless for DICKA protocols. We hope that this fundamental question can spark the interest of the community towards more conclusive results.

Any DICKA protocol is characterized by two essential ingredients: a violation of a multipartite Bell inequality

to ensure secrecy of Alice's outcomes and correlated outcomes among all the parties yielding the conference key. Since a part of Alice's outcomes form the secret key, one of the measurements she uses to assess the violation of the inequality must be the same used for key generation [34,43,72]. Note that, unlike Alice, the other parties are equipped with an additional measurement option solely used for key generation.

It is known that every full-correlator Bell inequality with two dichotomic observables per party is maximally violated by the GHZ state [56]. Moreover, the only multiqubit state leading to perfectly correlated and random outcomes among all the parties is the GHZ state, when the parties measure in the  $Z$  basis [21].

However, a GHZ state maximally violates a full-correlator Bell inequality when the measurements are chosen such that the resulting inequality (modulo rearrangements) is only composed of expectation values of GHZ stabilizers, which acquire the extremal value 1. Moreover, the stabilizers appearing in the inequality do not act trivially on any qubit—i.e., do not contain the identity—due to the full-correlator structure of the inequality. We call such stabilizers “full stabilizers” for ease of comprehension.

The problem is that none of the  $N$ -partite GHZ state full stabilizers, for  $N$  odd, contains the  $Z$  operator [73]. This implies that, in order to maximally violate the inequality, Alice's measurement directions are orthogonal to  $Z$ . Since one of these measurements is also used to generate her raw key, she would obtain totally uncorrelated outcomes with the rest of the parties (perfect correlations are only obtained with a GHZ state when measuring  $Z$ ). This causes the unwanted situation of having maximal violation and perfect correlations among the parties' key bits as mutually exclusive conditions. Since both conditions are required in a DICKA protocol, the above argument constitutes an initial evidence that full-correlator Bell inequalities are not suited for DICKA protocols.

A similar argument holds when the number of parties  $N$  is even ( $N > 2$ ). As a matter of fact, in this case there exists only one GHZ full stabilizer, which contains the  $Z$  operator, namely  $Z^{\otimes N}$ . If  $\langle Z^{\otimes N} \rangle$  were to appear in the rearranged inequality expression, there should be at least another correlator containing at least one  $Z$  operator. Indeed, if each observable in a correlator never appears again in any other term of the inequality, that correlator is useless since Eve could assign to it any value (Eve is supposed to know the inequality being tested). The lack of any other full stabilizer containing the  $Z$  operator prevents having a second correlator containing  $Z$ , thus excluding the term  $\langle Z^{\otimes N} \rangle$  in the first place. Therefore, also in the  $N$ -even case Alice's measurements leading to maximal violation are orthogonal to  $Z$ , yielding uncorrelated raw-key bits. We remark that the  $N = 2$  case is peculiar since the low number of parties allows  $\langle ZZ \rangle$  [obtained from the term

$\langle A_1(B_0 - B_1) \rangle$  in the inequality] to appear just once in the CHSH inequality [55].

It is worth mentioning that in Ref. [72] the apparent incompatibility of the MABK inequality with a DICKA protocol was already discussed. In particular, it is shown in the tripartite case that there exists no honest implementation such that the parties' outcomes are perfectly correlated and at the same time the MABK inequality is violated above the GME threshold, which is a necessary condition as we pointed out above.

Despite the concerns on the use of MABK inequalities in DICKA protocols, the results of this paper are still of fundamental interest for DIRG [44–50] based on multi-party nonlocality. As a further application, in the following we improve the bound on Eve's uncertainty of Alice and Bob's outcomes derived in Ref. [50].

## VI. TWO-OUTCOME CONDITIONAL ENTROPY BOUND

Consider the same DI scenario of Fig. 1 and suppose that Eve wishes to jointly guess the measurement outcomes  $X$  and  $Y$  of Alice and Bob, respectively. This scenario may occur in DIRG protocols where the parties are assumed to be colocated and collaborate to generate global secret randomness [50,69]. We estimate Eve's uncertainty by providing a lower bound on the conditional von Neumann entropy  $H(XY|E)_{\rho_\alpha}$ , as a function of the MABK violation  $m_\alpha$ . The entropy is computed on the following quantum state:

$$\rho_{XYE}^\alpha = (\mathcal{E}_X \otimes \mathcal{E}_Y \otimes \text{id}_E) \text{Tr}_C[|\phi_{ABCE}^\alpha\rangle\langle\phi_{ABCE}^\alpha|], \quad (32)$$

where the maps  $\mathcal{E}_X$  and  $\mathcal{E}_Y$  represent Alice's and Bob's measurements, respectively, defined by the eigenvectors:

$$\begin{aligned} |a\rangle_X &= \frac{1}{\sqrt{2}}[|0\rangle + (-1)^a e^{i\varphi_X} |1\rangle] \quad a \in \{0, 1\}, \\ |b\rangle_Y &= \frac{1}{\sqrt{2}}[|0\rangle + (-1)^b e^{i\varphi_Y} |1\rangle] \quad b \in \{0, 1\}. \end{aligned} \quad (33)$$

For definiteness, we select  $\varphi_X = \varphi_{A_0}$  and  $\varphi_Y = \varphi_{B_0}$  and define the optimization problem:

$$\begin{aligned} \min_{\{\rho_{ijk,t,\vec{\varphi}}\}} & H(X_{A_0} Y_{B_0} | E)_{\rho_\alpha}(\rho_{ijk}, t, \varphi_{A_0}, \varphi_{B_0}) \\ \text{subject to} & \langle M_3 \rangle_{\rho_\alpha}(\rho_{ijk}, t, \vec{\varphi}) \geq m_\alpha; \quad \rho_{0jk} \geq \rho_{1jk}; \quad (34) \\ & \sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned}$$

whose solution yields a tight lower bound on  $H(X_{A_0} Y_{B_0} | E)_{\rho_\alpha}$ . Nonetheless, due to the MABK symmetries, the lower bounds on  $H(X_{A_i} Y_{B_j} | E)_{\rho_\alpha}$ ,  $H(X_{A_i} Z_{C_j} | E)_{\rho_\alpha}$ , and  $H(Y_{B_i} Z_{C_j} | E)_{\rho_\alpha}$  coincide (for  $i, j \in \{0, 1\}$ ). Thus, the solution of Eq. (34) actually provides the tight lower bound

on the conditional entropy  $H(XY|E)_{\rho_\alpha}$  of any pair of outcomes  $X$  and  $Y$  belonging to distinct parties.

Similarly to the case of  $H(X|E)_{\rho_\alpha}$ , we analytically solve the following simplified optimization problem (details in Appendix F):

$$\begin{aligned} \min_{\{\rho_{ijk,t,\varphi_X,\varphi_Y}\}} & H(XY|E)_{\rho_\alpha}(\rho_{ijk}, t, \varphi_X, \varphi_Y) \\ \text{subject to} & \mathcal{M}_\alpha^\dagger(\rho_{ijk}) \geq m_\alpha; \quad \rho_{0jk} \geq \rho_{1jk}; \quad (35) \\ & \sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned}$$

which yields a lower bound on the solution of the original optimization problem, Eq. (34). The lower bound on  $H(XY|E)_{\rho_\alpha}$  obtained by solving Eq. (35) reads

$$H(XY|E)_{\rho_\alpha} \geq G(m_\alpha), \quad (36)$$

where

$$G(m_\alpha) := 2 - H[\{1 - 3f(m_\alpha), f(m_\alpha), f(m_\alpha), f(m_\alpha)\}], \quad (37)$$

and where the function  $f$  is defined as

$$f(m_\alpha) = \frac{1}{4} - \frac{\sqrt{3}}{24} \sqrt{m_\alpha^2 - 4}. \quad (38)$$

Similarly to the case of  $H(X|E)_{\rho_\alpha}$ , we can exploit the convexity of the function in Eq. (37) to lower bound the conditional entropy of the global state prepared by Eve:

$$H(XY|E_{\text{tot}}) \geq G(m), \quad (39)$$

where  $m$  is the violation observed by Alice, Bob, and Charlie and  $G(m)$  is the function defined in Eq. (37).

The bound in Eq. (39) is plotted in Fig. 3, together with the tight lower bound on the correspondent min-entropy obtained in Ref. [50] and a numerical optimization of Eq. (34). As already mentioned in Sec. IV, the tight bound on  $H(XY|E_{\text{tot}})$  must lie between the convex hull of the numerical curve and our analytical bound, Eq. (39). Figure 3 suggests that our analytical bound is close to the ideal tight bound.

We point out the dramatic improvement in certifying device independently the privacy of two parties' outcomes with our lower bound on the conditional von Neumann entropy  $H(XY|E_{\text{tot}})$ , as opposed to bounding the conditional min-entropy  $H_{\min}(XY|E_{\text{tot}})$  [50].

The min-entropy is often used to lower bound the von-Neumann entropy in DI protocols, since it can be directly estimated using the statistics of the measurement outcomes [51,52]. In general it holds that  $H \geq H_{\min}$  [74]. However, bounding the von Neumann entropy with the min-entropy



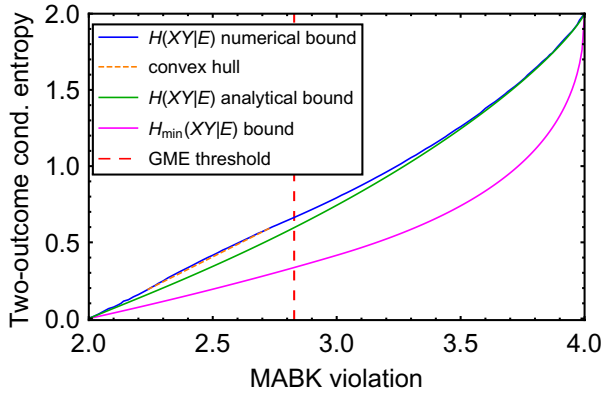


FIG. 3. Analytical lower bound on the conditional von Neumann entropy  $H(XY|E_{\text{tot}})$  [green line, Eq. (39)] as a function of the MABK violation observed by three parties. We compare it to the lower bound on the conditional min-entropy  $H_{\min}(XY|E_{\text{tot}})$  derived in Ref. [50] (magenta line) and to the numerical solution of Eq. (34) (blue line), whose convex hull (dashed orange) yields an upper limit on the lowest value of  $H(XY|E_{\text{tot}})$ . Our bound dramatically improves the one in Ref. [50] since it directly bounds the von Neumann entropy. Unlike the case of  $H(X|E_{\text{tot}})$  in Fig. 2, Eve’s uncertainty on outcomes  $X$  and  $Y$  is nonzero even for violations below the GME threshold.

can be far from optimal, as in the case analyzed here (see Fig. 3).

From Fig. 3 we also observe that the joint conditional entropy of two parties’ outcomes  $H(XY|E_{\text{tot}})$  is nonzero for violations below the GME threshold, unlike the entropy of a single party’s outcome  $H(X|E_{\text{tot}})$  (cf. Sec. IV).

## VII. CONCLUSION

The security of DI cryptographic protocols is based on the ability to bound the entropy of the protocols’ outcomes, conditioned on the eavesdropper’s knowledge, by a Bell inequality violation. To this aim, we consider a DI scenario where  $N$  parties test a generic full-correlator Bell inequality, with two measurement settings and two outcomes per party. We prove, in this context, that it is not restrictive to reduce the most general quantum state tested by the parties to simple  $N$ -qubit states. Our result reduces to the only other one of this kind [34] when  $N = 2$ .

In order to obtain the entropy bounds, we prove an analytical upper bound on the maximal violation of the MABK inequality achieved by a given  $N$ -qubit state, when the parties perform rank-one projective measurements. The bound is tight on certain classes of states and has general validity (i.e., independent of the parties’ measurements) for states whose maximal violation is above the GME threshold. Our bound generalizes the known result [60] valid for the CHSH inequality to an arbitrary number of parties. This is the first bound on the maximal violation of a  $N$ -partite Bell inequality achievable by a given state, expressed in terms of the state’s parameters.

These results enabled us to derive an analytical lower bound on the conditional von Neumann entropy of a party’s outcome, when Alice, Bob, and Charlie test the tripartite MABK inequality. We also derive an analytical lower bound on the conditional von Neumann entropy of any pair of outcomes from distinct parties, which dramatically improves a similar estimation made in Ref. [50] in terms of the corresponding min-entropy. The improvement gained by directly bounding the von Neumann entropy has direct implications for randomness generation protocols, inasmuch as it increases the fraction of random bits guaranteed to be private.

Moreover, both analytical bounds perform well when compared to the numerical estimation of the corresponding tight bounds, leaving little room for improvement.

By proving that our bound on the conditional entropy of a party’s outcome is tight at the GME threshold, we deduce that genuine multipartite entanglement (GME) is necessary to guarantee the privacy of a party’s random outcome in any device-independent scenario based on the MABK inequality. It is an open question whether GME is a fundamental requirement for DI conference-key agreement (DICKA). In this regard, we heuristically argue that full-correlator Bell inequalities with two binary observables per party, such as the MABK inequality, are unlikely to be employed in any DICKA protocol. We envision further and more conclusive results in this direction from the scientific community interested in this topic.

The bounds on the conditional entropies derived in this work can find potential application in DI randomness generation based on multipartite nonlocality. Depending on the application, such protocols would generate local randomness for one party or global randomness for two or more parties. In all cases, the privacy of the generated random data would be ensured by entropy bounds like the ones we derive.

Furthermore, the techniques developed in proving Theorem 1 can inspire analogous analytical reductions of the quantum state for other Bell inequalities. Indeed, of particular interest are the Bell inequalities employed in the existing DICKA protocols [42,43], for which a result like Theorem 1 would be the first step towards a tight security analysis, which is still lacking.

## VIII. METHODS

Here we present the proofs of Theorems 1 and 2.

### A. Proof of Theorem 1

The proof of Theorem 1 is based on three main ingredients: (i) the fact that each party has only two inputs with two outputs allows the reduction of the analysis to qubits and rank-one projective measurements; (ii) the symmetries of the MABK inequality allow us to set all the marginals to

zero, without changing the MABK violation or the information available to the eavesdropper; (iii) the freedom in the definition of the local axes is used to further reduce the number of free parameters. Our proof is inspired by a similar proof given in Ref. [34]. However, our result is valid for an arbitrary number of parties  $N$  in the generic  $(N, 2, 2)$  DI scenario described in the main text. Notably, for  $N = 2$  we recover the result of Ref. [34].

In order to prove Theorem 1, we make use of the following Lemma 1, which is a consequence of a result given in Ref. [75] and whose proof is reported in Appendix A.

**Lemma 1.** *Let  $\{P_0, P_1\}$  and  $\{Q_0, Q_1\}$  be two projective measurements acting on a Hilbert space  $\mathcal{H}$ , such that  $P_0, P_1, Q_0$ , and  $Q_1$  are projectors and  $P_0 + P_1 = \text{id}$  and  $Q_0 + Q_1 = \text{id}$ . There exists an orthonormal basis in an enlarged Hilbert space  $\mathcal{H}^*$  such that the four projectors are simultaneously block diagonal, in blocks of size  $2 \times 2$ . Moreover, within a  $2 \times 2$  block, each projector has rank one.*

*Proof.* The first step consists in reducing the state distributed by Eve to a convex combination of  $N$ -qubit states. To start with, every generalized measurement (POVM) can be viewed as a projective measurement in a larger Hilbert space. Since we did not fix the Hilbert space to which the shared quantum state belongs, we can assume without loss of generality that the parties' measurements are binary projective measurements on a given Hilbert space  $\mathcal{H}$ . In particular, the projectors  $P_0^{(i)}$  and  $P_1^{(i)}$  ( $Q_0^{(i)}$  and  $Q_1^{(i)}$ ) correspond to Alice <sub>$i$</sub> 's binary observable  $A_0^{(i)}$  ( $A_1^{(i)}$ ) relative to input  $x_i = 0$  ( $x_i = 1$ ).

Now we can apply Lemma 1 to the projective measurements of Alice <sub>$i$</sub>  for  $i = 1, \dots, N$  and state that, at every round of the protocol, the Hilbert space on which, e.g., Alice<sub>1</sub>'s measurements are acting is decomposed as

$$\mathcal{H}^* = \bigoplus_{\alpha} \mathcal{H}_{\alpha}^2, \quad (40)$$

where every subspace  $\mathcal{H}_{\alpha}^2$  is two-dimensional and both Alice<sub>1</sub>'s measurements act within  $\mathcal{H}_{\alpha}^2$  as rank-one projective measurements. From Alice<sub>1</sub>'s point of view, the measurement process consists of a projection in one of the two-dimensional subspaces followed by a projective measurement in that subspace (selected according to Alice<sub>1</sub>'s input). Therefore, Eve is effectively distributing to Alice<sub>1</sub> a direct sum of qubits at every round. Alice<sub>1</sub>'s measurement then selects one of the qubit subspaces and performs a projective measurement within that subspace. Of course, since Eve fabricates the measurement device, the projective measurements occurring in every subspace can be predefined by Eve. Since this argument holds for every party, Eve is effectively distributing a direct sum of  $N$ -qubit states in each round.

Certainly, it cannot be worse for Eve to learn the flag  $\alpha$  of the subspace selected in a particular round before sending the direct sum of  $N$ -qubit states to the parties. For this reason, we can reformulate the state preparation and measurement in a generic round of the protocol as Eve preparing a mixture

$$\rho_{A_1 \dots A_N \Xi} = \sum_{\alpha} p_{\alpha} \rho_{\alpha} \bigotimes_{i=1}^N |\alpha\rangle\langle\alpha|_{\xi_i} \quad (41)$$

of  $N$ -qubit states  $\rho_{\alpha}$ , together with a set of ancillae  $\Xi := \{\xi_i\}_{i=1}^N$  (known to her), which fixes the rank-one projective measurements that each party can select on  $\rho_{\alpha}$ .

Let us now focus on one specific occurrence defined by a given  $\alpha$ , i.e., on one of the  $N$ -qubit states  $\rho_{\alpha}$ . For ease of notation, in the following we omit the symbol  $\alpha$ .

We define the plane induced by the two rank-one projective measurements of each party to be the  $(x, y)$  plane of the Bloch sphere. Now, we assume without loss of generality that the statistics observed by the parties is such that every marginal is random:

$$\left\langle \prod_{i \in P} A^{(i)} \right\rangle = 0, \quad (42)$$

where  $A^{(i)}$  is any dichotomic observable of Alice <sub>$i$</sub>  and  $P$  is any nonempty strict subset of all the parties:  $P \subsetneq \{1, \dots, N\}$ . Indeed, if this is not the case, the parties can perform the following classical procedure on their outcomes, which enforces the requirement in Eq. (42): "Alice<sub>1</sub> and Alice <sub>$i$</sub>  flip their outcome with probability  $\frac{1}{2}$ ," repeated for every  $i = 2, \dots, N$ . This procedure does not change the observed Bell violation since an even number of flips occurs at every time, thus leaving the correlators, Eq. (1), composing the Bell inequality unchanged. Moreover, it requires classical communication between the parties, which we assume to be known by Eve.

Since the observed statistics always satisfies Eq. (42), we can imagine that it is Eve herself who performs the classical flipping on the outputs in place of the parties. To this aim, Eve could apply the following map to the state  $\rho$  she prepared, before distributing it:

$$\rho \mapsto \bar{\rho} = \circ_{i=2}^N \mathcal{D}_i(\rho), \quad (43)$$

where the composition operator in Eq. (43) represents the successive application of the following operations:

$$\mathcal{D}_i(\rho) = \frac{1}{2}\rho + \frac{1}{2}Z_i Z_i \rho Z_i^{\dagger} Z_i^{\dagger}, \quad (44)$$

with  $Z_i$  representing the third Pauli operator applied on Alice <sub>$i$</sub> 's qubit. Note that the application of  $Z$  prior to measurement flips the outcome of a measurement in the  $(x, y)$

plane. Thus, by applying the map in Eq. (43), Eve is distributing a state, which automatically satisfies Condition (42). We can safely assume that Eve implements the map in Eq. (43) since this is not disadvantageous to her. As a matter of fact, her uncertainty on the parties' outcomes, quantified by the conditional von Neumann entropy, does not increase when she sends the state  $\bar{\rho}$  instead of  $\rho$ . We provide a detailed proof of this fact in Appendix B. Therefore, it is not restrictive to assume that the parties receive the state, Eq. (43), from Eve, which can be recast as

$$\bar{\rho} = \frac{1}{2^{N-1}} \sum_{n=0}^{\lfloor N/2 \rfloor} \sum_{\mathbf{x} \in I(n)} Z^{\mathbf{x}} \rho Z^{\mathbf{x}}, \quad (45)$$

with

$$I(n) = \{\mathbf{x} \in \{0,1\}^N : \omega(\mathbf{x}) = 2n\}, \quad (46)$$

$$Z^{\mathbf{x}} = \bigotimes_{j=1}^N Z_j^{x_j}, \quad (47)$$

where the Hamming weight  $\omega(\mathbf{x})$  of a bit string  $\mathbf{x}$  returns the total number of bits that are equal to one and  $\lfloor y \rfloor$  returns the greatest integer smaller or equal to  $y$ .

By expressing the initial generic state  $\rho$  in the GHZ basis:

$$\rho = \sum_{\vec{u}, \vec{v} \in \{0,1\}^{N-1}} \sum_{\sigma, \tau=0}^1 \rho_{(\sigma\vec{u})(\tau\vec{v})} |\psi_{\sigma, \vec{u}}\rangle \langle \psi_{\tau, \vec{v}}|, \quad (48)$$

where  $\rho_{(\sigma\vec{u})(\tau\vec{v})} \in \mathbb{C}$  and by substituting it into Eq. (45), we notice that the state  $\bar{\rho}$  is greatly simplified in the GHZ basis. In particular, all the coherences between states of the GHZ basis relative to different vectors  $\vec{u}$  are null:

$$\bar{\rho} = \sum_{\vec{u} \in \{0,1\}^{N-1}} \sum_{\sigma, \tau=0}^1 \rho_{(\sigma\vec{u})(\tau\vec{u})} |\psi_{\sigma, \vec{u}}\rangle \langle \psi_{\tau, \vec{u}}|. \quad (49)$$

This means that the matrix representation of  $\bar{\rho}$  is block diagonal in the GHZ basis. By relabeling the nonzero matrix coefficients, we represent  $\bar{\rho}$  as follows:

$$\bar{\rho} = \bigoplus_{\vec{u} \in \{0,1\}^{N-1}} \begin{bmatrix} \lambda_{0\vec{u}} & r_{\vec{u}} + \mathfrak{i}s_{\vec{u}} \\ r_{\vec{u}} - \mathfrak{i}s_{\vec{u}} & \lambda_{1\vec{u}} \end{bmatrix}, \quad (50)$$

where  $\lambda_{j\vec{u}}$ ,  $r_{\vec{u}}$ , and  $s_{\vec{u}}$  are real numbers. The number of free parameters characterizing Eq. (50) can be further reduced by exploiting the remaining degrees of freedom in the parties' local reference frames [34]. Indeed, although we identify the plane containing the measurement directions

to be the  $(x, y)$  plane for every party, they can still choose the orientation of the axes by applying rotations  $R(\theta)$  along the  $z$  direction. Consequently, the state distributed by Eve without loss of generality is given by

$$\bar{\rho}_+ = \bigotimes_{i=1}^N R_i(\theta_i) \bar{\rho} \bigotimes_{i=1}^N R_i^\dagger(\theta_i), \quad (51)$$

where the rotation  $R_i(\theta_i)$  acts on the Hilbert space of party number  $i$  and reads

$$R_i(\theta_i) = \cos \frac{\theta_i}{2} \text{id} + \mathfrak{i} \sin \frac{\theta_i}{2} Z_i, \quad (52)$$

where “id” is the identity operator. Similarly to  $\bar{\rho}$ , even the global rotation operator is block diagonal in the GHZ basis:

$$\bigotimes_{i=1}^N R_i(\theta_i) = \bigoplus_{\vec{u} \in \{0,1\}^{N-1}} \begin{bmatrix} \cos \frac{\beta(\vec{\theta}, \vec{u})}{2} & \mathfrak{i} \sin \frac{\beta(\vec{\theta}, \vec{u})}{2} \\ \mathfrak{i} \sin \frac{\beta(\vec{\theta}, \vec{u})}{2} & \cos \frac{\beta(\vec{\theta}, \vec{u})}{2} \end{bmatrix}, \quad (53)$$

where  $\vec{\theta}$  is the vector defined by the rotation angles  $\{\theta_1, \dots, \theta_N\}$  and  $\beta$  is a function of  $\theta$  and  $\vec{u}$  defined as

$$\beta(\vec{\theta}, \vec{u}) = \theta_1 + \sum_{j=1}^{N-1} (-1)^{u_j} \theta_{j+1}. \quad (54)$$

This fact greatly simplifies the calculation in Eq. (51), as it allows to multiply the matrices, Eqs. (50) and (53), block by block. The resulting block-diagonal matrix representing the state distributed by Eve reads

$$\bar{\rho}_+ = \bigoplus_{\vec{u} \in \{0,1\}^{N-1}} \begin{bmatrix} \lambda'_{0\vec{u}} & r_{\vec{u}} + \mathfrak{i}s'_{\vec{u}} \\ r_{\vec{u}} - \mathfrak{i}s'_{\vec{u}} & \lambda'_{1\vec{u}} \end{bmatrix}, \quad (55)$$

where the new matrix coefficients are given by

$$\lambda'_{0\vec{u}} = \frac{1}{2} \left[ \lambda_{0\vec{u}} + \lambda_{1\vec{u}} + (\lambda_{0\vec{u}} - \lambda_{1\vec{u}}) \cos \beta(\vec{\theta}, \vec{u}) + 2s_{\vec{u}} \sin \beta(\vec{\theta}, \vec{u}) \right] \quad (56)$$

$$s'_{\vec{u}} = s_{\vec{u}} \cos \beta(\vec{\theta}, \vec{u}) - \frac{1}{2} (\lambda_{0\vec{u}} - \lambda_{1\vec{u}}) \sin \beta(\vec{\theta}, \vec{u}) \quad (57)$$

$$\lambda'_{1\vec{u}} = \frac{1}{2} \left[ \lambda_{0\vec{u}} + \lambda_{1\vec{u}} - (\lambda_{0\vec{u}} - \lambda_{1\vec{u}}) \cos \beta(\vec{\theta}, \vec{u}) - 2s_{\vec{u}} \sin \beta(\vec{\theta}, \vec{u}) \right]. \quad (58)$$

From Eq. (57) we deduce that choosing the rotation angles  $\theta_1, \dots, \theta_N$  such that the following linear constraint is

verified:

$$\theta_1 + \sum_{j=1}^{N-1} (-1)^{u_j} \theta_{j+1} = \arctan \frac{2s_{\vec{u}}}{\lambda_{0\vec{u}} - \lambda_{1\vec{u}}}, \quad (59)$$

sets the corresponding imaginary part in Eq. (55) to zero:  $s'_{\vec{u}} = 0$ . However, we can only impose  $N$  constraints like Eq. (59) on the  $N$  rotation angles, thus we are able to arbitrarily set to zero  $N$  terms like  $s_{\vec{u}}$  in Eq. (55). Moreover, by applying further rotations (note that the composition of rotations is still a rotation) such that

$$\tilde{\theta}_1 + \sum_{j=1}^{N-1} (-1)^{u_j} \tilde{\theta}_{j+1} = \pi, \quad (60)$$

we can exchange the diagonal terms in Eq. (55):  $\lambda'_{0\vec{u}} = \lambda_{1\vec{u}}$  and  $\lambda'_{1\vec{u}} = \lambda_{0\vec{u}}$ . This allows us to order up to  $N$  pairs  $(\lambda_{0\vec{u}}, \lambda_{1\vec{u}})$ , for the same argument as above. Note that the blocks with ordered pairs must be the same blocks with null imaginary parts. Indeed, if a block identified by  $\vec{u}$  with null imaginary part undergoes a rotation such that  $\beta(\vec{\theta}, \vec{u}) \neq \{0, \pm\pi\}$ , it will acquire a nonzero imaginary part  $s'_{\vec{u}} = -(1/2)(\lambda_{0\vec{u}} - \lambda_{1\vec{u}}) \sin \beta(\vec{\theta}, \vec{u})$ , (see Eq. (57)).

Finally, we construct the state  $\bar{\rho}_-$  starting from  $\bar{\rho}_+$  given in Eq. (55) by replacing  $r_{\vec{u}}$  with  $-r_{\vec{u}}$ :

$$\bar{\rho}_- = \bigoplus_{\vec{u} \in \{0,1\}^{N-1}} \begin{bmatrix} \lambda'_{0\vec{u}} & -r_{\vec{u}} + \mathfrak{i}s'_{\vec{u}} \\ -r_{\vec{u}} - \mathfrak{i}s'_{\vec{u}} & \lambda'_{1\vec{u}} \end{bmatrix}. \quad (61)$$

We observe that the two states  $\bar{\rho}_{\pm}$  yield the same measurement statistics and provide Eve with the same information—i.e., their conditional entropies coincide. Additionally, it is not disadvantageous for Eve to prepare a balanced mixture of  $\bar{\rho}_+$  and  $\bar{\rho}_-$  given by  $(\bar{\rho}_+ + \bar{\rho}_-)/2$ , rather than preparing one of the two states with certainty. A detailed proof of these observations is given in Appendix C.

We conclude that it is not restrictive to assume that Eve distributes to the parties a mixture of  $N$ -qubit states  $\rho_{\alpha}$  together with an ancillary system fixing the parties' measurements. Each state  $\rho_{\alpha}$  is represented by the following block diagonal matrix in the GHZ basis:

$$\rho_{\alpha} = \frac{\bar{\rho}_+ + \bar{\rho}_-}{2} = \bigoplus_{\vec{u} \in \{0,1\}^{N-1}} \begin{bmatrix} \lambda_{0\vec{u}} & \mathfrak{i}s_{\vec{u}} \\ -\mathfrak{i}s_{\vec{u}} & \lambda_{1\vec{u}} \end{bmatrix}, \quad (62)$$

where the diagonal elements of  $N$  arbitrary blocks are ordered and the corresponding off-diagonal elements are zero. This concludes the proof. ■

## B. Proof of Theorem 2

We present the proof of Theorem 2, which generalizes the analogous result valid in the bipartite case for the CHSH inequality [60]. This is the only existing upper bound on the violation of the  $N$ -partite MABK inequality by rank-one projective measurements on an arbitrary  $N$ -qubit state, expressed as a function of the state's parameters. Note that an analogous upper bound on the violation of the tripartite MABK inequality was recently derived in Ref. [63]. However, here we show that our bound is tight on a broader class of states and valid for an arbitrary number of parties. In order to prove Theorem 2 we make use of the following Lemma 2, which generalizes an analogous result in Ref. [60] to rectangular matrices of arbitrary dimensions. The proof of Lemma 2 is reported in Appendix D.

**Lemma 2.** *Let  $Q$  be an  $m \times n$  real matrix and let  $\|\vec{v}\|$  be the Euclidean norm of vectors  $\vec{v} \in \mathbb{R}^k$ , for  $k = m, n$ . Finally, let “ $\cdot$ ” indicate both the scalar product and the matrix-vector multiplication. Then*

$$\max_{\substack{\vec{c}, \vec{c}' \text{ such that} \\ \|\vec{c}\| = \|\vec{c}'\| = 1}} [\|Q \cdot \vec{c}\|^2 + \|Q \cdot \vec{c}'\|^2] = u_1 + u_2, \quad (63)$$

where  $u_1$  and  $u_2$  are the largest and second-to-the-largest eigenvalues of  $U \equiv Q^T Q$ , respectively.

For illustration purposes, here we report the proof of Theorem 2 for the case of  $N = 3$  parties. The full proof is given in Appendix D.

*Proof.* By assumption we restrict the description of the parties' observables to rank-one projective measurements on their respective qubit [56]. Hence they can be represented as follows:

$$A_x = \vec{a}_x \cdot \vec{\sigma}, \quad B_y = \vec{b}_y \cdot \vec{\sigma}, \quad \text{and} \quad C_z = \vec{c}_z \cdot \vec{\sigma}, \quad (64)$$

where  $\vec{a}_x, \vec{b}_y, \vec{c}_z$  are unit vectors in  $\mathbb{R}^3$  and where  $\sigma_1 = X, \sigma_2 = Y$ , and  $\sigma_3 = Z$ . We can then express the tripartite MABK operator, Eq. (8), as follows:

$$M_3 = \sum_{i,j,k=1}^3 M_{ijk} \sigma_i \otimes \sigma_j \otimes \sigma_k, \quad (65)$$

where we define

$$M_{ijk} \equiv a_0 b_j c_k + a_0 b_j c_k + a_1 b_j c_k - a_1 b_j c_k. \quad (66)$$

A generic three-qubit state can be expressed in the Pauli basis as follows:

$$\rho = \frac{1}{8} \sum_{\mu, \nu, \gamma=0}^3 \Lambda_{\mu\nu\gamma} \sigma_{\mu} \otimes \sigma_{\nu} \otimes \sigma_{\gamma}, \quad (67)$$



with  $\Lambda_{\mu\nu\gamma} = \text{Tr}[\rho\sigma_\mu \otimes \sigma_\nu \otimes \sigma_\gamma]$  and  $\sigma_0 = \text{id}$ . With the MABK operator in Eq. (65), the MABK expectation value on the generic three-qubit state in Eq. (67) is given by

$$\begin{aligned} \langle M_3 \rangle_\rho &= \text{Tr}(M_3\rho) \\ &= \frac{1}{8} \sum_{i,j,k=1}^3 \sum_{\mu,\nu,\gamma=0}^3 M_{ijk} \Lambda_{\mu\nu\gamma} \underbrace{\text{Tr}(\sigma_i\sigma_\mu \otimes \sigma_j\sigma_\nu \otimes \sigma_k\sigma_\gamma)}_{8\delta_{i,\mu}\delta_{j,\nu}\delta_{k,\gamma}} \\ &= \sum_{i,j,k=1}^3 M_{ijk} \Lambda_{ijk}. \end{aligned} \quad (68)$$

By recalling the correlation matrix of a tripartite state (cf. Definition 3), the MABK expectation value in Eq. (68) can be recast as follows:

$$\begin{aligned} \langle M_3 \rangle_\rho &= (\vec{a}_0 \otimes \vec{b}_1 + \vec{a}_1 \otimes \vec{b}_0)^T \cdot T_\rho \cdot \vec{c}_0 \\ &\quad + (\vec{a}_0 \otimes \vec{b}_0 - \vec{a}_1 \otimes \vec{b}_1)^T \cdot T_\rho \cdot \vec{c}_1. \end{aligned} \quad (69)$$

Finally, the maximum violation  $\mathcal{M}_\rho$  of the MABK inequality achieved by an arbitrary three-qubit state is obtained by optimizing Eq. (69) over all possible observables that the parties can choose to measure:

$$\begin{aligned} \mathcal{M}_\rho &= \max_{\substack{\vec{a}_i, \vec{b}_i, \vec{c}_i \text{ such that} \\ \|\vec{a}_i\| = \|\vec{b}_i\| = \|\vec{c}_i\| = 1}} (\vec{a}_0 \otimes \vec{b}_1 + \vec{a}_1 \otimes \vec{b}_0)^T \cdot T_\rho \cdot \vec{c}_0 \\ &\quad + (\vec{a}_0 \otimes \vec{b}_0 - \vec{a}_1 \otimes \vec{b}_1)^T \cdot T_\rho \cdot \vec{c}_1. \end{aligned} \quad (70)$$

Let us now evaluate the norm of the composite vectors in Eq. (70):

$$\begin{aligned} \left\| \vec{a}_0 \otimes \vec{b}_1 + \vec{a}_1 \otimes \vec{b}_0 \right\|^2 &= 2 + 2 \underbrace{\cos \theta_a \cos \theta_b}_{\equiv \cos \theta_{ab}} \\ &= 4 \cos^2 \left( \frac{\theta_{ab}}{2} \right), \end{aligned} \quad (71)$$

where  $\theta_a$  ( $\theta_b$ ) is the angle between vectors  $\vec{a}_0$  and  $\vec{a}_1$  ( $\vec{b}_0$  and  $\vec{b}_1$ ). Similarly,

$$\left\| \vec{a}_0 \otimes \vec{b}_0 - \vec{a}_1 \otimes \vec{b}_1 \right\|^2 = 4 \sin^2 \left( \frac{\theta_{ab}}{2} \right). \quad (72)$$

We then define normalized vectors  $\vec{v}_0$  and  $\vec{v}_1$  such that

$$\vec{a}_0 \otimes \vec{b}_1 + \vec{a}_1 \otimes \vec{b}_0 = 2 \cos \left( \frac{\theta_{ab}}{2} \right) \vec{v}_0, \quad (73)$$

$$\vec{a}_0 \otimes \vec{b}_0 - \vec{a}_1 \otimes \vec{b}_1 = 2 \sin \left( \frac{\theta_{ab}}{2} \right) \vec{v}_1. \quad (74)$$

It can be easily checked that the normalized vectors  $\vec{v}_0$  and  $\vec{v}_1$  are orthogonal. By substituting the definitions,

Eqs. (73) and (74), into the maximal violation of the MABK inequality, Eq. (70), we can upper bound the latter as follows:

$$\begin{aligned} \mathcal{M}_\rho &\leq \max_{\substack{\vec{c}_i, \vec{v}_i, \theta_{ab} \text{ such that} \\ \|\vec{c}_i\| = \|\vec{v}_i\| = 1 \wedge \vec{v}_0 \perp \vec{v}_1}} 2 \cos \left( \frac{\theta_{ab}}{2} \right) \vec{v}_0^T \cdot T_\rho \cdot \vec{c}_0 \\ &\quad + 2 \sin \left( \frac{\theta_{ab}}{2} \right) \vec{v}_1^T \cdot T_\rho \cdot \vec{c}_1. \end{aligned} \quad (75)$$

The inequality in Eq. (75) is due to the fact that now the optimization is over arbitrary orthonormal vectors  $\vec{v}_0, \vec{v}_1$  and angle  $\theta_{ab}$ , while originally the optimization was over variables satisfying the structure imposed by Eqs. (73) and (74). We now simplify the rhs of Eq. (75) to obtain the theorem claim. In particular, we optimize over the unit vectors  $\vec{c}_0$  and  $\vec{c}_1$  by choosing them in the directions of  $T_\rho^T \cdot \vec{v}_0$  and  $T_\rho^T \cdot \vec{v}_1$ , respectively, and we also optimize over  $\theta_{ab}$  by exploiting the fact that the general expression  $A \cos \theta + B \sin \theta$  is maximized to  $\sqrt{A^2 + B^2}$  for  $\theta = \arctan B/A$ :

$$\begin{aligned} \mathcal{M}_\rho &\leq \max_{\substack{\vec{v}_i, \theta_{ab} \text{ such that} \\ \|\vec{v}_i\| = 1 \wedge \vec{v}_0 \perp \vec{v}_1}} 2 \left[ \cos \left( \frac{\theta_{ab}}{2} \right) \|T_\rho^T \cdot \vec{v}_0\| \right. \\ &\quad \left. + \sin \left( \frac{\theta_{ab}}{2} \right) \|T_\rho^T \cdot \vec{v}_1\| \right] \\ &= \max_{\substack{\vec{v}_i \text{ such that} \\ \|\vec{v}_i\| = 1 \wedge \vec{v}_0 \perp \vec{v}_1}} 2 \sqrt{\|T_\rho^T \cdot \vec{v}_0\|^2 + \|T_\rho^T \cdot \vec{v}_1\|^2}. \end{aligned} \quad (76)$$

Finally, by applying the result of Lemma 2, we know that the maximum in Eq. (76) is achieved when  $\vec{v}_0$  and  $\vec{v}_1$  are chosen in the direction of the eigenstates of  $T_\rho T_\rho^T$  corresponding to the two largest eigenvalues. This concludes the proof for the  $N = 3$  case:

$$\mathcal{M}_\rho \leq 2\sqrt{t_0 + t_1}, \quad (77)$$

where  $t_0$  and  $t_1$  are the two largest eigenvalues of  $T_\rho T_\rho^T$ . ■

### 1. Tightness conditions

The bound, Eq. (10), is tight if the correlation matrix  $T_\rho$  of the considered state satisfies certain conditions, i.e., for certain classes of states. Here we report the tightness conditions valid in the  $N = 3$  case, while the ones for general  $N$  and their derivation are given in Appendix D.

The upper bound, Eq. (10), on the maximal violation of the tripartite MABK inequality by a given state  $\rho$  is tight, that is there exists a quantum implementation achieving the bound, if there exist unit vectors  $\vec{a}_0, \vec{a}_1, \vec{b}_0,$  and  $\vec{b}_1$  in

$\mathbb{R}^3$  such that the following identities are satisfied:

$$\begin{aligned}\vec{a}_0 \otimes \vec{b}_1 + \vec{a}_1 \otimes \vec{b}_0 &= 2\sqrt{\frac{t_0}{t_0+t_1}}\vec{t}_0, \\ \vec{a}_0 \otimes \vec{b}_0 - \vec{a}_1 \otimes \vec{b}_1 &= 2\sqrt{\frac{t_1}{t_0+t_1}}\vec{t}_1,\end{aligned}\quad (78)$$

where  $\vec{t}_0$  and  $\vec{t}_1$  are the normalized eigenvectors of  $T_\rho T_\rho^T$  corresponding to the two largest eigenvalues  $t_0$  and  $t_1$ . The tightness conditions in Eq. (78) are sufficient conditions such that the equality sign holds in Eq. (77) and can be directly deduced from the theorem's proof.

We point out that by repeating the proof with different definitions of correlation matrix, one can potentially end up with alternative MABK violation upper bounds together with their own set of tightness conditions. This depends on the symmetries of the state  $\rho$ .

More concretely, the correlation matrix of a tripartite state  $\rho$  used in the proof above is a  $9 \times 3$  matrix expressed as follows (cf. Definition 3):

$$[T_\rho]_{ij} = \text{Tr}[\rho(\sigma_{\lceil \frac{i}{3} \rceil} \otimes \sigma_{i-3(\lceil \frac{i}{3} \rceil-1)} \otimes \sigma_j)], \quad (79)$$

where  $i \in \{1, \dots, 9\}$  and  $j \in \{1, 2, 3\}$ . With the definition, Eq. (79), we express the MABK expectation value as in Eq. (69). This leads to the MABK violation upper bound, Eq. (77), and to the tightness conditions, Eq. (78). However, nothing prevents us from defining the tripartite correlation matrix as

$$[T'_\rho]_{ij} = \text{Tr}[\rho(\sigma_{\lceil \frac{i}{3} \rceil} \otimes \sigma_j \otimes \sigma_{i-3(\lceil \frac{i}{3} \rceil-1)})], \quad (80)$$

or as

$$[T''_\rho]_{ij} = \text{Tr}\left[\rho\left(\sigma_j \otimes \sigma_{\lceil \frac{i}{3} \rceil} \otimes \sigma_{i-3(\lceil \frac{i}{3} \rceil-1)}\right)\right]. \quad (81)$$

The alternative definitions of the correlation matrix lead to similar proofs of the MABK violation upper bound. In particular, we obtain an analogous MABK violation upper bound, Eq. (77), and analogous tightness conditions, Eq. (78), except that the eigenvalues  $t_0, t_1$  and eigenvectors  $\vec{t}_0, \vec{t}_1$  of  $T_\rho T_\rho^T$  are replaced by the corresponding eigenvalues and eigenvectors of  $T'_\rho T'^T_\rho$  or  $T''_\rho T''^T_\rho$ , depending on the chosen correlation matrix.

An example showing the importance of this remark is given by the family of states  $\tau(v)$  defined in Eq. (25). Indeed, the MABK violation upper bound obtained for  $\tau(v)$  by using the correlation matrices  $T_\rho$ ,  $T'_\rho$ , and  $T''_\rho$  reads the same and is given in Eq. (26). However, the tightness conditions related to  $T_\rho$  and  $T'_\rho$  are satisfied, while those related to  $T''_\rho$  are not. Thus, the use of different correlation matrices in the above proof can lead to tighter MABK violation upper bounds or to a successful verification of their tightness.

It is interesting to compare the tightness of our bound with the bound derived in Ref. [63]. The major difference is that our bound can be saturated even when the matrix  $T_\rho T_\rho^T$  has no degenerate eigenvalues, opposed to Ref. [63], which requires the degeneracy of the largest eigenvalue of  $T_\rho T_\rho^T$ . When the matrix  $T_\rho T_\rho^T$  is degenerate in its largest eigenvalue (i.e.,  $t_0 = t_1$ ), we recover the same tightness conditions of Ref. [63]. For this reason, our bound is tight on a larger set of states compared to the bound in Ref. [63].

## ACKNOWLEDGMENTS

We thank Timo Holz and Flavien Hirsch for helpful discussions, and Peter Brown for clarifying contributions regarding the tightness of the entropy bounds. This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - Cluster of Excellence Matter and Light for Quantum Computing (ML4Q), by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 675662, and by the Federal Ministry of Education and Research BMBF (Project Q.Link.X and HQS).

## APPENDIX A: REDUCTION TO RANK-ONE PROJECTIVE MEASUREMENTS

Here we provide a detailed proof of Lemma 1, by building on a result proved in Ref. [75]. We report the lemma's statement for clarity.

**Lemma 1.** *Let  $\{P_0, P_1\}$  and  $\{Q_0, Q_1\}$  be two projective measurements acting on a Hilbert space  $\mathcal{H}$ , such that  $P_0, P_1, Q_0$ , and  $Q_1$  are projectors,  $P_0 + P_1 = \text{id}$  and  $Q_0 + Q_1 = \text{id}$ . There exists an orthonormal basis in an enlarged Hilbert space  $\mathcal{H}^*$  such that the four projectors are simultaneously block diagonal, in blocks of size  $2 \times 2$ . Moreover, within a  $2 \times 2$  block, each projector has rank one.*

*Proof.* Let us consider the following three positive operators  $P_0$ ,  $P_0 Q_0 P_0$ , and  $P_0 Q_1 P_0$ . One can check that they commute and therefore can be simultaneously diagonalized. Let  $|v\rangle$  be one of their simultaneous eigenvector. Since  $P_1 \cdot P_0 = 0$ , then  $P_1 |v\rangle = 0$ . So  $|v\rangle$  is also an eigenvector of  $P_1$  with eigenvalue zero. Now, because  $Q_0 + Q_1 = I$ , we cannot have that  $Q_0 |v\rangle = 0$  and  $Q_1 |v\rangle = 0$ . Therefore, one of the following cases hold.

(a) If  $Q_0 |v\rangle = 0$ , then  $Q_1 |v\rangle = |v\rangle$ , and the span of  $|v\rangle$  corresponds to a  $1 \times 1$  block in which  $P_0, P_1, Q_0, Q_1$  have  $|v\rangle$  as a common eigenvector with respective eigenvalues  $1, 0, 0, 1$ .

(b) If  $Q_1 |v\rangle = 0$ , then similarly we have a  $1 \times 1$  block in which  $P_0, P_1, Q_0, Q_1$  have  $|v\rangle$  as a common eigenvector with respective eigenvalues  $1, 0, 1, 0$ .

(c) If  $Q_0|v\rangle \neq 0$  and  $Q_1|v\rangle \neq 0$ , then we define the orthogonal vectors  $|u_0\rangle = Q_0|v\rangle$  and  $|u_1\rangle = Q_1|v\rangle$  and the two-dimensional subspace  $E_v = \{c_0|u_0\rangle + c_1|u_1\rangle : c_0, c_1 \in \mathbb{C}\}$ . We have that  $|v\rangle \in E_v$  since  $|v\rangle = |u_0\rangle + |u_1\rangle$ . Because  $|v\rangle$  is also an eigenvector of  $P_0Q_0P_0$  and  $P_0Q_1P_0$ , then  $P_0|u_0\rangle = P_0Q_0|v\rangle = P_0Q_0P_0|v\rangle \propto |v\rangle$ , similarly  $P_0|u_1\rangle \propto |v\rangle$ . Therefore,  $\exists|w\rangle \in E_v$  such that  $P_0|w\rangle = 0$  and then  $P_1|w\rangle = |w\rangle$ . So the vectors  $|u_0\rangle, |u_1\rangle \in E_v$  are simultaneous eigenvectors of  $Q_0$  and  $Q_1$ , and the vectors  $|v\rangle, |w\rangle \in E_v$  are simultaneous eigenvectors of  $P_0$  and  $P_1$ . And the subspace  $E_v$  corresponds to a  $2 \times 2$  simultaneous diagonal block for the measurements operators  $P_0, P_1, Q_0, Q_1$ .

This procedure can be performed on all the simultaneous eigenvectors of  $P_0, P_0Q_0P_0$ , and  $P_0Q_1P_0$ , and similarly on the remaining simultaneous eigenvectors of  $P_1, P_1Q_0P_1$ , and  $P_1Q_1P_1$ .

Now, if we restrict to a  $2 \times 2$  subspace  $E_v$  with  $\Pi_v$  being the projector on the subspace  $E_v$ , the projectors  $\Pi_v P_0 \Pi_v, \Pi_v P_1 \Pi_v, \Pi_v Q_0 \Pi_v, \Pi_v Q_1 \Pi_v$  are given by

$$\begin{aligned} \Pi_v P_0 \Pi_v &= \frac{|v\rangle\langle v|}{\langle v|v\rangle}, \\ \Pi_v P_1 \Pi_v &= \frac{|w\rangle\langle w|}{\langle w|w\rangle}, \\ \Pi_v Q_0 \Pi_v &= \frac{|u_0\rangle\langle u_0|}{\langle u_0|u_0\rangle}, \\ \Pi_v Q_1 \Pi_v &= \frac{|u_1\rangle\langle u_1|}{\langle u_1|u_1\rangle}, \end{aligned} \quad (\text{A1})$$

i.e., they are all rank-one projectors.

Within a  $1 \times 1$  block, the two measurements defined by  $\{P_0, P_1\}$ , and  $\{Q_0, Q_1\}$  have fixed outputs. Let  $|\tilde{v}\rangle$  be a normalized simultaneous eigenvector of  $P_0, P_0Q_0P_0$ , and  $P_0Q_1P_0$  and consider the case  $Q_0|\tilde{v}\rangle = 0$ , which leads to a block of size  $1 \times 1$  formed by the span of the vector  $|\tilde{v}\rangle$ . We can now artificially enlarge the system dimension by embedding this block into a block of size  $2 \times 2$ . Let  $|\tilde{w}\rangle\langle\tilde{w}|$  be a projector on the extra artificial dimension, with  $|\tilde{w}\rangle$  a normalized vector. Then we can define the two-dimensional subspace  $E_{\tilde{v}} = \{c_0|\tilde{v}\rangle + c_1|\tilde{w}\rangle : c_0, c_1 \in \mathbb{C}\}$ , and we define the projectors within this subspace to be given by  $\Pi_{\tilde{v}}P_0\Pi_{\tilde{v}} = |\tilde{v}\rangle\langle\tilde{v}|$ ,  $\Pi_{\tilde{v}}P_1\Pi_{\tilde{v}} = |\tilde{w}\rangle\langle\tilde{w}|$ ,  $\Pi_{\tilde{v}}Q_0\Pi_{\tilde{v}} = |\tilde{w}\rangle\langle\tilde{w}|$ , and  $\Pi_{\tilde{v}}Q_1\Pi_{\tilde{v}} = |\tilde{v}\rangle\langle\tilde{v}|$ . One can perform a similar embedding for the other case that leads to a  $1 \times 1$  block, that is  $Q_1|\tilde{v}\rangle = 0$ . Note that the new projective measurements defined on  $\mathcal{H}^*$ , when applied to a quantum state  $\rho$  on  $\mathcal{H}$  that has no components in the artificial dimensions, have still fixed outcomes in the enlarged subspaces like  $E_{\tilde{v}}$ .

With this artificial construction, the representation of the four projectors  $P_0, P_1, Q_0$ , and  $Q_1$  in the artificially

enlarged Hilbert space  $\mathcal{H}^*$  is only composed of  $2 \times 2$  diagonal blocks. Moreover, if we restrict to one of these blocks, the two measurements defined by  $\{P_0, P_1\}$  and  $\{Q_0, Q_1\}$  are rank-one projective measurements. ■

## APPENDIX B: EVE'S UNCERTAINTY IS NONINCREASING UNDER SYMMETRIZATION OF THE OUTCOMES

In proving Theorem 1, we argue that all the marginals are random without loss of generality. This can be enforced by assuming that Eve flips the classical outcomes of the measurements in specific combinations. Otherwise, Eve could also provide the parties with a state that inherently leads to the symmetrized marginals, which is the mixture  $\bar{\rho}$  given in Eq. (45). However, Eve would provide such a state in place of the original (unknown) state  $\rho$  only if her uncertainty on the parties' outcomes does not increase.

We quantify Eve's uncertainty via the von Neumann entropy of the classical outcomes conditioned on Eve's quantum side information  $E$ . The specific outcomes that we consider depend on the cryptographic application that is being addressed. For instance, in the main text we employ Theorem 1 to tightly estimate Eve's uncertainty on Alice's random outcome  $X$  by computing  $H(X|E)$ , when Alice, Bob, and Charlie test the MABK inequality. This result finds potential application in DICKA and DIRG protocols. Indeed, in a DICKA scheme Bob and Charlie would correct their raw-key bits to match Alice's bits represented by  $X$ , while in a DIRG protocol the goal is to ensure that Alice's random outcome  $X$  is unknown to Eve. Additionally, we employ Theorem 1 to estimate Eve's uncertainty on the outcomes of Alice ( $X$ ) and Bob ( $Y$ ) jointly, by computing  $H(XY|E)$ .

For illustration purposes, here we provide the full proof that Eve's uncertainty of Alice's outcome  $X$  is nonincreasing if she distributes the state  $\bar{\rho}$  in place of  $\rho$  to  $N = 3$  parties. However, we remark that an analogous proof would hold for any number of parties and any number of outcomes. Therefore, we must verify that the following condition is met:

$$H(X|E)_\rho \geq H(X|E_{\text{tot}})_{\bar{\rho}}, \quad (\text{B1})$$

where Eve's quantum system  $E_{\text{tot}} = ETT'$  contains the quantum side information  $E$ , the outcome of the random variable  $T$  indicating to Eve which of the four states in the mixture  $\bar{\rho}$  to distribute, and the purifying system  $T'$ . Indeed, Eve preparing  $\bar{\rho}$  can be interpreted as she preparing one of the four states:

$$\begin{aligned} \rho, \quad & (Z \otimes Z \otimes \text{id})\rho(Z \otimes Z \otimes \text{id}), \\ & (Z \otimes \text{id} \otimes Z)\rho(Z \otimes \text{id} \otimes Z), \\ & (\text{id} \otimes Z \otimes Z)\rho(\text{id} \otimes Z \otimes Z), \end{aligned} \quad (\text{B2})$$

depending on the outcome  $t$  of a random variable stored in the register  $T$ . Since Eve holds the purification of every state in Eq. (B2):  $\{|\phi_{ABCE}^t\rangle\}_{t=1}^4$ , the global state prepared by Eve is

$$\bar{\rho}_{ABCET} = \frac{1}{4} \sum_t |\phi_{ABCE}^t\rangle \langle \phi_{ABCE}^t| \otimes |t\rangle \langle t|_T. \quad (\text{B3})$$

Finally, we assume that Eve holds the purifying system of the global state, thus the state she prepares is

$$|\bar{\phi}_{ABCETT'}\rangle = \frac{1}{2} \sum_t |\phi_{ABCE}^t\rangle \otimes |t\rangle_T \otimes |t\rangle_{T'}, \quad (\text{B4})$$

which is a purification of Eq. (B3), where both registers  $T$  and  $T'$  are held by Eve and thus appear in  $E_{\text{tot}}$ .

In order to prove Eq. (B1), we start by using the strong subadditivity property:

$$H(X|E_{\text{tot}})_{\bar{\rho}} \leq H(X|ET)_{\bar{\rho}}, \quad (\text{B5})$$

where the rhs entropy is computed on the following state:

$$\begin{aligned} \bar{\rho}_{XET} &= (\mathcal{E}_X \otimes \text{id}_{ET}) \text{Tr}_{BC} [\bar{\rho}_{ABCET}] \\ &= \frac{1}{4} (\mathcal{E}_X \otimes \text{id}_{ET}) \text{Tr}_{BC} \left[ \sum_t |\phi_{ABCE}^t\rangle \langle \phi_{ABCE}^t| \otimes |t\rangle \langle t|_T \right] \\ &\equiv \frac{1}{4} \sum_t \rho_{XE}^t \otimes |t\rangle \langle t|_T, \end{aligned} \quad (\text{B6})$$

where the quantum map

$$\mathcal{E}_X(\sigma) = \sum_{a=0}^1 |a\rangle \langle a| \langle a| \sigma |a\rangle$$

represents the projective measurement performed by Alice. Being the state in Eq. (B6) a c.q. state, its entropy simplifies to

$$H(X|ET)_{\bar{\rho}} = \frac{1}{4} \sum_t H(X|E)_{\rho^t}. \quad (\text{B7})$$

The last part of the proof shows that  $H(X|E)_{\rho^t}$  is actually independent of  $t$  and equal to conditional entropy of the original state  $H(X|E)_{\rho}$ . This is clear if the state  $\rho_{XE}^t$  is made explicit. From Eq. (B6) we have that

$$\rho_{XE}^t = (\mathcal{E}_X \otimes \text{id}_{ET}) \text{Tr}_{BC} [|\phi_{ABCE}^t\rangle \langle \phi_{ABCE}^t|], \quad (\text{B8})$$

where  $|\phi_{ABCE}^t\rangle$  is the purification of one of the four states in Eq. (B2) prepared by Eve according to the random variable  $T$ . For definiteness, let us fix that state to be

$(Z \otimes Z \otimes \text{id})\rho(Z \otimes Z \otimes \text{id})$ , although an analogous reasoning holds for any other state in Eq. (B2). By writing  $\rho$  in its spectral decomposition:

$$\rho = \sum_{\lambda} \lambda |\lambda\rangle \langle \lambda|, \quad (\text{B9})$$

we can immediately explicit  $|\phi_{ABCE}^t\rangle$  as follows:

$$|\phi_{ABCE}^t\rangle = \sum_{\lambda} \sqrt{\lambda} |\lambda^t\rangle_{ABC} \otimes |e_{\lambda}\rangle_E, \quad (\text{B10})$$

where the eigenstates of the operator  $(Z \otimes Z \otimes \text{id})\rho(Z \otimes Z \otimes \text{id})$  read  $|\lambda^t\rangle = (Z \otimes Z \otimes \text{id})|\lambda\rangle$ . By substituting Eq. (B10) into Eq. (B8) and by expliciting the map  $\mathcal{E}_X$  we obtain the following expression:

$$\begin{aligned} \rho_{XE}^t &= \sum_{a=0}^1 |a\rangle \langle a| \otimes \sum_{\lambda, \sigma} \sqrt{\lambda \sigma} \text{Tr}_{BC} [\langle a| \lambda^t\rangle \langle \sigma^t| a\rangle] |e_{\lambda}\rangle \langle e_{\sigma}| \\ &= \sum_{a=0}^1 |a\rangle \langle a| \otimes \sum_{\lambda, \sigma} \sqrt{\lambda \sigma} \text{Tr}_{BC} [\langle \bar{a}| \lambda\rangle \langle \sigma| \bar{a}\rangle] |e_{\lambda}\rangle \langle e_{\sigma}| \\ &= \sum_{a=0}^1 |\bar{a}\rangle \langle \bar{a}| \otimes \sum_{\lambda, \sigma} \sqrt{\lambda \sigma} \text{Tr}_{BC} [\langle a| \lambda\rangle \langle \sigma| a\rangle] |e_{\lambda}\rangle \langle e_{\sigma}| \\ &\equiv \sum_{a=0}^1 |\bar{a}\rangle \langle \bar{a}| \otimes \rho_E^a, \end{aligned} \quad (\text{B11})$$

where in the second equality we use the fact that Alice's measurement lies in the  $(x, y)$  plane hence the  $Z$  operator flips its outcome ( $a \rightarrow \bar{a}$ ) and the cyclic property of the trace. In the third equality we relabel the classical outcomes:  $a \leftrightarrow \bar{a}$ . Finally, by comparing Eq. (B11) with the analogous state  $\rho_{XE}$  obtained from the original state  $\rho$  [i.e., in the case where Eve does not prepare the mixture of states in Eq. (B2)]:

$$\rho_{XE} = \sum_{a=0}^1 |a\rangle \langle a| \otimes \rho_E^a, \quad (\text{B12})$$

we observe that  $\rho_{XE}^t$  and  $\rho_{XE}$  are the same state up to a permutation of the classical outcomes, thus their conditional entropies coincide:

$$H(X|E)_{\rho^t} = H(X|E)_{\rho} \quad \forall t. \quad (\text{B13})$$

In conclusion, by combining Eqs. (B13), (B7), and (B5), we obtain the claim given in Eq. (B1). This concludes the proof.



**APPENDIX C: EQUIVALENCE OF  $\bar{\rho}_+$  AND  $\bar{\rho}_-$** 

In the proof of Theorem 1 we claim that it is not restrictive to assume that Eve distributes the following mixture:

$$\rho_\alpha = \frac{\bar{\rho}_+ + \bar{\rho}_-}{2}, \quad (\text{C1})$$

in place of the state  $\bar{\rho}_+$  given in Eq. (51). For illustration purposes we prove the claim in the case where three parties, Alice, Bob, and Charlie, test a (3, 2, 2) full-correlator Bell inequality and are interested in bounding Eve's uncertainty about Alice's outcome  $X$ , quantified by the conditional von Neumann entropy  $H(X|E)$ . Nevertheless, an analogous proof would hold for any number of parties and joint entropies.

In the first part of the proof, we verify that the states  $\bar{\rho}_+$  and  $\bar{\rho}_-$  are equivalent from the viewpoint of the protocol. Precisely, the statistics generated by the two states coincides, as well as Eve's uncertainty about Alice's outcome, quantified by the conditional entropy  $H(X|E)$ . In the second part we show that Eve's uncertainty does not increase if she prepares a balanced mixture of the two states, Eq. (C1), instead of preparing one of the two states singularly.

We start by computing the statistics generated by the states  $\bar{\rho}_+$  in Eq. (55) and  $\bar{\rho}_-$ , which read as follows for  $N = 3$ :

$$\begin{aligned} \bar{\rho}_\pm &= \sum_{i,j,k=0}^1 \lambda_{ijk} |\psi_{i,j,k}\rangle \langle \psi_{i,j,k}| \\ &\pm \sum_{j,k=0}^1 r_{jk} (|\psi_{0,j,k}\rangle \langle \psi_{1,j,k}| + \text{h.c.}) \\ &+ \text{is} (|\psi_{0,1,1}\rangle \langle \psi_{1,1,1}| - \text{h.c.}), \end{aligned} \quad (\text{C2})$$

where h.c. indicates the Hermitian conjugate of the term appearing alongside it. Note that we arbitrarily assume three out of four off-diagonal elements to be purely real, according to the prescription characterizing  $\bar{\rho}_+$  and  $\bar{\rho}_-$ .

Since we fix the parties' measurements to be in the  $(x, y)$  plane, their observables and the relative eigenstates can be written as follows:

$$\begin{aligned} A &= \cos(\varphi_A)X + \sin(\varphi_A)Y, \\ |a\rangle_A &= \frac{1}{\sqrt{2}}[|0\rangle + (-1)^a e^{i\varphi_A} |1\rangle], \\ B &= \cos(\varphi_B)X + \sin(\varphi_B)Y, \\ |b\rangle_B &= \frac{1}{\sqrt{2}}[|0\rangle + (-1)^b e^{i\varphi_B} |1\rangle], \end{aligned}$$

$$C = \cos(\varphi_C)X + \sin(\varphi_C)Y,$$

$$|c\rangle_C = \frac{1}{\sqrt{2}}[|0\rangle + (-1)^c e^{i\varphi_C} |1\rangle], \quad (\text{C3})$$

where  $X, Y$ , and  $Z$  are the Pauli operators,  $A, B$ , and  $C$  are the observables of Alice, Bob, and Charlie, respectively, and the measurement outcomes are defined to be  $a, b, c \in \{0, 1\}$  (where  $a = 0$  corresponds to eigenvalue  $+1$  and  $a = 1$  to eigenvalue  $-1$ ). Then, the statistics generated by the states  $\bar{\rho}_+$  and  $\bar{\rho}_-$  read

$$\begin{aligned} &\Pr[A = a, B = b, C = c]_{\bar{\rho}_\pm} \\ &= \sum_{i,j,k=0}^1 \lambda_{ijk} \langle \psi_{i,j,k} | a, b, c \rangle \langle a, b, c | \psi_{i,j,k} \rangle \\ &\pm 2 \sum_{j,k=0}^1 r_{jk} \text{Re}[\langle \psi_{0,j,k} | a, b, c \rangle \langle a, b, c | \psi_{1,j,k} \rangle] \\ &- 2s \text{Im}[\langle \psi_{1,1,1} | a, b, c \rangle \langle a, b, c | \psi_{0,1,1} \rangle]. \end{aligned} \quad (\text{C4})$$

Therefore, the two statistics coincide if and only if the coefficients of the terms  $r_{jk}$  are all identically null:

$$\text{Re}[\langle \psi_{0,j,k} | a, b, c \rangle \langle a, b, c | \psi_{1,j,k} \rangle] = 0 \quad \forall j, k, a, b, c. \quad (\text{C5})$$

A straightforward calculation of the coefficients of  $r_{jk}$ , by using the expressions in Eq. (C3) and the GHZ-basis definition, leads to the following result:

$$\begin{aligned} &\langle \psi_{0,j,k} | a, b, c \rangle \langle a, b, c | \psi_{1,j,k} \rangle \\ &= \frac{2(-1)^{a+b+c} \text{Im}[e^{i\varphi_A} e^{i\varphi_B} (-1)^j e^{i\varphi_C} (-1)^k]}{16}, \end{aligned} \quad (\text{C6})$$

which is indeed purely imaginary. This proves the condition, Eq. (C5), and thus that the statistics of  $\bar{\rho}_+$  and  $\bar{\rho}_-$  are identical.

The next step of the proof consists in showing that Eve's uncertainty about Alice's outcome is unchanged if she distributes  $\bar{\rho}_+$  or  $\bar{\rho}_-$ , i.e., the following condition must be verified:

$$H(X|E)_{\bar{\rho}_+} = H(X|E)_{\bar{\rho}_-}. \quad (\text{C7})$$

In order to show Eq. (C7), we compute the conditional entropy produced by each state as follows:

$$H(X|E) = H(E|X) + H(X) - H(E) \quad (\text{C8})$$

and verify that each term in Eq. (C8) is identical for the two states  $\bar{\rho}_+$  and  $\bar{\rho}_-$ . To begin with, we know that the

Shannon entropy  $H(X)$  is given by

$$H(X) = h(\Pr[A = 0]), \quad (\text{C9})$$

where  $h(\cdot)$  is the binary entropy, defined as  $h(p) = -p \log_2 p - (1-p) \log_2(1-p)$ . Since we prove that the statistics generated by  $\bar{\rho}_+$  and  $\bar{\rho}_-$  are the same, it follows that

$$H(X)_{\bar{\rho}_+} = H(X)_{\bar{\rho}_-}. \quad (\text{C10})$$

In order to compute the other two terms in Eq. (C8), we write  $\bar{\rho}_+$  and  $\bar{\rho}_-$  in their spectral decomposition:

$$\bar{\rho}_\pm = \sum_{i,j,k=0}^1 \rho_{ijk} |\rho_{ijk}^\pm\rangle \langle \rho_{ijk}^\pm|, \quad (\text{C11})$$

where  $\rho_{ijk}$  are the states' eigenvalues, which one can easily verify to be identical for the two states, while  $|\rho_{ijk}^\pm\rangle$  are the normalized eigenvectors, expressed for simplicity in terms of the following non-normalized eigenvectors:

$$\begin{aligned} |\tilde{\rho}_{ijk}^\pm\rangle &= \frac{\lambda_{0jk} - \lambda_{1jk} - (-1)^i \sqrt{4r_{jk}^2 + (\lambda_{0jk} - \lambda_{1jk})^2}}{\pm 2r_{jk}} |\psi_{0,j,k}\rangle + |\psi_{1,j,k}\rangle \\ &\equiv \pm f_{jk}^i |\psi_{0,j,k}\rangle + |\psi_{1,j,k}\rangle \quad (j,k) \neq (1,1) \end{aligned} \quad (\text{C12})$$

$$\begin{aligned} |\tilde{\rho}_{i11}^\pm\rangle &= (\pm r_{11} + \text{is}) \frac{\lambda_{011} - \lambda_{111} - (-1)^i \sqrt{4r_{11}^2 + 4s^2 + (\lambda_{011} - \lambda_{111})^2}}{2(r_{11}^2 + s^2)} |\psi_{0,1,1}\rangle + |\psi_{1,1,1}\rangle \\ &\equiv (\pm g_{11}^i + \text{ih}_{11}^i) |\psi_{0,1,1}\rangle + |\psi_{1,1,1}\rangle. \end{aligned} \quad (\text{C13})$$

Since  $\bar{\rho}_+$  and  $\bar{\rho}_-$  have the same eigenvalues, it holds that

$$H(ABC)_{\bar{\rho}_+} = H(ABC)_{\bar{\rho}_-}. \quad (\text{C14})$$

Assuming that Eve holds the purification

$$|\tilde{\phi}_{ABCE}^\pm\rangle = \sum_{i,j,k=0}^1 \sqrt{\rho_{ijk}} |\rho_{ijk}^\pm\rangle \otimes |e_{ijk}\rangle \quad (\text{C15})$$

of the parties' state, where  $\{|e_{ijk}\rangle\}$  is an orthonormal basis in  $E$ , it follows that

$$H(E)_{\bar{\rho}_+} = H(E)_{\bar{\rho}_-}. \quad (\text{C16})$$

The remaining term in Eq. (C8) is  $H(E|X)$ , which is computed on the c.q. state:

$$\begin{aligned} \bar{\rho}_{XE}^\pm &= \sum_{a=0}^1 |a\rangle \langle a| \otimes \sum_{\substack{i,j,k=0 \\ l,m,n=0}}^1 \sqrt{\rho_{ijk} \rho_{lmn}} \text{Tr}_{BC} \left[ \langle a | \rho_{ijk}^\pm \rangle \langle \rho_{lmn}^\pm | a \rangle \right] |e_{ijk}\rangle \langle e_{lmn}| \\ &\equiv \sum_{a=0}^1 \Pr[A = a] |a\rangle \langle a| \otimes \rho_E^{a,\pm}, \end{aligned} \quad (\text{C17})$$

where  $\rho_E^{a,\pm}$  is the conditional state of Eve, given that Alice obtained outcome  $a$ . By employing the expressions in Eqs. (C12) and (C13), one can verify that the operators  $\rho_E^{a,\pm}$  are one the transpose of the other:  $\rho_E^{a,+} = (\rho_E^{a,-})^T$ .

Thus  $\rho_E^{a,+}$  and  $\rho_E^{a,-}$  have the same eigenvalues, which implies that

$$H(\rho_E^{a,+}) = H(\rho_E^{a,-}). \quad (\text{C18})$$

Finally, since the conditional entropy  $H(E|X)$  is computed as follows on the classical quantum states in Eq. (C17):

$$H(E|X)_{\bar{\rho}_{\pm}} = \sum_{a=0}^1 \Pr[A = a] H(\rho_E^{a,\pm}), \quad (\text{C19})$$

we conclude that

$$H(E|X)_{\bar{\rho}_+} = H(E|X)_{\bar{\rho}_-}. \quad (\text{C20})$$

By combining the results in Eqs. (C10), (C16), and (C20) into (C7), we verify that the states  $\bar{\rho}_+$  and  $\bar{\rho}_-$  lead to the same conditional entropy.

The final part of the proof shows that Eve's uncertainty in preparing the mixture  $\rho_\alpha$ , Eq. (C1), does not increase with respect to preparing one of the two states  $\bar{\rho}_{\pm}$ :

$$H(X|E_{\text{tot}})_{\rho_\alpha} \leq H(X|E)_{\bar{\rho}_+}. \quad (\text{C21})$$

In this way we can guarantee that it is not restrictive to assume that Eve prepares the mixture, Eq. (C1). In giving Eve maximum power, we assume that she prepares the following global pure state (similarly to Appendix B):

$$|\phi_{ABCEMM'}\rangle \frac{1}{\sqrt{2}} \sum_{m=+,-} |\bar{\phi}_{ABCE}^m\rangle \otimes |m\rangle_M \otimes |m\rangle_{M'}, \quad (\text{C22})$$

where  $|\bar{\phi}_{ABCE}^{\pm}\rangle$  are the purifications of the individual states  $\bar{\rho}_{\pm}$  defined in Eq. (C15), while  $M$  is an ancillary system informing Eve on which of the two purified states she prepared and  $M'$  is the purifying system of the global state. Therefore, Eve has maximum power and her quantum system comprises  $E_{\text{tot}} = EMM'$ . Naturally, it holds that

$$\rho_\alpha = \text{Tr}_{E_{\text{tot}}} [|\phi_{ABCEMM'}\rangle \langle \phi_{ABCEMM'}|]. \quad (\text{C23})$$

For the strong subadditivity property, we have that

$$\begin{aligned} H(X|E_{\text{tot}})_{\rho_\alpha} &= H(X|EMM')_{\rho_\alpha} \leq H(X|EM)_{\rho_\alpha} \\ &= \frac{1}{2} \sum_{m=+,-} H(X|E)_{\bar{\rho}_m}, \end{aligned} \quad (\text{C24})$$

where the last equality is due to the fact that the state  $\text{Tr}_{M'} [|\phi_{ABCEMM'}\rangle \langle \phi_{ABCEMM'}|]$  is classical on  $M$ . Finally, by employing the result, Eq. (C7) into Eq. (C24), we obtain the claim in Eq. (C21). This concludes the proof. The same argument can be used to generalize the proof for the case of  $N$  parties and for the conditional entropy of the joint outcome of more than one party.

## APPENDIX D: MAXIMAL MABK VIOLATION BY AN $N$ -QUBIT STATE: PROOF

Here we provide the full proof of Theorem 2 and of Lemma 2, which combined provide an analytical upper bound on the maximal violation of the  $N$ -partite MABK inequality by an arbitrary  $N$ -qubit state, for rank-one projective measurements. This is the only existing upper bound on the violation of an  $N$ -partite Bell inequality by an  $N$ -qubit state, expressed as a function of the state's parameters. In Ref. [63], the authors only conjectured a bound for the  $N$ -party case based on their result valid for three parties. Analogously to the three-party case (see Sec. VIII), our  $N$ -partite bound is tight on a broader class of states than the bound conjectured in Ref. [63].

We start by proving Lemma 2, which plays an important role in the proof of Theorem 2. We report the lemma's statement for clarity.

**Lemma 2.** *Let  $Q$  be an  $m \times n$  real matrix and let  $\|\vec{v}\|$  be the Euclidean norm of vectors  $\vec{v} \in \mathbb{R}^k$ , for  $k = m, n$ . Finally, let “ $\cdot$ ” indicate both the scalar product and the matrix-vector multiplication. Then*

$$\max_{\substack{\vec{c} \perp \vec{c}' \text{ such that} \\ \|\vec{c}\| = \|\vec{c}'\| = 1}} \left[ \|Q \cdot \vec{c}\|^2 + \|Q \cdot \vec{c}'\|^2 \right] = u_1 + u_2, \quad (\text{D1})$$

where  $u_1$  and  $u_2$  are the largest and second-to-the-largest eigenvalues of  $U \equiv Q^T Q$ , respectively.

*Proof.* Note that  $U$  is a symmetric  $n \times n$  real matrix, thus it can be diagonalized. The eigenvalue equation for  $U$  reads

$$U \cdot \vec{u}_i = u_i \vec{u}_i \quad i = 1, \dots, n, \quad (\text{D2})$$

where the set of eigenvectors forms an orthonormal basis of  $\mathbb{R}^n$ :  $\vec{u}_i^T \cdot \vec{u}_j = \delta_{ij}$  and without loss of generality we order the eigenvalues as  $u_1 \geq u_2 \geq \dots \geq u_n \geq 0$ . Note that every eigenvalue is non-negative:

$$u_i = \vec{u}_i^T \cdot U \cdot \vec{u}_i = \vec{u}_i^T \cdot Q^T Q \cdot \vec{u}_i = \|Q \cdot \vec{u}_i\|^2 \geq 0.$$

By considering that  $\|Q \cdot \vec{c}\|^2 = \vec{c}^T \cdot Q^T Q \cdot \vec{c} = \vec{c}^T \cdot U \cdot \vec{c}$  and by expressing the vectors  $\vec{c}$  and  $\vec{c}'$  in the eigenbasis of  $U$ :

$$\begin{aligned} \vec{c} &= \sum_{i=1}^n c_i \vec{u}_i, \\ \vec{c}' &= \sum_{i=1}^n c'_i \vec{u}_i, \end{aligned}$$

we can recast the claim in Eq. (D1) as follows:

$$\max_{\substack{\vec{c} \perp \vec{c}' \text{ such that} \\ \|\vec{c}\| = \|\vec{c}'\| = 1}} \sum_{i=1}^n u_i (c_i^2 + c_i'^2) = u_1 + u_2. \quad (\text{D3})$$

Let us consider the most general scenario in which some of the eigenvalues of  $U$  are degenerate:  $u_1 \geq u_2 = u_3 = \dots = u_d > u_{d+1} \geq \dots u_n \geq 0$ , where  $d = 2, \dots, n$ . Note that we also account for the possibility that  $u_1 = u_2$ .

We are now going to prove Eq. (D3) by showing that for any couple of mutually orthogonal unit vectors  $\vec{c}$  and  $\vec{c}'$  the lhs of Eq. (D3) is upper bounded by  $u_1 + u_2$  and that the bound is tight.

We start by considering two unit vectors in  $\mathbb{R}^n$ :

$$\begin{cases} \vec{c}^T = (c_1, \dots, c_n) & \text{such that } \|\vec{c}\|^2 = 1 \\ \vec{c}'^T = (c'_1, \dots, c'_n) & \text{such that } \|\vec{c}'\|^2 = 1, \end{cases} \quad (\text{D4})$$

and we define two unit vectors  $\vec{v}, \vec{w} \in \mathbb{R}^{d-1}$  along the directions individuated by  $(c_2, \dots, c_d)$  and  $(c'_2, \dots, c'_d)$ , i.e.,

$$\begin{aligned} c_v \vec{v}^T &\equiv (c_2, \dots, c_d), \\ c'_w \vec{w}^T &\equiv (c'_2, \dots, c'_d), \end{aligned} \quad (\text{D5})$$

where  $c_v$  and  $c'_w$  are the norms of  $(c_2, \dots, c_d)$  and  $(c'_2, \dots, c'_d)$ , respectively. For  $d = 2$  we simply have that  $c_v \vec{v}^T = c_2$  and  $c'_w \vec{w}^T = c'_2$ .

With an abuse of notation, we can rewrite Eq. (D4) as

$$\begin{cases} \vec{c}^T = (c_1, c_v \vec{v}^T, c_{d+1}, \dots, c_n) & \text{such that } c_1^2 + c_v^2 + r = 1 \\ \vec{c}'^T = (c'_1, c'_w \vec{w}^T, c'_{d+1}, \dots, c'_n) & \text{such that } c_1'^2 + c_w'^2 + r' = 1, \end{cases} \quad (\text{D6})$$

where  $r \equiv \sum_{i=d+1}^n c_i^2$  and  $r' \equiv \sum_{i=d+1}^n c_i'^2$  and for both it holds that  $0 \leq r \leq 1$  and  $0 \leq r' \leq 1$ . From the orthogonal condition  $\vec{c}^T \cdot \vec{c}' = 0$  we get that

$$|c_1 c'_1| = \left| \sum_{i=2}^n c_i c'_i \right|, \quad (\text{D7})$$

and from the Cauchy-Schwarz inequality we deduce that

$$\left| \sum_{i=2}^n c_i c'_i \right| \leq \sqrt{(c_v^2 + r)(c_w'^2 + r')}. \quad (\text{D8})$$

By employing Eqs. (D7), (D8) and the normalization conditions in Eq. (D6), we show that  $c_1^2 + c_1'^2 \leq 1$  holds:

$$\begin{aligned} (c_1^2 + c_1'^2)^2 &= c_1^4 + c_1'^4 + 2c_1^2 c_1'^2 \\ &\leq c_1^4 + c_1'^4 + 2(1 - c_1^2)(1 - c_1'^2) \\ &= 2 - 2(c_1^2 + c_1'^2) + (c_1^2 + c_1'^2)^2. \end{aligned} \quad (\text{D9})$$

By comparing the lhs with the rhs one gets the desired result:

$$c_1^2 + c_1'^2 \leq 1. \quad (\text{D10})$$

We now prove the claim in Eq. (D3) through the following chain of equalities and inequalities:

$$\begin{aligned} \sum_{i=1}^n u_i (c_i^2 + c_i'^2) &= u_1 (c_1^2 + c_1'^2) + u_2 (c_v^2 + c_w'^2) + \sum_{i=d+1}^n u_i (c_i^2 + c_i'^2) \\ &\leq u_1 (c_1^2 + c_1'^2) + u_2 (1 - r - c_1^2 + 1 - r' - c_1'^2) + u_{d+1} (r + r') \\ &= u_2 + (u_1 - u_2)(c_1^2 + c_1'^2) + u_2 - (r + r')(u_2 - u_{d+1}) \\ &\leq u_1 + u_2, \end{aligned} \quad (\text{D11})$$



where we use the normalization conditions and the fact that the eigenvalues are ordered in descending order for the first inequality, and we use Eq. (D10) together with the fact that  $r, r' \geq 0$  for the second inequality.

We are left to show that Eq. (D11) is tight, that is there exist unit vectors  $\vec{c}$  and  $\vec{c}'$  for which the equality sign holds. If  $u_1 = u_2$ , the upper bound is attained when  $r = r' = 0$ . Thus the most general pair of vectors satisfying Eq. (D3) is given by

$$\begin{cases} \vec{c}^T = (\vec{V}^T, 0, \dots, 0) \\ \vec{c}'^T = (\vec{W}^T, 0, \dots, 0) \end{cases} \quad (\text{D12})$$

with  $\vec{V}, \vec{W} \in \mathbb{R}^d$  such that  $\|\vec{V}\| = \|\vec{W}\| = 1$  and  $\vec{V} \cdot \vec{W} = 0$ .

If instead  $u_1 > u_2$ , the upper bound is attained when  $r = r' = 0$  and  $c_1^2 + c_1'^2 = 1$ . The second condition is verified when the equality holds in Eq. (D8), which in turn happens when the unit vectors  $\vec{v}$  and  $\vec{v}'$  are parallel. Thus the most general pair of vectors satisfying Eq. (D3) is given by

$$\begin{cases} \vec{c}^T = (c_1, c_v \vec{v}^T, 0, \dots, 0) \\ \vec{c}'^T = (c_1', c_w' \vec{v}^T, 0, \dots, 0) \end{cases}, \vec{v} \in \mathbb{R}^{d-1} \wedge c_1^2 + c_1'^2 = 1, \quad (\text{D13})$$

and where the orthogonality and normalization conditions hold:  $c_1 c_1' + c_v c_w' = 0$ ,  $c_1^2 + c_v^2 = 1$  and  $c_1'^2 + c_w'^2 = 1$ . Such solutions can always be parametrized as follows:

$$\begin{cases} \vec{c}^T = (\cos \alpha, \sin \alpha \vec{v}^T, 0, \dots, 0) \\ \vec{c}'^T = (-\sin \alpha, \cos \alpha \vec{v}^T, 0, \dots, 0) \end{cases}, \quad \alpha \in \mathbb{R}. \quad (\text{D14})$$

This concludes the proof.  $\blacksquare$

We are now ready to prove Theorem 2.

*Proof.* Firstly, we present closed expressions for the  $N$ -partite MABK operator, defined recursively in Definition 2. In particular, in Ref. [72] the explicit expression of the  $N$ -partite MABK operator when  $N$  is odd is given:

$$M_N^{\text{odd}} = \frac{1}{2^{(N-3)/2}} \sum_{\mathbf{x} \in \mathcal{L}_N} (-1)^{(1/2)[(N-1)/2 - \omega(\mathbf{x})]} \bigotimes_{i=1}^N A_{x_i}^{(i)}, \quad (\text{D15})$$

where  $A_0^{(i)}$  and  $A_1^{(i)}$  are the two binary observables of Alice<sub>*i*</sub>, while  $\mathbf{x} = (x_1, \dots, x_N)$  is a bit string with Hamming weight given by

$$\omega(\mathbf{x}) = |\{1 \leq i \leq N | x_i = 1\}|, \quad (\text{D16})$$

and the set  $\mathcal{L}_N$  is defined as follows:

$$\mathcal{L}_N = \left\{ \mathbf{x} \in \{0, 1\}^N \mid \omega(\mathbf{x}) = \frac{N-1}{2} \pmod{2} \right\}. \quad (\text{D17})$$

By applying once the MABK recursive formula of Definition 2 on Eq. (D15), one obtains an explicit expression of the  $N$ -partite MABK operator for  $N$  even. We

distinguish the case  $N/2$  even:

$$M_N^{\text{even}} = \frac{1}{2^{(N-2)/2}} \sum_{\mathbf{x} \in \{0,1\}^N} (-1)^{N/4 - \lceil \omega(\mathbf{x})/2 \rceil} \bigotimes_{i=1}^N A_{x_i}^{(i)}, \quad (\text{D18})$$

and the case  $N/2$  odd:

$$M_N^{\text{even}} = \frac{1}{2^{(N-2)/2}} \sum_{\mathbf{x} \in \{0,1\}^N} (-1)^{(N-2)/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^N A_{x_i}^{(i)}, \quad (\text{D19})$$

where  $\lceil a \rceil$  and  $\lfloor a \rfloor$  are the ceiling and floor functions, respectively.

We now derive an explicit expression of the MABK expectation value for a generic  $N$ -qubit state. As shown above, the  $N$ -party MABK operator can be written in explicit form as follows:

$$M_N = \frac{1}{\mathcal{N}_N} \sum_{\mathbf{x} \in \mathcal{S}_N} (-1)^{\xi_N(\mathbf{x})} \bigotimes_{i=1}^N A_{x_i}^{(i)}, \quad (\text{D20})$$

where the normalization factor  $\mathcal{N}_N$ , the set of  $N$ -bit strings  $\mathcal{S}_N$  and the exponent  $\xi_N(\mathbf{x})$  depend on the parity of  $N$  (e.g.,  $\mathcal{S}_N = \{0, 1\}^N$  for  $N$  even and  $\mathcal{S}_N = \mathcal{L}_N$  for  $N$  odd). By assumption we restrict to rank-one projective measurements, hence every observable  $A_{x_i}^{(i)}$  can be individuated by a unit vector  $\vec{a}_{x_i}^i \in \mathbb{R}^3$  such that

$$A_{x_i}^{(i)} = \vec{a}_{x_i}^i \cdot \vec{\sigma} = \sum_{v_i=1}^3 a_{x_i, v_i}^i \sigma_{v_i}, \quad (\text{D21})$$

where  $\vec{\sigma} = (X, Y, Z)^T$ . By substituting Eq. (D21) into Eq. (D20) and by rearranging the terms we get

$$\begin{aligned} M_N &= \frac{1}{\mathcal{N}_N} \sum_{v_1, \dots, v_N=1}^3 \left[ \sum_{\mathbf{x} \in \mathcal{S}_N} (-1)^{\xi_N(\mathbf{x})} \prod_{i=1}^N a_{x_i, v_i}^i \right] \\ &\quad \sigma_{v_1} \otimes \dots \otimes \sigma_{v_N} \\ &\equiv \frac{1}{\mathcal{N}_N} \sum_{v_1, \dots, v_N=1}^3 M_{v_1, \dots, v_N} \sigma_{v_1} \otimes \dots \otimes \sigma_{v_N}. \end{aligned} \quad (\text{D22})$$

We now employ Eq. (D22) and the expression for a generic  $N$ -qubit state:

$$\rho = \frac{1}{2^N} \sum_{\mu_1 \dots \mu_N=0}^3 \Lambda_{\mu_1 \dots \mu_N} \sigma_{\mu_1} \otimes \dots \otimes \sigma_{\mu_N}, \quad (\text{D23})$$

to derive an explicit expression for the MABK expectation value as follows:

$$\begin{aligned} \langle M_N \rangle_\rho &= \text{Tr}[M_N \rho] = \frac{1}{\mathcal{N}_N} \sum_{v_1, \dots, v_N=1}^3 M_{v_1, \dots, v_N} \Lambda_{v_1 \dots v_N} \\ &= \frac{1}{\mathcal{N}_N} \sum_{v_1, \dots, v_N=1}^3 \left[ \sum_{\mathbf{x} \in \mathcal{S}_N} (-1)^{\xi_N(\mathbf{x})} a_{x_1, v_1}^1 \cdots a_{x_N, v_N}^N \Lambda_{v_1 \dots v_N} \right], \end{aligned} \quad (\text{D24})$$

where we use the fact that  $\text{Tr}[\sigma_i \sigma_j] = 2\delta_{ij}$ .

We now specify the expressions for  $\mathcal{N}_N$ ,  $\mathcal{S}_N$ , and  $\xi_N(\mathbf{x})$  when  $N/2$  is even and prove the theorem's statement in this particular case. However, a similar procedure applies to the  $N/2$  odd and  $N$  odd cases and leads to the same final result.

We thus have the following expression for the MABK expectation value:

$$\langle M_N \rangle_\rho = \frac{1}{2^{(N-2)/2}} \sum_{v_1, \dots, v_N=1}^3 \left[ \sum_{\mathbf{x} \in \{0,1\}^N} (-1)^{N/4 - \lceil \omega(\mathbf{x})/2 \rceil} a_{x_1, v_1}^1 \cdots a_{x_N, v_N}^N \Lambda_{v_1 \dots v_N} \right], \quad (\text{D25})$$

and we rearrange it as follows:

$$\begin{aligned} \langle M_N \rangle_\rho &= \frac{1}{2^{(N-2)/2}} \sum_{v_1, \dots, v_N=1}^3 \left[ \sum_{\mathbf{x} \in \{0,1\}^{N/2}} \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{N/4 - \lceil [\omega(\mathbf{x}) + \omega(\mathbf{y})]/2 \rceil} a_{x_1, v_1}^1 \cdots a_{x_{N/2}, v_{N/2}}^{N/2} \Lambda_{v_1 \dots v_N} a_{y_1, v_{N/2}}^{N/2} \cdots a_{y_{N/2}, v_N}^N \right] \\ &= \frac{1}{2^{(N-2)/2}} \sum_{v_1, \dots, v_N=1}^3 \left[ \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{N/4 - \lceil [\omega(\mathbf{x}) + \omega(\mathbf{y})]/2 \rceil} a_{x_1, v_1}^1 \cdots a_{x_{N/2}, v_{N/2}}^{N/2} \Lambda_{v_1 \dots v_N} a_{y_1, v_{N/2}}^{N/2} \cdots a_{y_{N/2}, v_N}^N \right. \\ &\quad \left. + \sum_{\mathbf{x} \in \mathcal{O}_{N/2}} \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{N/4 - \lceil [\omega(\mathbf{x}) + \omega(\mathbf{y})]/2 \rceil} a_{x_1, v_1}^1 \cdots a_{x_{N/2}, v_{N/2}}^{N/2} \Lambda_{v_1 \dots v_N} a_{y_1, v_{N/2}}^{N/2} \cdots a_{y_{N/2}, v_N}^N \right], \end{aligned} \quad (\text{D26})$$

where the sets  $\mathcal{E}_{N/2}$  and  $\mathcal{O}_{N/2}$  are defined as follows:

$$\mathcal{E}_{N/2} = \{ \mathbf{x} \in \{0,1\}^{N/2} \mid \omega(\mathbf{x}) \bmod 2 = 0 \} \quad (\text{D27})$$

$$\mathcal{O}_{N/2} = \{ \mathbf{x} \in \{0,1\}^{N/2} \mid \omega(\mathbf{x}) \bmod 2 = 1 \}. \quad (\text{D28})$$

We basically split the bit strings  $\mathbf{x}$  into those with an even Hamming weight and those with an odd Hamming weight. Now considering that the following identity holds:

$$\left\lceil \frac{\omega(\mathbf{x}) + \omega(\mathbf{y})}{2} \right\rceil = \begin{cases} \omega(\mathbf{x}) \text{ even:} & \lfloor \omega(\mathbf{x})/2 \rfloor + \lfloor \omega(\mathbf{y})/2 \rfloor + [\omega(\mathbf{y}) \bmod 2] \\ \omega(\mathbf{x}) \text{ odd:} & \left\lfloor \frac{\omega(\mathbf{x})}{2} \right\rfloor + \left\lfloor \frac{\omega(\mathbf{y})}{2} \right\rfloor + 1, \end{cases} \quad (\text{D29})$$

we can recast the MABK expectation value in Eq. (D26) as follows:

$$\begin{aligned} \langle M_N \rangle_\rho &= \frac{1}{2^{(N-2)/2}} \sum_{v_1, \dots, v_N=1}^3 \left\{ \left[ \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} (-1)^{N/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} a_{x_1, v_1}^1 \cdots a_{x_{N/2}, v_{N/2}}^{N/2} \right] \right. \\ &\quad \left. \Lambda_{v_1 \dots v_N} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor + (\omega(\mathbf{y}) \bmod 2)} a_{y_1, v_{N/2}}^{N/2} \cdots a_{y_{N/2}, v_N}^N \right] \right\} \end{aligned}$$

$$\begin{aligned}
 & + \left[ \sum_{\mathbf{x} \in \mathcal{O}_{N/2}} (-1)^{N/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} a_{x_1, v_1}^1 \cdots a_{x_{N/2}, v_{N/2}}^{N/2} \right] \Lambda_{v_1, \dots, v_N} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor + 1} a_{y_1, v_{N/2}}^{N/2} \cdots a_{y_{N/2}, v_N}^N \right] \Big\} \\
 & \equiv \frac{1}{2^{(N-2)/2}} [\vec{v}_0^T \cdot T_\rho \cdot \vec{u}_0 + \vec{v}_1^T \cdot T_\rho \cdot \vec{u}_1]. \tag{D30}
 \end{aligned}$$

In the expression, Eq. (D30), we define the vectors

$$\vec{v}_0 = \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} (-1)^{N/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{x_i}^i, \tag{D31}$$

$$\vec{v}_1 = \sum_{\mathbf{x} \in \mathcal{O}_{N/2}} (-1)^{N/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{x_i}^i, \tag{D32}$$

$$\vec{u}_0 = \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor + \lfloor \omega(\mathbf{y}) \bmod 2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{y_i}^{N/2+i}, \tag{D33}$$

$$\vec{u}_1 = \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor + 1} \bigotimes_{i=1}^{N/2} \vec{a}_{y_i}^{N/2+i}, \tag{D34}$$

and we use Definition 3 of the correlation matrix of an  $N$ -qubit state. The  $3^{N/2}$ -dimensional vectors in Eqs. (D31)–(D34) are heavily constrained by their tensor-product structure and satisfy the following properties:

$$\text{Prop. 1: } \|\vec{v}_0\|^2 + \|\vec{v}_1\|^2 = 2^{N/2}, \tag{D35}$$

$$\text{Prop. 2: } \|\vec{u}_0\|^2 = \|\vec{u}_1\|^2 = 2^{N/2}, \tag{D36}$$

$$\text{Prop. 3: } \vec{v}_0 \cdot \vec{v}_1 = 0. \tag{D37}$$

These properties play a fundamental role in deriving a meaningful upper bound on the MABK expectation value.

We prove the first property, Eq. (D35), by directly computing the lhs:

$$\begin{aligned}
 \|\vec{v}_0\|^2 + \|\vec{v}_1\|^2 & = \vec{v}_0 \cdot \vec{v}_0 + \vec{v}_1 \cdot \vec{v}_1 \\
 & = \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{N/2 - \lfloor \omega(\mathbf{x})/2 \rfloor - \lfloor \omega(\mathbf{y})/2 \rfloor} \prod_{i=1}^{N/2} (\cos \theta_i)^{x_i \oplus y_i} \\
 & \quad + \sum_{\mathbf{x}, \mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{N/2 - \lfloor \omega(\mathbf{x})/2 \rfloor - \lfloor \omega(\mathbf{y})/2 \rfloor} \prod_{i=1}^{N/2} (\cos \theta_i)^{x_i \oplus y_i}, \tag{D38}
 \end{aligned}$$

where we use the fact that  $\vec{a}_{x_i}^i$  are unit vectors and we call  $\theta_i$  the angle between the two measurement directions of party number  $i$ :  $\cos \theta_i = \vec{a}_0^i \cdot \vec{a}_1^i$ . Note that the symbol  $\oplus$  is the binary operation XOR. We now define the bit string:  $\mathbf{r} = \mathbf{x} \oplus \mathbf{y}$ , whose Hamming weight can be computed as

$$\omega(\mathbf{r}) = \omega(\mathbf{x} \oplus \mathbf{y}) = \omega(\mathbf{x}) + \omega(\mathbf{y}) - 2\omega(\mathbf{x} \wedge \mathbf{y}), \tag{D39}$$

where  $\wedge$  is the binary operation AND. From Eq. (D39) it follows immediately that the Hamming weight of the string  $\mathbf{r}$  is always even, since the Hamming weights of  $\mathbf{x}$  and  $\mathbf{y}$  are either both even or both odd in Eq. (D38). With this information, we can recast Eq. (D38) as follows:

$$\|\vec{v}_0\|^2 + \|\vec{v}_1\|^2 = \sum_{\mathbf{r} \in \mathcal{E}_{N/2}} \left[ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\lfloor \omega(\mathbf{r} \oplus \mathbf{y})/2 \rfloor + \lfloor \omega(\mathbf{y})/2 \rfloor} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\lfloor \omega(\mathbf{r} \oplus \mathbf{y})/2 \rfloor + \lfloor \omega(\mathbf{y})/2 \rfloor} \right] \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i}, \tag{D40}$$

where we use the fact that  $N/2$  is even and where the string  $\mathbf{x}$  is completely fixed once  $\mathbf{r}$  and  $\mathbf{y}$  are fixed:  $\mathbf{x} = \mathbf{r} \oplus \mathbf{y}$ . Now we employ the relation, Eq. (D39), in Eq. (D40) and we make use of the information on the parity of the Hamming weights appearing in the two sums:

$$\|\vec{v}_0\|^2 + \|\vec{v}_1\|^2 = \sum_{\mathbf{r} \in \mathcal{E}_{N/2}} \left\{ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r})/2 + \omega(\mathbf{y}) - \omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\lfloor \omega(\mathbf{r}) + \omega(\mathbf{y}) - 2\omega(\mathbf{r} \wedge \mathbf{y}) - 1 \rfloor / 2 + \frac{\omega(\mathbf{y}) - 1}{2}} \right\} \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i}. \tag{D41}$$

Note that  $\lfloor a/2 \rfloor = (a - 1)/2$  if  $a$  is an odd number. The expression in Eq. (D41) can be further simplified by considering that even addends in the exponents of  $(-1)$  can be ignored:

$$\begin{aligned}
\|\vec{v}_0\|^2 + \|\vec{v}_1\|^2 &= \sum_{\mathbf{r} \in \mathcal{E}_{N/2}} \left[ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r})/2 - \omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\omega(\mathbf{r})/2 - \omega(\mathbf{r} \wedge \mathbf{y})} \right] \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i} \\
&= \sum_{\mathbf{r} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r})/2} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} \right] \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i} \\
&= 2^{N/2} + \sum_{\substack{\mathbf{r} \in \mathcal{E}_{N/2} \\ \mathbf{r} \neq \mathbf{0}}} (-1)^{\omega(\mathbf{r})/2} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} \right] \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i}, \tag{D42}
\end{aligned}$$

where we extract the term  $\mathbf{r} = \mathbf{0}$  from the sum in the last equality.

The last step to prove the first property, Eq. (D35), is to show that every term in the remaining sum in Eq. (D42) is identically zero, i.e., we want to show that

$$\sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} = 0 \quad \forall \mathbf{r} \neq \mathbf{0}. \tag{D43}$$

In order for Eq. (D43) to be verified, there must be as many  $(-1)$  terms as  $+1$  terms, and since there are in total  $2^{N/2}$  terms, there must be exactly  $2^{N/2-1}$  terms (half of the total) that are  $(-1)$ . We can count the number of  $(-1)$  terms in Eq. (D43) as follows:

$$\sum_{\mathbf{y} \in \{0,1\}^{N/2}} [\omega(\mathbf{r} \wedge \mathbf{y}) \bmod 2], \tag{D44}$$

and check whether it equals  $2^{N/2-1}$ , as claimed. Note that  $\omega(\mathbf{r} \wedge \mathbf{y})$  represents the number of ones in  $\mathbf{r}$  that are also in  $\mathbf{y}$ . The parity of this number is then summed over all the possible bit strings  $\mathbf{y}$  of length  $N/2$ . We can thus recast the sum, as a sum over the number of ones that  $\mathbf{r}$  and  $\mathbf{y}$  have in common ( $k$ ), times the number of bit strings  $\mathbf{y}$  that share  $k$  ones with  $\mathbf{r}$ :

$$\begin{aligned}
&\sum_{\mathbf{y} \in \{0,1\}^{N/2}} [\omega(\mathbf{r} \wedge \mathbf{y}) \bmod 2] \\
&= \sum_{k=0}^{\omega(\mathbf{r})} (k \bmod 2) \binom{\omega(\mathbf{r})}{k} 2^{N/2 - \omega(\mathbf{r})}. \tag{D45}
\end{aligned}$$

Note that the number of bit strings  $\mathbf{y}$  that have  $k$  ones in common with a fixed string  $\mathbf{r}$ , is given by the number of possible combinations of  $k$  ones from the total number of ones  $[\omega(\mathbf{r})]$  populating the string  $\mathbf{r}$ , times the number of possibilities ( $2^{N/2 - \omega(\mathbf{r})}$ ) that we have to fill

the remaining bits of  $\mathbf{y}$  that are not part of the  $k$  ones in common with  $\mathbf{r}$ .

We can now adjust the rhs of Eq. (D45) to the following computable form:

$$\begin{aligned}
2^{N/2 - \omega(\mathbf{r})} \sum_{\substack{k=0 \\ k \text{ odd}}}^{\omega(\mathbf{r})} \binom{\omega(\mathbf{r})}{k} &= 2^{N/2 - \omega(\mathbf{r})} 2^{\omega(\mathbf{r}) - 1} \\
&= 2^{N/2 - 1}, \tag{D46}
\end{aligned}$$

where the first equality is obtained by combining two known facts about the binomial coefficient, namely,

$$\sum_{k=0}^n \binom{n}{k} = 2^n \tag{D47}$$

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = 0. \tag{D48}$$

Indeed, by subtracting Eq. (D48) from Eq. (D47) one gets that

$$\sum_{k \text{ odd}} \binom{n}{k} = 2^{n-1}, \tag{D49}$$

which is used in the first equality in Eq. (D46).

Combining Eqs. (D45) and (D46) we conclude that Eq. (D43) is verified. We thus show the validity of the first property, Eq. (D35).

We move on to prove the second property, Eq. (D36). We start from Eq. (D33) and use the fact that  $\lfloor \omega(\mathbf{x})/2 \rfloor + [\omega(\mathbf{x}) \bmod 2] = \{\omega(\mathbf{x}) + [\omega(\mathbf{x}) \bmod 2]\}/2$ . We obtain

$$\begin{aligned}
 \|\vec{u}_0\|^2 &= \sum_{\mathbf{x}, \mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\{\omega(\mathbf{x}) + [\omega(\mathbf{x}) \bmod 2]\}/2 + \{\omega(\mathbf{y}) + [\omega(\mathbf{y}) \bmod 2]\}/2} \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{x_i \oplus y_i} \\
 &= \sum_{\mathbf{r} \in \{0,1\}^{N/2}} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\{\omega(\mathbf{r}) + \omega(\mathbf{y}) - 2\omega(\mathbf{r} \wedge \mathbf{y}) + [\omega(\mathbf{r}) + \omega(\mathbf{y}) \bmod 2]\}/2 + \frac{\omega(\mathbf{y}) + [\omega(\mathbf{y}) \bmod 2]}{2}} \right] \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i}, \quad (\text{D50})
 \end{aligned}$$

where we define  $\mathbf{r} = \mathbf{x} \oplus \mathbf{y}$  and we use the relation, Eq. (D39). We proceed to simplify Eq. (D50) by splitting the sum over  $\mathbf{y}$  over the strings with even and odd Hamming weight:

$$\begin{aligned}
 \|\vec{u}_0\|^2 &= \sum_{\mathbf{r} \in \{0,1\}^{N/2}} \left[ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\{\omega(\mathbf{r}) + [\omega(\mathbf{r}) \bmod 2]\}/2 - \omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\{\omega(\mathbf{r}) + 1 + [\omega(\mathbf{r}) + 1 \bmod 2]\}/2 - \omega(\mathbf{r} \wedge \mathbf{y}) + \omega(\mathbf{y})} \right] \\
 &\quad \times \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i}. \quad (\text{D51})
 \end{aligned}$$

By employing the following identities:

$$\frac{\omega(\mathbf{r}) + [\omega(\mathbf{r}) \bmod 2]}{2} = \left\lceil \frac{\omega(\mathbf{r})}{2} \right\rceil \quad (\text{D52})$$

$$(-1)^a = (-1)^1 \quad a \text{ odd} \quad (\text{D53})$$

$$\frac{\omega(\mathbf{r}) + 1 + [\omega(\mathbf{r}) + 1 \bmod 2]}{2} = \left\lceil \frac{\omega(\mathbf{r}) + 1}{2} \right\rceil = \left\lceil \frac{\omega(\mathbf{r})}{2} \right\rceil + 1 - [\omega(\mathbf{r}) \bmod 2] \quad (\text{D54})$$

into Eq. (D51) we obtain

$$\begin{aligned}
 \|\vec{u}_0\|^2 &= \sum_{\mathbf{r} \in \{0,1\}^{N/2}} (-1)^{\lceil \omega(\mathbf{r})/2 \rceil} \left\{ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y}) + [\omega(\mathbf{r}) \bmod 2]} \right\} \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i} \\
 &= 2^{N/2} + \sum_{\mathbf{r} \in \{0,1\}^{N/2}, \mathbf{r} \neq \mathbf{0}} (-1)^{\lceil \omega(\mathbf{r})/2 \rceil} \left\{ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y}) + [\omega(\mathbf{r}) \bmod 2]} \right\} \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i} \\
 &= 2^{N/2} + \sum_{\substack{\mathbf{r} \in \mathcal{E}_{N/2} \\ \mathbf{r} \neq \mathbf{0}}} (-1)^{\lceil \omega(\mathbf{r})/2 \rceil} \left[ \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} \right] \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i} \\
 &\quad + \sum_{\mathbf{r} \in \mathcal{O}_{N/2}} (-1)^{\lceil \omega(\mathbf{r})/2 \rceil} \left[ \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} - \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} \right] \prod_{i=1}^{N/2} [\cos(\theta_{N/2+i})]^{r_i}, \quad (\text{D55})
 \end{aligned}$$

where we isolate the  $\mathbf{r} = \mathbf{0}$  term in the second equality and we split the sum over  $\mathbf{r}$  in two sums over the strings with even and odd Hamming weights in the third equality.

The first sum in Eq. (D55) is zero thanks to Eq. (D43). From Eq. (D43) we also deduce that

$$\sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} + \sum_{\mathbf{y} \in \mathcal{O}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} = 0, \quad (\text{D56})$$



which means that the term in square brackets in the second sum can be reduced to

$$2 \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{y})} = 0. \quad (\text{D57})$$

The proof that Eq. (D57) holds is analogous to that of Eq. (D43). In particular, Eq. (D57) is verified if the number of  $(-1)$  terms is exactly half the total number of terms, that is  $2^{N/2-2}$ . We show that this is true by computing the number of  $(-1)$  terms as follows:

$$\begin{aligned} & \sum_{\mathbf{y} \in \mathcal{E}_{N/2}} [\omega(\mathbf{r} \wedge \mathbf{y}) \bmod 2] \\ &= \sum_{k=0}^{\omega(\mathbf{r})} (k \bmod 2) \binom{\omega(\mathbf{r})}{k} 2^{N/2-\omega(\mathbf{r})-1} \\ &= 2^{N/2-\omega(\mathbf{r})-1} \sum_{\substack{k=0 \\ k \text{ odd}}}^{\omega(\mathbf{r})} \binom{\omega(\mathbf{r})}{k}. \end{aligned} \quad (\text{D58})$$

Note that this time, compared to Eq. (D45), the number of possibilities  $[2^{N/2-\omega(\mathbf{r})-1}]$  to fill the nonfixed bits of  $\mathbf{y}$  is halved. The reason is that in this case  $\mathbf{y}$  is constrained to have an even number of ones, thus after fixing  $N/2 - 1$  of its bits, no degree of freedom is left.

By employing again the result on binomial distributions, Eq. (D49), in Eq. (D58), we obtain

$$\sum_{\mathbf{y} \in \mathcal{E}_{N/2}} [\omega(\mathbf{r} \wedge \mathbf{y}) \bmod 2] = 2^{N/2-2}, \quad (\text{D59})$$

which proves Eq. (D57).

We thus show that both the sums in Eq. (D55) are zero, thus proving the second property, Eq. (D36), for  $\vec{u}_0$ . The proof of Eq. (D36) for  $\vec{u}_1$  is analogous and we omit it.

Finally, we show that the third property, Eq. (D37), is satisfied by direct computation:

$$\begin{aligned} \vec{v}_0 \cdot \vec{v}_1 &= \sum_{\substack{\mathbf{x} \in \mathcal{E}_{N/2} \\ \mathbf{y} \in \mathcal{O}_{N/2}}} (-1)^{[\omega(\mathbf{x})/2] + [\omega(\mathbf{y})/2]} \prod_{i=1}^{N/2} (\cos \theta_i)^{x_i \oplus y_i} \\ &= \sum_{\substack{\mathbf{x} \in \mathcal{E}_{N/2} \\ \mathbf{y} \in \mathcal{O}_{N/2}}} (-1)^{\omega(\mathbf{x})/2 + [\omega(\mathbf{y})-1]/2} \prod_{i=1}^{N/2} (\cos \theta_i)^{x_i \oplus y_i} \\ &= \sum_{\mathbf{r} \in \mathcal{O}_{N/2}} \left\{ \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} (-1)^{[\omega(\mathbf{x})-1 + \omega(\mathbf{x}) + \omega(\mathbf{r}) - 2\omega(\mathbf{r} \wedge \mathbf{x})]/2} \right\} \\ & \quad \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i} \end{aligned} \quad (\text{D60})$$

where we define  $\mathbf{r} = \mathbf{x} \oplus \mathbf{y}$  and use Eq. (D39). By simplifying the last expression we get

$$\begin{aligned} \vec{v}_0 \cdot \vec{v}_1 &= \sum_{\mathbf{r} \in \mathcal{O}_{N/2}} (-1)^{[\omega(\mathbf{r})-1]/2} \left[ \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} (-1)^{\omega(\mathbf{r} \wedge \mathbf{x})} \right] \\ & \quad \times \prod_{i=1}^{N/2} (\cos \theta_i)^{r_i} = 0, \end{aligned} \quad (\text{D61})$$

where we use Eq. (D57) to prove the final equality.

Thanks to the properties, Eqs. (D35)–(D37), we can express the vectors  $\vec{v}_k$  and  $\vec{u}_k$  ( $k = 0, 1$ ) as follows:

$$\vec{v}_0 = 2^{N/4} \cos \theta \hat{v}_0, \quad (\text{D62})$$

$$\vec{v}_1 = 2^{N/4} \sin \theta \hat{v}_1, \quad (\text{D63})$$

$$\vec{u}_k = 2^{N/4} \hat{u}_k, \quad (\text{D64})$$

where  $\hat{v}_k$  and  $\hat{u}_k$  are unit vectors in the directions of  $\vec{v}_k$  and  $\vec{u}_k$ , respectively, and where  $\theta$  is a real number. With the expressions, Eqs. (D62), (D63), and (D64), we recast the MABK expectation value, Eq. (D30), as follows:

$$\begin{aligned} \langle M_N \rangle_\rho &= \frac{2^{N/2}}{2^{(N-2)/2}} [\cos \theta \hat{v}_0^T \cdot T_\rho \cdot \hat{u}_0 + \sin \theta \hat{v}_1^T \cdot T_\rho \cdot \hat{u}_1] \\ &= 2 [\cos \theta \hat{v}_0^T \cdot T_\rho \cdot \hat{u}_0 + \sin \theta \hat{v}_1^T \cdot T_\rho \cdot \hat{u}_1]. \end{aligned} \quad (\text{D65})$$

The maximal violation  $\mathcal{M}_\rho$  of the  $N$ -partite MABK inequality is then obtained by maximizing Eq. (D65) over all the parties' measurements directions  $\vec{a}_0^i$  and  $\vec{a}_1^i$  (for  $i = 1, \dots, N$ ). A valid upper bound on the maximal violation  $\mathcal{M}_\rho$  is thus given by

$$\mathcal{M}_\rho \leq \max_{\substack{\hat{v}_k, \hat{u}_k, \theta \\ \hat{v}_0 \perp \hat{v}_1}} 2 [\cos \theta \hat{v}_0^T \cdot T_\rho \cdot \hat{u}_0 + \sin \theta \hat{v}_1^T \cdot T_\rho \cdot \hat{u}_1], \quad (\text{D66})$$

where the inequality is due to the fact that we are now optimizing the expectation value over all the possible unit vectors  $\hat{v}_k$  (such that  $\hat{v}_0 \cdot \hat{v}_1 = 0$ ) and  $\hat{u}_k$ , and freely over  $\theta$ , ignoring the more stringent structures, Eqs. (D31)–(D34), characterizing these vectors and their relation to  $\theta$ . By choosing  $\hat{u}_0$  and  $\hat{u}_1$  in the direction of  $\hat{v}_0^T \cdot T_\rho$  and  $\hat{v}_1^T \cdot T_\rho$ , respectively, and by fixing  $\theta$  such that

$$\tan \theta = \frac{\|T_\rho^T \cdot \hat{v}_1\|}{\|T_\rho^T \cdot \hat{v}_0\|}, \quad (\text{D67})$$

we can simplify the maximization in Eq. (D66) as follows:

$$\begin{aligned}
 \mathcal{M}_\rho &\leq \max_{\substack{\hat{v}_k, \hat{u}_k, \theta \\ \hat{v}_0 \perp \hat{v}_1}} 2 \left[ \cos \theta \hat{v}_0^T \cdot T_\rho \cdot \hat{u}_0 + \sin \theta \hat{v}_1^T \cdot T_\rho \cdot \hat{u}_1 \right] \\
 &= \max_{\substack{\hat{v}_k, \theta \\ \hat{v}_0 \perp \hat{v}_1}} 2 \left[ \cos \theta \|T_\rho^T \cdot \hat{v}_0\| + \sin \theta \|T_\rho^T \cdot \hat{v}_1\| \right] \\
 &= \max_{\substack{\hat{v}_k \\ \hat{v}_0 \perp \hat{v}_1}} 2 \sqrt{\|T_\rho^T \cdot \hat{v}_0\|^2 + \|T_\rho^T \cdot \hat{v}_1\|^2}. \quad (\text{D68})
 \end{aligned}$$

Finally, by employing the result of Lemma 2 in Eq. (D68), we obtain the statement, Eq. (10), of the theorem:

$$\mathcal{M}_\rho \leq \sqrt{t_0 + t_1}, \quad (\text{D69})$$

where  $t_0$  and  $t_1$  are the two largest eigenvalues of  $T_\rho T_\rho^T$ . This concludes the proof.  $\blacksquare$

### 1. Tightness conditions

Here we derive the conditions for which the upper bound on the MABK violation given in Eq. (10) is tight. That is, there exist observables for the  $N$  parties such that the violation achieved on the state  $\rho$  is exactly given by the rhs of Eq. (10). We first address the case  $N/2$  even since it is the one explicitly derived in the proof, then we present the tightness conditions valid in the other cases.

The bound is tight when equality holds in Eq. (D66). Considering that we made specific choices for the unit vectors  $\hat{v}_i$  and  $\hat{u}_i$  and for  $\theta$ , the vectors in Eqs. (D31)–(D34) should comply with these specific choices. In particular, consider the eigenvalue equation for  $T_\rho T_\rho^T$  with normalized eigenvectors and where  $t_0$  and  $t_1$  are the two largest eigenvalues:

$$T_\rho T_\rho^T \hat{t}_k = t_k \hat{t}_k. \quad (\text{D70})$$

In order to use Lemma 2 in Eq. (D68), it must hold that

$$\hat{v}_k = \frac{\vec{v}_k}{\|\vec{v}_k\|} = \hat{t}_k \quad k = 0, 1, \quad (\text{D71})$$

where  $\vec{v}_k$  ( $k = 0, 1$ ) are defined in Eqs. (D31) and (D32). Employing Eq. (D71) into the relation, Eq. (D67), that fixes  $\theta$  we get

$$\frac{\|\vec{v}_1\|}{\|\vec{v}_0\|} = \tan \theta = \frac{\|T_\rho^T \cdot \hat{t}_1\|}{\|T_\rho^T \cdot \hat{t}_0\|} = \sqrt{\frac{t_1}{t_0}}, \quad (\text{D72})$$

where the last equality is due to Eq. (D70). Combining Eq. (D72) with property Eq. (D35) we completely fix the norms of vectors  $\vec{v}_0$  and  $\vec{v}_1$ , while their direction is already fixed by Eq. (D71). In conclusion we get the following tightness conditions for  $\vec{v}_0$  and  $\vec{v}_1$ , which we recall

being specific combinations, Eqs. (D31) and (D32), of the parties' measurement directions:

$$\vec{v}_k = 2^{N/4} \sqrt{\frac{t_k}{t_0 + t_1}} \hat{t}_k \quad k = 0, 1. \quad (\text{D73})$$

In addition to this, we also fix the directions  $\hat{u}_0$  and  $\hat{u}_1$  to those of  $T_\rho^T \cdot \hat{v}_0$  and  $T_\rho^T \cdot \hat{v}_1$ , respectively. Due to Eq. (D71) and recalling property, Eq. (D36), we derive the following tightness conditions on  $\vec{u}_0$  and  $\vec{u}_1$ :

$$\vec{u}_k = \frac{2^{N/4}}{\sqrt{t_k}} T_\rho^T \hat{t}_k \quad k = 0, 1. \quad (\text{D74})$$

One can verify that upon substituting the tightness conditions, Eqs. (D73) and (D74), into the MABK expectation value, Eq. (D30), the theorem claim is obtained.

Here we recapitulate the tightness conditions of Theorem 2 for the two cases  $N$  even and  $N$  odd. The bound in Eq. (10) is tight if there exist unit vectors  $\vec{a}_0^i, \vec{a}_1^i$  (with  $i = 1, \dots, N$ ) such that

(a)  $N$  even:

$$\vec{v}_k = 2^{N/4} \sqrt{\frac{t_k}{t_0 + t_1}} \hat{t}_k, \quad \vec{u}_k = \frac{2^{N/4}}{\sqrt{t_k}} T_\rho^T \hat{t}_k \quad (k = 0, 1), \quad (\text{D75})$$

where vectors  $\vec{v}_k$  and  $\vec{u}_k$  are defined in Eqs. (D31)–(D34) if  $N/2$  is even, or as follows if  $N/2$  is odd:

$$\vec{v}_0 = \sum_{\mathbf{x} \in \mathcal{E}_{N/2}} (-1)^{(N-2)/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{x_i}^i, \quad (\text{D76})$$

$$\vec{v}_1 = \sum_{\mathbf{x} \in \mathcal{O}_{N/2}} (-1)^{(N-2)/4 - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{x_i}^i, \quad (\text{D77})$$

$$\vec{u}_0 = \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor} \bigotimes_{i=1}^{N/2} \vec{a}_{y_i}^{N/2+i}, \quad (\text{D78})$$

$$\vec{u}_1 = \sum_{\mathbf{y} \in \{0,1\}^{N/2}} (-1)^{\lceil \omega(\mathbf{y})/2 \rceil} \bigotimes_{i=1}^{N/2} \vec{a}_{y_i}^{N/2+i}, \quad (\text{D79})$$

where the sets  $\mathcal{E}_{N/2}$  and  $\mathcal{O}_{N/2}$  are defined in Eqs. (D27) and (D28), respectively.

(b)  $N$  odd:

$$\vec{v}_k = 2^{(N+1)/4} \sqrt{\frac{t_k}{t_0 + t_1}} \hat{t}_k, \quad \vec{u}_k = \frac{2^{(N-3)/4}}{\sqrt{t_k}} T_\rho^T \hat{t}_k \quad (k = 0, 1), \quad (\text{D80})$$

where vectors  $\vec{v}_k$  and  $\vec{u}_k$  are defined as follows:

$$\vec{v}_0 = \sum_{\mathbf{x} \in \mathcal{E}_{(N+1)/2}} (-1)^{\lfloor (N-1)/4 \rfloor - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{(N+1)/2} \vec{a}_{x_i}^i, \quad (\text{D81})$$

$$\vec{v}_1 = \sum_{\mathbf{x} \in \mathcal{O}_{(N+1)/2}} (-1)^{\lfloor (N-1)/4 \rfloor - \lfloor \omega(\mathbf{x})/2 \rfloor} \bigotimes_{i=1}^{(N+1)/2} \vec{a}_{x_i}^i, \quad (\text{D82})$$

$$\vec{u}_0 = \sum_{\mathbf{y} \in \mathcal{J}_{(N-1)/2}} (-1)^{\lfloor \omega(\mathbf{y})/2 \rfloor} \bigotimes_{i=1}^{(N-1)/2} \vec{a}_{y_i}^{(N+1)/2+i}, \quad (\text{D83})$$

$$\vec{u}_1 = \sum_{\mathbf{y} \in \bar{\mathcal{J}}_{(N-1)/2}} (-1)^{\lceil \omega(\mathbf{y})/2 \rceil} \bigotimes_{i=1}^{(N-1)/2} \vec{a}_{y_i}^{(N+1)/2+i}, \quad (\text{D84})$$

with the sets  $\mathcal{J}_{(N-1)/2}$  and  $\bar{\mathcal{J}}_{(N-1)/2}$  fixed as

$$\mathcal{J}_{(N-1)/2} = \left\{ \mathbf{x} \in \{0, 1\}^{(N-1)/2} \mid \omega(\mathbf{x}) = \frac{N-1}{2} \bmod 2 \right\} \quad (\text{D85})$$

$$\bar{\mathcal{J}}_{(N-1)/2} = \left\{ \mathbf{x} \in \{0, 1\}^{(N-1)/2} \mid \omega(\mathbf{x}) + 1 = \frac{N-1}{2} \bmod 2 \right\}. \quad (\text{D86})$$

Note that, similarly to the  $N=3$  case discussed in Sec. VIII, one can potentially obtain tighter MABK violation upper bounds accompanied by the corresponding tightness conditions if one employs variations (in terms of row and column definitions) of the correlation matrix given in Definition 3.

#### APPENDIX E: ANALYTICAL PROOF OF THE LOWER BOUND ON $H(X|E)_{\rho_\alpha}$

In this appendix we derive the analytical solution of the optimization problem in Eq. (24), which we report here for clarity:

$$\begin{aligned} H(X|E)_{\rho_\alpha}^\downarrow(m) &= \min_{\{\rho_{ijk}\}} 1 - H(\{\rho_{ijk}\}) + H(\{\rho_{ijk} + \rho_{ij\bar{k}}\}) \\ &\text{subject to } \mathcal{M}_\alpha^\uparrow \geq m; \quad \rho_{0jk} \geq \rho_{1jk}; \\ &\quad \sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned} \quad (\text{E1})$$

where the upper bound on the MABK violation is given in Corollary 1, where the second constraint is given in Eq. (19) and where  $m \geq 2\sqrt{2}$ , otherwise the conditional entropy is null (see Fig. 2). For ease of notation, we drop the subscript  $\alpha$  in the observed violation and we indicate the objective function of the optimization problem as  $H(X|E)_{\rho_\alpha}$ .

Because of the symmetry of the problem, we can assume without loss of generality that the largest element in  $\{\rho_{ijk}\}$  is  $\rho_{000}$ . Then, a necessary condition such that  $\mathcal{M}_\alpha^\uparrow \geq 2\sqrt{2}$  is given by  $\rho_{000} \geq 1/2$ . Indeed, the following upper bound on  $\mathcal{M}_\alpha^\uparrow$ :

$$\begin{aligned} \mathcal{M}_\alpha^\uparrow &= 4 \sqrt{\sum_{j,k=0}^1 (\rho_{0jk} - \rho_{1jk})^2} \leq 4 \sqrt{\sum_{j,k=0}^1 \rho_{0jk}^2} \\ &\leq 4 \sqrt{\sum_{j,k=0}^1 \rho_{000} \cdot \rho_{0jk}} = 4 \sqrt{\rho_{000} \left( \sum_{j,k=0}^1 \rho_{0jk} \right)} \\ &\leq 4\sqrt{\rho_{000}}, \end{aligned} \quad (\text{E2})$$

is greater than or equal to  $2\sqrt{2}$  when  $\rho_{000} \geq 1/2$ .

Note that, by definition, the minimal entropy  $H(X|E)_{\rho_\alpha}^\downarrow(m)$  in Eq. (E1) is monotonically nondecreasing in  $m$ .

The upper bound on the maximal MABK violation, Eq. (12), is tight on the following class of states (the tightness conditions are verified):

$$\tau(v) = v |\psi_{0,0,0}\rangle \langle \psi_{0,0,0}| + (1-v) |\psi_{0,1,1}\rangle \langle \psi_{0,1,1}|, \quad (\text{E3})$$

and reads in this case

$$\mathcal{M}_\tau(v) = \mathcal{M}_\tau^\uparrow(v) = 4\sqrt{v^2 + (1-v)^2}. \quad (\text{E4})$$

It is straightforward to verify that

$$\mathcal{M}_\tau(\rho_{000}) \geq \mathcal{M}_\alpha^\uparrow \quad \forall \{\rho_{ijk}\}. \quad (\text{E5})$$

Moreover, the objective function of the minimization, when evaluated on the states, Eq. (E3), reads

$$H(X|E)_{\tau(v)} = 1 - h(v), \quad (\text{E6})$$

where we use the binary entropy  $h(p) = -p \log p - (1-p) \log(1-p)$ . Here and in the following, “log” represents the logarithm in base 2.

By definition, the entropy minimized over all the states with  $\mathcal{M}_\alpha^\uparrow \geq m$ , Eq. (E1), is upper bounded by the entropy of any particular state with  $\mathcal{M}_\alpha^\uparrow = m$ :

$$H(X|E)_{\rho_\alpha}^\downarrow(m) \leq H(X|E)_{\tau(v_m)}, \quad (\text{E7})$$

where  $v_m$  is fixed such that the maximal violation of the state  $\tau(v_m)$  is given by  $m$ :

$$\mathcal{M}_\tau^\uparrow(v_m) = 4\sqrt{v_m^2 + (1-v_m)^2} = m. \quad (\text{E8})$$

On the other hand, in the following we prove that

$$H(X|E)_{\rho_\alpha} \geq H(X|E)_{\tau(\rho_{000})} \quad \forall \{\rho_{ijk}\}, \quad (\text{E9})$$

where  $\rho_{000} \geq 1/2$  is the largest element in  $\{\rho_{ijk}\}$ . In particular, the last expression holds for the state  $\rho_\alpha^*$ , which is the

solution of the minimization in Eq. (E1):

$$\begin{aligned} H(X|E)_{\rho_\alpha}^\downarrow(m) &= H(X|E)_{\rho_\alpha^*} \geq H(X|E)_{\tau(\rho_{000}^*)} \\ &\geq H(X|E)_{\tau(v_m)}. \end{aligned} \quad (\text{E10})$$

The last inequality in Eq. (E10) is due to a couple of observations. Firstly, by applying Eq. (E5) to the state  $\rho_\alpha^*$  we obtain  $\mathcal{M}_\tau(\rho_{000}^*) \geq m$ , which combined with Eq. (E8) implies that  $\rho_{000}^* \geq v_m$  (in the interval of interest  $\rho_{000}^*, v_m \geq 1/2$ ). Then, we observe that the entropy of the states  $\tau$  in Eq. (E6) is monotonically increasing in the interval  $v \in [1/2, 1]$ . The two observations lead to the second inequality in Eq. (E10).

By combining Eq. (E10) with Eq. (E7), we obtain the desired lower bound:

$$H(X|E)_{\rho_\alpha}^\downarrow(m) = H(X|E)_{\tau(v_m)}. \quad (\text{E11})$$

Note that the family of states  $\tau(v)$  in Eq. (E3) minimizes the entropy for every observed violation  $m$ . The bound in Eq. (E11) can be expressed in terms of the violation  $m$  by reverting Eq. (E8) and by using it in Eq. (E6), thus obtaining Eq. (28).

We are thus left to prove the inequality in Eq. (E9), which can be recast as follows:

$$h(\rho_{000}) + H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}) - H(\{\rho_{ijk}\}) \geq 0. \quad (\text{E12})$$

To start with, we simplify the difference of the following entropies:

$$\begin{aligned} h(\rho_{000}) - H(\{\rho_{ijk}\}) &= -(1 - \rho_{000}) \log(1 - \rho_{000}) \\ &\quad + \sum_{(i,j,k) \neq (0,0,0)} \rho_{ijk} \log \rho_{ijk}. \end{aligned} \quad (\text{E13})$$

By substituting Eq. (E13) into the lhs of Eq. (E12), we get

$$\begin{aligned} H(X|E)_{\rho_\alpha} - H(X|E)_{\tau(\rho_{000})} &= H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}) \\ &\quad + \sum_{(i,j,k) \neq (0,0,0)} \rho_{ijk} \log \rho_{ijk} - (1 - \rho_{000}) \log(1 - \rho_{000}). \end{aligned} \quad (\text{E14})$$

We then apply Jensen's inequality

$$f(x+y) \geq \frac{f(2x) + f(2y)}{2}, \quad (\text{E15})$$

where  $f(x) = -x \log x$  is a concave function, to the last three terms of the first entropy in Eq. (E14):

$$\begin{aligned} H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}) &= -(\rho_{000} + \rho_{011}) \log(\rho_{000} + \rho_{011}) \\ &\quad - (\rho_{001} + \rho_{010}) \log(\rho_{001} + \rho_{010}) \\ &\quad - (\rho_{100} + \rho_{111}) \log(\rho_{100} + \rho_{111}) \\ &\quad - (\rho_{101} + \rho_{110}) \log(\rho_{101} + \rho_{110}), \end{aligned} \quad (\text{E16})$$

such that we get

$$\begin{aligned} H(\{\rho_{ijk} + \rho_{i\bar{j}\bar{k}}\}) &\geq -(\rho_{000} + \rho_{011}) \log(\rho_{000} + \rho_{011}) \\ &\quad + \sum_{(i,j,k) \neq \begin{smallmatrix} (0,0,0) \\ (0,1,1) \end{smallmatrix}} -\rho_{ijk} \log(2\rho_{ijk}) \\ &= -(\rho_{000} + \rho_{011}) \log(\rho_{000} + \rho_{011}) - (1 - \rho_{000} - \rho_{011}) \\ &\quad + \sum_{(i,j,k) \neq \begin{smallmatrix} (0,0,0) \\ (0,1,1) \end{smallmatrix}} -\rho_{ijk} \log \rho_{ijk}. \end{aligned} \quad (\text{E17})$$

With this result, the difference of entropies in Eq. (E14) can be estimated by

$$\begin{aligned} H(X|E)_{\rho_\alpha} - H(X|E)_{\tau(\rho_{000})} &\geq -(\rho_{000} + \rho_{011}) \log(\rho_{000} + \rho_{011}) \\ &\quad - (1 - \rho_{000}) \log[2(1 - \rho_{000})] \\ &\quad + \rho_{011} \log(2\rho_{011}) \\ &=: g(\rho_{000}, \rho_{011}). \end{aligned} \quad (\text{E18})$$

In the function  $g$  the first two terms are positive and the last is negative. We further analyze and estimate the function  $g(x, y)$  in the range of interest, i.e.,  $1/2 \leq x \leq 1$ ,  $0 \leq y \leq 1 - x$ . In this range  $g(x, y)$  is concave in  $x$  because its second derivative is always negative:

$$\frac{\partial^2 g(x, y)}{\partial x^2} = -\frac{1}{\ln(2)} \left[ \frac{1}{(1-x)} + \frac{1}{(x+y)} \right] < 0. \quad (\text{E19})$$

Consider the boundary  $x + y = 1$  of  $g(x, y)$  for which we get  $g(1 - y, y) = 0$ . Due to the concavity it holds for  $0 \leq p \leq 1$  that

$$\begin{aligned} &g \left[ p \frac{1}{2} + (1-p)(1-y), y \right] \\ &\geq pg \left( \frac{1}{2}, y \right) + (1-p)g(1-y, y) \end{aligned}$$

or equivalently that

$$g(x, y) \geq \left( \frac{1-x-y}{\frac{1}{2}-y} \right) g \left( \frac{1}{2}, y \right). \quad (\text{E20})$$

Note that from the parameter regimes of  $x$  and  $y$  it follows that

$$0 \leq \left( \frac{1-x-y}{\frac{1}{2}-y} \right) \leq 1. \quad (\text{E21})$$

We finally analyze the properties of  $g(\frac{1}{2}, y)$ , which is convex in  $y$  as its second derivative is always positive:

$$\frac{\partial^2 g(\frac{1}{2}, y)}{\partial y^2} = \frac{1}{y \ln(2) + y^2 \ln(4)} > 0. \quad (\text{E22})$$

A convex function has a unique minimum if it exists in the parameter regime. In our case this is given by

$$\frac{\partial g(\frac{1}{2}, y)}{\partial y} = \log(2y) - \log\left(\frac{1}{2} + y\right) = 0 \quad \Leftrightarrow \quad y = \frac{1}{2} \quad (\text{E23})$$

for which  $g(\frac{1}{2}, \frac{1}{2}) = 0$  holds. Thus in general it holds

$$g\left(\frac{1}{2}, y\right) \geq 0. \quad (\text{E24})$$

By combining these considerations we obtain the desired inequality (E9):

$$\begin{aligned} H(X|E)_{\rho_\alpha} - H(X|E)_{\tau(\rho_{000})} &\stackrel{(\text{E18})}{\geq} g(\rho_{000}, \rho_{011}) \\ &\stackrel{(\text{E20})}{\geq} \left(\frac{1 - \rho_{000} - \rho_{011}}{\frac{1}{2} - \rho_{011}}\right) g\left(\frac{1}{2}, \rho_{011}\right) \\ &\geq 0, \end{aligned} \quad (\text{E25})$$

where in the last inequality we use the fact that the prefactor is positive, Eq. (E21) and that  $g(\frac{1}{2}, \rho_{011})$  is lower bounded by zero, Eq. (E24).

## APPENDIX F: ANALYTICAL PROOF OF THE LOWER BOUND ON $H(XY|E)_{\rho_\alpha}$

In order to derive an analytical lower bound on  $H(XY|E)_{\rho_\alpha}$ , we solve the simplified optimization problem in Eq. (35), where we can independently minimize the entropy over  $t, \varphi_X$  and  $\varphi_Y$  without affecting the MABK violation. We report Eq. (35) for clarity:

$$\begin{aligned} H(XY|E)_{\rho_\alpha}^\downarrow(m) &= \min_{\{\rho_{ijk}, t, \varphi_X, \varphi_Y\}} H(XY|E)_{\rho_\alpha} \\ &\text{subject to } \mathcal{M}_\alpha^\uparrow \geq m; \quad \rho_{0jk} \geq \rho_{1jk}; \\ &\sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned} \quad (\text{F1})$$

where  $\mathcal{M}_\alpha^\uparrow$  is the upper bound on the MABK violation derived in Corollary 1, while  $\varphi_X$  and  $\varphi_Y$  are the measurement directions of the outcomes  $X$  and  $Y$  in the  $(x, y)$  plane. We drop the subscript in  $m_\alpha$  for ease of notation.

Eve is assumed to hold the purifying system  $E$  of the state  $\rho_\alpha$  shared by Alice, Bob, and Charlie. The purification of  $\rho_\alpha$  can thus be written as follows:

$$|\Phi_{ABCE}^\alpha\rangle = \sum_{ijk} \sqrt{\rho_{ijk}} |\rho_{ijk}\rangle \otimes |e_{ijk}\rangle, \quad (\text{F2})$$

where  $|\rho_{ijk}\rangle$  are the eigenstates of  $\rho_\alpha$  defined in Eq. (17), while  $\{|e_{ijk}\rangle\}$  is an orthonormal basis of Eve's eight-dimensional Hilbert space  $\mathcal{H}_E$ .

We restrict our proof to states  $\rho_\alpha$  with a non-negative off-diagonal term  $s \geq 0$ , which corresponds to  $t \geq 0$  [see Eq. (18)]. Since  $t \in [-\pi/2, \pi/2]$  by definition, this means that we restrict ourselves to the region where  $\sin t \geq 0$  and  $\cos t \geq 0$ . The complementary case corresponds to states  $\rho_\alpha^*$ , which would lead to the same result. For this, we employ a parametrization of the eigenstates slightly different from Eq. (17), which reads as follows:

$$\begin{aligned} |\rho_{ijk}\rangle &= |\psi_{i,j,k}\rangle, \quad \text{for } (j, k) \neq (1, 1) \\ |\rho_{011}\rangle &= \sqrt{(1-p)} |\psi_{0,1,1}\rangle - i\sqrt{p} |\psi_{1,1,1}\rangle \\ |\rho_{111}\rangle &= \sqrt{p} |\psi_{0,1,1}\rangle + i\sqrt{(1-p)} |\psi_{1,1,1}\rangle, \end{aligned} \quad (\text{F3})$$

where  $|\psi_{i,j,k}\rangle$  are the GHZ basis states (Definition 1) and where  $p$  is completely fixed by  $t$  through the relation:

$$p = \frac{(\tan t)^2}{1 + (\tan t)^2}, \quad (\text{F4})$$

from which we deduce that  $0 \leq p \leq 1$  and that  $p = 0$  when  $t = 0$ .

From now on, we omit the subscript  $\rho_\alpha$  in the entropy symbol for ease of notation. We thus have that the conditional entropy  $H(XY|E)$  can be expressed as

$$\begin{aligned} H(XY|E) &= H(XY) + H(E|XY) - H(E) \\ &= 2 + H(E|XY) - H(\{\rho_{ijk}\}), \end{aligned} \quad (\text{F5})$$

where the last equation follows from the fact that all marginals have been symmetrized and from the fact that the state on  $ABCE$  is pure, Eq. (F2), thus  $H(E) = H(ABC) = H(\{\rho_{ijk}\})$ .

The proof of the analytical lower bound on  $H(XY|E)$  as a function of the MABK violation is subdivided in three parts: (i) we first derive an analytical expression for  $H(E|XY)$ ; (ii) we minimize  $H(E|XY)$  with respect to  $t, \varphi_X$  and  $\varphi_Y$ ; (iii) we proceed minimizing the resulting expression of Eq. (F5) given a fixed MABK violation. Note that we are allowed to minimize  $H(E|XY)$  over  $t, \varphi_X$  and  $\varphi_Y$  independently of  $H(E)$ , since the latter is independent of the mentioned optimization variables.

### Step 1. Analytical expression for $H(E|XY)$ .

In order to derive the analytical expression for  $H(E|XY)$ , we use the following lemma.



**Lemma 3.** *The following equality holds:*

$$H(E|XY) = H(C|XY). \quad (\text{F6})$$

*Proof.* The proof follows from the fact that the state shared by Charlie and Eve conditioned on the outcomes  $X = a$  of Alice and  $Y = b$  of Bob, is a pure state. Indeed, if projective measurements are applied to a pure state, the resulting state, conditioned on a specific outcome, remains pure. Moreover, for a pure state, the entropies of its subsystems are equal, which implies

$$H(E|X = a, Y = b) = H(C|X = a, Y = b). \quad (\text{F7})$$

Therefore,

$$H(E|XY) = \sum_{a,b} \Pr(a, b) H(E|X = a, Y = b) \quad (\text{F8})$$

$$= \sum_{a,b} \Pr(a, b) H(C|X = a, Y = b) \quad (\text{F9})$$

$$= H(C|XY). \quad (\text{F10})$$

■

Lemma 3 is of great use as Eve's system is described by an eight-dimensional Hilbert space, whereas Charlie is only in possession of a single qubit. So the computation of  $H(C|XY)$  is significantly simpler.

We obtain Charlie's state, conditioned on the outcomes  $X = a$  and  $Y = b$ , by partially tracing over Eve's degrees of freedom

$$\rho_{C_{ab}}^\alpha = \text{Tr}_E \left( |\phi_{CE_{ab}}^\alpha\rangle \langle \phi_{CE_{ab}}^\alpha| \right), \quad (\text{F11})$$

where  $|\phi_{CE_{ab}}^\alpha\rangle$  is the state of Charlie and Eve given that Alice and Bob obtain outcomes  $X = a$  and  $Y = b$ , respectively, which is determined by

$$\begin{aligned} & (P_{|a\rangle} \otimes P_{|b\rangle} \otimes \text{id}_{CE}) |\phi_{ABCE}^\alpha\rangle \langle \phi_{ABCE}^\alpha| (P_{|a\rangle} \otimes P_{|b\rangle} \otimes \text{id}_{CE}) \\ &= \frac{1}{4} P_{|a\rangle} \otimes P_{|b\rangle} \otimes |\phi_{CE_{ab}}^\alpha\rangle \langle \phi_{CE_{ab}}^\alpha|, \end{aligned} \quad (\text{F12})$$

where  $P_{|a\rangle} = |a\rangle\langle a|$  and  $P_{|b\rangle} = |b\rangle\langle b|$ . The projected state  $|\phi_{CE_{ab}}^\alpha\rangle$  can be computed using the definition of the purification given in Eq. (F2), the definition of the eigenstates in Eq. (F3), and the fact that the measurements performed by Alice and Bob are restricted to the  $(x, y)$  plane. Indeed, the measurements are defined by the following projectors:

$$\begin{aligned} |a\rangle_X &= \frac{1}{\sqrt{2}} [ |0\rangle + (-1)^a e^{i\varphi_X} |1\rangle ] \quad a \in \{0, 1\}, \\ |b\rangle_Y &= \frac{1}{\sqrt{2}} [ |0\rangle + (-1)^b e^{i\varphi_Y} |1\rangle ] \quad b \in \{0, 1\}. \end{aligned} \quad (\text{F13})$$

In the following we abbreviate  $\xi_a = (-1)^a e^{i\varphi_X}$  and  $\xi_b = (-1)^b e^{i\varphi_Y}$ . We then have that

$$\begin{aligned} |\phi_{CE_{ab}}^\alpha\rangle &= \sum_{\substack{ijk \\ jk \neq 11}} \frac{1}{\sqrt{2}} [ (\delta_{0j} + \delta_{1j} \xi_b) |k\rangle + (\delta_{0j} \xi_a + \delta_{1j} \xi_a \xi_b) (-1)^l |\bar{k}\rangle ] \otimes |e_{ljk}\rangle \sqrt{\rho_{ljk}} \\ &+ \left\{ \left[ \sqrt{(1-p)} - i\sqrt{p} \right] \xi_b |1\rangle + \left[ \sqrt{(1-p)} + i\sqrt{p} \right] \xi_a |0\rangle \right\} \otimes |e_{011}\rangle \sqrt{\rho_{011}} \\ &+ \left\{ \left[ \sqrt{p} + i\sqrt{(1-p)} \right] \xi_b |1\rangle + \left[ \sqrt{p} - i\sqrt{(1-p)} \right] \xi_a |0\rangle \right\} \otimes |e_{111}\rangle \sqrt{\rho_{111}}. \end{aligned} \quad (\text{F14})$$

Finally, the partial trace over Eve results in

$$\rho_{C_{ab}}^\alpha = \text{Tr}_E \left( |\phi_{CE_{ab}}^\alpha\rangle \langle \phi_{CE_{ab}}^\alpha| \right) \quad (\text{F15})$$

$$\begin{aligned} &= \left\{ \sum_{\substack{ijk \\ jk \neq 11}} \frac{1}{\sqrt{2}} [ (\delta_{0j} + \delta_{1j} \xi_b) |k\rangle + (\delta_{0j} \xi_a + \delta_{1j} \xi_a \xi_b) (-1)^l |\bar{k}\rangle ] \right\} \cdot (\text{h.c.}) \rho_{ijk} \\ &+ \left\{ \left[ \sqrt{(1-p)} - i\sqrt{p} \right] \xi_b |1\rangle + \left[ \sqrt{(1-p)} + i\sqrt{p} \right] \xi_a |0\rangle \right\} \cdot (\text{h.c.}) \rho_{011} \\ &+ \left\{ \left[ \sqrt{p} + i\sqrt{(1-p)} \right] \xi_b |1\rangle + \left[ \sqrt{p} - i\sqrt{(1-p)} \right] \xi_a |0\rangle \right\} \cdot (\text{h.c.}) \rho_{111}. \end{aligned} \quad (\text{F16})$$

As  $\rho_{C_{ab}}^\alpha$  is a qubit state, we can now analytically calculate its eigenvalues, which can be reduced to

$$\lambda_{1,2}(\rho_{C_{ab}}^\alpha) = \frac{1}{2} (1 \pm |C|), \quad (\text{F17})$$

where

$$\begin{aligned} C &= (\rho_{000} - \rho_{100}) \xi_a^2 + (\rho_{001} - \rho_{101}) (\xi_b^2)^* + (\rho_{010} - \rho_{110}) \xi_a^2 (\xi_b^2)^* + (\rho_{011} - \rho_{111}) \left[ 1 - 2p - 2i\sqrt{p(1-p)} \right] \\ &= (\rho_{000} - \rho_{100}) e^{i2\varphi_X} + (\rho_{001} - \rho_{101}) e^{-i2\varphi_Y} + (\rho_{010} - \rho_{110}) e^{i2(\varphi_X - \varphi_Y)} + (\rho_{011} - \rho_{111}) e^{i\varphi_3}, \end{aligned} \quad (\text{F18})$$

where  $\varphi_3$  is a function of the parameter  $p$ . We see that the eigenvalues do not depend on the measurement outcomes  $a$  and  $b$  of Alice and Bob. The entropy is then given by

$$H(E|XY) = H(C|XY) = h \left[ \frac{1}{2} (1 + |C|) \right], \quad (\text{F19})$$

where  $h(x) = -x \log x - (1-x) \log(1-x)$  is the binary entropy.

### Step 2. Minimization of $H(E|XY)$

Minimizing the binary entropy in Eq. (F19), with respect to the measurement directions and the parameter  $p$ , is equivalent to maximizing the largest eigenvalue of  $\rho_{C_{ab}}^\alpha$ . The optimum can directly be deduced from Eq. (F17). Since it holds that  $(\rho_{0jk} - \rho_{1jk}) \geq 0 \forall j, k$ , the largest eigenvalue is maximized if

$$e^{i2\varphi_X} = e^{-i2\varphi_Y} = e^{i2(\varphi_X - \varphi_Y)} = e^{i\varphi_3}, \quad (\text{F20})$$

which holds, e.g., for  $\varphi_X = \varphi_Y = \varphi_3 = 0$ . Since  $\varphi_3 = 0$  implies  $p = t = 0$ , we verify that in the minimization of the conditional entropy of two parties' outcomes,  $H(XY|E)$ , it would be optimal for Eve to distribute a GHZ-diagonal state, which Alice and Bob measure in the  $X$  basis. The largest eigenvalue of  $\rho_{C_{ab}}^\alpha$  is then given by

$$\bar{\lambda} := \sum_{jk} \rho_{0jk}, \quad (\text{F21})$$

where we use the normalization of the eigenvalues to eliminate the terms  $\rho_{1jk}$ . The lower bound on the conditional entropy  $H(E|XY)$  is thus given by

$$H(E|XY) \geq h(\bar{\lambda}). \quad (\text{F22})$$

### Step 3. Minimization of $H(XY|E)$ with given MABK violation

By using Eq. (F22) in Eq. (F5), we can concentrate on minimizing the following expression:

$$\begin{aligned} H(E|XY) - H(E) &\geq h(\bar{\lambda}) - H(\{\rho_{ijk}\}) \\ &= -\bar{\lambda} \log \bar{\lambda} - (1 - \bar{\lambda}) \log(1 - \bar{\lambda}) \\ &\quad + \sum_{ijk} \rho_{ijk} \log \rho_{ijk} \\ &= \bar{\lambda} \sum_{jk} \frac{\rho_{0jk}}{\bar{\lambda}} \log \frac{\rho_{0jk}}{\bar{\lambda}} + (1 - \bar{\lambda}) \sum_{jk} \frac{\rho_{1jk}}{(1 - \bar{\lambda})} \log \frac{\rho_{1jk}}{(1 - \bar{\lambda})} \\ &= -\bar{\lambda} H \left( \left\{ \frac{\rho_{0jk}}{\bar{\lambda}} \right\} \right) - (1 - \bar{\lambda}) H \left( \left\{ \frac{\rho_{1jk}}{(1 - \bar{\lambda})} \right\} \right), \end{aligned} \quad (\text{F23})$$

where we use the definition of  $\bar{\lambda}$  in Eq. (F21). We now use the concavity of the Shannon entropy over probability distributions  $\vec{u}$  and  $\vec{v}$ , i.e.,

$$\bar{\lambda} H(\vec{u}) + (1 - \bar{\lambda}) H(\vec{v}) \leq H[\bar{\lambda} \vec{u} + (1 - \bar{\lambda}) \vec{v}], \quad (\text{F24})$$

in Eq. (F23) and obtain

$$H(E|XY) - H(E) \geq -H(\{\rho_{0jk} + \rho_{1jk}\}). \quad (\text{F25})$$

With the lower bound obtained in Eq. (F25), the optimization problem we have to solve is now the following:

$$\begin{aligned} &\max_{\{\rho_{ijk}\}} H(\{\rho_{0jk} + \rho_{1jk}\}) \\ &\text{subject to} \quad \frac{m^2}{16} \leq \sum_{jk} (\rho_{0jk} - \rho_{1jk})^2; \\ &\sum_{ijk} \rho_{ijk} = 1; \quad \rho_{ijk} \geq 0, \end{aligned} \quad (\text{F26})$$

where  $m$  is the observed MABK violation. Now notice that for every solution  $\{\rho_{0jk}, \rho_{1jk}\}$  of the maximization problem, there exists another equivalent solution—i.e., that leads to the same value for  $H(\{\rho_{0jk} + \rho_{1jk}\})$ —of the form  $\{\rho'_{0jk} = \rho_{0jk} + \rho_{1jk}, \rho'_{1jk} = 0\}$ . Therefore, we can restrict

the optimization to the solutions of that form:

$$\begin{aligned}
 & \max_{\{\rho_{0jk}\}} H(\{\rho_{0jk}\}) \\
 \text{subject to} & \quad \frac{m^2}{16} \leq \sum_{jk} \rho_{0jk}^2 \\
 & \quad \sum_{jk} \rho_{0jk} = 1 \\
 & \quad \rho_{000} \geq \rho_{001} \geq \rho_{010} \geq \rho_{011} \geq 0,
 \end{aligned} \tag{F27}$$

where we impose the ordering of the four remaining eigenvalues  $\{\rho_{0jk}\}$  without loss of generality, since the optimization problem is symmetric with respect to their permutations.

We thus reduce the problem to the constrained maximization of  $H(\{\rho_{0jk}\})$ , as described in Eq. (F27). In the following calculations, we rescale the function  $H(\{\rho_{0jk}\})$  by  $\ln 2$ , so that it is expressed in terms of natural logarithms instead of the logarithm in base 2. This simplifies the notation when computing its derivatives but does not change the solution of the optimization problem.

We use the Karush-Kuhn-Tucker multipliers method [76,77] to identify necessary conditions for extremal points of the optimization problem in Eq. (F27). The Lagrangian for our maximization problem is then given by

$$\begin{aligned}
 \mathcal{L}(\rho_{000}, \rho_{001}, \rho_{010}, \rho_{011}, u, v) &= H(\rho_{000}, \rho_{001}, \rho_{010}, \rho_{011}) \\
 &+ u \left( \rho_{000}^2 + \rho_{001}^2 + \rho_{010}^2 + \rho_{011}^2 - \frac{m^2}{16} \right) \\
 &+ v (\rho_{000} + \rho_{001} + \rho_{010} + \rho_{011} - 1).
 \end{aligned}$$

The necessary conditions to have an extremal point are given by the solution of the following system:

$$\begin{cases}
 \nabla_{\rho_{0jk}} \mathcal{L} = 0 \\
 \rho_{000}^2 + \rho_{001}^2 + \rho_{010}^2 + \rho_{011}^2 \geq \frac{m^2}{16} \\
 \rho_{000} + \rho_{001} + \rho_{010} + \rho_{011} = 1 \\
 u \geq 0 \\
 u \left( \rho_{000}^2 + \rho_{001}^2 + \rho_{010}^2 + \rho_{011}^2 - \frac{m^2}{16} \right) = 0.
 \end{cases} \tag{F28}$$

The last equation in Eq. (F28) implies that either  $u = 0$  or the inequality constraint holds with the equal sign. Let us first consider the case that  $u = 0$  and compute the derivative of  $\mathcal{L}$  with respect to  $\rho_{0jk}$  in the first equation of Eq. (F28):

$$\frac{\partial \mathcal{L}}{\partial \rho_{0jk}} = -\ln \rho_{0jk} + v - 1 = 0 \quad \forall \rho_{0jk}. \tag{F29}$$

Since the logarithm is a monotonic function, the set of equations in the last expression imply one of the following cases:

- (a)  $\rho_{000} = \rho_{001} = \rho_{010} = \rho_{011}$ ,
- (b)  $\rho_{000} = \rho_{001} = \rho_{010}$  and  $\rho_{011} = 0$ ,
- (c)  $\rho_{000} = \rho_{001}$  and  $\rho_{010} = \rho_{011} = 0$ ,
- (d)  $\rho_{001} = \rho_{010} = \rho_{011} = 0$ ,

where we account for the border conditions, i.e., when one or more  $\rho_{0jk}$  are equal to zero.

By combining the equality conditions with the constraint that  $\rho_{0jk}$  sum to one, we can easily obtain the solution of the system, Eq. (F28), for each of the above cases. Note that the inequality constraint is still valid, therefore the derived solutions will only hold for certain values of  $m$ :

- (a)  $H(\{\rho_{0jk}\}) = 2$ , valid for  $m \leq 2$ ,
- (b)  $H(\{\rho_{0jk}\}) = \log 3$ , valid for  $m \leq 4/\sqrt{3}$ ,
- (c)  $H(\{\rho_{0jk}\}) = 1$ , valid for  $m \leq 2\sqrt{2}$ ,
- (d)  $H(\{\rho_{0jk}\}) = 0$ , valid for  $m \leq 4$ .

Cases (a) and (d) are useless since the former is never valid in the range of interest for the observed violation (i.e., above the classical bound), while the latter leads to zero entropy, which is definitely not the solution of our maximization problem.

Let us consider now the case  $u > 0$ , which implies that the inequality constraint becomes an equality [the last equation in Eq. (F28) must be satisfied]. We compute the derivatives in the first equation of Eq. (F28):

$$\frac{\partial \mathcal{L}}{\partial \rho_{0jk}} = 2\rho_{0jk}u - \ln \rho_{0jk} + v - 1 = 0 \quad \forall \rho_{0jk}. \tag{F30}$$

Notice that the function  $g(x) = ax - \ln x + b$  can have at most two roots (zero points), because

$$g'(x) = a - \frac{1}{x}, \tag{F31}$$

has at most a single root (zero point), corresponding to one extremum for  $g(x)$ . It follows that there can be at most a single  $y \neq x$  such that  $g(x) = g(y) = 0$ . The potential critical points of the Lagrangian  $\mathcal{L}$  are hence restricted to the following cases (remember we use the ordering  $\rho_{000} \geq \rho_{001} \geq \rho_{010} \geq \rho_{011} \geq 0$ )

- (i)  $\rho_{000} = \rho_{001} = \rho_{010} = \rho_{011}$ ,
- (ii)  $\rho_{000} = \rho_{001} = \rho_{010} > \rho_{011}$ ,

- (iii)  $\rho_{000} > \rho_{001} = \rho_{010} = \rho_{011}$ ,
- (iv)  $\rho_{000} = \rho_{001} > \rho_{010} = \rho_{011}$ .

We again account for the border conditions, and analog conditions directly follow in case some  $\rho_{0jk}$  are zero:

- (v)  $\rho_{000} = \rho_{001} = \rho_{010}$  and  $\rho_{011} = 0$ ,
- (vi)  $\rho_{000} = \rho_{001} > \rho_{010}$  and  $\rho_{011} = 0$ ,
- (vii)  $\rho_{000} > \rho_{001} = \rho_{010}$  and  $\rho_{011} = 0$ ,
- (viii)  $\rho_{000} > \rho_{001}$  and  $\rho_{010} = \rho_{011} = 0$ ,
- (ix)  $\rho_{000} = \rho_{001}$  and  $\rho_{010} = \rho_{011} = 0$ ,
- (x)  $\rho_{001} = \rho_{010} = \rho_{011} = 0$ .

Note that in all the listed cases there are a maximum of two distinct eigenvalues, which are thus completely fixed by the two equality constraints. Moreover, we observe that the cases (i), (v), (ix), and (x) correspond to the already investigated cases (a), (b), (c), and (d), respectively.

Analyzing the resulting entropy  $H$  as a function of the MABK violation  $m$  for each of the ten possible extremal points, we conclude that the maximum is achieved for the case (iii) for every value of  $m$ . In this case, the eigenvalues are fixed to

$$\rho_{000} = \frac{1}{8} \left( 2 + \sqrt{3} \sqrt{m^2 - 4} \right) =: v_m \quad (\text{F32})$$

$$\rho_{0jk} = \frac{(1 - v_m)}{3} \quad (j, k) \neq (0, 0). \quad (\text{F33})$$

The solution of the optimization problem in Eq. (F27) then reads

$$H(\{\rho_{0jk}\}) = H\left(\left\{v_m, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}\right\}\right). \quad (\text{F34})$$

The lower bound on the entropy difference (F25) is thus given by

$$\begin{aligned} & H(E|XY) - H(E) \\ & \geq -H\left(\left\{v_m, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}\right\}\right). \end{aligned} \quad (\text{F35})$$

Finally, we can lower bound the entropy of Alice and Bob's outcomes given Eve's quantum side information by

$$\begin{aligned} H(XY|E) &= 2 + H(E|XY) - H(E) \\ &\geq 2 - H\left(\left\{v_m, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}, \frac{1 - v_m}{3}\right\}\right), \end{aligned} \quad (\text{F36})$$

with

$$v_m = \frac{1}{4} + \frac{\sqrt{3}}{8} \sqrt{m^2 - 4}. \quad (\text{F37})$$

The rhs of Eq. (F36) is the bound reported in Eq. (36).

- [1] E. Commission, The quantum flagship. <https://qt.eu>. [Online].
- [2] I. Q. Technology, Quantum key distribution (qkd) markets: 2019–2028. <https://www.insidequantumtechnology.com/product/quantum-key-distribution-qkd-markets-2019-2028>. [Online].
- [3] C. H. Bennett and G. Brassard, in *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (Steering Committee, Bangalore, 1984), p. 175.
- [4] D. Bruß, Optimal Eavesdropping in Quantum Cryptography with six States, *Phys. Rev. Lett.* **81**, 3018 (1998).
- [5] A. K. Ekert, Quantum Cryptography Based on Bell's Theorem, *Phys. Rev. Lett.* **67**, 661 (1991).
- [6] R. Renner, Security of quantum key distribution, *Int. J. Quantum Inf.* **6**, 1 (2008).
- [7] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dusek, N. Lütkenhaus, and M. Peev, The security of practical quantum key distribution, *Rev. Mod. Phys.* **81**, 1301 (2009).
- [8] H.-K. Lo, M. Curty, and K. Tamaki, Secure quantum key distribution, *Nat. Photonics* **8**, 595 (2014).
- [9] E. Diamanti, H.-K. Lo, B. Qi, and Z. Yuan, Practical challenges in quantum key distribution, *npj Quantum Inf.* **2**, 16025 (2016).
- [10] S. Pirandola, U. L. Andersen, L. Banchi, M. Berta, D. Bunandar, R. Colbeck, D. Englund, T. Gehring, C. Lupo, C. Ottaviani, *et al.*, Advances in quantum cryptography, [arXiv:quant-ph/1906.01645](https://arxiv.org/abs/quant-ph/1906.01645) (2019).
- [11] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, Quantum random number generation, *npj Quantum Inf.* **2**, 16021 (2016).
- [12] M. Herrero-Collantes and J. C. Garcia-Escartin, Quantum random number generators, *Rev. Mod. Phys.* **89**, 015004 (2017).
- [13] M. Epping, H. Kampermann, and D. Bruß, Large-scale quantum networks based on graphs, *New J. Phys.* **18**, 053036 (2016).
- [14] M. Epping, H. Kampermann, and D. Bruß, Robust entanglement distribution via quantum network coding, *New J. Phys.* **18**, 103052 (2016).
- [15] A. Pirker, J. Wallnöfer, and W. Dür, Modular architectures for quantum networks, *New J. Phys.* **20**, 053054 (2018).
- [16] F. Hahn, A. Pappa, and J. Eisert, Quantum network routing and local complementation, *npj Quantum Inf.* **5**, 76 (2019).
- [17] V. Krutyanskiy, M. Meraner, J. Schupp, V. Krcmarsky, H. Hainzer, and B. P. Lanyon, Light-matter entanglement over 50 km of optical fibre, *npj Quantum Inf.* **5**, 72 (2019).
- [18] A. Tchebotareva, S. L. N. Hermans, P. C. Humphreys, D. Voigt, P. J. Harmsma, L. K. Cheng, A. L. Verlaan, N. Dijkhuizen, W. de Jong, A. Dréau, *et al.*, Entanglement between a Diamond Spin Qubit and a Photonic Time-Bin Qubit at Telecom Wavelength, *Phys. Rev. Lett.* **123**, 063601 (2019).
- [19] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, *et al.*, Satellite-Relayed Intercontinental Quantum Network, *Phys. Rev. Lett.* **120**, 030501 (2018).
- [20] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, *Science* **362**, 6412 (2018).

- [21] M. Epping, H. Kampermann, C. Macchiavello, and D. Bruß, Multi-partite entanglement can speed up quantum key distribution in networks, *New J. Phys.* **19**, 093012 (2017).
- [22] F. Grasselli, H. Kampermann, and D. Bruß, Finite-key effects in multipartite quantum key distribution protocols, *New J. Phys.* **20**, 113014 (2018).
- [23] F. Grasselli, H. Kampermann, and D. Bruß, Conference key agreement with single-photon interference, *New J. Phys.* **21**, 123002 (2019).
- [24] Y. Wu, J. Zhou, X. Gong, Y. Guo, Z.-M. Zhang, and G. He, Continuous-variable measurement-device-independent multipartite quantum communication, *Phys. Rev. A* **93**, 022325 (2016).
- [25] Z. Zhang, R. Shi, and Y. Guo, Multipartite continuous variable quantum conferencing network with entanglement in the middle, *Appl. Sci.* **8**, 00 (2018).
- [26] R. L. C. Ottaviani, C. Lupo, and S. Pirandola, Modular network for high-rate quantum conferencing, *Commun. Phys.* **2**, 00 (2019).
- [27] G. Murta, F. Grasselli, H. Kampermann, and D. Bruß, Quantum conference key agreement: A review, *arXiv:quant-ph/2003.10186* (2020).
- [28] M. Proietti, J. Ho, F. Grasselli, P. Barrow, M. Malik, and A. Fedrizzi, Experimental quantum conference key agreement, *arXiv:quant-ph/2002.01491* (2020).
- [29] F. Hahn, J. de Jong, C. Thalacker, B. Demirel, S. Barz, and A. Pappa, Anonymous conference key agreement in quantum networks, *arXiv:quant-ph/2007.07995* (2020).
- [30] J. S. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, Cambridge, 2004).
- [31] A. Yao and D. Mayers, in *2013 IEEE 54th Annual Symposium on Foundations of Computer Science* (Los Alamitos, CA, USA, IEEE Computer Society, 1998), p. 503.
- [32] A. Acín, N. Gisin, and L. Masanes, From Bell's Theorem to Secure Quantum key Distribution, *Phys. Rev. Lett.* **97**, 120405 (2006).
- [33] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, *Phys. Rev. Lett.* **98**, 230501 (2007).
- [34] S. Pironio, A. Acín, N. Brunner, N. Gisin, S. Massar, and V. Scarani, Device-independent quantum key distribution secure against collective attacks, *New J. Phys.* **11**, 045021 (2009).
- [35] L. Masanes, S. Pironio, and A. Acín, Secure device-independent quantum key distribution with causally independent measurement devices, *Nat. Commun.* **2**, 238 (2011).
- [36] U. Vazirani and T. Vidick, Fully Device-Independent Quantum key Distribution, *Phys. Rev. Lett.* **113**, 140501 (2014).
- [37] S. Pironio, L. Masanes, A. Leverrier, and A. Acín, Security of Device-Independent Quantum key Distribution in the Bounded-Quantum-Storage Model, *Phys. Rev. X* **3**, 031007 (2013).
- [38] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, Practical device-independent quantum cryptography via entropy accumulation, *Nat. Commun.* **9**, 459 (2018).
- [39] T. Holz, H. Kampermann, and D. Bruß, Device-independent secret-key-rate analysis for quantum repeaters, *Phys. Rev. A* **97**, 012337 (2018).
- [40] V. Scarani and N. Gisin, Quantum Communication between  $n$  Partners and Bell's Inequalities, *Phys. Rev. Lett.* **87**, 117901 (2001).
- [41] V. Scarani and N. Gisin, Quantum key distribution between  $n$  partners: Optimal eavesdropping and bell's inequalities, *Phys. Rev. A* **65**, 012311 (2001).
- [42] J. Ribeiro, G. Murta, and S. Wehner, Reply to "comment on 'fully device-independent conference key agreement'", *Phys. Rev. A* **100**, 026302 (2019).
- [43] T. Holz, H. Kampermann, and D. Bruß, A genuine multipartite bell inequality for device-independent conference key agreement, *arXiv:quant-ph/1910.11360* (2019).
- [44] R. Colbeck, PhD thesis, University of Cambridge, 2007. Also available at: *arXiv:quant-ph/0911.3814*.
- [45] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, *et al.*, Random numbers certified by bell's theorem, *Nature* **464**, 1021 (2010).
- [46] R. Colbeck and A. Kent, Private randomness expansion with untrusted devices, *J. Phys. A: Math. Theor.* **44**, 095305 (2011).
- [47] C. A. Miller and Y. Shi, Robust protocols for securely expanding randomness and distributing keys using untrusted quantum devices, *J. ACM* **63**, 00 (2016).
- [48] S. Pironio and S. Massar, Security of practical private randomness generation, *Phys. Rev. A* **87**, 012336 (2013).
- [49] S. Fehr, R. Gelles, and C. Schaffner, Security and composability of randomness expansion from bell inequalities, *Phys. Rev. A* **87**, 012335 (2013).
- [50] E. Woodhead, B. Bourdoncle, and A. Acín, Randomness versus nonlocality in the Mermin-Bell experiment with three parties, *Quantum* **2**, 82 (2018).
- [51] M. Navascués, S. Pironio, and A. Acín, Bounding the set of Quantum Correlations, *Phys. Rev. Lett.* **98**, 010401 (2007).
- [52] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, *New J. Phys.* **10**, 073013 (2008).
- [53] O. Nieto-Silleras, S. Pironio, and J. Silman, Using complete measurement statistics for optimal device-independent randomness evaluation, *New J. Phys.* **16**, 013035 (2014).
- [54] J.-D. Bancal, L. Sheridan, and V. Scarani, More randomness from the same data, *New J. Phys.* **16**, 033011 (2014).
- [55] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, *Phys. Rev. Lett.* **23**, 880 (1969).
- [56] R. F. Werner and M. M. Wolf, All-multipartite bell-correlation inequalities for two dichotomic observables per site, *Phys. Rev. A* **64**, 032112 (2001).
- [57] N. D. Mermin, Extreme Quantum Entanglement in a Superposition of Macroscopically Distinct States, *Phys. Rev. Lett.* **65**, 1838 (1990).
- [58] M. Ardehali, Bell inequalities with a magnitude of violation that grows exponentially with the number of particles, *Phys. Rev. A* **46**, 5375 (1992).
- [59] A. V. Belinskii and D. N. Klyshko, Interference of light and bell's theorem, *Phys. Rev. A* **36**, 653 (1993).



- [60] R. Horodecki, P. Horodecki, and M. Horodecki, Violating bell inequality by mixed spin-12 states: Necessary and sufficient condition, *Phys. Lett. A* **200**, 340 (1995).
- [61] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Bell-Type Inequalities to Detect True  $n$ -body Nonseparability, *Phys. Rev. Lett.* **88**, 170405 (2002).
- [62] J. Ribeiro, G. Murta, and S. Wehner, Fully device-independent conference key agreement, *Phys. Rev. A* **97**, 022307 (2018).
- [63] M. A. Siddiqui and S. Sazim, Tight upper bound for the maximal expectation value of the mermin operators, *Quantum Inf. Process.* **18**, 131 (2019).
- [64] T. Vértesi and E. Bene, Two-qubit bell inequality for which positive operator-valued measurements are relevant, *Phys. Rev. A* **82**, 062115 (2010).
- [65] Y.-C. Liang and A. C. Doherty, Bounds on quantum correlations in bell-inequality experiments, *Phys. Rev. A* **75**, 042103 (2007).
- [66] M. Żukowski and C. Brukner, Bell's Theorem for General  $N$ -qubit States, *Phys. Rev. Lett.* **88**, 210401 (2002).
- [67] W. Laskowski, T. Paterek, M. Żukowski, and I. C. V. Brukner, Tight Multipartite Bell's Inequalities Involving Many Measurement Settings, *Phys. Rev. Lett.* **93**, 200401 (2004).
- [68] M. Li and S.-M. Fei, Bell inequalities for multipartite qubit quantum systems and their maximal violation, *Phys. Rev. A* **86**, 052119 (2012).
- [69] F. Dupuis and O. Fawzi, Entropy accumulation with improved second-order term, *IEEE Trans. Inf. Theory* **65**, 7596 (2019).
- [70] P. J. Brown, S. Ragy, and R. Colbeck, A framework for quantum-secure device-independent randomness expansion, *IEEE Trans. Inf. Theory* **66**, 2964 (2020).
- [71] G. Carrara, H. Kampermann, D. Bruß, and G. Murta, Genuine multipartite entanglement is not a precondition for secure conference key agreement, *arXiv:quant-ph/2007.11553* (2020).
- [72] T. Holz, D. Miller, H. Kampermann, and D. Bruß, Comment on “fully device-independent conference key agreement”, *Phys. Rev. A* **100**, 026301 (2019).
- [73] G. Tóth and O. Gühne, Entanglement detection in the stabilizer formalism, *Phys. Rev. A* **72**, 022340 (2005).
- [74] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, *IEEE Trans. Inf. Theory* **55**, 5840 (2009).
- [75] L. Masanes, Asymptotic Violation of Bell Inequalities and Distillability, *Phys. Rev. Lett.* **97**, 050503 (2006).
- [76] W. Karush, in *Traces and Emergence of Nonlinear Programming* (Springer Basel, Basel, 2014), p. 217.
- [77] H. W. Kuhn and A. W. Tucker, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, (Berkeley, California, University of California Press, 1951), p. 481.