Elemental and tight monogamy relations in nonsignalling theories

R. Augusiak,1 M. Demianowicz,1 M. Pawłowski,2 J. Tura,1 and A. Acín1,3

1ICFO–Institut de Ciències Fotòniques, 08860 Castelldefels (Barcelona), Spain
2Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański, PL-80-952 Gdańsk, Poland
3ICREA–Institut de Ciències Fotòniques, 08860 Castelldefels (Barcelona), Spain

Physical principles constrain the way nonlocal correlations can be distributed among distantly separated parties. These constraints are usually expressed by monogamy relations that bound the amount of Bell inequality violation observed among a set of parties by the violation observed by a different set of parties. We prove here that much stronger monogamy relations are possible for nonsignalling correlations by showing how nonlocal correlations among a set of parties limit any form of correlations, not necessarily nonlocal, shared among other parties. In particular, we provide tight bounds between the violation of a family of Bell inequalities among an arbitrary number of parties and the knowledge an external observer can gain about outcomes of any single measurement performed by the parties. Finally, we show how the obtained monogamy relations offer an improvement over the existing protocols for device-independent quantum key distribution and randomness amplification.

**Introduction.** It is a well established fact that entanglement and nonlocal correlations (cf. Refs. [1, 2]), i.e., correlations violating a Bell inequality [3], are fundamental resources of quantum information theory. It has been confirmed by many instances that, when distributed among spatially separated observers, they give an advantage over classical correlations at certain information-theoretic tasks, many of them being considered in the multipartite scenario. For instance, nonlocal correlations outperform their classical counterpart at communication complexity problems [4], and allow for security not achievable within classical theory [5].

Physical principles impose certain constraints on the way these resources can be distributed among separated parties; these are commonly referred to as monogamy relations. For instance, in any three–qubit pure state one party cannot share large amount of entanglement, as measured by concurrence, simultaneously with both remaining parties [6]. Analogous monogamy relations, both in qualitative [7–10] and quantitative [11, 12] form, were demonstrated for nonlocal correlations, with the measure of nonlocality being the violation of specific Bell inequalities. In particular, Toner and Verrus [11] and later Toner [12] showed that if three parties A, B, and C share, respectively, quantum and general nonsignalling correlations, then only a single pair can violate the Clauser–Horne–Shimony–Holt (CHSH) Bell inequality [13]. These findings were generalized to more complex scenarios [14, 15] (see also Ref. [16]), and in particular in [14] a general construction of monogamy relations for nonsignalling correlations from any bipartite Bell inequality was proposed.

In this work, we demonstrate that nonsignalling correlations are monogamous in a much stronger sense: the amount of nonlocality observed by a set of parties may imply severe limitations on any form of correlations with other parties. That is, instead of comparing nonlocality between distinct groups of parties, we rather relate it to the knowledge that external parties can gain on outcomes of any of the measurements performed by the parties (see Fig. 1). To be more illustrative, consider again parties A, B, and C performing a Bell experiment with M observables and d outcomes. We construct tight bounds between the violation of certain Bell inequalities [9] among any pair of parties, say A and B, and classical correlations that the third party C can establish with outcomes of any measurement performed by A or B. This means that the amount of any correlations — classical or nonlocal — that C could share with A or B is bounded by the strength of the Bell inequality violation between A and B. Our monogamies are further generalized to the scenario with an arbitrary number of parties N [(N, M, d) scenario] with nonlocality measured by the recent generalization of the Bell inequalities [9] presented in Ref. [10]. The obtained monogamy relations are logically independent from, and in fact stronger than, the existing relations involving only nonlocal correlations, as a bound on nonlocal correlations does not necessarily imply any nontrivial constraint on the amount of classical correlations.

Our new monogamy relations prove useful in device-independent protocols [17–21]. First, we show that they impose tight bounds on the guessing probability, the commonly used measure of randomness, that are significantly better than the existing ones [9, 10]. We then argue that this translates into superior performance in protocols for device-independent quantum key distribution (DIQKD) [29] using measurements with more than two outputs. Finally, we show that they allow for a generalization of the results of [19] on randomness amplification to any number of parties and outcomes, demon-

---

**FIG. 1.** (a) The usual monogamies compare nonlocality (measured by the value of some Bell inequality I) between different groups of parties (here between two pairs of parties AB and AC). Instead, our monogamy relations compare nonlocality observed by a group of parties (here AB) to the knowledge, represented by the probability $p(A = C)$, that the third party C can have about outcomes observed by either of the parties. As such, they are qualitatively different, and in fact stronger than those of type (a).
strating, in particular, that arbitrary amount of arbitrarily good randomness can be amplified in a bipartite setup.

Before turning to the results, we provide some background. Consider $N$ parties $A^{(1)}, \ldots, A^{(N)}$ (for $N = 3$ denoted by $A, B, C$), each measuring one of $M$ possible observables $A^{(i)}_{x_i}$ ($x_i = 1, \ldots, M$) with $d$ outcomes (enumerated by $a_i = 1, \ldots, d$) on their local physical systems. The produced correlations are described by a collection of probabilities $p(A^{(i)}_{x_i} = a_1, \ldots, A^{(N)}_{x_N} = a_N) \equiv p(a_1 \ldots a_N|x_1 \ldots x_N) \equiv p(a|x)$ of obtaining results $a \equiv a_1 \ldots a_N$ upon measuring $x \equiv x_1 \ldots x_N$. One then says that the correlations $\{p(a|x)\}$ are (i) nonsignalling (NC) if any of the marginals describing a subset of parties is independent of the measurements choices made by the remaining parties and (ii) quantum (QC) if they arise by local measurements on quantum states (cf. [2]).

Elemental and tight monogamies for nonsignalling correlations. We start with the derivation of our monogamy relations in the case of nonsignalling correlations. For clarity, we begin with the simplest tripartite scenario. We will use the Bell inequality introduced by Barrett, Kent, and Pironio (BKP) [9]. Denoting by $\langle \Omega \rangle$ the mean value of a random variable $\Omega$, that is, $\langle \Omega \rangle = \sum_{\Omega} p(\Omega = i)$, it reads

$$I_{AB}^{2,M,d} := \sum_{\Omega = 1}^{M} \left( \langle A_{-\Omega} - B_{\Omega} \rangle + \langle B_{\Omega} - A_{\Omega+1} \rangle \right) \geq d - 1$$

(1)

with $[\Omega]$ being $\Omega$ modulo $d$, and $\Omega_{M+1} := [\Omega_1 + 1]$. For $d = 2$, Ineq. (1) reproduces the chained Bell inequalities [23], while for $M = 2$ the Collins-Gisin-Linden-Massar-Popescu (CGLMP) inequalities [24]. The maximal nonsignalling violation of (1) is $I_{AB}^{2,M,d} = 0$.

The only monogamy relations for (1) have been formulated in terms of its violations between Alice and $M$ Bobs [14], which is a natural quantitative extension of the concept of $M$-shareability [7]. In the following theorem we show that the BKP Bell inequalities allow one to introduce elemental monogamies obeyed by any NC.

**Theorem 1.** For any tripartite NC $\{p(abc|x_{xy})\}$ with $M$-outcome measurements, the inequality

$$I_{AB}^{2,M,d} + \langle [X_i - C_j] \rangle + \langle [C_j - X_i] \rangle \geq d - 1$$

holds for any pair $i, j = 1, \ldots, M$ and $X$ denoting $A$ or $B$.

Interestingly, all these inequalities are tight in the sense that for any values of $I_{AB}^{2,M,d}$ and $\langle [X_i - C_j] \rangle + \langle [C_j - X_i] \rangle$ saturating (2), one can find NC realizing these values. Take, for instance, a probability distribution $p(a, b, c|x, y, z) = p(a, b|x, y)p(c|z)$, with $\{p(a, b|x, y)\}$ being a mixture of a nonlocal model maximally violating (1) and a local deterministic one saturating it. Then, $\{p(c|z)\}$ is the same distribution as the one used by $A$ or $B$ in the local model saturating (1).

The physical interpretation of our monogamies can be now concluded if we rewrite them in a bit different form. Using the fact that for any variable $\Omega$, $\langle [\Omega] \rangle + \langle [-\Omega] \rangle = dP(\Omega) \neq 0 = d[1 - P(\Omega)] = 0$ [25], Ineqs. (2) transform to

$$I_{AB}^{2,M,d} + 1 \geq dp(X_i = C_j)$$

(3)

for $X = A, B$, and any pair $i, j = 1, \ldots, M$. These relations hold if $AB$ is replaced by any pair of parties and if any $m = 1, \ldots, d - 1$ is added modulo $d$ to the argument of probability. The meaning of the introduced monogamy relations is now transparent. The probability $p(X_i = C_j)$ that parties $X$ and $Y$ can obtain the same results upon measuring the $i$th and $j$th observables is a measure of how the outcomes of these measurements are classically correlated. Consequently, Ineqs. (2) establish trade-offs between nonlocality, as measured by (1), that can be generated between any two parties and classical correlations that the third party can share with the results of any measurement performed by any of these two parties. Furthermore, they are tight. In fact, it is known that the maximal NC violation of (1), $I_{AB}^{2,M,d} = 0$, implies $p(X_i = C_j) = 1/d$ for any $i, j = 1, \ldots, M$, meaning that at the point of maximal violation $C$ cannot share any correlations with any other party’s measurement outcomes [9]. On the other hand, it is well known that at the point of no violation $C$ can be arbitrarily correlated with $A$ and $B$. For intermediate violations, the best one can hope for is a linear interpolation between these two extreme values and this is precisely what our monogamy relations predict, see Fig. 2.

Let us now move to the general case of an arbitrary number of parties each having $M$-outcome observables at their disposal. We will utilize the generalization of the Bell inequality (1) introduced in Ref. [10], which can be stated as

$$I_{A}^{N,M,d} \geq d - 1$$

(4)

with $A = A^{(1)} \ldots A^{(N)}$. Since the form of $I_{A}^{N,M,d}$ is rather lengthy and actually not relevant for further considerations, for clarity, we omit presenting it here (see [25]). We only mention that it can be recursively determined from $I_{AB}^{2,M,d}$ and that its minimal nonsignalling value is $I_{A}^{N,M,d} = 0$. Then, the generalization of Theorem 1 to arbitrary $N$ goes as follows.

**Theorem 2.** For any $(N + 1)$–partite NC $\{p(a|x)\}$ with $M$-outcome measurements per site, the following inequality

$$I_{A}^{N,M,d} + \langle [A_{z_k} - A_{z_{N+1}}] \rangle + \langle [A_{z_{N+1}} - A_{z_k}] \rangle \geq d - 1$$

(5)

is satisfied for any $x_k, x_{N+1} = 1, \ldots, M$ and $k = 1, \ldots, N$.

All the properties of the three-party monogamy relations persist for any $N$. In particular, all inequalities (5) are tight. Moreover, they can be rewritten as

$$I_{A}^{N,M,d} + 1 \geq dp(A_k = [A_{z_{N+1}} + m])$$

(6)

for any $x_k, x_{N+1} = 1, \ldots, M$, $k = 1, \ldots, N$ and $m = 0, \ldots, d - 1$ and remain valid if the nonlocality is tested among any $N$-element subset of $N + 1$ parties. Analogously to the three-party case, Ineqs. (6) tightly relate the nonlocality observed by any $N$ parties, as measured by $I_{A}^{N,M,d}$, and correlations that party $(N + 1)$ can share between measurement outcomes of any of these $N$ parties. It is worth pointing out that
for $d = 2$ it holds \(\langle X - Y \rangle = \langle Y - X \rangle\), and Ineqs. (5) simplify to \(I_A^{N,M,2} + 2(|A_x^{(k)} - A_{x+1}^{(N+1)}|) \geq 1\) which can be rewritten in a more familiar form as \(|\langle A_x^{(k)} A_{x+1}^{(N+1)} \rangle| \leq I_A^{N,M,2}\), where \(A_x^{(k)}\) stand now for dichotomic observables with outcomes $\pm 1$, while \(\langle XY \rangle = P(X = Y) - P(X \neq Y)\). Thus, the strength of violation of (24) imposes tight bounds on a single mean value \(|\langle A_x^{(k)} A_{x+1}^{(N+1)} \rangle|\) for any \(x_k, x_{N+1} \) and \(k = 1, \ldots, N\), which is also a measure of how outcomes of a measurement performed by the external party \(A^{(N+1)}\) are correlated to those of \(A^{(k)}\) for any \(k\). In particular, when \(I_A^{N,M,2} = 0\) (maximal non-signalling violation), all these means are zero, while maximal correlations between a single pair of measurements, i.e., \(|\langle A_x^{(k)} A_{x+1}^{(N+1)} \rangle| = \pm 1\) for some \(x_k, x_{N+1}\), make the \(N \) parties unable to violate \(I_A^{N,M,2} \geq 1\).

**Bounds on randomness.** Our monogamies are of particular importance for device-independent applications since they imply upper bounds on the guessing probability (GP) of the outcomes of any measurement performed by any of the \(N\) parties by the additional party, here called \(E\). To be precise, assume that \(E\) has full knowledge about all parties devices and their measurement choices and wishes to guess the outcomes of, say \(A_x^{(k)}\). The best \(E\) can do for this purpose is to simply measure one of its observables, say the \(z\)th one, and, irrespectively of the obtained result, deliver the most probable outcome of \(A_x^{(k)}\). Then, \(\max_{a_k} p(A_x^{(k)} = a_k) = p(E_z = \hat{A}_x^{(k)}(z))\), and Ineqs. (6) imply that for any \(x_k\) and \(k\), GP is bounded as

\[
\max_{a_k} p(a_k|x_k) \equiv \max_{a_k} p(A_x^{(k)} = a_k) \leq \frac{1}{d}(1 + I_A^{N,M,d}). \quad (7)
\]

These bounds are tight and significantly stronger than the previously existing one,

\[
\max_{a_k} p(a_k|x_k) \leq \frac{1}{d} \left(1 + \frac{dN}{4}(N - 1)I_A^{N,M,d}\right) \quad (8)
\]

derived in Refs. [9, 10] (see Fig. 2).

Let us now discuss how the bound (7) performs in comparison to (8) in security proofs of DIQKD against non-signalling eavesdroppers. At the moment, a general security proof in this scenario is missing and the strongest proof requires the assumption that the eavesdropper \(E\) is not only limited by the no-signalling principle but also lacks a long-term quantum memory (so-called bounded-storage model) [29]. Assume that Alice and Bob share a two-quito maximally entangled state and they use it to maximally violate (1) by performing the optimal measurements for this setup (see, e.g., [9]). To generate the secure key, Bob performs one more measurement that is perfectly correlated to one of Alice’s measurements. The key rate of this protocol is lower-bounded as \(R \geq -\log_2(\tau(I_A^{N,M,d}) - H(A|B))\) [29], where \(\tau\) is any upper bound on GP for non-signalling correlations. and \(H(A|B)\) is the conditional Shannon entropy between Alice and Bob for the measurements used to generate the secret key. As the state is maximally entangled, this term is equal to zero. Fig. 2 compares bounds on the secret key obtained by using our bound (7) and the previous bound (8) in this protocol. We fix the key rate and compute the minimal number of measurements needed to attain this rate using these bounds as a function of the number of outputs. As shown in Fig. 2, the number of measurements when using our bound is much smaller and, in particular, decreases with the number of outputs.

**Randomness amplification.** Let us finally show the usefulness of our monogamy relations in randomness amplification. Assume that each party is given a sequence of bits produced by the Santha–Vazirani (SV) source (or the \(\varepsilon\)-source). Its working is defined as follows: it produces a sequence \(y_1, y_2, \ldots, y_n\) of bits according to

\[
\frac{1}{2} - \varepsilon \leq p(y_k|w) \leq \frac{1}{2} + \varepsilon, \quad k = 1, \ldots, n, \quad (9)
\]

where \(w\) denotes any space-time variable that could be the cause of \(y_k\). Thus the bits are possibly correlated with each other retaining, however, some intrinsic randomness — we say that they are \(\varepsilon\)-free. The goal is now to obtain a perfectly random bit (or more generally \(d\) bits) from an arbitrarily long sequence of \(\varepsilon\)-free bits by using quantum correlations that violate the Bell inequality (24). This procedure is called randomness amplification (RA).

It is useful to recall this task in the adversarial picture [19], in which one assumes that an adversary \(E\), using the \(\varepsilon\)-sources, wants to simulate the quantum violation of (24) by NC, in particular the local ones. The random variable \(W\) is now held by \(E\) who uses it to control both the \(\varepsilon\)-sources and the physical devices possessed by the parties. That is, for every value \(w\) of \(W\) the former provides settings \(x\) with probabilities obeying (45), while these devices generate the \(N\)-partite probability distribution \(\{p(a|x, w)\}_{a, x}\). Using (7), we can now restate and generalize Lemma 1 of [19] (see [25]).

**Theorem 3.** Let \(\{p(a|x, w)\}_{a, x}\) be a non-signalling probabil-
ity distribution for any $w$. Then for any $x$ and $k = 1, \ldots, N$: 
\[
\sum_{a_k, w} |p(a_k, w|x) - \bar{p}(a_k)p(w|x)| \leq \frac{2^{d-1}+1}{d} Q_M(x) I_{A}^{N,M,d},
\]
where $\bar{p}(a) = 1/d$ for any $a$. \{p(a_k, w|x)\}_{a_k, w} describes correlations between outcomes obtained by party $k$ and the random variable $W$ for the measurements choice $x$, and $I_{A}^{N,M,d}$ is taken in the probability distribution \{p(a|x)\} observed by the parties. Finally, $Q_M(x) = \max_w \epsilon_{a_k}[p(w|x)/p_{\min}(w)]$, where $p_{\min}(w) = \min_x \{p(w|x)\}$ with minimum taken over those measurement settings $x$ that appear in $I_{A}^{N,M,d}$.

It then follows that if correlations \{p(a|x)\} violate maximally the Bell inequality (24), then the $dits$ observed by the parties are perfectly random and uncorrelated from $W$ [19].

Let us now show that one can amplify partially random input bits to almost perfectly random $dits$ by using QC that produce arbitrarily high violation of $I_{A}^{N,M,d}$. To generate one of the $M$ measurement settings, each party uses its SV source $r = \lfloor \log_2 M \rfloor$ times. Hence for any $x$, $Q_r(x) \leq \lfloor (1+2\varepsilon)/(1-2\varepsilon) \rfloor^{N \varepsilon}$ (cf. Ref. [19]). Then, there is a state and measurement settings [9, 10] such that for large $M$, 
\[
I_{A}^{N,M,d} \approx \lambda(d)/M \leq \lambda(d)/2^{r-1},
\]
where $\lambda(d)$ is a function of $d$. After plugging everything into (10), one checks that its r.h.s. tends to zero for $M \to \infty$ iff $\varepsilon < \varepsilon_N := (2^{1/N} - 1)/(2^{2^{1/N} + 1})$. As a result, QC violating (11) can be used to amplify randomness of any $\varepsilon$-source provided $\varepsilon < \varepsilon_N$. In particular, for $N = 2$, the above reproduces the value $\varepsilon_2 = (\sqrt{2} - 1)^2/2$ found in [19], and, because $\varepsilon_N$ is a strictly decreasing function of $N$, the larger the critical epsilon $\varepsilon_N$ for this method to work. Notice, however, that $\varepsilon_N$ is independent of $d$, so almost perfectly random $dits$ are obtained from partially random bits. This means that using the setup from Ref. [19] we can in fact achieve both amplification and expansion of randomness simultaneously.

Recently, with the same Bell inequality but for $N = d = 2$, the critical epsilon was shifted from $\varepsilon_2 \approx 0.086$ to $\tilde{\varepsilon}_2 \approx 0.0961$ [20]. We will now show that by using a slightly different approach the critical epsilon can be almost doubled. To this end, we exploit the fact that only $2M^{N-1}$ measurement settings out of all possible $M^N$ appear in $I_{A}^{N,M,d}$. However, to generate them a common source has to be used. Assumption then that this is the case, $R = \log_2(2M^{N-1}) = 1 + (N-1)r$ (instead of $Nr$) uses of the SV source are enough to generate all measurement settings in $I_{A}^{N,M,d}$. Thus, $Q_r(x) \leq \lfloor (1+2\varepsilon)/(1-2\varepsilon) \rfloor^{N \varepsilon}$, which together with (11) imply that the right-hand side of (10) vanishes for $M \to \infty$ iff $\varepsilon < \varepsilon_N'(2^{1/(N-1)} - 1)/(2^{2^{1/(N-1)} + 1})$, and in particular $\varepsilon_2' = 1/6 > \tilde{\varepsilon}_2$.

Conclusions. We have presented a novel class of monogamy relations, obeyed by any nonsignalling physical theory. They tightly relate the amount of nonlocality, as quantified by the violation of Bell inequalities [9, 10], that $N$ parties have generated in an experiment to the classical correlations an external party can share with outcomes of any measurement performed by the parties. Such trade-offs find natural applications in device-independent protocols and here we have discussed how they apply in quantum key distribution (cf. also Ref. [26]) and generation and amplification of randomness. We have finally showed that bipartite quantum correlations allow one to amplify $\varepsilon$-free $dits$ for any $\varepsilon < 1/6$.

Our results provoke further questions. First, it is natural to ask if analogous monogamies hold for quantum correlations, and, in fact, such elemental monogamies can be derived in the simplest (3,2,2) scenario (see [25]). From a more fundamental perspective, it is of interest to understand what is the (minimal) set of of monogamy relations generating the same set of multipartite correlations as the no-signalling principle.

Acknowledgments. Discussions with Gonzalo De La Torre are gratefully acknowledged. This work is supported by NFN grant 2013/08/M/ST2/00262, FNP TEAM, EU project SIQS, ERC grants QITBOX, QOLAPS and QUAGATUA, the Spanish project Chist-Era DIQIP. This publication was made possible through the support of a grant from the John Templeton Foundation. R. A. also acknowledges the Spanish MINECO for the support through the Juan de la Cierva program.
Here we present detailed proofs of Theorems 1, 2, and 3 of the main text. Also, in the simplest (3, 2, 2) scenario we provide elemental monogamies for quantum correlations.

**APPENDIX A: MONOGAMY RELATIONS**

Monogamy relations for nonsignalling correlations

Let us start with a simple fact. Recall for this purpose that $\langle \Omega \rangle$ is the standard mean value of a random variable $\Omega$, that is, $\langle \Omega \rangle = \sum_{i=1}^{d-1} iP(\Omega = i)$ and $[\Omega]$ stands for $\Omega$ modulo $d$.

**Fact 1.** It holds that for any random variable $\Omega$,

(a) $\langle [\Omega]\rangle + \langle [-\Omega - 1]\rangle = d - 1,$ (12)

(b) $\langle \Omega \rangle + \langle -[\Omega]\rangle = d [1 - p(\Omega = 0)].$ (13)

**Proof.** Both equations follow from the very definition of $\langle [\cdot] \rangle$. To prove (a) we notice that $[-\Omega - 1] + [\Omega] = d - 1$, and hence

$$\langle [-\Omega - 1]\rangle = \sum_{i=1}^{d-1} ip([\Omega] = d - i - 1)$$

$$= \sum_{i=0}^{d-2} (d - i - 1)p([\Omega] = i)$$

$$= (d - 1) \sum_{i=0}^{d-2} p([\Omega] = i) - \sum_{i=0}^{d-2} ip([\Omega] = i)$$

$$= (d - 1) \sum_{i=0}^{d-1} p([\Omega] = i) - \langle \Omega \rangle$$

$$= (d - 1) - \langle \Omega \rangle,$$ (14)

where the second equality is a consequence of changing of the summation index, the fourth one stems from the definition of $\langle \Omega \rangle$ and rearranging terms, and the last equality follows from normalization.

To prove (b), we write

$$\langle [\Omega]\rangle + \langle [-\Omega]\rangle = \sum_{i=1}^{d-1} ip([\Omega] = i) + p([-\Omega] = i)$$

$$= \sum_{i=1}^{d-1} ip([\Omega] = i) + p([\Omega] = d - i)$$

$$= \sum_{i=1}^{d-1} ip([\Omega] = i) + \sum_{i=1}^{d-1} (d - i)p([\Omega] = i)$$

$$= d \sum_{i=1}^{d-1} p([\Omega] = i)$$

$$= d[1 - p([\Omega] = 0)],$$ (15)

where the second equality is a consequence of the fact that $[\Omega] + [-\Omega] = d$, while the third equality follows from shifting of the summation index in the second sum.

To prove (b), we write

$$\langle [\Omega]\rangle + \langle [-\Omega]\rangle = \sum_{i=1}^{d-1} ip([\Omega] = i) + p([-\Omega] = i)$$

$$= \sum_{i=1}^{d-1} ip([\Omega] = i) + p([\Omega] = d - i)$$

$$= \sum_{i=1}^{d-1} ip([\Omega] = i) + \sum_{i=1}^{d-1} (d - i)p([\Omega] = i)$$

$$= d \sum_{i=1}^{d-1} p([\Omega] = i)$$

$$= d[1 - p([\Omega] = 0)],$$ (15)

where the second equality is a consequence of the fact that $[\Omega] + [-\Omega] = d$, while the third equality follows from shifting of the summation index in the second sum.

Let us now move to the proofs of the monogamy relations. In the tripartite case we make use of the Barrett, Kent, and Pironio (BKP) [9] inequality

$$I^{A,B}_{AB} = \sum_{\alpha=1}^{M} (|A_{\alpha} - B_{\alpha}| + |B_{\alpha} - A_{\alpha+1}|) \geq d - 1,$$ (16)

where the convention that $X_{M+1} = [X_{1} + 1]$ is assumed.

**Theorem 1.** For any three-partite nonsignalling correlations $\{p(a, b | x, y, z)\}$ with $M$ measurements and $d$ outcomes per site and any pair $\{i, j\}$ $(i, j = 1, \ldots, M)$, the following inequality

$$I^{A,B}_{AB} + \langle [X_{i} - C_{j}]\rangle + \langle [C_{j} - X_{i}]\rangle \geq d - 1$$ (17)

is satisfied with $X$ denoting either $A$ or $B$.

**Proof.** Let us start with the case of $X = A$ and then notice that for a random variable $\Omega$ it holds that $\langle [\Omega]\rangle + \langle [-\Omega - 1]\rangle = d - 1$ (see Fact 1). Consequently,

$$\sum_{\beta=1}^{M} (|C_{j} - A_{\beta} - 1| + |A_{\beta} - C_{j}|) - (M - 1)(d - 1)$$ (18)

is equal to zero. The fact that for any $\beta$ and $j$ it holds that $\langle [C_{j} - A_{\beta} - 1]\rangle + \langle [A_{\beta} - C_{j}]\rangle = d - 1 = \langle [A_{\beta} - C_{j} - 1]\rangle + \langle [C_{j} - A_{\beta}]\rangle$ allows us to rewrite (18) in the following way

$$\sum_{\beta=1}^{i-1} (|C_{j} - A_{\beta} - 1| + |A_{\beta+1} - C_{j}|)$$

$$+ \sum_{\beta=i+1}^{M} ((|A_{\beta} - C_{j} - 1| + |C_{j} - A_{\beta}|)) - (M - 1)(d - 1).$$ (19)
Then, by adding \( \langle [A_i - C_j] \rangle + \langle [C_j - A_i] \rangle \) to both sides of the above and rearranging some terms in the resulting expression, one obtains
\[
\langle [A_i - C_j] \rangle + \langle [C_j - A_i] \rangle = \sum_{\beta=1}^{i-1} \left( \langle [C_j - A_{\beta} - 1] \rangle + \langle [A_{\beta+1} - C_j] \rangle \right) + \sum_{\beta=i}^{M-1} \left( \langle [A_{\beta+1} - C_j - 1] \rangle + \langle [C_j - A_{\beta}] \rangle \right) + \langle [A_i - C_j] \rangle + \langle [C_j - A_M] \rangle - (M - 1)(d - 1). \tag{20}
\]

In an analogous way, we may decompose \( I_{AB}^{2,M,d} \):
\[
I_{AB}^{2,M,d} = \sum_{\alpha=1}^{i-1} \left( \langle [A_\alpha - B_{\alpha}] \rangle + \langle [B_\alpha - A_{\alpha+1}] \rangle \right) + \sum_{\alpha=i}^{M-1} \left( \langle [A_\alpha - B_{\alpha}] \rangle + \langle [B_\alpha - A_{\alpha+1}] \rangle + \langle [A_{\alpha+1} - C_j - 1] \rangle \right) + \langle [C_j - A_M] \rangle + \langle [A_M - B_M] \rangle + \langle [B_M - A_1 - 1] \rangle + \langle [A_1 - C_j] \rangle - (M - 1)(d - 1). \tag{21}
\]

In the last step of these manipulations, we add line by line Eqs. (20) and (21) in order to finally obtain
\[
I_{AB}^{2,M,d} + \langle [A_i - C_j] \rangle + \langle [C_j - A_i] \rangle = \sum_{\alpha=1}^{i-1} \left( \langle [C_j - A_\alpha - 1] \rangle + \langle [A_\alpha - B_\alpha] \rangle + \langle [B_\alpha - A_{\alpha+1}] \rangle + \langle [A_{\alpha+1} - C_j] \rangle \right) + \sum_{\alpha=i}^{M-1} \left( \langle [C_j - A_\alpha] \rangle + \langle [A_\alpha - B_\alpha] \rangle + \langle [B_\alpha - A_{\alpha+1}] \rangle + \langle [A_{\alpha+1} - C_j - 1] \rangle \right) + \langle [C_j - A_M] \rangle + \langle [A_M - B_M] \rangle + \langle [B_M - A_1 - 1] \rangle + \langle [A_1 - C_j] \rangle - (M - 1)(d - 1). \tag{22}
\]

What we have arrived at is basically the sum of \( M \) Bell expressions \( I_{AB}^{2,2,d} \) but ‘distributed’ among three parties in such a way that Bob and Charlie measure only a single observable. It was shown in [14] that the minimal value such an expression can achieve over nonsignalling correlations is precisely its classical bound \( d - 1 \). As a result, \( I_{AB}^{2,M,d} + \langle [A_i - C_j] \rangle + \langle [C_j - A_i] \rangle \geq M(d - 1) - (M - 1)(d - 1) = d - 1 \), finishing the proof for the case \( X = A \).

If \( X = B \) in Ineq. (17), then it suffices to rewrite the Bell expression from (16) as
\[
I_{AB}^{2,M,d} = \sum_{\alpha=1}^{M} \left( \langle [B_\alpha - A_{\alpha+1}] \rangle + \langle [A_{\alpha+1} - B_\alpha] \rangle \right), \tag{23}
\]
add to it the zero expression (18) with \( A \) replaced by \( B \), and repeat the above manipulations. This completes the proof.

Now let us move to the general \((N,M,d)\) scenario. The inequality of interest is now the one from Ref. [10], namely:
\[
I_{A}^{N,M,d} = \frac{1}{M} \sum_{\alpha_{N-1}=1}^{M} I_{A^{(1)}...A^{(N-1)}}^{N-1,M,d} (\alpha_{N-1}) \circ A_{\alpha_{N-1}}^{(N)} \geq d - 1. \tag{24}
\]

where \( A = A^{(1)} \ldots A^{(N)} \). The notation \( \circ A_{\gamma}^{(\delta)} \) means insertion of \( A^{(\delta)} \) to the average \( \langle \cdot \rangle \) with the opposite sign to the one of \( A_{\gamma}^{(\delta)} \) with any \( \gamma, \delta \), while \( I_{A^{(1)}...A^{(N-1)}}^{N-1,M,d} (\alpha_{N-1}) \) is the same Bell expression as in (24), but for \( N = 1 \) parties, and with observables of the last party relabeled as \( \alpha_{N-1} \rightarrow \alpha_{N-2} + \alpha_{N-1} - 1 \) with \( \alpha_N = 1, \ldots, M \).

Theorem 2. For any \((N+1)\)-partite nonsignalling correlations \( \{ p(a|x) \} \) with \( M \) outcome measurements per site, the following inequality
\[
I_{A}^{N,M,d} + \langle [A^{(k)}_x - A^{(N+1)}_x] \rangle + \langle [A^{(N+1)}_x - A^{(K)}_x] \rangle \geq d - 1 \tag{25}
\]
is satisfied for any \( x_k, x_{N+1} = 1, \ldots, M \) and \( k = 1, \ldots, N \).

Proof. The recursive formula in Ineq. (24), which for convenience we restate here
\[
I_{A}^{N,M,d} = \frac{1}{M} \sum_{\alpha_{N-1}=1}^{M} I_{A^{(1)}...A^{(N-1)}}^{N-1,M,d} (\alpha_{N-1}) \circ A_{\alpha_{N-1}}^{(N)}, \tag{26}
\]
allows us to demonstrate the theorem inductively. The case of \( N = 2 \) has already been proved as Theorem 1, so we consider \( N = 3 \). Exploiting Eq. (26), one can express \( I_{A^{(1)}A^{(2)}A^{(3)}}^{3,M,d} \) as
\[
I_{A^{(1)}A^{(2)}A^{(3)}}^{3,M,d} = \frac{1}{M} \sum_{\alpha_{2}=1}^{M} I_{A^{(1)}A^{(2)}}^{2,M,d} (\alpha_{2}) \circ A_{\alpha_{2}}^{(3)}. \tag{27}
\]
It is clear that for every \( \alpha_{2} = 1, \ldots, M \)
\[
I_{A^{(1)}A^{(2)}A^{(3)}}^{2,M,d} (\alpha_{2}) = \sum_{\alpha_{1}=1}^{M} (\langle [A_{\alpha_{1}}^{(1)} - A_{\alpha_{1}+\alpha_{2}-1}^{(2)}] \rangle + \langle [A_{\alpha_{1}+\alpha_{2}-1}^{(1)} + A_{\alpha_{1}}^{(2)}] \rangle) \geq d - 1 \tag{28}
\]
is a Bell inequality equivalent to (16), in which the observables of the second party \( A^{(2)} \) have been relabeled according to \( \alpha_{1} \rightarrow \alpha_{1} + \alpha_{2} - 1 \). It must then fulfill the monogamy relations (17) (with \( N = 2 \)) independently of the value of \( \alpha_{2} \). In order to see it in a more explicit way, let us consider the case
\(k = 1\), and in Eq. \((22)\) just rename \(A \to A^{(1)}\), \(B \to A^{(2)}\), and \(C \to A^{(3)}\), and also \(\alpha \to \alpha_1\) for the first party, while \(\alpha \to \alpha_1 + \alpha_2 - 1\) for the second one. Then, for those observables \(A^{(2)}_{\alpha_1+\alpha_2-1}\) for which \(\alpha_1 + \alpha_2 - 1 > M\) we use the rule \(X_iX_jM+\gamma = [X_\gamma + i]\) to get \([A^{(2)}_\gamma + i]\) with some \(\gamma\) and \(i\), and later replace the latter by another variable \(A^{(3)}_{\alpha_2}\) (this is just \(A^{(1)}_\gamma\) with outcomes shifted by a constant). With the aid of formula \((23)\) the same reasoning can be repeated for \(k = 2\).

Now, we prove that each term in Eq. \((27)\) fulfills \((25)\) for \(N = 3\), that is, the inequalities

\[
I^{M,d}_{A^{(1)},A^{(2)},A^{(3)}}(\alpha_2) \circ A^{(3)}_{\alpha_2} \geq \frac{1}{M} \sum_{\alpha_2=1}^{M} I^{M,d}_{A^{(1)},A^{(2)}}(\alpha_2) \circ A^{(1)}_{\alpha_2}. \tag{31}
\]

Now, it is enough to repeat the above reasoning to complete the proof of the monogamy relations \((25)\) for \(N = 3\).

Having it proven for \(N = 3\), let us now assume that the theorem is true for \(N\) parties (any \(N\)-partite non-signalling probability distribution). In order to complete the proof we again refer to the recursive formula \((26)\). By grouping together the last two parties, each term in the sum in Eq. \((26)\) is effectively an \((N - 1)\)-partite Bell expression for which we have just assumed \((25)\) to hold for any \(x_k, x_N\) and \(k = 1, \ldots, N\). Performing the summation over \(\alpha_{N-1}\) and dividing further by \(M^{N-2}\) we obtain \((25)\) for any \(i, j\) and \(k = 1, \ldots, N - 1\). The case \(k = N\) can be reached by using the fact that \(T^{N,M,d}\) is invariant under exchange of the last and the \((N - 2)\)th party \([10]\).

**Elemental monogamies for quantum correlations**

Let us now discuss the case of quantum correlations in which case similar monogamy relations are also expected to hold. Their derivation, however, is much more cumbersome and we only consider the simplest \((3,2,2)\) scenario and derive quantum analogs of the non-signalling monogamies \((17)\). To this end, we use a one-parameter modification of the CHSH Bell inequality \([13]\) with the latter being a particular case of \((16)\) with \(M = d = 2\). Here, for convenience, we write it down in its “standard” form:

\[
\tilde{\Gamma}_{AB} := \alpha(\langle A_1B_1 \rangle + \langle A_1B_2 \rangle) + \langle A_2B_1 \rangle - \langle A_2B_2 \rangle \leq 2\alpha \tag{32}
\]

with \(\alpha \geq 1\). Here, \(A_i\) and \(B_i\) are local quantum observables with eigenvalues \(\pm 1\) and \(\langle XY \rangle = \text{Tr}\{\rho [X \otimes Y]\}\) for some state \(\rho\) and local observables \(X, Y\). Actually, one proves the following more general theorem, generalizing the result of Ref. \([11]\) for the Bell inequality \((32)\).

**Theorem 3.** Any three-partite quantum correlations with two dichotomic measurements per site must satisfy the following inequalities

\[
\alpha^2 \max\{\tilde{\Gamma}_{AB}^2, \tilde{\Gamma}_{AC}^2\} + \min\{\tilde{\Gamma}_{AB}^2, \tilde{\Gamma}_{AC}^2\} \leq 4\alpha^2(1 + \alpha^2) \tag{33}
\]

and

\[
(\tilde{\Gamma}_{AB}^2) + 4(A_iC_j)^2 \leq 4(1 + \alpha^2) \tag{34}
\]

for any \(\alpha \geq 1\) and \(i, j = 1, 2\).

**Proof.** The proof is nothing more than a slight modification of the considerations of Ref. \([11]\) (see also Ref. \([27]\)). Nevertheless, we attach it here for completeness.

We start by noting that the monogamy regions, that is, the two-dimensional sets of allowed (realizable) within quantum theory pairs \(\{\tilde{\Gamma}_{AB}, \tilde{\Gamma}_{AC}\}\) for Ineq. \((33)\) and \(\{\tilde{\Gamma}_{AB}, \tilde{\Gamma}_{AC}\}\) with fixed \(i\) and \(j\) for Ineq. \((34)\), must be convex. Therefore, as it is shown in Ref. \([11]\) (see also Ref. \([28]\)), every point of their boundaries can be realized with a real three-qubit pure state and real local one-qubit measurements. Recall that the latter assumes the form

\[
X = x \cdot \sigma \tag{35}
\]

with \(x \in \mathbb{R}^2\) being a unit vector and \(\sigma = [\sigma_x, \sigma_z]\) denoting a vector consisting of the standard Pauli matrices \(\sigma_x\) and \(\sigma_z\).

Then, it follows from a series of papers \([11, 22, 27]\) that for a given two-qubit state \(\rho_{AB}\), the maximal value of \(\tilde{\Gamma}_{AB}\) over local, real, and traceless observables [i.e., those of the form \((35)\)] measured by Alice \(A_i\) and Bob \(B_i\), amounts to

\[
\max_{A_i,B_i} \tilde{\Gamma}_{AB} = 2\sqrt{\alpha^2 \lambda_1 + \lambda_2}. \tag{36}
\]

Here, \(\lambda_i (i = 1, 2)\) denote the eigenvalues of \(T_{AB}T_{AB}^T\) put in a decreasing order, i.e., \(\lambda_1 \geq \lambda_2\), and \(T_{AB}\) is the following
reduced correlation matrix

\[
T_{AB} = \begin{pmatrix}
    \langle \sigma_x \otimes \sigma_x \rangle_{AB} & \langle \sigma_x \otimes \sigma_z \rangle_{AB} \\
    \langle \sigma_z \otimes \sigma_x \rangle_{AB} & \langle \sigma_z \otimes \sigma_z \rangle_{AB}
\end{pmatrix}.
\] (37)

We added the subscript \(AB\) in (37) to indicate that the mean values are taken in the state \(\rho_{AB}\). In particular, one can similarly compute the maximal value of a single average \(\langle AB \rangle\) in the state \(\rho_{AB}\) over local observables \(A\) and \(B\) of the form (35) to be

\[
\max_{A,B} \langle AB \rangle = \lambda_1.
\] (38)

Equipped with these facts, we can now turn to the proof of the inequalities (33) and (34). We start from the first one and note that it suffices to demonstrate it in the case of \(\tilde{T}_{AB}^\alpha \geq \tilde{T}_{AC}^\alpha\), in which it becomes

\[
\alpha^2 (\tilde{T}_{AB}^\alpha)^2 + (\tilde{T}_{AC}^\alpha)^2 \leq 4\alpha^2.
\] (39)

The opposite case will follow immediately by exchanging \(B \leftrightarrow C\).

Let then \(\psi_{ABC}\) be a pure real three-qubit state. By \(\rho_{AB}\) and \(\rho_{AC}\) we denote its subsystems arising by tracing out the third and the second party, respectively, and by \(T_{AB}\) and \(T_{AC}\) the corresponding correlation matrices [cf. Eq. (37)]. Finally, let \(\lambda_i\) and \(\hat{\lambda}_i\) \((i = 1, 2)\) be eigenvalues of \(T_{AB}T_{AB}^T\) and \(T_{AC}T_{AC}^T\), respectively, where we keep the convention that \(\lambda_1 \geq \lambda_2\) and \(\hat{\lambda}_1 \geq \hat{\lambda}_2\). It was pointed out in Ref. [11] that the latter matrices are diagonal in the same basis, which allows one to simultaneously maximize both \(\tilde{T}_{AB}^\alpha\) and \(\tilde{T}_{AC}^\alpha\) with the same observables on Alice site. This, together with Eq. (36), means that

\[
\max_{\alpha \in [1, 2], \beta \in [0, \pi]} \alpha^2 (\tilde{T}_{AB}^\alpha)^2 + (\tilde{T}_{AC}^\alpha)^2 = 4[\alpha^2 (\lambda_1 + \lambda_2) + \alpha^2 \hat{\lambda}_1 + \hat{\lambda}_2]
\]

\[
= 4[\alpha^2 (\lambda_1 + \lambda_2) + 2\alpha^2]
\]

\[
= 4[\lambda_1 + 2\alpha^2].
\] (40)

In order to complete the proof, we use the result of the Toner-Verstraete monogamy relation for the CHSH Bell inequality [11], which we state here in terms of \(\lambda_i\) and \(\hat{\lambda}_i\) as

\[
\lambda_2 + \hat{\lambda}_1 \leq 2 - \lambda_1 - 2\alpha^2.
\] (41)

When applied to (40), it leads us to

\[
\max_{\alpha \in [1, 2], \beta \in [0, \pi]} \alpha^2 (\tilde{T}_{AB}^\alpha)^2 + (\tilde{T}_{AC}^\alpha)^2 \leq 4[\alpha^2 (\lambda_1 + \lambda_2 - 2\alpha^2) + 2\alpha^2]
\]

\[
= 4(\alpha^2 - 1)(\lambda_1 - \lambda_2) + 2\alpha^2)
\]

\[
= 4\alpha^2 (1 + \alpha^2),
\] (42)

where the second line follows from the facts that \(\lambda_1 \leq 1\), \(\lambda_2 \geq 0\), and \(\alpha \geq 1\).

To prove Ineq. (34), we follow the above reasoning to obtain

\[
\max_{\alpha \in [1, 2], \beta \in [0, \pi]} (\tilde{T}_{AB}^\alpha)^2 + 4(\sigma_i C_i)^2 = 4(\alpha^2 \lambda_1 + \lambda_2 + 2\alpha^2)
\]

\[
= 4\alpha^2 \lambda_1 + 4(\lambda_2 + 2\alpha^2)
\] (43)

for \(k = 1, 2\). Subsequent application of (41) to the term in parentheses in the second line of the above directly gives Ineq. (34), completing the proof.

For \(i = 1\) and \(j = 1, 2\), the relations (34) are tight as any pair of values of \(\tilde{T}_{AB}^\alpha\) and \(\langle A_i C_j \rangle\) saturating them can be realized with the state \(\langle \beta_+ |01\rangle + \beta_-|10\rangle|0\rangle\), where \(\beta_\pm = \frac{1}{2}(1 \pm \sqrt{2} \sin \theta)^{1/2}\) and \(\theta \in [0, \pi/4]\). It is, however, no longer true for \(i = 2\). In this case we numerically found tight monogamy relations for particular values of \(\alpha\) (see Fig. 3).

![Graph](image_url)

FIG. 3. (a) Guessing probability (and simultaneously the tight analogs of monogamies in Theorem 3) for \(i = 2\) as a function of \((\tilde{T}_{AB}^\alpha - 2\alpha)/(\sqrt{1 + \alpha^2 - \alpha})\) for various values of \(\alpha\). All curves were found using two methods. First, we maximized the guessing probability for a given value of \(\tilde{T}_{AB}^\alpha\) over two-ququart states and one-ququart dichotomic measurements. Then, we used the hierarchy of Ref. [11] and with its third level we arrived at curves that coincide with those obtained with the first method with precision 10^{-7}. For comparison (b) presents the corresponding nontight monogamies proven in theorem 3 (\(i = 2\)) for \(\alpha = 1, 3\) (the curves for \(\alpha = 1.5, 2\) fall in between these two). The black curve is the same on both plots.

Let us finally notice that the quantum elemental monogamies impose the following upper bounds on the guess-
ing probability
\[
\max_j p(X_i = j) \leq \frac{1}{2} \left[ 1 + [1 + \alpha^2 - (\bar{T}_A^B/2)^2]^{1/2} \right]
\] (44)
with \( X = A, B, i = 1, 2, \) and \( \alpha \geq 1 \). This bound was already derived in Ref. [22], and, as already said, it is tight only for \( i = 1 \). In the case \( i = 2 \), we determined the tight bounds numerically for few \( \alpha \)'s and they are presented in Fig. 3.

APPENDIX B: RANDOMNESS AMPLIFICATION

Let us begin with recalling the description of the Santha–Vazirani (SV) source (or the \( \varepsilon \)-source). Its working is defined as follows: it produces a sequence \( y_1, y_2, \ldots, y_n \) of bits according to
\[
\frac{1}{2} - \varepsilon \leq p(y_k | w) \leq \frac{1}{2} + \varepsilon \quad (k = 1, \ldots, n),
\] (45)
where \( w \) denotes any space-time variable that could be the cause of \( y_k \). In particular, \( y_k \) can depend on \( y_1, \ldots, y_{k-1} \).

Let now \( W \) be any random variable used by an adversary to control the \( \varepsilon \)-sources and the physical systems held by the parties. The random variable can be thought of a device, held by a villain \( E \), with a knob that when set to a particular value \( w \) of \( W \) makes (i) the SV sources produce bits with certain probabilities obeying (45) and (ii) the devices held by the parties generate a concrete nonsignalling probability distribution represented by \( \{ p(a|w, x, y) \}_{x,y} \). Let us then by \( \{ p(a_k, w|x) \}_{a_k,w} \) denote correlations between outcomes obtained by party \( k \) and the random variable \( W \) for a particular choice of measurement settings \( x \). Also, let \( \{ \tilde{p}(a) \} \) be the one-party uniform probability distribution, i.e., \( \tilde{p}(a) = 1/d \) for any \( a \). Introducing then the variational distance
\[
D(\{ p(x) \}, \{ q(x) \}) = \frac{1}{2} \sum_x |p(x) - q(x)|
\] (46)
between two probability distributions \( \{ p(x) \} \) and \( \{ q(x) \} \), we can prove the following.

Theorem 4. Let for any \( w \), \( \{ p(a|x, w) \}_{a,x} \) be an \( N \)-partite nonsignalling probability distribution. Then for any \( k = 1, \ldots, N \) and any choice of measurement settings \( x \):
\[
D(\{ p(a_k, w|x) \}_{a_k,w}, \{ \tilde{p}(a_k)p(w|x) \}_{a_k,w}) = \frac{1}{2} \sum_{a_k,w} |p(a_k, w|x) - \tilde{p}(a_k)p(w|x)|
\leq \frac{(d - 1)^2 + 1}{2d} Q_M(x) I_{A}^{N,M,d},
\] (47)
where \( I_{A}^{N,M,d} \) is taken in the probability distribution observed by the parties \( \{ p(a|x) \} \). Then
\[
Q_M(x) = \max_w \left[ \frac{p(w|x)}{p_{\min}(w)} \right],
\] (48)
where \( p_{\min}(w) = \min_a \{ p(w|x) \} \) with the minimum taken over all measurement settings \( x \) appearing in the Bell inequality (24).

Proof. For simplicity, but without any loss of generality, we prove this theorem for the bipartite case. The generalization to the multipartite case is straightforward.

As before, we denote the parties by \( A \) and \( B \), while the adversary by \( E \). Then, the corresponding inputs and outputs are denoted by \( x, y, z, a, b, \) and \( e \), respectively.

Let us start by noting that for any probability distribution \( \{ p(a,b|x,y,w) \}_{a,b,x,y} \), the maximal probability of local outcomes obtained by any of the parties, say for simplicity Alice, must obey the inequalities on the guessing probability [see Ineq. (7) in the main text]. That is
\[
\max_a p(a|x, w) \leq \frac{1}{d} \left( 1 + I_{w}^{2,M,d} \right)
\] (49)
for any \( x = 1, \ldots, M \), where by \( I_{w}^{2,M,d} \) we have denoted the value of the Bell expression (16) computed for the probability distribution \( \{ p(a,b|x,y,w) \}_{a,b,x,y} \). Clearly, this bound holds also for any \( p(a|x, w) \) which together with the normalization
\[
p(a|x, w) = 1 - \sum_{a \neq a} p(a|x, w),
\] (50)
means that \( p(a|x, w) \geq (1/d)[1 - (d - 1)I_{w}^{2,M,d}] \), and therefore the inequality
\[
\left| p(a|x, w) - \frac{1}{d} \right| \leq \frac{d - 1}{d} I_{w}^{2,M,d}
\] (51)
holds for any \( a \) and \( x \). Using then the inequality (49) for \( \max_a p(a|x, w) \) and (51) for the rest of \( p(a|x, w) \), we obtain that for any strategy \( w \) and a measurement setting \( x \),
\[
D(\{ p(a|x, w) \}_{a}, \{ \tilde{p}(a) \}) = \frac{1}{2} \sum_{a} |p(a|x, w) - \tilde{p}(a)|
\leq \frac{(d - 1)^2 + 1}{2d} I_{w}^{2,M,d}.
\] (52)

The remainder of the proof goes along exactly the same lines as in Ref. [19], however, for completeness, we will recall it here.

Due to the fact that the observers do not have access to the variable \( W \), one has to average Ineq. (52) over the probability distribution \( \{ p(w|x,y) \}_{w} \) for a particular choice of measurements \( x \) and \( y \). Together with the facts that \( p(a|x, w) = p(a|x, y, w) \) (no-signalling) and \( p(w|x, y)p(a|x, y, w) = p(a|x, w, y) \), this allows one to write
\[
D(\{ p(a,x,y,w) \}_{a,w}, \{ \tilde{p}(a)p(w|x,y) \}_{a,w}) = \frac{1}{2} \sum_{a,w} |p(a,x,y,w) - \tilde{p}(a)p(w|x,y)|
\leq \frac{(d - 1)^2 + 1}{2d} \sum_{w} p(w|x,y) I_{w}^{2,M,d}
\] (53)

Let us now concentrate on the right-and side of Ineq. (53). By using Eq. (16), we can bound it from above in the follow-
\[
\sum_w p(w|x,y) I^2_{w,M,d} = \\
\sum_{w,\alpha} p(w|x,y) ([A_\alpha - B_\alpha]_w + [B_\alpha - A_{\alpha+1}]_w) \\
= \sum_{w,\alpha} \left( p(w|\alpha,\alpha) p(w|x,y) ([A_\alpha - B_\alpha]_w \\
+ p(w|\alpha+1,\alpha) p(w|x,y) ([B_\alpha - A_{\alpha+1}]_w) \right) \\
\leq Q_M(x,y) \sum_{w,\alpha} [p(w|\alpha,\alpha) ([A_\alpha - B_\alpha]_w \\
+ p(w|\alpha+1,\alpha) ([B_\alpha - A_{\alpha+1}]_w) \\
= Q_M(x,y) \sum_{\alpha} ([A_\alpha - B_\alpha] + [B_\alpha - A_{\alpha+1}]) \\
= Q_M(x,y) I^2_{A,B}, \tag{54}\]

where the subscript \(w\) in the expectation values \([A_\alpha - B_\alpha]_w\) and \([B_\alpha - A_{\alpha+1}]_w\) means that they are computed for the probability distribution \(\{p(a,b|x,y,w)\}_{a,b},x,y,w\), and also the convention \(p(M+1,M|w) \equiv p(1,M|w)\) is used. Then, \(I^2_{A,B}\) is computed for the probability distribution \(\{p(a,b|x,y,w)\}_{a,b},x,y,w\) observed by \(A\) and \(B\).

By substituting Ineq. (54) to Ineq. (53), one finally obtains Ineq. (47), completing the proof. \(\square\)

One then recovers the inequality of Ref. [19] from Ineq. (47) by exploiting the fact that \((d-1)^2 + 1/d \leq d-1\) \((d \geq 2)\). Let us also notice that one can derive Ineq. (47) using a slightly different approach, which, for completeness, we present below.

**Theorem 5.** Let \(\{p(a|x,w)\}_{a,x}\) be a nonsignalling probability distribution for any \(w\) and let the probabilities \(p(x)\) be all equal. Then for any \(k = 1,\ldots,N\) and any choice of measurement settings \(x:\)

\[
D(\{p(a_k,w|x)\}_{a_k,w}, \{\tilde{p}(a_k)p(w|x)\}_{a_k,w}) = \frac{1}{2} \sum_{a_k,w} |p(a_k,w|x) - \tilde{p}(a_k)p(w|x)| \\
\leq \frac{(d-1)^2 + 1}{2d} Q_M(x) I^N_{A,M,d}, \tag{55}\]

where \(I^N_{A,M,d}\) is taken in the probability distribution observed by the parties \(\{p(a|x)\}\)

\[
\tilde{Q}_M(x) = \max_w \left[ \frac{p(x|w)}{\tilde{p}(w)} \right], \tag{56}\]

where \(\tilde{p}_{\min}(w) = \min_x \{p(x|w)\}\) with the minimum taken over all measurement settings \(x\) appearing in the Bell inequality (24).

**Proof.** For simplicity but without any loss of generality, we prove this theorem for the bipartite case. The generalization to the multipartite case is straightforward.

As before, we denote the parties by \(A\) and \(B\), while the adversary by \(E\). Then, the corresponding inputs and outputs are denoted by \(x, y, z\), and \(a, b\), and \(e\), respectively.

Let us start by noting that, by analogy to the case considered in the main text [see Ineq. (6) there], for any \(w\), the probability distribution \(\{p(a,b|x,y,w)\}_{a,b,x,y}\) satisfies the following monogamy relations

\[
\frac{I^2_{w,M,d}}{\tilde{p}_{\min}(w)} + 1 \geq dp(X_i = E_j|w) \quad (X = A,B) \tag{57}\]

for any pair \((i,j)\) \((i,j = 1,\ldots,M)\). In the above

\[
I^2_{w,M,d} = \sum_{\alpha=1}^M [p(\alpha,\alpha|w) ([A_\alpha - B_\alpha]_w \\
+ p(\alpha+1,\alpha[p(\alpha+1,\alpha) ([B_\alpha - A_{\alpha+1}]_w)]_w, \tag{58}\]

is a modified BKP Bell expression taking into account that the inputs \(x, y\) are generated with the biased probabilities \(p(x,y|w)\), all correlators \([A_\alpha - B_\alpha]_w\) and \([B_\alpha - A_{\alpha+1}]_w\) are computed for the distribution \(\{p(a,b|x,y,w)\}_{a,b,x,y,w}\), and now

\[
\tilde{p}_{\min}(w) = \min_{\alpha=1,\ldots,M} \{p(a,\alpha|w), p(\alpha+1,\alpha|w)\}, \tag{59}\]

where the convention \(p(M+1,M|w) \equiv p(1,M|w)\) is used.

The monogamy relations (57) imply (see the main text for the argument in favor of this fact) the bound on the probability of the adversary when using the strategy \(w\) to guess the outcomes of any of the measurements performed by one of the parties, say for concreteness Alice (but the same bound holds for outcomes of party \(B\)):

\[
\max_a p(a|x,w) \leq \frac{1}{d} \left( 1 + \frac{I^2_{w,M,d}}{\tilde{p}_{\min}(w)} \right) \quad (x = 1,\ldots,M). \tag{60}\]

Clearly, this bound holds also for any \(p(a|x,w)\) which together with the normalization

\[
p(a|x,w) = 1 - \sum_{\alpha \neq a} p(\alpha|x,w), \tag{61}\]

mean that \(p(a|x,w) \geq (1/d) - (d-1)(I^2_{w,M,d}/d\tilde{p}_{\min}(w))\), and therefore the inequality

\[
|p(a|x,w) - \frac{1}{d}| \leq \frac{d-1}{d} \frac{I^2_{w,M,d}}{\tilde{p}_{\min}(w)}. \tag{62}\]

holds for any \(a\) and \(x\). Using then the inequality (60) for \(\max_a p(a|x,w)\) and (62) for the rest of \(p(a|x,w)\), we obtain that for any strategy \(w\),

\[
D(\{p(a|x,w)\}_a, \{\tilde{p}(a)\}) = \frac{1}{2} \sum_a |p(a|x,w) - \tilde{p}(a)| \\
\leq \frac{(d-1)^2 + 1}{2d} \frac{I^2_{w,M,d}}{\tilde{p}_{\min}(w)}. \tag{63}\]

Now, since the parties do not have access to \(W\), one needs further to average Ineq. (63) over the probability distribution
\( \{p(w|x,y)\}_w \) for a particular choice of measurements \( x \) and \( y \). This, together with the facts that \( p(a|x,w) = p(a|x,y,w) \) (no-signalling) and \( p(w|x,y) = p(x,y|w)p(w)/p(x,y) \) implying that \( p(w|x,y)p(a|x,y,w) = p(a,w|x,y) \), allows one to write

\[
D(\{p(a,w|x,y)\}_{a,w}, \{\bar{p}(a)p(w|x,y)\}_{a,w}) = \frac{1}{2} \sum_{a,w} |p(a,w|x,y) - \bar{p}(a)p(w|x,y)|
\]

\[
\leq \frac{(d-1)^2 + 1}{2d} \sum_w p(x,y|w) \frac{p(w)}{\tilde{p}_{\min}(w)} f_{w}^{2,M,d}
\]

\[
\leq \frac{(d-1)^2 + 1}{2d} \tilde{Q}_M(x,y) \sum_w p(w) \frac{p(x,y)}{\tilde{p}_{\min}(w)} f_{w}^{2,M,d}, \tag{64}
\]

with \( \tilde{Q}_M(x,y) = \max_w [p(x,y|w)/\tilde{p}_{\min}(w)] \). In order to obtain Ineq. (55) from Ineq. (64) it is enough to notice that

\[
p(a,b|x,y) = \sum_w p(w|x,y)p(a,b|x,y,w) \tag{65}
\]

which, with the aid of the assumption that all the probabilities \( p(x,y) \) are equal, further translates to

\[
I_{AB}^{2,M,d} = \sum_w p(w) f_{w}^{2,M,d}, \tag{66}
\]

where \( I_{AB}^{2,M,d} \) is computed for the observed probability distribution \( \{p(a,b|x,y)\} \) and the probabilities \( p(x,y) = \sum_w p(w)p(x,y|w) \) are assumed to be equal for all \( x,y \). This completes the proof.

Let us finally notice that under the assumption, which we make above, that all \( p(x,y) \) are equal, it holds that \( Q_M(x) = \tilde{Q}_M(x) \).