Certifying the Absence of Apparent Randomness under Minimal Assumptions

Gonzalo de la Torre, Chirag Dhara, and Antonio Acin

ICFO-Institut de Ciencies Fotoniques
Av. Carl Friedrich Gauss, 3, 08860 Castelldefels, Barcelona, Spain

Abstract

Contrary to classical physics, the predictions of quantum theory for measurement outcomes are of a probabilistic nature. Questions about the completeness of such predictions lie at the core of quantum physics and can be traced back to the foundations of the field. Recently, the completeness of quantum probabilistic predictions could be established based on the assumption of freedom of choice. Here we ask when can events be established to be as unpredictable as we observe them to be relying only on minimal assumptions, ie. distrusting even the free choice assumption but assuming the existence of an arbitrarily weak (but non-zero) source of randomness. We answer the latter by identifying a sufficient condition weaker than the monogamy of correlations which allow us to provide a family of finite scenarios based on GHZ paradoxes where quantum probabilistic predictions are as accurate as they can possibly be. Our results can be used for a protocol of full randomness amplification, without the need of privacy amplification, in which the final bit approaches a perfect random bit exponentially fast on the number of parties.

1998 ACM Subject Classification E.4 Coding and Information Theory, H.1.1 Systems and Information Theory

Keywords and phrases randomness, Bell nonlocality, free choice

Digital Object Identifier 10.4230/LIPIcs.TQC.2013.207

1 Introduction

Physical theories aim at providing the best possible predictions for both natural phenomena taking place in the universe and on the controlled environment of our laboratories. Interestingly, the type of predictions theories can make has been changing over time, depending heavily on the specific physical theory considered. Moreover, the latter happens not only at the less surprising quantitative level, ie. general relativistic predictions about the perihelion precession shift of mercury are more accurate than those of newtonian theory of gravity, but more strikingly, also at a qualitative level ie. the uncertainty on predictions hold fundamentally different statuses in classical and quantum theory.

Classical mechanics, the theory governing our physical understanding until the XIX century, is a deterministic theory by construction. The latter neither does imply that probabilistic predictions do not play any role nor that we cannot observe physical phenomena behaving as random and yet being governed by classical mechanical laws. Instead, it means that all uncertainty in the predictions of the theory can be traced back to a lack of knowledge about all the relevant degrees of freedom of the physical phenomena considered. As an example, accurate knowledge of the applied force and torque, viscosity and gravitational potential would make the outcome prediction of a coin flip fully predictable. Thus, no room for intrinsic unpredictability is available within classical theory and the best possible predictions are deterministic.

8th Conference on Theory of Quantum Computation, Communication and Cryptography.
Editors: Simone Severini and Fernando Brandao; pp. 207–219
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

With the advent of quantum mechanics, the former intuitions had to change dramatically. Contrary to classical mechanics, quantum theory is a probabilistic theory as dictated by its axioms. This means that, in general, the predictions of the theory for measurement outcomes have an inevitable amount of uncertainty, even when full knowledge and control over all relevant degrees of freedom is assumed.

Such a striking change on the role of predictability greatly shaked the foundations of physics. The completeness of quantum predictions has indeed been widely debated by some of the most eminent physicists that contributed to its development [1, 2]. However, John Bell was the first to derive consequences on the issue of completeness from experimentally fasifiable predictions under rigorous assumptions [3]. He proved that according to the predictions of quantum theory and under the assumption of locality and freedom of choice, the outcomes of some quantum experiments would be incompatible with an underlying deterministic theory. Very recently, completeness of quantum theory could be established under the assumptions of locality, the correctness of quantum theory and the crucial assumption of freedom of choice [4]. However, one may consider cases where the freedom of choice assumption cannot be fully trusted and ask whether it is truly a necessary requirement in order to exclude all apparent randomness. In the present letter we pose such question, that is, under what conditions and minimal assumptions can we certify that an event is as intrinsically unpredictable as it is observed to be. In other words, when can we exclude all possible apparent randomness of an event. Interestingly, in a recent work the full unpredictability of an event could be certified under minimal assumptions [5]. Nevertheless, this result required a complex scenario on the infinite number of parties limit. The proof was based on the monogamy of correlations in such limit. The main result of this letter is to identify a sufficient condition weaker than the monogamy of correlations [6] that certify events without any apparent randomness under the assumptions of locality and the existence of a source of arbitrarily deterministic bits. Using this condition, we construct a family of finite scenarios based on GHZ paradoxes [7] where events are indeed as intrinsically random as they appear to be. Moreover, our results imply a perfect free random bit can be approached exponentially fast in the number of parties and is therefore suitable for a full randomness amplification protocol without privacy amplification[5].

1.1 Geometric interpretation of the problem

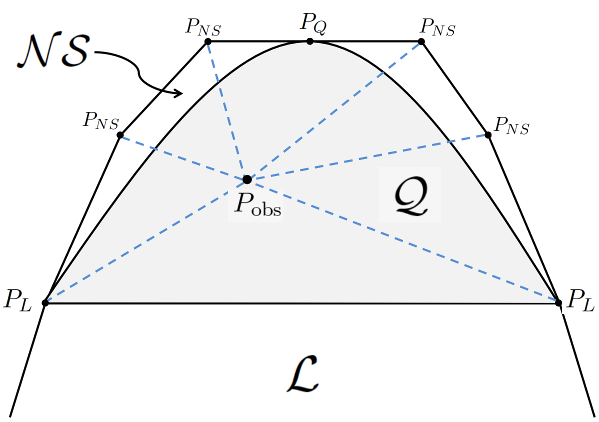

Figure 1 Qualitative picture of local, quantum and no-sigalling sets.

Fig. 1 is a useful qualitative geometric picture which serves to clarify the general idea and to explain the scenario we work with. Given some non-local distribution P_{obs}, its intrinsic randomness content is quantitatively dependent on whether we use the quantum or non-signalling framework. For example, the Tsirelson correlations [8] in the $(2, 2, 2)$ scenario considered strictly within the quantum set yields 1.23 bits of randomness [9]. However, its randomness in the larger non-signalling set is a much smaller 0.34 bits. Another example is the GHZ correlations [10] which contain (considering the tripartite states in particular) 3 bits of randomness within the quantum set. However, in the non-signalling set it reduces to just 1 bit since the extremal points are fully characterized in [11]. In fact, it is generally the case that the intrinsic randomness of a point considered to be embedded in the non-signalling set is lower than its intrinsic randomness within the quantum set. The reason is simply that there are more general decompositions possible within the non-signalling set which increases our ignorance about its underlying preparation. It is in this context that we can finally pose the question that is the theme of this work. *Is it possible to guarantee that the observed correlations do not contain any classical randomness for some correlations P_{obs} even allowing the largest possible ignorance by embedding it in the non-signalling set?*

The challenge to answering this question in full generality is that the definition of intrinsic randomness in such scenarios is defined as the optimization over all possible preparations of P_{obs}. However, this computation requires a complete characterization of the corresponding no-signalling polytope. This is known only for the smallest dimension and thus the computation is infeasible for anything but the smallest systems. What we show here is that despite the infeasibility of calculating the intrinsic randomness in full generality it is possible to choose scenarios carefully in which the computation is rendered feasible. We not only demonstrate one such case but also certify that the observed randomness is fully intrinsic in our chosen scenario. What make the result counter-intuitive is that our results are valid for a whole class of non-extremal distributions.

There is a further layer of subtlety which we additionally address in our work. This is related to a paradox of randomness certification using Bell inequalities, which is the freedom of choice assumption. The assumption of freedom of choice may be regarded a reasonable assumption in many cases but it is particularly problematic for randomness certification. Recently there has been a significant body of work in deriving Bell inequalities with relaxations of this assumption [12, 13, 14, 15, 5]. A significant feature of our results are that they are valid even under a complete (non-zero) relaxation of the measurement assumption. For this reason, these results may also be interpreted as an alternative approach for full randomness amplification with the benefit of significantly easier techniques.

2 Preliminaries

Suppose that a Bell test is performed repeatedly among N parties and the resulting statistics is given by $P_{\text{obs}}(\mathbf{a}|\mathbf{x})$, where $\mathbf{a} = (a_1, \ldots, a_N)$ and $\mathbf{x} = (x_1, \ldots, x_N)$ are the string of outcomes and measurement inputs of the parties involved. Let g be a function acting on the measurement results $\mathbf{a}$. As previously explained, there are different physically relevant notions of randomness.

First, the *observed randomness* of g for measurements $\mathbf{x}$ is the randomness computed directly from the statistics. Operationally, this may be defined as the optimal probability of guessing the outcome of g for input $\mathbf{x}$,

$$G_{\text{obs}}(g, \mathbf{x}, P_{\text{obs}}) = \max_{k \in \text{Im}(g)} P_{\text{obs}}(g(\mathbf{a}) = k|\mathbf{x}). \tag{1}$$

where $\mathrm{Im}(g)$ is the image of function g.

Moving to the definition of the *intrinsic randomness*, one should consider all possible preparations of the observed statistics in terms of no-signalling probability distributions. In our context, a particular preparation reads

$$P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x}) = \sum_e p(e|\mathbf{x}) P_e^{\mathrm{ex}}(\mathbf{a}|\mathbf{x}) \tag{2}$$

where the P_e^{ex} are extremal points of the no-signaling set [16]. The terms $p(e|\mathbf{x})$ may depend on $\mathbf{x}$, which accounts for possible correlations between the preparation e and the measurement settings $\mathbf{x}$, given that the choice of measurements are not assumed to be free. Hence, we define the intrinsic randomness of a function g by optimizing over all possible non-signalling preparations of P_{obs} so as to minimize the randomness of g. In other words,

$$G_{\mathrm{int}}(g, \mathbf{x}, P_{\mathrm{obs}}) = \max_{\{p(e|\mathbf{x}), P_e^{\mathrm{ex}}\}} \sum_e p(e|\mathbf{x}) G_{\mathrm{obs}}(g, \mathbf{x}, P_e^{\mathrm{ex}})$$

subject to:

$$\sum_e p(e|\mathbf{x}) P_e^{\mathrm{ex}}(\mathbf{a}|\mathbf{x}) = P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x}) \tag{3}$$

$$p(\mathbf{x}|e) \geq \delta \quad \text{with} \quad \delta > 0; \ \forall\ \mathbf{x}, e \tag{4}$$

where $G_{\mathrm{obs}}(g, \mathbf{x}, P_e^{\mathrm{ex}}) = \max_k \ P_e^{\mathrm{ex}}(g(\mathbf{a}) = k|\mathbf{x})$ is also the intrinsic randomness of P_e^{ex}, since intrinsic and observed randomness must coincide for extremal points of the non-signalling set. Note that condition $p(\mathbf{x}|e) \geq \delta > 0$ allows for an arbitrary (but not absolute) relaxation of the freedom of choice assumption by allowing for arbitrary (yet not complete) correlations between the preparation and the measurement settings. Physically, this condition ensures that all measurement combinations appear for all possible preparations e. An example of a source of randomness fulfilling this condition is a Santha-Vazirani source [17]. Note however that our definition allows sources more general than the Santha-Vazirani sources.

From a cryptographic point of view, the observed randomness is the one perceived by the parties performing the Bell test, whereas the intrinsic randomness is that perceived by a non-signalling eavesdropper possessing knowledge of the preparation of the observed correlations and with the ability to arbitrarily (yet not fully) bias the choice of the measurement settings.

In general, G_{obs} is strictly larger than G_{intr}, as the set of non-signalling correlations is larger than the quantum. The results in [6, 18] provide a Bell test in which G_{intr} approaches G_{obs} (and to 1/2) in the limit of an infinite number of measurements and assuming free choices, that is, $p(\mathbf{x}|e)$ in (2) is independent of e. The results in [19] allow some relaxation of this last condition. The results in [5] arbitrarily relaxed the free-choice condition and give a Bell test in which G_{intr} tends to G_{obs} (and both tend to 1/2) in the limit of an infinite number of parties. Here, we provide a significantly stronger proof, as we allow the same level of relaxation on free choices and provide Bell tests in which $G_{\mathrm{intr}} = G_{\mathrm{obs}}$ for any number of parties. Moreover, a perfect random bit is obtained in the limit of an infinite number of parties.

3 Scenario

Our scenario consists of N parties where each performs two measurements of two outcomes. In what follows, we adopt a spin-like notation and label the outputs by ± 1. Then, any non-signalling probability distribution can be written as (for simplicity we give the expression for three parties, but it easily generalizes to an arbitrary number)

$$\begin{aligned} &P(a_1, a_2, a_3|x_1, x_2, x_3) = \\ &\quad \frac{1}{8}\Big(1 + a_1\langle A_1^{(x_1)}\rangle + a_2\langle A_2^{(x_2)}\rangle + a_3\langle A_3^{(x_3)}\rangle + \\ &\quad a_1a_2\langle A_1^{(x_1)}A_2^{(x_2)}\rangle + a_1a_3\langle A_1^{(x_1)}A_3^{(x_3)}\rangle + \\ &\quad a_2a_3\langle A_2^{(x_2)}A_3^{(x_3)}\rangle + a_1a_2a_3\langle A_1^{(x_1)}A_2^{(x_2)}A_3^{(x_3)}\rangle\Big), \end{aligned} \tag{5}$$

where $A_i^{(x_i)}$ denotes the outputs of measurement x_i by each party i. In this scenario, we consider Mermin Bell inequalities, whose Bell operator reads

$$M_N = \frac{1}{2}M_{N-1}(A_N^{(0)} + A_N^{(1)}) + \frac{1}{2}M'_{N-1}(A_N^{(0)} - A_N^{(1)}), \tag{6}$$

where M_2 is the Clauser-Horne-Shimony-Holt operator and M'_{N-1} is obtained from M_{N-1} after swapping $A_i^{(0)} \leftrightarrow A_i^{(1)}$. We study probability distributions that give the maximal non-signalling violation of the Mermin inequalities and focus our analysis on a function f that maps the N measurement results into one bit as follows:

$$f(\mathbf{a}) = \begin{cases} +1 & n_-(\mathbf{a}) = (4j+2); \quad \text{with } j \in \{0,1,2,\dots\} \\ \\ -1 & \text{otherwise} \end{cases} \tag{7}$$

where $n_-(\mathbf{a})$ denotes the number of results in $\mathbf{a}$ that are equal to -1.

4 Results

Our goal in what follows is to quantify the intrinsic randomness of the bit defined by $f(\mathbf{a})$ for those distributions maximally violating the Mermin inequality for odd N. We first prove the following

▶ **Lemma 1.** *Let $P_{\mathrm{M}}(\mathbf{a}|\mathbf{x})$ be an N-partite (odd N) non-signalling probability distribution maximally violating the corresponding Mermin inequality. Then, for any input $\mathbf{x}$ appearing in the inequality*

$$P_{\mathrm{M}}(f(\mathbf{a}) = h_N|\mathbf{x}) \geq 1/2, \text{ with } h_N = \sqrt{2}\cos\left(\frac{\pi(N+4)}{4}\right). \tag{8}$$

Note that, as N is odd, $h_N = \pm 1$. Operationally, the Lemma implies that, for all points maximally violating the Mermin inequality, the bit defined by f is biased towards the same value h_N. Since the proof of the Lemma for arbitrary odd N is convoluted, we give the explicit proof for $N = 3$ here, which already conveys the main ingredients of the general proof, and relegate the generalization to the Supplementary Information.

Proof for three parties. With some abuse of notation, the tripartite Mermin inequality may be expressed as,

$$M_3 = \langle 001\rangle + \langle 010\rangle + \langle 100\rangle - \langle 111\rangle \leq 2, \tag{9}$$

where $\langle x_1x_2x_3\rangle = \langle A_1^{(x1)}A_2^{(x2)}A_3^{(x3)}\rangle$ and similar for the other terms. The maximal non-signalling violation assigns $M_3 = 4$ which can only occur when the first three correlators in (9) take their maximum value of $+1$ and the last takes its minimum of -1.

Take any input combination appearing in the inequality (9), say, $\mathbf{x_m} = (0, 0, 1)$. Maximal violation of M_3 imposes the following conditions:

1. $\langle 001\rangle = 1$. This further implies $\langle 0\rangle_1 = \langle 01\rangle_{23}$, $\langle 0\rangle_2 = \langle 01\rangle_{13}$ and $\langle 1\rangle_3 = \langle 00\rangle_{12}$.
2. $\langle 010\rangle = 1$ implying $\langle 0\rangle_1 = \langle 10\rangle_{23}$, $\langle 1\rangle_2 = \langle 00\rangle_{13}$ and $\langle 0\rangle_3 = \langle 01\rangle_{12}$.
3. $\langle 100\rangle = 1$ implying $\langle 1\rangle_1 = \langle 00\rangle_{23}$, $\langle 0\rangle_2 = \langle 10\rangle_{13}$ and $\langle 0\rangle_3 = \langle 10\rangle_{12}$.
4. $\langle 111\rangle = -1$ implying $\langle 1\rangle_1 = -\langle 11\rangle_{23}$, $\langle 1\rangle_2 = -\langle 11\rangle_{13}$ and $\langle 1\rangle_3 = -\langle 11\rangle_{12}$

Imposing these relations on (5) for input $\mathbf{x_m} = (0, 0, 1)$ one gets

$$P_{\mathrm{M}}(a_1, a_2, a_3|0, 0, 1) = \frac{1}{8}(1 + a_1a_2a_3 + (a_1 + a_2a_3)\langle 0\rangle_1 + (a_2 + a_1a_3)\langle 0\rangle_2 + (a_3 + a_1a_2)\langle 1\rangle_3) \tag{10}$$

Using all these constraints and the definition of the function (20), Eq. (8) can be expressed as

$$\begin{aligned} &P_M(f(\mathbf{a}) = +1|\mathbf{x_m}) \\ &= P_{\mathrm{M}}(1, -1, -1|\mathbf{x_m}) + P_{\mathrm{M}}(-1, 1, -1|\mathbf{x_m}) \\ &+ P_{\mathrm{M}}(-1, -1, 1|\mathbf{x_m}) \\ &= \frac{1}{4}(3 - \langle 0\rangle_1 - \langle 0\rangle_2 - \langle 1\rangle_3) \end{aligned} \tag{11}$$

Proving that $P(f(\mathbf{a}) = +1|\mathbf{x_m}) \geq 1/2$ then amounts to showing that $\langle 0\rangle_1 + \langle 0\rangle_2 + \langle 1\rangle_3 \leq 1$. This form is very convenient since it reminds one of a positivity condition of probabilities.

We then consider the input combination $\bar{\mathbf{x}}_\mathbf{m}$ such that all the bits in $\bar{\mathbf{x}}_\mathbf{m}$ are different from those in $\mathbf{x}_m$. We call this the swapped input, which in the previous case is $\bar{\mathbf{x}}_\mathbf{m} = (1, 1, 0)$. Note that this is *not* an input appearing in the Mermin inequality. However, using the previous constraints derived for distributions P_{M} maximally violating the inequality, one has

$$\begin{aligned} &P_{\mathrm{M}}(a_1, a_2, a_3|1, 1, 0) \\ =&\frac{1}{8}(1 + a_1\langle 1\rangle_1 + a_2\langle 1\rangle_2 + a_3\langle 0\rangle_3 + a_1a_2\langle 11\rangle_{12} \\ &+ a_1a_3\langle 10\rangle_{13} + a_2a_3\langle 10\rangle_{23} + a_1a_2a_3\langle 110\rangle_{123}) \\ =&\frac{1}{8}(1 + a_1\langle 1\rangle_1 + a_2\langle 1\rangle_2 + a_3\langle 0\rangle_3 - a_1a_2\langle 1\rangle_3 \\ &+ a_1a_3\langle 0\rangle_2 + a_2a_3\langle 0\rangle_1 + a_1a_2a_3\langle 110\rangle_{123}), \end{aligned} \tag{12}$$

where the second equality results from the relations $\langle 11\rangle_{12} = -\langle 1\rangle_3$, $\langle 10\rangle_{13} = \langle 0\rangle_2$ and $\langle 10\rangle_{23} = \langle 0\rangle_1$.

It can be easily verified that summing the two positivity conditions $P_{\mathrm{M}}(1, 1, -1|\bar{\mathbf{x}}_\mathbf{m}) \geq 0$ and $P_{\mathrm{M}}(-1, -1, 1|\bar{\mathbf{x}}_\mathbf{m}) \geq 0$ gives the result we seek, namely $1 - \langle 0\rangle_1 - \langle 0\rangle_2 - \langle 1\rangle_3 \geq 0$, which completes the proof. ◀

Using the previous Lemma, it is rather easy to prove the following

▶ **Theorem 2.** *Let $P_{\mathrm{obs}}(\mathbf{a}|\mathbf{x})$ be an N-partite (odd N) non-signalling probability distribution maximally violating the corresponding Mermin inequality. Then the intrinsic and the observed randomness of the function f are equal for any input $\mathbf{x}$ appearing in the Mermin inequality:*

$$G_{\mathrm{int}}(f, \mathbf{x}, P_{\mathrm{obs}}) = G_{\mathrm{obs}}(f, \mathbf{x}, P_{\mathrm{obs}})$$

where

$$G_{\mathrm{obs}}(f, \mathbf{x}, P_{\mathrm{obs}}) \quad = \quad \max_{k\in\{+1,-1\}} P_{\mathrm{obs}}(f(\mathbf{a}) = k|\mathbf{x})$$

Proof of Theorem 1. Since $P_{\rm obs}$ maximally and algebraically violates the Mermin inequality, all the extremal distributions $P_e^{\rm ex}$ appearing in its decomposition must also necessarily lead to the maximal violation of the Mermin inequality (see Supplementary Information for details). Hence, the randomness of f in these distributions as well satisfies Eqn. (8) of Lemma 1. Using this, we find,

$$\begin{aligned} G_{\rm obs}(f, \mathbf{x}, P_e^{\rm ex}) &= \max_{k\in\{+1,-1\}} P_e^{\rm ex}(f(\mathbf{a}) = k|\mathbf{x}) \\ &= |P_e^{\rm ex}(f(\mathbf{a}) = h_N|\mathbf{x}) - 1/2| + 1/2 \\ &= P_e^{\rm ex}(f(\mathbf{a}) = h_N|\mathbf{x}), \end{aligned} \tag{13}$$

for every e. Therefore,

$$\begin{aligned} G_{\rm int}(f, \mathbf{x}, P_{\rm obs}) &= \max_{\{p(e|\mathbf{x}), P_e^{\rm ex}\}} \sum_e p(e|\mathbf{x}) G_{\rm obs}(f, \mathbf{x}, P_e^{\rm ex}) \\ &= \max_{\{p(e|\mathbf{x}), P_e^{\rm ex}\}} \sum_e p(e|\mathbf{x}) P_e^{\rm ex}(f(\mathbf{a}) = h_N|\mathbf{x}) \\ &= P_{\rm obs}(f(\mathbf{a}) = h_N|\mathbf{x}), \end{aligned} \tag{14}$$

where the last equality follows from the constraint $\sum_e p(e|\mathbf{x}) P_e(\mathbf{a}|\mathbf{x}) = P_{\rm obs}(\mathbf{a}|\mathbf{x})$. On the other hand the observed randomness for f is, $G_{\rm obs}(f, \mathbf{x}, P_{\rm obs}) = P_{\rm obs}(f(\mathbf{a}) = h_N|\mathbf{x})$. ◀

The previous technical results are valid for any non-signalling distribution maximally violating the Mermin inequality. For odd N this maximal violation can be attained by a unique quantum distribution, denoted by $P_{\rm ghz}(\mathbf{a}|\mathbf{x})$, resulting from measurements on a Greenberger-Horne-Zeilinger (GHZ) state. When applying Theorem 2 to this distribution, one gets

Main result: Let $P_{\rm ghz}(\mathbf{a}|\mathbf{x})$ be the N-partite (odd N) quantum probability probability distribution attaining the maximal violation of the Mermin inequality. The intrinsic and observed randomness of f for a Mermin input satisfy

$$G_{\rm int/obs}(f, \mathbf{x}, P_{\rm ghz}) = \frac{1}{2} + \frac{1}{2^{(N+1)/2}} \tag{15}$$

This follows straightforwardly from Theorem 2, since $P_{\rm ghz}(\mathbf{a}|\mathbf{x}) = 1/2^{N-1}$ for outcomes $\mathbf{a}$ with an even number of results equal to -1 and for those measurements appearing in the Mermin inequality.

It is important to remark that $f(\mathbf{a}|\mathbf{x}_m)$ approaches a perfect random bit exponentially with the number of parties. In fact, this bit defines a process in which full randomness amplification takes place. Yet, it is not a complete protocol as, contrary to the existing proposal in [5], no estimation part is provided.

5 Discussion

We have seen that for the choice of our function, the observed randomness in distributions maximally violating the Mermin inequality is wholly intrinsic. This includes the physically realizable GHZ correlations. For the latter, the randomness of the function approaches that of a perfect bit exponentially fast in the size of the system. In adversarial terms, this implies that no non-signalling adversary has additional knowledge or can predict the outcome of f better than the parties performing the Bell test.

In the context of the GHZ correlations (being the only correlations in the class we have defined that may be attained by quantum systems), our result bears a resemblance to those in [4, 18] where the completeness of quantum theory was discussed. These results show that the predictive power of quantum theory is maximal. However, our scenario departs significantly from the one considered there. For one thing, we do not assume quantum theory is correct at the level of the dynamics, i.e we do not assume the unitarity of the dynamics, but only at the level of correlations. Besides, we consider a function of the outcomes. Most important of all, our setup allows us to relax the critical free choice assumption arbitrarily, as long as it is not absolute. This was not possible in [4, 18], except perhaps in a very limited sense due to the results of [19].

Furthermore, our results bear a deep relationship with full randomness amplification [20]. Since the free choice can be relaxed and we find our function approaching a perfect random bit with increasing system size, this is precisely the task set out to full randomness amplification. The missing link is the full protocol including estimation, which we do not provide here.

Future directions of work include exploiting such relations to upper bound the classical randomness where exact relations are not possible. Moreover, an interesting line of work is to extend these techniques for distributions non-maximally violating Bell inequalities. These could perhaps lead to experimentally viable tests of fully general device independence randomness certification.

References

1 Niels Bohr. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 48:696–702, 1935.

2 A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

3 John Bell. On the einstein podolsky rosen paradox. *Physics*, 1:195–200, 1964.

4 Roger Colbeck and Renato Renner. No extension of quantum theory can have improved predictive power. *Nat Commun*, 2:411–, 2011.

5 Rodrigo Gallego, Lluis Masanes, Gonzalo de la Torre, Chirag Dhara, Leandro Aolita, and Antonio Acín. Full randomness from arbitrarily deterministic events. *arXiv*, arXiv:1210.6514 [quant-ph], 2012.

6 Jonathan Barrett, Adrian Kent, and Stefano Pironio. Maximally nonlocal and monogamous quantum correlations. *Phys. Rev. Lett.*, 97(17):170409–4, 2006.

7 N. David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.*, 65(15):1838–1840, 1990.

8 B.S. Tsirelson. Quantum analogues of the bell inequalities. the case of two spatially separated domains. *Journal of Soviet Mathematics*, 36:557, 1987.

9 Antonio Acín, Serge Massar, and Stefano Pironio. Randomness versus nonlocality and entanglement. *Phys. Rev. Lett.*, 108(10):100402–, 2012.

10 D. M. Greenberger, M. A. Horne, and A. Zeilinger. *Bell's Theorem, Quantum Theory, and Conceptions of the Universe.* Kluwer, 1989.

11 Stefano Pironio, Jean-Daniel Bancal, and Valerio Scarani. Extremal correlations of the tripartite no-signaling polytope. *Journal of Physics A: Mathematical and Theoretical*, 44(6):065303–, 2011.

12 Johannes Kofler, Tomasz Paterek, and Caslav Brukner. Experimenter'ss freedom in bell's theorem and quantum cryptography. *Phys. Rev. A*, 73(2):022104–, 2006.

13 J. Barrett and N. Gisin. How much measurement independence is needed in order to demonstrate nonlocality? *Arxiv*, arXiv:1008.3612, 2010.

14 Nicolas Gisin. Is realism compatible with true randomness? *arXiv*, arXiv:1012.2536 [quant-ph], 2010.
15 Michael J. W. Hall. Relaxed bell inequalities and kochen-specker theorems. *Phys. Rev. A*, 84(2):022102–, 2011.
16 J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, 2005.
17 Miklos Santha and Umesh V. Vazirani. Generating quasi-random sequences from semi-random sources. *J. Comput. Syst. Sci.*, 33(1):75–87, 1986.
18 R. Colbeck and R. Renner. The completeness of quantum theory for predicting measurement outcomes. *arXiv*, arXiv:1208.4123 [quant-ph], 2012.
19 Roger Colbeck and Renato Renner. Free randomness can be amplified. *Nat Phys*, 8(6):450–454, 2012.
20 Rodrigo Gallego, Lars Erik Würflinger, Antonio Acín, and Miguel Navascues. Operational framework for nonlocality. *Phys. Rev. Lett.*, 109(7):070401–, 2012.
21 Arthur T. Benjamin, Bob Chen, and Kimberly Kindred. Sums of evenly spaced binomial coefficients. *Mathematics Magazine*, 83:370–373, 2010.

A Proof of the main theorem

Here we prove the principal theorem of the main text. It is basically a generalization of the the proof for $N = 3$. We would like to prove that the function f defined in the main text, satisfies the property:

$$P(f(\mathbf{a}) = h_N|\mathbf{x_m}) \geq 1/2 \tag{16}$$

for any N-partite distribution (odd N) that maximally violates the Mermin inequality. As in the tripartite case, in order to prove the result we (I) express condition (16) in terms of some correlators and (II) use positivity conditions from the swapped input to prove the inequality.

An N-partite no-signalling probability distribution $P(\mathbf{a}|\mathbf{x})$ with inputs $\mathbf{x} \in \{0,1\}^N$ and outputs $\mathbf{a} \in \{+1,-1\}^N$ can be parameterized in terms of correlators as,

$$P(\mathbf{a}|\mathbf{x}) = \frac{1}{2^N}\left(1 + \sum_{i=1}^{N} a_i\langle x_i\rangle + \sum_{i<j} a_i a_j \langle x_i x_j\rangle \sum_{i<j<k} a_i a_j a_k \langle x_i x_j x_k\rangle + \cdots + a_1 a_2 \ldots a_N \langle x_1 x_2 \ldots x_N\rangle\right) \tag{17}$$

Restricting $P(\mathbf{a}|\mathbf{x})$ to those maximally violating the N-partite Mermin inequality is equivalent to requiring all correlators of input strings of odd parity to take their extremal values. Namely, we have,

$$\langle x_1 x_2 \ldots x_N\rangle = (-1)^{(-1+\sum_{i=1}^{N} x_i)/2}, \tag{18}$$

for all N-point correlators satisfying $\sum_{i=0}^{N} x_i = 1 \bmod 2$. For instance, $\langle 0,0,\ldots,1\rangle = 1$ and similarly for all permutations. Also, $\langle 0,0,\ldots,0,1,1,1\rangle = -1$ as well as for for all permutations, etc. In the following we will use the notation $\langle . \rangle_k$ to denote a k-point correlator. The input combination used to extract randomness is a generalization of the tripartite case and denoted by $\mathbf{x_m} = (0,0,\ldots,0,1)$. The corresponding N-point correlator satisfies $\langle 0,0,\ldots,0,1\rangle = 1$ for all N. The latter implies two useful relations:

1. Half the total outcomes vanish. In particular these are the terms for which the product of outcomes is -1 *i.e.* $P(\prod_{i=1}^{N} a_i = -1|\mathbf{x_m}) = 0$.
2. $\langle . \rangle_{N-k} = \langle . \rangle_k$ for all $1 \leq k \leq (N-1)/2$ where the correlators $\langle . \rangle_{N-k}$ and $\langle . \rangle_k$ are complementary in the input $\mathbf{x_m}$.

One can use these in Eqn. (17) to express $P(\mathbf{a}|\mathbf{x_m})$ in terms of only the first $(N-1)/2$-point correlators as,

$$P(\mathbf{a}|\mathbf{x_m}) = \frac{1}{2^{N-1}} \left(1 + \sum a_i \langle x_i \rangle + \sum a_i a_j \langle x_i x_j \rangle + \cdots + \sum a_i a_j \cdots a_p \langle x_i x_j \cdots x_p \rangle_{(N-1)/2}\right). \tag{19}$$

where $a_1 \cdot a_2 \cdot a_3 \ldots a_N = +1$ since $P(\mathbf{a}|\mathbf{x_m}) = 0$ when $a_1 \cdot a_2 \cdot a_3 \ldots a_N = -1$.

B Expressing the inequality in terms of correlators

As mentioned, our first goal is to express Eq.(16) as a function of some correlators. Let us recall the function we use in our main theorem,

$$f(\mathbf{a}) = \begin{cases} +1 & n_-(\mathbf{a}) = (4j+2); \quad \text{with } j \in \{0,1,2,\ldots\} \\ -1 & \text{otherwise} \end{cases} \tag{20}$$

where $n_-(\mathbf{a})$ denotes the number of results in $\mathbf{a}$ that are equal to -1.

It turns out that the quantity (Eq. (16)) we would like to calculate, namely, $P(f(\mathbf{a}) = h_N|\mathbf{x}_m) - 1/2$ can be equivalently expressed as $h_N \cdot (P(f(\mathbf{a}) = +1|\mathbf{x}_m) - 1/2)$. The latter form is convenient since the function only takes value $+1$ for all N.

We proceed to express the latter in terms of correlators (as in the proof for three parties in the main text),

$$(h_N \cdot P(f(\mathbf{a}) = +1|\mathbf{x}_m) - 1/2) = 2^{-(N-1)} \boldsymbol{\alpha}' \cdot \mathbf{c}, \tag{21}$$

where

$$\begin{aligned} \boldsymbol{\alpha}' &= h_N \cdot (\alpha_0 - 2^{N-2}, \alpha_1, \alpha_2, \ldots, \alpha_{(N-1)/2}) \\ \mathbf{c} &= \left(1, \sum_{\mathcal{S}^1} \langle . \rangle_1, \sum_{\mathcal{S}^2} \langle . \rangle_2, \ldots, \sum_{\mathcal{S}^{(N-1)/2}} \langle . \rangle_{(N-1)/2}\right) \end{aligned} \tag{22}$$

Note that, since the function f symmetric under permutations, the vector $\mathbf{c}$ consists of the different sums of all k-point correlators, denoted by $\mathcal{S}^k$, where k ranges from 0 to $(N-1)/2$ because of Eq. (19). The vector $\boldsymbol{\alpha}'$ is the vector of coefficients for each sum of correlators. Our next goal is to compute this vector.

Recall that function f is such that $f(\mathbf{a}) = +1$ if $n_-(\mathbf{a}) = 4j+2$ for any $j \in \mathbb{N} \cup \{0\}$. By inspection, the explicit values of α_i can be written as

$$\alpha_i = \sum_{r=0}^{i} (-1)^r \binom{i}{r} \sum_{j \geq 0} \binom{n-i}{4j+2-r}. \tag{23}$$

For example, $\alpha_0 = \sum_{j \geq 0} \binom{n}{4j+2}$ as one would expect since α_0 simply counts the total number of terms $P(\mathbf{a}|\mathbf{x}_m)$ being summed to obtain $P(f(\mathbf{a}) = +1|\mathbf{x}_m)$.

Making use of the closed formula $\sum_{j \geq 0} \binom{n}{rj+a} = \frac{1}{r} \sum_{k=0}^{r-1} \omega^{-ka} (1+\omega^k)^n$ [21], where $\omega = e^{i2\pi/r}$ is the r^{th} root of unity, we can simplify the second sum appearing in Eq. 23. Finally we recall that the phase h_N was defined (in the main text) to be $h_N = \sqrt{2}\cos(N+4)\pi/4$. Putting all this together and performing the first sum in Eq. (23) gives us,

$$\alpha'_i = 2^{\frac{N-3}{2}}\left(-2\cos\frac{(N-2i)\pi}{4}\cos\frac{(N+4)\pi}{4}\right) \tag{24}$$

Notice that the term in the parenthesis is a phase taking values in the set $\{+1,-1\}$ since N is odd while the amplitude is independent of N. Thus, we can simplify Eqn. (24) for even and odd values of i as,

$$\alpha'_i = \begin{cases} 2^{(N-3)/2}(-1)^{\frac{N-i}{2}} & i \text{ odd} \\ 2^{(N-3)/2}(-1)^{\frac{i}{2}} & i \text{ even} \end{cases} \tag{25}$$

Thus, to prove that f possesses the property $h_N \cdot (P(f(\mathbf{a}) = +1|\mathbf{x_m}) - 1/2) \geq 0$ necessary to proving the main theorem is equivalent to proving

$$\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0, \tag{26}$$

for $\mathbf{c}$ as defined in Eqn. (22) and for the values of $\boldsymbol{\alpha}'$ given by Eqn. (25). This is the task of the following section, where we show that it follows from positivity constraints on $P(\mathbf{a}|\mathbf{x})$.

C Proving the inequality from positivity constraints

We show that positivity conditions derived from the swapped input $\bar{\mathbf{x}}_\mathbf{m} = (1,1,\ldots,1,0)$ may be used to show $\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0$. Notice that the components of $\bar{\mathbf{x}}_\mathbf{m}$ and $\mathbf{x_m}$ are opposite, ie. $\{\bar{\mathbf{x}}_\mathbf{m}\}_i = \{\mathbf{x_m}\}_i \oplus 1$ for all i. In the following we will repeatedly use the Mermin conditions of Eqn. (18).

We start by summing the positivity conditions $P(++\,+\cdots+-|\bar{\mathbf{x}}_\mathbf{m}) \geq 0$ and $P(--\,-\cdots-+|\bar{\mathbf{x}}_\mathbf{m}) \geq 0$. Using Eqn. (17), one can easily see that upon summing, all k-point correlators for *odd* k are cancelled out since these are multiplied by coefficients (products of a_is) that appear with opposite signs in the two positivity expressions. In contrast, k-point correlators for *even* k add up since they are multiplied by coefficients that appear with the same sign in the two expressions. For example, N being odd, the full correlator always cancels out while the $(N-1)$-point correlators always appear.

This leaves us with an expression containing only the even-body correlators,

$$1 + \sum_{i<j} a_i a_j \langle x_i x_j \rangle + \sum_{i<j<k<l} a_i a_j a_k a_l \langle x_i x_j x_k x_l \rangle + \cdots + \sum a_i \ldots a_p \underbrace{\langle x_i \ldots x_p \rangle}_{(N-1)\text{-pt. corr}} \geq 0. \tag{27}$$

Note once again, that this inequality is derived from the so-called swapped input $\bar{\mathbf{x}}_\mathbf{m}$. We aim to cast it in a form that can be compared directly with Eqn. (22), which comes from the chosen Mermin input $\mathbf{x_m}$. To this end, we need to convert Eqn. (27) to an expression of the form,

$$(\beta_0, \beta_1, \ldots \beta_{(N-1)/2}). \qquad \left(1, \sum\langle.\rangle_1, \ldots, \sum\langle.\rangle_{(N-1)/2}\right) \geq 0 \tag{28}$$

We first highlight the similarities and differences between the two preceding expressions, namely, the one we have *i.e.* Eqn. (27) and the one we want, *i.e.* Eqn. (28). Each contains $(N-1)/2$ distinct classes of terms. However the former contains only even k-point correlators for $k = 2$ to $(N-1)$ while the latter contains all terms from $k = 1$ to $(N-1)/2$. Thus, terms of Eq. (27) must be mapped to ones in Eqn. (28). Moreover, since the point of making this mapping is to finally compare with Eqn. (22), we also note that the correlators appearing in Eqn. (27) are locally swapped relative to those appearing in Eqn. (22). Thus, our mapping must also convert correlators of the swapped input into those corresponding to the chosen input.

We demonstrate next that one may indeed transform the inequality (27) into the inequality (28) satisfying both the demands above. To this end, all the *even* k-point correlators (for $k \geq \frac{N-1}{2}$) appearing in Eqn. (27) are mapped to odd $(N-k)$-point correlators in Eqn. (28). Likewise, all the even k-point correlators (for $k < \frac{N-1}{2}$) of the swapped input appearing in Eqn. (27) are mapped to the corresponding k-point correlators of the chosen input in Eqn. (28).

These mappings make systematic use of the Mermin conditions Eqn. (18) and are made explicit in the following section.

C.1 Even-point correlators

Consider a $2k$-point correlator where $2k \leq (N-1)/2$. The correlators are of two forms and we show how they are transformed in each case:

- $\langle 11\ldots1\rangle_{2k}$. We would like to map this to the correlator $\langle 00\ldots0\rangle_{2k}$ appearing in $\mathbf{x_m}$. We achieve the mapping by completing each to the corresponding Mermin full-correlators $\langle\underbrace{11\ldots1}_{2k}\underbrace{100\ldots0}_{(N-2k)}\rangle_N = (-1)^k$ and $\langle\underbrace{00\ldots0}_{2k}\underbrace{100\ldots0}_{(N-2k)}\rangle_N = (-1)^0 = 1$. From the signs, we have the relation, $\langle 11\ldots1\rangle_{2k} = (-1)^k\langle 00\ldots0\rangle_{2k}$
- $\langle 11\ldots10\rangle_{2k}$, which we would like to map to $\langle 00\ldots01\rangle_{2k}$. Using the same ideas we get $\langle\underbrace{11\ldots10}_{2k}\underbrace{110\ldots0}_{(N-2k)}\rangle_N = (-1)^k$ and $\langle\underbrace{00\ldots01}_{2k}\underbrace{110\ldots0}_{(N-2k)}\rangle_N = (-1)^1 = -1$. Thus, giving us the relation $\langle 11\ldots10\rangle_{2k} = (-1)^{k+1}\langle 00\ldots01\rangle_{2k}$.

By inspection one can write the relationship

$$\underbrace{a_1a_2\ldots a_{2k}}_{\text{even}}\underbrace{\langle x_1x_2\ldots x_{2k}\rangle}_{\text{cor in }\bar{\mathbf{x}}_\mathbf{m}} = (-1)^k \underbrace{\langle x_1x_2\ldots x_{2k}\rangle}_{\text{cor in }\mathbf{x_m}(desired)}$$

for correlators of either form discussed above on multiplying with their corresponding coefficients. Since we have finally converted to the desired correlators of the chosen input $\bar{\mathbf{x}}$, we can read off β_i as the corresponding phase. Thus, $\beta_i = (-1)^{i/2}$ for even i.

C.2 Odd-point correlators

Consider now a $2k$-point correlator where $2k \geq (N-1)/2$. The correlators are again of two forms and may be transformed to the required $(N-2k)$-point correlators in each case. The only difference from before is that the two correlators are now complementary to each other in the swapped input. Since the details are similar, we simply state the final result $\beta_i = (-1)^{(N-i)/2}$ for odd i.

The final expression thus reads,

$$\beta_i = \begin{cases} (-1)^{\frac{N-i}{2}} & i \text{ odd} \\ \\ (-1)^{\frac{i}{2}} & i \text{ even} \end{cases} \tag{29}$$

Thus, the values of β given in Eqs. (29) exactly match the ones for α'_i (up to the constant factor) given in Eqn. 25. Together with the correlators matching those in $\mathbf{c}$, it proves that f satisfies the required $\boldsymbol{\alpha}' \cdot \mathbf{c} \geq 0$ and hence the full result.

D Proof that all distributions in decomposition maximally violate the Mermin inequality

We end by proving the claim made in the main text that if an observed probability distribution $P_{\text{obs}}(\mathbf{a}|\mathbf{x})$ violates maximally and algebraically the corresponding Mermin inequality, all the no-signaling components $P_e^{\text{ex}}(\mathbf{a}|\mathbf{x})$ present in its preparation must also algebraically violate the inequality.

We recall that the decomposition appears in the definition of intrinsic randomness given by,

$$G_{\text{int}}(g, \mathbf{x}, P_{\text{obs}}) = \max_{\{p(e|\mathbf{x}), P_e^{\text{ex}}\}} \sum_e p(e|\mathbf{x}) G_{\text{obs}}(g, \mathbf{x}, P_e^{\text{ex}})$$

subject to:

$$\sum_e p(e|\mathbf{x}) P_e^{\text{ex}}(\mathbf{a}|\mathbf{x}) = P_{\text{obs}}(\mathbf{a}|\mathbf{x}) \tag{30}$$

$$p(\mathbf{x}|e) \geq \delta \;\text{ with } \delta > 0 \;\; \forall\, \mathbf{x}, e \tag{31}$$

Since P_{obs} algebraically violates the Mermin inequality, this definition imposes stringent conditions on the correlators of P_{obs} satisfying the Mermin condition (18), namely that,

$$\langle x_1 \dots x_N \rangle_{P_{\text{obs}}} = \pm 1 = \sum_e p(e|x_1, \dots, x_N)\, \langle x_1 \dots x_N \rangle_{P_e^{\text{ex}}} \tag{32}$$

where by normalization $\sum_e p(e|x_1, \dots, x_N) = +1$ and $-1 \leq \langle x_1 \dots x_N \rangle_{P_e^{\text{ex}}} \leq +1$. Note that condition $p(\mathbf{x}|e) \geq \delta$ for all $\mathbf{x}, e$ for $\delta > 0$ can be inverted using the Bayes' rule to obtain $p(e|\mathbf{x}) > 0$ for all $\mathbf{x}, e$. Now is clear by convexity that the condition $p(\mathbf{x}|e) \geq \delta$ (denying *absolute* relaxation of freedom of choice) implies that all the correlator $\langle x_1 \dots x_N \rangle_{P_e^{\text{ex}}}$ appearing in the Mermin inequality must also necessarily satisfy $\langle x_1 \dots x_N \rangle_{P_e^{\text{ex}}} = \pm 1$ for all e thus maximally violating the Mermin inequality. In fact it is also clear that this constraint on $p(\mathbf{x}|e)$ is strictly necessary to ensure that the decomposition correlations satisfy maximal Mermin violation. To see this, suppose $p(\mathbf{x}|e_0) = 0$, then the corresponding $\langle x_1 \dots x_N \rangle_{P_{e_0}^{\text{ex}}}$ is fully unconstrained while satisfying Eq. (32).