Maximally Nonlocal Theories Cannot Be Maximally Random

Gonzalo de la Torre,^{1,*} Matty J. Hoban,^{1,3,†} Chirag Dhara,¹ Giuseppe Pretico,¹ and Antonio Acín^{1,2}

¹*ICFO-Institut de Ciències Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*

²*ICREA-Institució Catalana de Recerca i Estudis Avançats, E-08010 Barcelona, Spain*

³*Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom*

(Received 20 March 2014; revised manuscript received 10 October 2014; published 22 April 2015)

Correlations that violate a Bell inequality are said to be nonlocal; i.e., they do not admit a local and deterministic explanation. Great effort has been devoted to study how the amount of nonlocality (as measured by a Bell inequality violation) serves to quantify the amount of randomness present in observed correlations. In this work we reverse this research program and ask what do the randomness certification capabilities of a theory tell us about the nonlocality of that theory. We find that, contrary to initial intuition, maximal randomness certification cannot occur in maximally nonlocal theories. We go on and show that quantum theory, in contrast, permits certification of maximal randomness in all dichotomic scenarios. We hence pose the question of whether quantum theory is optimal for randomness; i.e., is it the most nonlocal theory that allows maximal randomness certification? We answer this question in the negative by identifying a larger-than-quantum set of correlations capable of this feat. Not only are these results relevant to understanding quantum mechanics' fundamental features, but also put fundamental restrictions on device-independent protocols based on the no-signaling principle.

DOI: 10.1103/PhysRevLett.114.160502

PACS numbers: 03.65.Ud, 03.67.Ac

From a physical perspective, all classical physics is deterministic and any apparent randomness is due to ignorance therefore not exhibiting *intrinsic* randomness. Quantum theory is open to such a possibility since it is a fundamentally probabilistic theory. However, since the early days of quantum theory, its seemingly “intrinsic” unpredictability has been heavily debated even by some of its founding fathers [1,2]. A great advance came when John Bell [3] identified limitations on any theory founded on the following two basic physical principles: impossibility of instantaneous signaling between distant locations (no-signaling principle); and the existence of a complete set of variables of a system which, if known, would allow for deterministic predictions. Correlations among a number of distant parties that satisfy both principles are called “local” and are constrained by the now-eponymous Bell inequalities. Thus Bell established a fundamental link between the unpredictability of quantum mechanics and the concept of nonlocality [3]. In particular, assuming the validity of the no-signalling principle, a violation of a Bell inequality implies and certifies intrinsic randomness.

Recently, this deep connection between nonlocality and randomness has been made quantitative and exploited for information processing tasks [4,5]. Nonlocality-certified randomness represents an information resource in the now well-established area of “device-independent quantum information processing” [6–11]. In a device-independent protocol, no assumption is made about the inner workings of the devices used and are thus regarded as black boxes. There is, however, a crucial assumption to every protocol and that is the assumption of the background theory

dictating the devices' behavior, e.g., whether the devices are quantum mechanical [10], or just compatible with the no-signaling principle [11].

The assumption about the background theory is *vital* given that quantum mechanics is not the most nonlocal theory respecting the no-signaling principle [12] and therefore capable of producing intrinsic randomness. Theories allowing for all nonlocal correlations *only* restricted by the no-signaling principle are termed “maximally nonlocal” since they produce the most nonlocality that a nonsignalling theory can produce. Given the eminent role of nonlocality for randomness certification, one would intuitively expect maximally nonlocal theories to have more powerful randomness certification capabilities than other theories. Indeed, there are occasions where maximally nonlocal theories can certify randomness and quantum mechanics cannot even certify any randomness at all [13]. On the other hand, for the Clauser-Horne-Shimony-Holt (CHSH) inequality [14], we can certify more randomness assuming only quantum mechanics (however, not the maximal amount possible) rather than allowing maximally nonlocal correlations [15].

The main goal of our work is to understand the relationship between the nonlocality and randomness of a theory and, in particular, what the randomness capabilities of a theory tell us about the nonlocality allowed within that theory. Notice that we do not consider misalignments, preparation errors, or detection efficiencies as those sources of randomness are not fundamental to the physical theory and can, in principle, be reduced below any finite threshold. We are hence interested in studying only the fundamental differences in randomness certification between theories.

The first result is to show that the previous intuition is wrong: were the set of achievable physical correlations not more restricted than what the no-signaling principle allows, maximal randomness could not be certified in any possible scenario irrespective of the number of parties, measurements or outcomes; i.e., maximally nonlocal theories cannot be maximally random.

Second, we focus on quantum theory and provide, in contrast, scenarios with an arbitrary number of parties where maximal randomness *can* be certified. This should be compared with other works that showed that if maximally nonlocal theories were permitted in nature we would have unimaginable computational and communicating power [16,17]. Here, being in a maximally nonlocal world *limits* our information processing capabilities. This observation leads us to ask if the nonlocality of quantum theory is in some sense optimal for randomness certification. That is, is quantum theory the most nonlocal theory capable of certifying maximal randomness? Our final result answers this question in the negative: we identify a set of correlations larger than the quantum set that also permits the certification of maximal randomness.

Boxes and Bell tests.—We use the scenario of a Bell test to study the correlations observed among spacelike separated measurements on systems within different physical theories. There are N distant parties and each party makes a choice of measurement upon their system. These processes are arranged so that they define spacelike separated events. The N users have no knowledge of how a system or its measurement apparatus are prepared, they can only make measurement choices and observe classical outcomes. We even allow the possibility that a malicious agent prepared the devices and holds information about how they prepared their systems. We then model these parties as black boxes with the measurement choice for the j th party (for $j \in \{1, 2, \dots, N\}$) being an input $x_j \in \{0, 1, \dots, (M-1)\}$ where there are M possible choices; the measurement outcome for this j th party is then the output $a_j \in \{0, 1, \dots, (d-1)\}$ where for every party there are d possible outcomes to the every measurement. Therefore, a string of “dits” (generalization of bits to d values) is produced in each round of a Bell test. The Bell test is then labeled by the parameters (N, M, d) .

After a suitable number of uses of the boxes, the conditional probabilities $p_{\text{obs}}(\mathbf{a}|\mathbf{x})$ for all values of $\mathbf{a} = (a_1, \dots, a_N)$ and $\mathbf{x} = (x_1, \dots, x_N)$ that describe the observed process are obtained. These conditional probabilities form a full distribution P_{obs} with elements $p_{\text{obs}}(\mathbf{a}|\mathbf{x})$. In general, we will use an upper-case P for a distribution and lower-case p for an element of that distribution.

As mentioned, we must make an assumption on the *theory* that governs the workings of our boxes. This has the effect of indicating whether our observed correlations P_{obs} belong to a particular set of possible correlations. For example, if we say that our observed correlations result from a classical system or quantum system then the

distribution P_{obs} belongs to the set C or Q of correlations resulting from all possible classical and quantum systems, respectively. We can also define the set NS of all maximally nonlocal correlations.

The set of quantum correlations Q is contained in NS . This does not imply that the latter is as random as the former. The important distinction is that by dictating which theory is permitted, we bound the power of the malicious agent that can, in principle, prepare our devices. Therefore, upon obtaining our observed statistics P_{obs} , if we vary the theory that describes the source of these correlations then we allow a malicious agent to prepare the devices using different (even supraquantum) resources. The agent can then use this knowledge of the preparation to improve their predictive power thus leading to different implications for randomness certification.

Every set is convex because we can always prepare a convex mixture (by tossing a biased coin) of systems thus giving a convex mixture of correlations resulting from each system. So, in addition to the Bell scenario dictated by (N, M, d) we stipulate the set of correlations T to which our observed correlations can belong.

Every convex set can be described in terms of its extreme points. In the case of probability distributions P , the extreme points are those distributions that cannot be expressed as a convex combination of other distributions in the set. An immediate corollary of this property of convex sets is that the observed correlations P_{obs} can be written as a convex combination of the extreme points of a set T ; therefore $P_{\text{obs}} = \sum_{\text{ext}} q_{\text{ext}} P_{\text{ext}}$ where P_{ext} is an extreme point of the set T and $q_{\text{ext}} \geq 0$ is a probability distribution over these points such that $\sum_{\text{ext}} q_{\text{ext}} = 1$.

Randomness certification.—In randomness certification we take the standard approach of using just one particular $\mathbf{x}_0$ from which to obtain a dit string of length N that is hopefully random. The (potentially malicious) provider of the boxes knows this string $\mathbf{x}_0$ in advance but, however, does not know *when* this string is input into the boxes. Correlations for the rest of the inputs $\{\mathbf{x}|\mathbf{x} \neq \mathbf{x}_0\}$, encapsulated by P_{obs} , are used to certify that the randomness obtained from $\mathbf{x}_0$ is intrinsically random. To measure the randomness of the outputs obtained from input $\mathbf{x}_0$ given observed correlations P_{obs} , we require the *guessing probability* $G^T(\mathbf{x}_0, P_{\text{obs}})$: the probability for a malicious agent to predict the most likely outcome for input $\mathbf{x}_0$ given that the agent has complete knowledge of how a box is prepared within a theory with a corresponding set of correlations T . The larger this guessing probability the less random are the outputs and thus we have a measure of randomness.

Since the user of the box has no knowledge of how it was prepared, we must assume that all possible ways of producing P_{obs} from extreme points P_{ext} can be utilized such that $P_{\text{obs}} = \sum_{\text{ext}} q_{\text{ext}} P_{\text{ext}}$ and the malicious agent knows q_{ext} perfectly. Indeed, the agent may know what is the most advantageous distribution q_{ext} to maximize his

chances of guessing the output. On the other hand, for each of these extreme points P_{ext} we can evaluate the guessing probability easily since there is a unique way of preparing this probability distribution (from the set of correlations T). Therefore, the guessing probability for extreme points is $G^T(\mathbf{x}_0, P_{\text{ext}}) = \max_{\mathbf{a}} p_{\text{ext}}(\mathbf{a}|\mathbf{x}_0)$ where $p_{\text{ext}}(\mathbf{a}|\mathbf{x}_0)$ is an element of P_{ext} . Collating all of this information, we obtain the following optimization:

$$G^T(\mathbf{x}_0, P_{\text{obs}}) = \max_{\{q_{\text{ext}}, P_{\text{ext}}\}} \sum_{\text{ext}} q_{\text{ext}} G^T(\mathbf{x}_0, P_{\text{ext}}),$$

subject to

$$P_{\text{obs}} = \sum_{\text{ext}} q_{\text{ext}} P_{\text{ext}}, \quad P_{\text{ext}} \in T. \quad (1)$$

Immediately we see that for classical correlations (when $T = C$) $G^C(\mathbf{x}_0, P_{\text{obs}}) = 1$ since $\max_{\mathbf{a}} p_{\text{ext}}(\mathbf{a}|\mathbf{x}_0) = 1$ for all P_{ext} . This follows from the well-known fact that any classical correlations can be decomposed as a mixture of deterministic points. This highlights the need for nonclassical, or nonlocal correlations for randomness certification.

It is worth noting that we are explicitly assuming the independence between the preparation components labeled by P_{exp} and the measurement settings $\mathbf{x}$. This is commonly known as the *freedom of choice* assumption. Recent work has shown that this assumption can even be relaxed by implementing randomness amplification protocols [8,9,18].

In what follows we will perform the optimization in Eq. (1) for different sets of correlations, in particular maximally nonlocal and quantum correlations labeled NS and Q , respectively. In particular, we ask the question of whether a theory with correlations T can certify maximal randomness, which exactly means if for any observed correlations $P_{\text{obs}} \in T$ in any scenario (N, M, d) , we can obtain $G^T(\mathbf{x}_0, P_{\text{obs}}) = (1/d^N)$.

Maximally nonlocal correlations.—The set NS of maximally nonlocal correlations is the set of multipartite correlations solely restricted by the no-signaling principle. Here we permit any valid normalized probability distribution P with all elements satisfying $1 \geq p(\mathbf{a}|\mathbf{x}) \geq 0$ where marginals are well defined. That is, the probabilities (correlations) satisfy $\sum_{\mathbf{a}} p(\mathbf{a}|\mathbf{x}) = 1$. To prevent instantaneous signaling it is important that

$$\sum_{a_k} p(a_1, \dots, a_k \dots a_N | x_1, \dots, x_k, \dots, x_N) \quad (2)$$

is independent of x_k for all k .

Now that this set is defined we present our first result.

Result 1: Maximally nonlocal theories can never be maximally random.

Were the physically achievable correlations solely restricted by the no-signaling principle, the maximum amount of certifiable randomness in an arbitrary Bell

scenario (N, M, d) would be bounded through the intrinsic predictability by

$$G^{NS}(\mathbf{x}_0, P_{\text{obs}}) \geq \frac{1}{d^N - (d-1)^N}, \quad (3)$$

for any probability distribution $P_{\text{obs}} \in NS$ and all inputs $\mathbf{x}_0$.

To prove this result we only need to consider the randomness of the extreme points P_{ext} of NS as indicated by Eq. (1). Our proof is based on the simple observation that if for a particular $\mathbf{x}_0$ of correlations $p(\mathbf{a}|\mathbf{x})$, n values are equal to zero then $\max_{\mathbf{a}} p(\mathbf{a}|\mathbf{x}_0) \geq (1/d^N - n)$. Result 1 then follows from Theorem 1 in the Supplemental Material [19], which proves that some given non-signalling correlations $p(\mathbf{a}|\mathbf{x})$ cannot be extreme if there exists a string of inputs $\mathbf{x}_0$ such that the number of terms $p(\mathbf{a}|\mathbf{x}_0)$ that are equal to zero is smaller than $(d-1)^N$.

It is worth mentioning two facts. First, this result indicates an important limitation on maximally nonlocal theories. In fact, the gap between the ideal maximal randomness and that achievable in maximally nonlocal theories is unbounded. Second, the derived bound is, in general, not tight. For instance, all extreme nonsignaling correlations in Bell test scenarios $(2, M, 2)$ were obtained in [20,21] and in this case $G^{NS}(\mathbf{x}_0, P_{\text{obs}}) \geq 1/2$ whereas our bound gives $1/3$. Interestingly, the same difference appears in the $(3, 2, 2)$ scenario: looking at all the extreme points, classified in Ref. [22], the maximal randomness is equal to $1/6$, while our bound predicts $1/7$. However, in the asymptotic limit of $d \rightarrow \infty$ our bound gives $[1/O(d^{N-1})]$, which can be shown to be tight by comparing it with the results in Ref. [23]. We now move to randomness certification in quantum theory.

Quantum correlations.—Let $\rho \geq 0$ be some quantum state and $O_{a_j}^{x_j}$ be some measurement operators (technically a positive operator valued measure, POVM) for input x_j and output a_j . We say a probability distribution $P_{\text{obs}} \in Q$ belongs to the quantum set of correlations if it can be written as $p_{\text{obs}}(\mathbf{a}|\mathbf{x}) = \text{tr}(\rho \bigotimes_{j=1}^N O_{a_j}^{x_j})$.

Characterizing the set of correlations achievable in this way is a great open problem in quantum information theory. Therefore, in what follows, rather than solving exactly the optimization problem [Eq. (1)], we consider a relaxation that provides a lower bound to the intrinsic randomness. Instead of considering all convex combinations of extreme points of Q that reproduce the observed statistics, we ask for a convex combination of extreme points that give an observed violation of a Bell inequality. Given that a Bell inequality is just a linear combination of probabilities $p(\mathbf{a}|\mathbf{x})$ over all inputs $\mathbf{a}$ and outputs $\mathbf{x}$, let us define the following inner product between correlations P_{obs} and Bell inequality B that computes the Bell violation $B \cdot P_{\text{obs}} \equiv \sum_{\mathbf{a}, \mathbf{x}} \beta_{\mathbf{a}, \mathbf{x}} p_{\text{obs}}(\mathbf{a}|\mathbf{x}) = q_{\text{obs}}$, where the real coefficients $\beta_{\mathbf{a}, \mathbf{x}}$ define the Bell inequality B .

Computing a lower bound to the intrinsic predictability $G^Q(\mathbf{x}_0, P_{\text{obs}})$, certified this time by an observed violation of a Bell inequality, then amounts to solving the following optimization problem, a relaxation of Eq. (1):

$$G^Q(\mathbf{x}_0, P_{\text{obs}}) \leq \max_{\{q_{\text{ext}}, P_{\text{ext}}\}} \sum_{\text{ext}} q_{\text{ext}} G^T(\mathbf{x}_0, P_{\text{ext}}),$$

subject to

$$\sum_{\text{ext}} q_{\text{ext}} (B \cdot P_{\text{ext}}) = q_{\text{obs}}, \quad P_{\text{ext}} \in Q. \quad (4)$$

Since we are interested in the maximal amount of randomness allowed by quantum mechanics, we will restrict our study to maximal quantum violation of a Bell inequality $q_{\text{obs}} \equiv q_{\text{max}}$. In Ref. [24] a method was provided to detect when the maximal quantum violation of a Bell inequality certifies that the outputs are maximally random. The method has the advantage that it can be easily applied, but unfortunately it only works under the assumption that the maximal quantum violation of the inequality is unique. The uniqueness of the maximal quantum violation is in general hard to prove. However, in what follows, we consider Bell inequalities for which the uniqueness of the maximal violation can be proven using the results of Refs. [25,26]. This then allows us to apply the simple method in Ref. [24] and prove the following result.

Result 2: Quantum theory is maximally random in all dichotomic scenarios.

Assuming the set of physically achievable correlations to be the quantum set, the maximum amount of certifiable randomness in the family of Bell test scenarios $(N, M, 2)$ is maximal: $G^Q(\mathbf{x}_0, P_{\text{obs}}) = (1/2^N)$.

We prove this result in Ref. [19] by generalizing the results of Ref. [24] to all N via a Bell inequality introduced in Ref. [27]. We actually prove Result 2 for the $(N, 2, 2)$ scenario but this trivially applies to the $(N, M, 2)$ since we can always ignore $(M - 2)$ of the inputs for each party. While our proof does not apply to the case of two parties, it has been shown in Ref. [28] that for the $(2, 2, 2)$ scenario an amount of randomness arbitrarily close to the maximum of 2 random bits can be certified in some limit. Additionally, numerical and analytical evidence indicates that exactly 2 bits of maximal randomness can be attained in the $(2, 3, 2)$ scenario [24]. All of this serves to show that quantum correlations certify maximal randomness even if maximally nonlocal theories can never do this.

We have shown the difference for randomness certification of two sets of correlations: the maximally nonlocal set and the quantum set. A natural question is whether this contrast highlights the *uniqueness* of quantum correlations. Just as various information-theoretic principles aim to highlight single-out quantum theory [16,17,29], is Q the only set capable of certifying maximal randomness? We now address this question.

Supraquantum correlations.—Navascués, Pironio, and Acín introduced a means to approximate the set of quantum correlations which was an infinite hierarchy of semidefinite programs [30]. For example, the first nontrivial level of this hierarchy is Q^1 and this set is provably larger than the set of quantum correlations Q [31]. Already in the work of Pironio *et al.* in Ref. [6] these first few levels in the hierarchy were used to lower bound the amount of randomness certified for quantum correlations. In the Supplemental Material [19], we introduce a modification to the set Q^1 in the tripartite setting called Q^{1+ABC} that is strictly larger than the quantum set. On the other hand, this set also allows for maximal randomness certification. This represents the third main result of this work.

Result 3: There exist postquantum theories that can also certify maximal randomness.

Were the physically achievable correlations those of the strictly larger than quantum set Q^{1+ABC} [19], maximal randomness could also be certified in the Bell test scenario $(3, M, 2)$, i.e., $G^{Q^{1+ABC}}(\mathbf{x}_0, P_{\text{obs}}) = \frac{1}{8}$.

The proof of this result is presented in the Supplemental Material [19]. The crucial element in this proof is showing that there is only one probability distribution in the set Q^{1+ABC} that maximally violates the Mermin inequality [32] allowing us to use the results in Ref. [24].

At first, this result may seem disappointing but there are other examples of limitations to recovering quantum correlations from information principles. For example, it is known that we need truly multipartite principles [33]. It has also been shown that other information principles will never recover quantum mechanical correlations [31] and our work fits squarely within this foundational research program.

Discussion.—We have shown that correlations in maximally nonlocal theories and quantum theory have drastically different consequences for randomness certification. Therefore, if we assume nature does not abide by a nonlocality-restricted theory such as quantum theory it could severely limit its randomness capabilities. One can see this as a result of maximally nonlocal correlations having correlations between the outputs *for all* inputs, but quantum theory cannot produce such strong correlations.

Let us illustrate this point by revisiting the CHSH scenario of $(2, 2, 2)$. Here all extremal correlations in maximally nonlocal theories are equivalent to the so-called Popescu-Rorhlich box [12]. This box always satisfies the condition $x_1 \cdot x_2 = a_1 \oplus a_2 \bmod 2$ [34]. Therefore, knowing the inputs and one of the outputs, we can *perfectly determine* the other output. Quantum correlations, however, cannot produce these perfect correlations thus introducing more randomness.

These results are not only of foundational interest but have application in randomness extraction, certification, and amplification. For example, in Ref. [6] a lower bound on certifiable randomness was obtained using only the

no-signaling principle, and this bound has found applications in other protocols (e.g., Ref. [35]). An interesting follow-up question is to determine the *exact* maximum randomness allowed just by the no-signalling principle, a fundamental number providing a quantitative link between randomness and no-signalling.

We also showed that certain supraquantum correlations can also exhibit maximal randomness. This last result indicates that quantum theory is not so special from an information-theoretic perspective (cf. Ref. [31]). The set of quantum correlations is notoriously difficult to define but maximally nonlocal theories have a simple description. The fact that there exists a set of correlations that has a relatively simple description but facilitates maximal randomness certification provides a “third way” for the design and analysis of future protocols.

We acknowledge financial support from the ERC Consolidator Grant QITBOX, the John Templeton Foundation, an SGR from the Generalitat de Catalunya, and the Spanish MINECO, through an FPI grant and projects Intrinqra and Chist-Era DIQIP. This work was also supported by the FQXi Large Grant “Thermodynamic vs information theoretic entropies in probabilistic theories” and by the EPSRC, through the NQIT Quantum Hub.

*gonzalo.delatorre@icfo.es

†matthew.hoban@icfo.es

- [1] A. Einstein, B. Podolsky, and N. Rosen, *Phys. Rev.* **47**, 777 (1935).
- [2] N. Bohr, *Phys. Rev.* **48**, 696 (1935).
- [3] J. S. Bell, *Physics* **1**, 195 (1964).
- [4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (2014).
- [5] V. Scarani, *Nature (London)* **464**, 988 (2010).
- [6] Pironio *et al.*, *Nature (London)* **464**, 1021 (2010).
- [7] R. Colbeck, PhD Thesis, University of Cambridge, 2006.
- [8] R. Colbeck and R. Renner, *Nat. Phys.* **8**, 450 (2012).
- [9] R. Gallego, L. Masanes, G. De La Torre, C. Dhara, L. Aolita, and A. Acín, *Nat. Commun.* **4**, 2654 (2013).
- [10] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, *Phys. Rev. Lett.* **98**, 230501 (2007); B. W. Reichardt, F. Unger, and U. Vazirani, *Nature (London)* **496**, 456 (2013).
- [11] J. Barrett, L. Hardy, and A. Kent, *Phys. Rev. Lett.* **95**, 010503 (2005).
- [12] S. Popescu and D. Rohrlich, *Found. Phys.* **24**, 379 (1994).
- [13] M. L. Almeida, J.-D. Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, *Phys. Rev. Lett.* **104**, 230404 (2010).
- [14] J. F. Clauser, M. A. Horne, and A. Shimony, R. A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
- [15] A. Acín, S. Massar, and S. Pironio, *Phys. Rev. Lett.* **108**, 100402 (2012).
- [16] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, *Nature (London)* **461**, 1101 (2009).
- [17] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, *Phys. Rev. Lett.* **96**, 250401 (2006).
- [18] L. P. Thinh, L. Sheridan, and V. Scarani, *Phys. Rev. A* **87**, 062121 (2013).
- [19] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.114.160502> for proofs of the three main results of this Letter.
- [20] N. S. Jones and L. I. Masanes, *Phys. Rev. A* **72**, 052312 (2005).
- [21] J. Barrett and S. Pironio, *Phys. Rev. Lett.* **95**, 140401 (2005).
- [22] S. Pironio, J.-D. Bancal, and V. Scarani, *J. Phys. A* **44**, 065303 (2011).
- [23] L. Aolita, R. Gallego, and A. Cabello, A. Acín, *Phys. Rev. Lett.* **108**, 100401 (2012).
- [24] C. Dhara, G. Pretico, and A. Acín, *Phys. Rev. A* **88**, 052116 (2013).
- [25] T. Franz, F. Furrer, and R. F. Werner, *Phys. Rev. Lett.* **106**, 250502 (2011).
- [26] C. A. Miller and Y. Shi, *8th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2013)* 22, 254 (2013).
- [27] M. J. Hoban, E. T. Campbell, K. Loukopoulos, and D. E. Browne, *New J. Phys.* **13**, 023014 (2011).
- [28] A. Acín, S. Massar, and S. Pironio, *Phys. Rev. Lett.* **108**, 100402 (2012).
- [29] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, *Nat. Commun.* **4**, 2263 (2013).
- [30] M. Navascués, S. Pironio, and A. Acín, *Phys. Rev. Lett.* **98**, 010401 (2007); M. Navascués, S. Pironio, and A. Acín, *New J. Phys.* **10**, 073013 (2008).
- [31] M. Navascués, Y. Guryanova, and M. J. Hoban, A. Acín, *Nat. Commun.* **6**, 6288 (2015).
- [32] N. D. Mermin, *Phys. Rev. Lett.* **65**, 1838 (1990).
- [33] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, *Phys. Rev. Lett.* **107**, 210403 (2011).
- [34] This condition can be seen as a rewriting of the Clauser-Horne-Shimony-Holt (CHSH) inequality [14], sometimes referred to as the CHSH game.
- [35] U. Vazirani and T. Vidick, *Phys. Rev. Lett.* **113**, 140501 (2014).