Hardy's paradox for high-dimensional systems

Jing-Ling Chen,[1,2,*] Adán Cabello,[3,†] Zhen-Peng Xu,[1] Hong-Yi Su,[1] Chunfeng Wu,[4] and L. C. Kwek[2,5,‡]
[1]*Theoretical Physics Division, Chern Institute of Mathematics, Nankai University, Tianjin 300071, People's Republic of China*
[2]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[3]*Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain*
[4]*Pillar of Engineering Product Development, Singapore University of Technology and Design, 20 Dover Drive, Singapore 138682*
[5]*National Institute of Education and Institute of Advanced Studies, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616*
(Received 22 August 2013; published 30 December 2013)

Hardy's proof is considered the simplest proof of nonlocality. Here we introduce an equally simple proof that (i) has Hardy's as a particular case, (ii) shows that the probability of nonlocal events grows with the dimension of the local systems, and (iii) is always equivalent to the violation of a tight Bell inequality. Our proof has all the features of Hardy's and adds the only ingredient of the Einstein-Podolsky-Rosen scenario missing in Hardy's proof: It applies to measurements with an arbitrarily large number of outcomes.

DOI: 10.1103/PhysRevA.88.062116 PACS number(s): 03.65.Ud, 03.67.Mn, 42.50.Xa

I. INTRODUCTION

Nonlocality, namely, the impossibility of describing correlations in terms of local hidden variables [1], is a fundamental property of nature. Hardy's proof [2,3], in any of its forms [4–7], provides a simple way to show that quantum correlations cannot be explained with local theories. Hardy's proof is usually considered "the simplest form of Bell's theorem" [8].

However, if one wants to study nonlocality in a systematic way, one must define the local polytope [9] corresponding to any possible scenario (i.e., for any given number of parties, settings, and outcomes) and check whether quantum correlations violate the inequalities defining the facets of the corresponding local polytope. These inequalities are the so-called *tight* Bell inequalities. In this sense, Hardy's proof has another remarkable property: It is equivalent to a violation of a tight Bell inequality, the Clauser-Horne-Shimony-Holt inequality [10]. This was observed in [5].

Hardy's proof requires two observers, each with two measurements, each with two possible outcomes. The proof has been extended to the case of more than two measurements [11,12] and more than two outcomes [13–15]. However, none of these extensions is equivalent to the violation of a tight Bell inequality. Hardy-like proofs can also be applied to contextuality [16].

Hardy's paradox brings together two features that no other proof of nonlocality has: (i) It proves Bell's theorem under the condition proposed by Einstein, Podolsky, and Rosen (EPR) that one party's measurement outcome allows this party to predict *with certainty* the other party's measurement outcome [17] and (ii) it is equivalent to a violation of the condition that *exactly* separates local from nonlocal correlations for the 2-2-2 scenario (i.e., the tight Bell inequality for the scenario with two parties, two settings, and two outcomes). However, Hardy's paradox has a drawback: The EPR scenario is not 2-2-2 but 2-2-n with n arbitrarily large.

The aim of this work is to introduce an alternative paradox that keeps all the virtues of Hardy's but has the only ingredient of the EPR scenario that is missing in Hardy's paradox: It applies to measurements with an arbitrary number of outcomes. The alternative paradox shows that the maximum probability of nonlocal events, which has a limit of 0.09 in Hardy's paradox (and in previously proposed extensions of Hardy's paradox), actually grows with the number of possible outcomes, tending asymptotically to a limit that is more than four times higher than the one in Hardy's paradox. Moreover, we show that, for any given number n of outcomes, the alternative paradox is equivalent to a violation of the condition that exactly separates local from nonlocal correlations for the 2-2-n scenario. Arguably, all these features make this paradox of fundamental importance.

II. ALTERNATIVE FORMULATION OF HARDY'S PARADOX

Let us consider two observers: Alice, who can measure either A_1 or A_2 on her subsystem, and Bob, who can measure B_1 or B_2 on his. Suppose that each of these measurements has d outcomes that we will number as $0,1,2,\dots,d-1$. Let us denote as $P(A_2 < B_1)$ the joint conditional probability that the result of A_2 is strictly smaller than the result of B_1, that is,

$$P(A_2 < B_1) = \sum_{m<n} P(A_2 = m, B_1 = n), \tag{1}$$

with $m,n \in \{0,1,\dots,d-1\}$. Explicitly, for $d=2$, $P(A_2 < B_1) = P(A_2 = 0, B_1 = 1)$; for $d=3$, $P(A_2 < B_1) = P(A_2 = 0, B_1 = 1) + P(A_2 = 0, B_1 = 2) + P(A_2 = 1, B_1 = 2)$; etc.

Then the proof follows from the fact that, according to quantum theory, there are two-qudit entangled states and local measurements satisfying, simultaneously, the following conditions:

$$P(A_2 < B_1) = 0, \tag{2a}$$

$$P(B_1 < A_1) = 0, \tag{2b}$$

$$P(A_1 < B_2) = 0, \tag{2c}$$

$$P(A_2 < B_2) > 0. \tag{2d}$$

*chenjl@nankai.edu.cn
†adan@us.es
‡cqtklc@nus.edu.sg

Therefore, if events $A_2 < B_1$, $B_1 < A_1$, and $A_1 < B_2$ never happen, then, in any local theory, event $A_2 < B_2$ must never happen either. However, this is in contradiction with (2d). If $d = 2$, the proof is exactly Hardy's [2,3].

III. BEYOND HARDY'S LIMIT

Let us define

$$P_{\rm Hardy} = \max P(A_2 < B_2) \tag{3}$$

satisfying conditions (2a)–(2c). For $d = 2$,

$$P_{\rm Hardy}^{(d=2)} = \frac{5\sqrt{5} - 11}{2} \approx 0.09, \tag{4}$$

which is achieved with two-qubit systems [2,3]. In previous extensions of Hardy's paradox to two-qudit systems [13–15], (4) is also the maximum probability of events that cannot be explained by local theories. For example, the extension considered in Ref. [13] is based on the following four probabilities: $P(A_1 = 0, B_1 = 0) = 0$, $P(A_1 \neq 0, B_2 = 0) = 0$, $P(A_2 = 0, B_1 \neq 0) = 0$, and $P(A_2 = 0, B_2 = 0) = P_{\rm KC} > 0$. Reference [14] proves that, for two-qutrit systems, $\max P_{\rm KC}$ equals (4) and conjectures that $\max P_{\rm KC}$ is always (4) for arbitrary dimension. Reference [15] provides a proof of this conjecture.

Interestingly, in the proof presented in the previous section, $P_{\rm Hardy}$ equals Hardy's limit (4) for $d = 2$, but this is no longer true for higher-dimensional systems. To show this, we will consider pure states satisfying the three conditions (2a)–(2c). An arbitrary two-qudit pure state can be written as

$$|\Psi\rangle = \sum_{i=0}^{d-1}\sum_{j=0}^{d-1} h_{ij}|i\rangle_A|j\rangle_B, \tag{5}$$

where the basis states $|i\rangle_A, |j\rangle_B \in \{|0\rangle, |1\rangle, \ldots, |d-1\rangle\}$ and h_{ij} are coefficients satisfying the normalization condition $\sum_{ij} |h_{ij}|^2 = 1$.

The coefficients h_{ij} completely determine the state $|\Psi\rangle$. We can associate any two-qudit state $|\Psi\rangle$ with a coefficient-matrix $H = (h_{ij})_{d\times d}$, where $i,j = 0,1,\ldots,d-1$ and h_{ij} is the ith row and the jth column element of the $d \times d$ matrix H. The connection between the coefficient matrix H and the two reduced density matrices of $|\Psi\rangle\langle\Psi|$ is

$$\rho_A = {\rm tr}_B(|\Psi\rangle\langle\Psi|) = HH^\dagger, \tag{6a}$$

$$\rho_B = {\rm tr}_A(|\Psi\rangle\langle\Psi|) = H^T(H^T)^\dagger, \tag{6b}$$

where T denotes the matrix transpose and $H^\dagger$ is the Hermitian conjugate matrix of H.

The probability $P(A_i = m, B_j = n)$ can be calculated as

$$P(A_i = m, B_j = n) = {\rm tr}\big[\big(\hat\Pi_{A_i}^m \otimes \hat\Pi_{B_j}^n\big)\rho\big], \tag{7}$$

where $\hat\Pi_{A_i}^m$ and $\hat\Pi_{B_j}^n$ are projectors and $\rho = |\Psi\rangle\langle\Psi|$. Explicitly, the projectors are given by

$$\hat\Pi_{A_1}^m = \mathcal{U}_1|m\rangle\langle m|\mathcal{U}_1^\dagger, \tag{8a}$$

$$\hat\Pi_{B_1}^n = \mathcal{V}_1|n\rangle\langle n|\mathcal{V}_1^\dagger, \tag{8b}$$

$$\hat\Pi_{A_2}^m = \mathcal{U}_2|m\rangle\langle m|\mathcal{U}_2^\dagger, \tag{8c}$$

$$\hat\Pi_{B_2}^n = \mathcal{V}_2|n\rangle\langle n|\mathcal{V}_2^\dagger, \tag{8d}$$

TABLE I. Values of $P_{\rm Hardy}^{\rm opt}$ and $P_{\rm Hardy}^{\rm app}$ for $d = 2,\ldots,7$.

d	$P_{\rm Hardy}^{\rm opt}$	$P_{\rm Hardy}^{\rm app}$	Error rates
2	0.090170	0.088889	0.014207
3	0.141327	0.138426	0.020527
4	0.176512	0.171533	0.028208
5	0.203057	0.195869	0.035399
6	0.224221	0.214825	0.0419051
7	0.241728	0.230172	0.047807

where $\mathcal{U}_1$, $\mathcal{V}_1$, $\mathcal{U}_2$, and $\mathcal{V}_2$ are, in general, SU(d) unitary matrices. Hereafter we denote by $P_{\rm Hardy}^{\rm opt}$ the optimal value of $P_{\rm Hardy}$ by ranging over all unitary matrices and the state $|\Psi\rangle$.

To calculate $P_{\rm Hardy}^{\rm opt}$, it is sufficient to choose the settings A_1 and B_1 as the standard bases, i.e., taking $\mathcal{U}_1 = \mathcal{V}_1 = \mathbb{1}$, where $\mathbb{1}$ is the identity matrix. Evidently, the condition (2b) leads to $h_{ij} = 0$ for $i > j$. This implies that the matrix H is an upper-triangular matrix.

In Table I we list the optimal values of $P_{\rm Hardy}^{\rm opt}$ for $d = 2,\ldots,7$. The corresponding optimal Hardy states $H^{\rm opt}$ are explicitly given in Appendix A.

The calculations for $d > 7$ are beyond our computers capability. However, we observe that $H^{\rm opt}$, written in the representation of H, have reflection symmetry with respect to the antidiagonal line, that is, $h_{ij} = h_{d-1-j,d-1-i}$. We use this to calculate approximately the maximum probability for nonlocal events $P_{\rm Hardy}^{\rm app}$ by using a special class of states $H^{\rm app}$. The explicit form of states $H^{\rm app}$ is given in Appendix B. This allows us to go beyond $d = 7$ and compute $P_{\rm Hardy}^{\rm app}$ from $d = 2$ to 28 000. In Fig. 1 we have plotted $P_{\rm Hardy}^{\rm app}$ from $d = 2$ to 1000, showing that $P_{\rm Hardy}^{\rm app}$ increases with the dimension. Values for higher dimensions are given in Appendix B.

In Table I we also compare the $P_{\rm Hardy}$ for the optimal states and the approximate optimal states. This allows us to speculate that the asymptotic limit may be a little higher than the one showed in Fig. 1. However, the limit $1/2$ can never be surpassed since $P(A_2 > B_2)$ is always bigger than $P(A_2 < B_2)$, as observed in the numerical computations. At

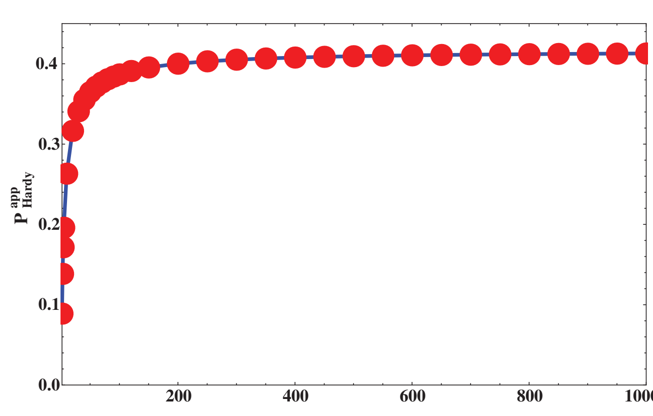

FIG. 1. (Color online) Plot of $P_{\rm Hardy}^{\rm app}$ from $d = 2$ to 1000.

TABLE II. Entanglement degrees of the optimal states and the approximate optimal states for $d = 2, \ldots, 7$.

d	Optimal states	Approximate states
2	0.763932	0.825885
3	0.793888	0.845942
4	0.813483	0.861735
5	0.827702	0.874459
6	0.838679	0.884926
7	0.847510	0.893695

this point, we do not know whether or not 1/2 may be the asymptotic limit.

IV. DEGREE OF ENTANGLEMENT

Hardy's proof does not work for maximally entangled states. The same is true for the proof introduced here. However, in our proof, as d increases, the degree of entanglement tends to 1. To show this we use the generalized concurrence or degree of entanglement [18] for two-qudit systems given by

$$\mathcal{C} = \sqrt{\frac{d}{d-1}\left[1 - \mathrm{tr}\left(\rho_A^2\right)\right]} = \sqrt{\frac{d}{d-1}\left[1 - \mathrm{tr}\left(\rho_B^2\right)\right]}. \quad (9)$$

In Table II we list $\mathcal{C}$ for the optimal Hardy states and the approximate Hardy states. From Table II we observe that, for $d = 2$, the optimal Hardy state occurs at $\mathcal{C}^{\mathrm{opt}} \approx 0.763\,932$ and this value increases to $\mathcal{C}^{\mathrm{opt}} \approx 0.827\,702$ when $d = 5$. For a fixed d, the corresponding $\mathcal{C}^{\mathrm{app}}$ is larger than that of $\mathcal{C}^{\mathrm{app}}$ and it also increases with the dimension d. For $d = 800$, $\mathcal{C}^{\mathrm{app}} \approx 0.998\,062$ and tends to 1 as d grows.

Finally, we can prove that the proof cannot work for two-qudit maximally entangled states

$$|\Psi\rangle_{\mathrm{MES}} = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j\rangle_A |j\rangle_B. \quad (10)$$

Proof. Here $\mathrm{tr}[(\hat{\Pi}_{A_1}^m \otimes \hat{\Pi}_{B_1}^n)|\Psi\rangle\langle\Psi|]$ can be expressed as

$$\mathrm{tr}[(|m\rangle\langle m| \otimes |n\rangle\langle n|)(\mathcal{U}_1^\dagger \otimes \mathcal{V}_1^\dagger)|\Psi\rangle\langle\Psi|(\mathcal{U}_1 \otimes \mathcal{V}_1)]. \quad (11)$$

We will use

$$H_{\mathrm{MES}} \mapsto |\Psi\rangle_{\mathrm{MES}}, \quad H' \mapsto (\mathcal{U}_1^\dagger \otimes \mathcal{V}_1^\dagger)|\Psi\rangle_{\mathrm{MES}}. \quad (12)$$

Taking into account that, (i) given a pure state $H \mapsto |\Psi\rangle_{AB}$ and a local action U acting on Alice (the first part) and V acting on Bob (the second part), then

$$H' \mapsto (U \otimes V)|\Psi\rangle_{AB} = U H V^T, \quad (13)$$

(ii) Eq. (2b) requires H' to be an upper-triangular matrix, and (iii) $H_{\mathrm{MES}} = \frac{1}{\sqrt{d}}\mathbb{1}$, we have the solution

$$\mathcal{U}_1 \mathcal{V}_1^T = \mathcal{D}_1, \quad (14)$$

where $\mathcal{D}_1 = \mathrm{diag}(e^{i\chi_0}, e^{i\chi_1}, \ldots, e^{i\chi_{d-1}})$. Similarly, from (2a) and (2c), we have

$$\mathcal{U}_1 \mathcal{V}_2^T = \mathcal{D}_2, \quad \mathcal{U}_2 \mathcal{V}_1^T = \mathcal{D}_3, \quad (15)$$

where $\mathcal{D}_2$ and $\mathcal{D}_3$ are diagonal matrices similar to $\mathcal{D}_1$. From (14) and (15) we have

$$\mathcal{U}_2 \mathcal{V}_2^T = \mathcal{D}_3 \mathcal{D}_1^\dagger \mathcal{D}_2, \quad (16)$$

which directly leads to $P(A_2 < B_2) = 0$ for $|\Psi\rangle_{\mathrm{MES}}$. This ends the proof.

V. CONNECTION TO TIGHT BELL INEQUALITIES

As can be easily seen, for any d, our proof can be transformed into the following Bell inequality:

$$P(A_2 < B_1) + P(B_1 < A_1) + P(A_1 < B_2) - P(A_2 < B_2) \stackrel{\mathrm{LHV}}{\geqslant} 0, \quad (17)$$

where LHV indicates that the bound is satisfied by local hidden variable theories. The interesting point is that, for any d, inequality (17) is equivalent to Zohren and Gill's version [19] of the Collins-Gisin-Linden-Massar-Popescu inequalities (plural because there is a different inequality for each d) [20], which are tight Bell inequalities for any d [21]. This feature distinguishes our proof from any previously proposed nonlocality proof having Hardy's as a particular case.

VI. CONCLUSION

We have introduced a simple proof of nonlocality for pairs of systems of arbitrary dimension that has all the features of the celebrated proof by Hardy but applies to many other scenarios, including the scenario originally considered by Einstein, Podolsky, and Rosen in which measurements have an arbitrarily large number of outcomes [22].

As in the case of Hardy's paradox, an experimental test of our paradox consists of observing that the probabilities of three events are zero while the probability of a fourth event is not zero. The fact that in our proof the value of this fourth probability is larger than in Hardy's (since it grows with the dimension of the physical system from the value it has in Hardy's proof) makes it more adequate for experimental observation of Hardy-like nonlocality and for applications based on this type of nonlocality.

ACKNOWLEDGMENTS

J.-L.C. was supported by the National Basic Research Program (973 Program) of China under Grant No. 2012CB921900 and the NSF of China (Grants No. 10975075 and No. 11175089). A.C. was supported by Project No. FIS2011-29400 with FEDER funds (MINECO, Spain). This work was also partly supported by the National Research Foundation and the Ministry of Education, Singapore.

APPENDIX A: OPTIMAL HARDY STATES

The optimal Hardy states H_d for $d = 2, \ldots ,7$ are

$$H_2 = \begin{pmatrix} 0.618\,034 & 0.485\,868 \\ 0 & 0.618\,034 \end{pmatrix}, \tag{A1a}$$

$$H_3 = \begin{pmatrix} 0.498\,328 & 0.316\,483 & 0.329\,301 \\ 0 & 0.441\,108 & 0.316\,483 \\ 0 & 0 & 0.498\,328 \end{pmatrix}, \tag{A1b}$$

$$H_4 = \begin{pmatrix} 0.429\,796 & 0.262\,169 & 0.224\,332 & 0.249\,934 \\ 0 & 0.376\,021 & 0.217\,224 & 0.224\,332 \\ 0 & 0 & 0.376\,021 & 0.262\,169 \\ 0 & 0 & 0 & 0.429\,796 \end{pmatrix}, \tag{A1c}$$

$$H_5 = \begin{pmatrix} 0.383\,613 & 0.230\,044 & 0.189\,636 & 0.175\,427 & 0.201\,533 \\ 0 & 0.334\,102 & 0.185\,035 & 0.157\,012 & 0.175\,427 \\ 0 & 0 & 0.330\,72 & 0.185\,035 & 0.189\,636 \\ 0 & 0 & 0 & 0.334\,102 & 0.230\,044 \\ 0 & 0 & 0 & 0 & 0.383\,613 \end{pmatrix}, \tag{A1d}$$

$$H_6 = \begin{pmatrix} 0.349\,686 & 0.207\,877 & 0.168\,45 & 0.150\,559 & 0.144\,455 & 0.168\,83 \\ 0 & 0.303\,795 & 0.165\,105 & 0.134\,967 & 0.125\,208 & 0.144\,455 \\ 0 & 0 & 0.299\,72 & 0.160\,666 & 0.134\,967 & 0.150\,559 \\ 0 & 0 & 0 & 0.299\,72 & 0.165\,105 & 0.168\,45 \\ 0 & 0 & 0 & 0 & 0.303\,795 & 0.207\,877 \\ 0 & 0 & 0 & 0 & 0 & 0.349\,686 \end{pmatrix}, \tag{A1e}$$

$$H_7 = \begin{pmatrix} 0.323\,377 & 0.191\,279 & 0.153\,539 & 0.135\,037 & 0.125\,45 & 0.122\,887 & 0.145\,233 \\ 0 & 0.280\,442 & 0.150\,851 & 0.121\,193 & 0.108\,665 & 0.104\,707 & 0.122\,887 \\ 0 & 0 & 0.276\,282 & 0.145\,271 & 0.117\,498 & 0.108\,665 & 0.125\,45 \\ 0 & 0 & 0 & 0.275\,414 & 0.145\,271 & 0.121\,193 & 0.135\,037 \\ 0 & 0 & 0 & 0 & 0.276\,282 & 0.150\,851 & 0.153\,539 \\ 0 & 0 & 0 & 0 & 0 & 0.280\,442 & 0.191\,279 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.323\,377 \end{pmatrix}. \tag{A1f}$$

APPENDIX B: APPROXIMATE OPTIMAL HARDY STATES

The form of H_d for $d = 2, \ldots ,7$ suggests to define the approximate optimal Hardy states as follows:

$$H_d^{\mathrm{app}} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{d-1} & \alpha_d \\ & \alpha_1 & \alpha_2 & \cdots & \alpha_{d-2} & \alpha_{d-1} \\ & & \ddots & \ddots & \vdots & \vdots \\ & & & \alpha_1 & \alpha_2 & \alpha_3 \\ & & & & \alpha_1 & \alpha_2 \\ & & & & & \alpha_1 \end{pmatrix}, \tag{B1}$$

where

$$\alpha_r = \frac{\beta_r}{\sqrt{d+1-r}}, \quad r = 1,2, \ldots ,d, \tag{B2}$$

with $\beta_r > 0$ satisfying the following relations:

$$\beta_1 : \beta_2 : \beta_3 : \cdots : \beta_d = 1 : \frac{1}{2} : \frac{1}{3} : \cdots : \frac{1}{d}, \tag{B3a}$$

$$\sum_{r=1}^{d} \beta_r^2 = 1. \tag{B3b}$$

In Table III we have listed $P_{\mathrm{Hardy}}^{\mathrm{app}}$ up to $d = 28\,000$.

TABLE III. Values of $P_{\mathrm{Hardy}}^{\mathrm{app}}$ from $d = 2$ to $28\,000$.

d	$P_{\mathrm{Hardy}}^{\mathrm{app}}$	d	$P_{\mathrm{Hardy}}^{\mathrm{app}}$	d	$P_{\mathrm{Hardy}}^{\mathrm{app}}$	d	$P_{\mathrm{Hardy}}^{\mathrm{app}}$
2	0.088889	300	0.405106	2000	0.414711	10000	0.416300
10	0.263168	400	0.407749	2200	0.414885	11000	0.416339
20	0.316491	500	0.409394	2400	0.415031	12000	0.416371
30	0.340836	600	0.410520	2600	0.415156	13000	0.416398
40	0.355158	700	0.411341	2800	0.415263	14000	0.416421
50	0.364700	800	0.411966	3000	0.415357	16000	0.416459
60	0.371554	900	0.412459	4000	0.415687	18000	0.416489
70	0.376736	1000	0.412857	5000	0.415889	20000	0.416513
80	0.380803	1200	0.413464	6000	0.416024	22000	0.416533
90	0.384085	1400	0.413903	6000	0.416024	24000	0.416549
100	0.386793	1600	0.414230	8000	0.416196	26000	0.416563
200	0.400116	1800	0.414499	9000	0.416254	28000	0.416575

[1] J. S. Bell, Physics (NY) **1**, 195 (1964).
[2] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).
[3] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).
[4] S. Goldstein, Phys. Rev. Lett. **72**, 1951 (1994).
[5] N. D. Mermin, Phys. Today **47**(6), 9 (1994); **47**(11), 119 (1994).
[6] N. D. Mermin, Am. J. Phys. **62**, 880 (1994).
[7] P. G. Kwiat and L. Hardy, Am. J. Phys. **68**, 33 (2000).
[8] N. D. Mermin, in *Fundamental Problems in Quantum Theory*, edited by D. M. Greenberger and A. Zeilinger, Special issue of Ann. N. Y. Acad. Sci. **755**, 616 (1995).
[9] I. Pitowsky, *Quantum Probability–Quantum Logic* (Springer, New York, 1989).
[10] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[11] L. Hardy, in *New Developments on Fundamental Problems in Quantum Physics*, edited by M. Ferrero and A. van der Merwe (Kluwer, Dordrecht, 1997), p. 163.
[12] D. Boschi, S. Branca, F. De Martini, and L. Hardy, Phys. Rev. Lett. **79**, 2755 (1997).
[13] S. Kunkri and S. K. Choudhary, Phys. Rev. A **72**, 022348 (2005).
[14] K. P. Seshadreesan and S. Ghosh, J. Phys. A: Math. Theor. **44**, 315305 (2011).
[15] R. Rabelo, L. Y. Zhi, and V. Scarani, Phys. Rev. Lett. **109**, 180401 (2012).
[16] A. Cabello, P. Badziąg, M. Terra Cunha, and M. Bourennane, Phys. Rev. Lett. **111**, 180404 (2013).
[17] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[18] S. Albeverio and S.-M. Fei, J. Opt. B: Quant. Semiclass. Opt. **3**, 223 (2001).
[19] S. Zohren and R. D. Gill, Phys. Rev. Lett. **100**, 120406 (2008).
[20] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[21] L. Masanes, Quantum Inf. Comput. **3**, 345 (2003).
[22] A subtlety here is that, albeit both are infinite, the outcomes in Hardy's scenario presented here are discrete variables, while those in the original EPR scenario are continuous ones.