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. Kwek2,5,‡
1Theoretical Physics Division, Chern Institute of Mathematics, Nankai University, Tianjin 300071, People’s Republic of China
2Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543
3Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain
4Pillar of Engineering Product Development, Singapore University of Technology and Design, 20 Dover Drive, Singapore 138682
5National Institute of Education and Institute of Advanced Studies, Nanyang Technological University, 1 Nanyang Walk, Singapore 637616

(I received 22 August 2013; published 30 December 2013)

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 A1 or A2 on her subsystem, and Bob, who can measure B1 or B2 on his. Suppose that each of these measurements has d outcomes that we will number as 0,1,2,...,d−1. Let us denote as P(A2<B1) the joint conditional probability that the result of A2 is strictly smaller than the result of B1, that is,

\[ P(A_2 < B_1) = \sum_{m<n} P(A_2 = m, B_1 = n), \]

with m,n ∈ {0, 1, 2, ..., d − 1}. Explicitly, for d = 2, P(A2 < B1) = P(A2 = 0, B1 = 1); for d = 3, P(A2 < B1) = P(A2 = 0, B1 = 1) + P(A2 = 0, B1 = 2) + P(A2 = 1, B1 = 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:

\[ P(A_2 < B_1) = 0, \]  
\[ P(B_1 < A_2) = 0, \]  
\[ P(A_1 < B_2) = 0, \]  
\[ P(A_2 < B_2) > 0. \]
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

$$P_{\text{Hardy}} = \max P(A_2 < B_2)$$

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

$$P_{\text{Hardy}}(d=2) = \frac{\sqrt{3} - 1}{2} \approx 0.241728$$

which is achieved with two-qubit systems [2,3]. In previous extensions of Hardy’s paradox to two-qutrit 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 $P(A_2 = 0, B_2 = 0) = P_{\text{KC}} > 0$. Reference [14] proves that, for two-qutrit systems, max $P_{\text{KC}}$ equals (4) and conjectures that max $P_{\text{KC}}$ is always (4) for arbitrary dimension. Reference [15] provides a proof of this conjecture.

Interestingly, in the proof presented in the previous section, $P_{\text{Hardy}}$ 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

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

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 $\sum_{i,j} |h_{ij}|^2 = 1$.

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

$$\rho_A = \text{tr}_B(|\Psi\rangle\langle\Psi|) = HH^\dagger,$$

$$\rho_B = \text{tr}_A(|\Psi\rangle\langle\Psi|) = H^T (H^T)^\dagger,$$

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

$$P(A_i = m, B_j = n) = \text{tr}\left( (\hat{\Pi}_{A_i}^m \otimes \hat{\Pi}_{B_j}^n ) \rho \right),$$

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

$$\hat{\Pi}_{A_i}^m = \mathcal{U}_i |m\rangle \langle m| \mathcal{U}_i^\dagger,$$

$$\hat{\Pi}_{B_j}^n = \mathcal{V}_j |n\rangle \langle n| \mathcal{V}_j^\dagger,$$

$$\hat{\Pi}_{A_i}^m = \mathcal{U}_2 |m\rangle \langle m| \mathcal{U}_2^\dagger,$$

$$\hat{\Pi}_{B_j}^n = \mathcal{V}_2 |n\rangle \langle n| \mathcal{V}_2^\dagger.$$
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

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

In Table II we list \( C \) 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 \( C_{\text{opt}} \approx 0.763932 \) and tends to 1 as \( d \) grows.

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

\[
|\Psi_{\text{MES}}\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |j_A\rangle |j_B\rangle. \tag{10}
\]

**Proof.** Here \( \tr(\hat{\Pi}_A \otimes \hat{\Pi}_B |\Psi\rangle \langle\Psi|) \) can be expressed as

\[
\tr(|m\rangle \langle m| \otimes |n\rangle \langle n| (\hat{U}_1 \otimes \hat{V}_1) |\Psi\rangle \langle\Psi| (\hat{U}_1 \otimes \hat{V}_1)). \tag{11}
\]

We will use

\[
H_{\text{MES}} \mapsto |\Psi_{\text{MES}}\rangle, \quad H' \mapsto (\hat{U}_1 \otimes \hat{V}_1) |\Psi_{\text{MES}}\rangle. \tag{12}
\]

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

\[
H' \mapsto (U \otimes V) |\Psi\rangle_{AB} = EUHV_T, \tag{13}
\]

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

\[
\hat{U}_1 V_T^T = D_1, \tag{14}
\]

where \( D_1 = \text{diag}(e^{i\tau_0}, e^{i\tau_1}, \ldots, e^{i\tau_{d-1}}) \). Similarly, from (2a) and (2c), we have

\[
\hat{U}_1 V_T^T = D_2, \quad \hat{U}_2 V_T^T = D_3, \tag{15}
\]

which directly leads to \( P(A_2 < B_2) = 0 \) for \( |\Psi_{\text{MES}}\rangle \). 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:

\[
P(A_2 < B_1) + P(B_1 < A_1) + P(A_1 < B_2) - P(A_2 < B_2) \geq 0, \tag{17}
\]

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

\[
H_2 = \begin{pmatrix}
0.618034 & 0.485868 \\
0 & 0.618034
\end{pmatrix}, \quad (A1a)
\]

\[
H_3 = \begin{pmatrix}
0.498328 & 0.316483 & 0.329301 \\
0 & 0.441108 & 0.316483 \\
0 & 0 & 0.498328
\end{pmatrix}, \quad (A1b)
\]

\[
H_4 = \begin{pmatrix}
0.429796 & 0.262169 & 0.224332 & 0.249934 \\
0 & 0.376021 & 0.217224 & 0.224332 \\
0 & 0 & 0.376021 & 0.262169 \\
0 & 0 & 0 & 0.429796
\end{pmatrix}, \quad (A1c)
\]

\[
H_5 = \begin{pmatrix}
0.383613 & 0.230044 & 0.189636 & 0.175427 & 0.201533 \\
0 & 0.334102 & 0.185035 & 0.157012 & 0.175427 \\
0 & 0 & 0.33072 & 0.185035 & 0.189636 \\
0 & 0 & 0 & 0.334102 & 0.230044 \\
0 & 0 & 0 & 0 & 0.383613
\end{pmatrix}, \quad (A1d)
\]

\[
H_6 = \begin{pmatrix}
0.349686 & 0.207877 & 0.16845 & 0.150559 & 0.144455 & 0.16883 \\
0 & 0.303795 & 0.165105 & 0.134967 & 0.125208 & 0.144455 \\
0 & 0 & 0.29972 & 0.160666 & 0.134967 & 0.150559 \\
0 & 0 & 0 & 0.29972 & 0.165105 & 0.16845 \\
0 & 0 & 0 & 0 & 0.303795 & 0.207877 \\
0 & 0 & 0 & 0 & 0 & 0.349686
\end{pmatrix}, \quad (A1e)
\]

\[
H_7 = \begin{pmatrix}
0.323377 & 0.191279 & 0.153539 & 0.135037 & 0.12545 & 0.122887 & 0.145233 \\
0 & 0.280442 & 0.150851 & 0.121193 & 0.108665 & 0.104707 & 0.122887 \\
0 & 0 & 0.276282 & 0.145271 & 0.117498 & 0.108665 & 0.12545 \\
0 & 0 & 0 & 0.275414 & 0.145271 & 0.121193 & 0.135037 \\
0 & 0 & 0 & 0 & 0.276282 & 0.150851 & 0.153539 \\
0 & 0 & 0 & 0 & 0 & 0.280442 & 0.191279 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.323377
\end{pmatrix}, \quad (A1f)
\]

APPENDIX B: APPROXIMATE OPTIMAL HARDY STATES

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

\[
H_d^{\text{app}} = \begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{d-1} & \alpha_d \\
\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{d-2} & \alpha_{d-1} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_1
\end{pmatrix}, \quad (B1)
\]

where

\[
\alpha_r = \frac{\beta_r}{\sqrt{d + 1 - r}}, \quad r = 1, 2, \ldots, d, \quad (B2)
\]

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

\[
\beta_1 : \beta_2 : \beta_3 : \cdots : \beta_d = 1 : \frac{1}{2} : \frac{1}{3} : \frac{1}{d} \quad (B3a)
\]

\[
\sum_{r=1}^{d} \beta_r^2 = 1. \quad (B3b)
\]

In Table III we have listed $P_{\text{Hardy}}^{\text{app}}$ up to $d = 28000$.

<table>
<thead>
<tr>
<th>$d$</th>
<th>$P_{\text{Hardy}}^{\text{app}}$</th>
<th>$d$</th>
<th>$P_{\text{Hardy}}^{\text{app}}$</th>
<th>$d$</th>
<th>$P_{\text{Hardy}}^{\text{app}}$</th>
<th>$d$</th>
<th>$P_{\text{Hardy}}^{\text{app}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.088889</td>
<td>300</td>
<td>0.405106</td>
<td>10</td>
<td>0.414711</td>
<td>300</td>
<td>0.416000</td>
</tr>
<tr>
<td>10</td>
<td>0.263168</td>
<td>400</td>
<td>0.407749</td>
<td>2200</td>
<td>0.414885</td>
<td>5000</td>
<td>0.416339</td>
</tr>
<tr>
<td>20</td>
<td>0.316491</td>
<td>500</td>
<td>0.409394</td>
<td>2400</td>
<td>0.415031</td>
<td>6000</td>
<td>0.416371</td>
</tr>
<tr>
<td>30</td>
<td>0.340836</td>
<td>600</td>
<td>0.410520</td>
<td>2600</td>
<td>0.415156</td>
<td>7000</td>
<td>0.416398</td>
</tr>
<tr>
<td>40</td>
<td>0.355158</td>
<td>700</td>
<td>0.411341</td>
<td>2800</td>
<td>0.415263</td>
<td>8000</td>
<td>0.416421</td>
</tr>
<tr>
<td>50</td>
<td>0.364700</td>
<td>800</td>
<td>0.411966</td>
<td>3000</td>
<td>0.415357</td>
<td>9000</td>
<td>0.416459</td>
</tr>
<tr>
<td>60</td>
<td>0.371554</td>
<td>900</td>
<td>0.412459</td>
<td>4000</td>
<td>0.415687</td>
<td>10000</td>
<td>0.416489</td>
</tr>
<tr>
<td>70</td>
<td>0.376736</td>
<td>1000</td>
<td>0.412857</td>
<td>5000</td>
<td>0.415889</td>
<td>12000</td>
<td>0.416513</td>
</tr>
<tr>
<td>80</td>
<td>0.380803</td>
<td>1200</td>
<td>0.413464</td>
<td>6000</td>
<td>0.416024</td>
<td>14000</td>
<td>0.416533</td>
</tr>
<tr>
<td>90</td>
<td>0.384085</td>
<td>1400</td>
<td>0.413903</td>
<td>7000</td>
<td>0.416024</td>
<td>16000</td>
<td>0.416549</td>
</tr>
<tr>
<td>100</td>
<td>0.386793</td>
<td>1600</td>
<td>0.414230</td>
<td>8000</td>
<td>0.416196</td>
<td>18000</td>
<td>0.416563</td>
</tr>
<tr>
<td>200</td>
<td>0.400116</td>
<td>1800</td>
<td>0.414499</td>
<td>9000</td>
<td>0.416254</td>
<td>20000</td>
<td>0.416575</td>
</tr>
</tbody>
</table>
[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.