Entangled three-state systems violate local realism more strongly than qubits: An analytical proof

Jing-Ling Chen,[1,3] Dagomir Kaszlikowski,[1,2] L. C. Kwek,[1,4] C. H. Oh,[1] and Marek Żukowski[5]
[1]*Department of Physics, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260*
[2]*Instytut Fizyki Doświadczalnej, Uniwersytet Gdański, PL-80-952, Gdańsk, Poland*
[3]*Laboratory of Computational Physics, Institute of Applied Physics and Computational Mathematics, P.O. Box 8009(26), Beijing 100088, People's Republic of China*
[4]*National Institute of Education, Nanyang Technological University, 1 Nanyang Walk, Singapore 639798*
[5]*Instytut Fizyki Teoretycznej i Astrofizyki, Uniwersytet Gdański, PL-80-952, Gdańsk, Poland*
(Received 16 April 2001; published 12 October 2001)

In Kaszlikowski *et al.* [Phys. Rev. Lett. **85**, 4418 (2000)], it has been shown numerically that the violation of local realism for two maximally entangled N-dimensional ($3 \leqslant N$) quantum objects is stronger than for two maximally entangled qubits and grows with N. In this paper, we present the analytical proof of this fact for $N=3$.

DOI: 10.1103/PhysRevA.64.052109 PACS number(s): 03.65.Ud, 42.50.Dv

I. INTRODUCTION

Since the formulation of the Bell theorem [1], various forms of Bell inequalities, see for instance [2], have been devised to investigate the possibility (or lack of such possibility) of a local realistic description of correlations observed in various quantum systems such as M entangled N-dimensional quantum objects. The main advantage of this approach is its simplicity. The drawback of this method is that, in general, Bell inequalities are only a necessary condition for the existence of a local and realistic description of the investigated quantum system. Only in a few cases, see for instance [3,4], has it been proved that some Bell inequalities are necessary and a sufficient condition for the existence of local realism. Moreover, these cases generally deal with two [3] or more than two [4] qubits and two local observables measured at each side of the Bell experiment.

In [5], a general approach to the problem was presented. Indeed, it is possible to find all relevant inequalities that must be satisfied by the probabilities obtained in the measurement of any number of local observables on the system consisting of an arbitrary number of quantum objects, each of which described by a Hilbert space of arbitrary dimension, so that it may be described in terms of local realism. However, the number of inequalities that have to be examined grows extremely fast with the dimension of the problem, i.e., the number of local observables, quantum objects, and dimension of Hilbert space describing given objects. This renders the method practically useless, as shown in [6,7].

Recent research shows that a different approach is possible. In [8,9], the method of numerical linear optimization has been successfully applied to two qubit correlations with up to ten local observables being measured at each side of the experiment [8] and for two N-dimensional objects ($2 \leqslant N \leqslant 16$) with two local observables at each side of the experiment [9,10]. In this approach, one does not obtain Bell inequalities but finds the conditions under which there exists a local hidden variable model reproducing quantum results for the given quantum system and the measured quantum observables. Moreover, this method may be applied directly to the analysis of experimental data.

The paper of Kaszlikowski *et al.* [9] is a good example of how important it is to know necessary and sufficient conditions for the existence of local realism in a given case. For instance, in [11], it was shown for two N-dimensional entangled systems that the Clauser-Horne-Shimony-Holt (CHSH) inequality [12] is maximally violated by the factor of $\sqrt{2}$. However, CHSH inequality is not a sufficient condition for the existence of local realism for two entangled objects each described by a Hilbert space of the dimension greater than two. Indeed, the results of [9] show that violations of local realism increase with the dimension of the systems.

In this paper, we prove analytically that the violation of local realism for two maximally entangled qutrits (objects described by a three-dimensional Hilbert space) observed via two unbiased three input and three output beam splitters [13] is stronger than for two maximally entangled qubits [9]. Earlier numerical computations advocated such a violation but rigorous analytical evidence has so far been lacking except for the trivial case of qubit. Thus, it is anticipated that analytical proofs should exist for higher-dimensional quantum systems. Our present paper on three-state systems (qutrits) therefore constitutes an attempt to confirm the previous numerical claim. Moreover, we also see that the extension from qubit to qutrit is clearly nontrivial. In fact, a comparison of our results with the separability condition for so-called generalized Werner states [14] may shed some light on the relation between local realism and separability of bipartite quantum systems.

In Sec. II, we briefly recall and describe the possible experimental setting for the observers using six-port beam splitters. In order to arrive at a symmetrical form for the correlation function, we have ascribed a special set of complex numbers to the results of the measurements at the detectors. In Sec. III, we prove analytically the explicit expression for the minimal noise admixture. Finally, in Sec. IV, we conclude with some relevant remarks.

II. TRITTERS AND MEASUREMENTS

We consider the Bell-type experiment in which two spatially separated observers Alice and Bob measure two noncommuting observables A_1, A_2 for Alice and B_1, B_2 for Bob on the maximally entangled state $|\psi\rangle$ of two qutrits

$$|\psi\rangle = \frac{1}{\sqrt{3}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B + |2\rangle_A|2\rangle_B), \tag{1}$$

where $|k\rangle_A$ and $|k\rangle_B$ describe the kth basis state of the qutrit A and B, respectively. Such a state may be prepared with pairs of photons with the aid of parametric down conversion (see [13]), in which case, kets $|k\rangle_A$ and $|k\rangle_B$ denote photons propagating to Alice and Bob in mode k.

In this paper, we consider the special case in which both observers measure observables defined by a six-port (three input and three output ports) beam splitter. The extended theory of such devices may be found in [13]. We give here only a brief description.

Unbiased six-port beam splitter (called a tritter by some [13]), is a device with the following property: if a photon enters any of the three single-input ports, its chances of exit are equally split among the three output ports. In fact, one may always build a six-port beam splitter with the distinguishing trait that the elements of its unitary transition matrix $\hat{T}$ are *solely* powers of the third root of unity $\alpha = \exp(i2\pi/3)$, namely $T_{kl} = (1/\sqrt{3})\alpha^{(k-1)(l-1)}$. In front of the ith input port of the six-port beam splitter, we put a phase shifter that changes the phase of the incoming photon by ϕ_i. These three phase shifts, which we denote for convenience as a "vector" of phase shifts $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$, are macroscopic local parameters that can be changed by the observer. Therefore, a six-port beam splitter together with the three phase shifters performs the unitary transformation $\hat{U}(\vec{\phi})$ with the entries $U_{kl} = T_{kl}\exp(i\phi_l)$.

Alice and Bob measure the following observables:

$$\begin{aligned} A(\vec{\phi}_i) &= \hat{U}(\vec{\phi}_i)|0\rangle\langle 0|\hat{U}(\vec{\phi}_i)^\dagger + \alpha\hat{U}(\vec{\phi}_i)|1\rangle\langle 1|\hat{U}^\dagger(\vec{\phi}_i) \\ &\quad + \alpha^2\hat{U}(\vec{\phi}_i)|2\rangle\langle 2|\hat{U}^\dagger(\vec{\phi}_i), \\ B(\vec{\theta}_j) &= \hat{U}(\vec{\theta}_j)|0\rangle\langle 0|\hat{U}(\vec{\theta}_j)^\dagger + \alpha\hat{U}(\vec{\theta}_j)|1\rangle\langle 1|\hat{U}^\dagger(\vec{\theta}_j) \\ &\quad + \alpha^2\hat{U}(\vec{\theta}_j)|2\rangle\langle 2|\hat{U}^\dagger(\vec{\theta}_j), \end{aligned} \tag{2}$$

where $i,j = 1,2$ and where, for instance, $\vec{\phi}_i$ denotes the vector of local phase shifts for Alice in the ith experiment. Note that we have ascribed complex numbers to the results of the measurements, i.e., to the "click" of the lth detector we have ascribed the number α^l. The justification of such an assignment may be found in [13]. This assignment results in a very symmetrical complex correlation function

$$\begin{aligned} E(\vec{\phi}_i, \vec{\theta}_j) &= \langle\psi|A(\vec{\phi}_i)B(\vec{\theta}_j)|\psi\rangle \\ &= \frac{1}{3}[\exp(\phi_i^1 - \phi_i^2 + \theta_j^1 - \theta_j^2) \\ &\quad + \exp(\phi_i^2 - \phi_i^3 + \theta_j^2 - \theta_j^3) \\ &\quad + \exp(\phi_i^3 - \phi_i^1 + \theta_j^3 - \theta_j^1)], \end{aligned}$$

where, for instance, ϕ_i^1 denotes the first phase shift at Alice's side in the ith experiment. This correlation function retains the information about the correlations observed in the experiment. In fact, according to quantum mechanics, the whole information that is accessible in the experiment are probabilities of coincidence firings of the detectors. It may be easily verified through the knowledge of the correlation function $E(\vec{\phi}_i, \vec{\theta}_j)$ that one is able to calculate the probabilities of these coincidence "clicks" and in this way obtain the whole information about the correlations observed in the system.

Following [9], we define the strength of violation of local realism as the minimal noise admixture F_{thr} to the state (1) below that the measured correlations cannot be described by local realism for the given observables. Therefore, we assume that Alice and Bob perform their measurements on the following mixed state ρ_F:

$$\rho_F = (1-F)|\psi\rangle\langle\psi| + F\rho_{noise}, \tag{3}$$

where $0 \leqslant F \leqslant 1$ and where ρ_{noise} is a diagonal matrix with entries equal to 1/9. This latter matrix denotes a totally chaotic mixture (noise) that admits a local and realistic description. For $F=0$ (pure maximally entangled state), a local realistic description does not exist, whereas for $F=1$ (pure noise) it does. Therefore, there exists some threshold value of F, which we denote by F_{thr}, such that for every $F \leqslant F_{thr}$, a local realistic description does not exist. The bigger the value of F_{thr}, the stronger is the violation of local realism. The correlation function for the state (3) reads $E^F(\vec{\phi}_i, \vec{\theta}_j) = (1-F)E(\vec{\phi}_i, \vec{\theta}_j)$.

III. PROOF OF MINIMAL NOISE ADMIXTURE

Let us now assume that Alice measures two observables defined by the following sets of phase shifts $\vec{\phi}_1 = (0, \pi/3, -\pi/3)$ and $\vec{\phi}_2 = (0,0,0)$, whereas Bob measures two observables defined by the sets of phase shifts $\vec{\theta}_1 = (0, \pi/6, -\pi/6)$ and $\vec{\theta}_2 = (0, -\pi/6, \pi/6)$. From numerical computations, it has been found [9,10] that these sets of phases provide the highest F_{thr}. Straightforward calculations give the following values of the correlations functions for each experiment: $E(\vec{\phi}_1, \vec{\theta}_1) = E(\vec{\phi}_2, \vec{\theta}_2) = Q_1 = (2\sqrt{3} + 1/6) - i(2 - \sqrt{3}/6)$, $E(\vec{\phi}_1, \vec{\theta}_2) = Q_1^*$, $E(\vec{\phi}_2, \vec{\theta}_1) = Q_2 = -1/3(1+2i)$. From these complex numbers, we can create a 2×2 matrix $\hat{Q}$ with the entries $(\hat{Q})_{ij} = E(\vec{\phi}_i, \vec{\theta}_j)$.

Local realism implies the following structure of the correlation function for reproducing the quantum correlation function defined above: $E_{LHV}(\vec{\phi}_i, \vec{\theta}_j) = \int d\lambda\, \rho(\lambda) A(\vec{\phi}_i, \lambda) B(\vec{\theta}_j, \lambda)$, where for trichotomic measurements $A(\vec{\phi}_i, \lambda) = \alpha^m$ and $B(\vec{\theta}_j, \lambda) = \alpha^n$, $(m,n = 1,2,3)$. Three-valued functions $A(\vec{\phi}_i, \lambda), B(\vec{\theta}_j, \lambda)$ represent the values of local measurements predetermined by local hidden variables, denoted by λ, for the specified local settings. This expression is an average over a certain local hidden variable distribution $\rho(\lambda)$ of certain factorizable matrices, namely, those with elements given by $H_\lambda^{ij}=A(\vec{\phi}_i,\lambda)B(\vec{\theta}_j,\lambda)$. The symbol λ may hide many parameters. However, since the only possible values of $A(\vec{\phi}_i,\lambda)$ and $B(\vec{\theta}_j,\lambda)$ are $1,\alpha,\alpha^2$, there are only nine different sequences of the values of $(A(\vec{\phi}_1,\lambda),A(\vec{\phi}_2,\lambda))$, and nine different sequences of the values of $(B(\vec{\theta}_1,\lambda),B(\vec{\theta}_2,\lambda))$, and consequently they form only 81 matrices $\hat{H}_\lambda$.

Therefore, the structure of the local hidden variable model of $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$ reduces to a discrete probabilistic model involving the average of all the 81 matrices $\hat{H}_\lambda$. Therefore, we replace the parameter λ by index k ($k=1,2,\ldots,81$) to which we ascribe the matrix $\hat{H}_k$ with entries $H_k^{ij}=\alpha^{k_i+l_j}$ $(i,j=1,2)$, where $k_1=[(k-1)/9]-1$, $k_2=[(k-1)/3]-1$, $l_1=1$, $l_2=k$, and $[x]$ denotes the integer part of the number x. It may be checked that only the first 27 matrices are different, which means that it suffices to consider only them. With this notation, the correlation function $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$ acquires the following simple form $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)=\Sigma_{k=1}^{27}p_kH_k^{ij}$, with, of course, the probabilities satisfying $p_k\geqslant 0$ and $\Sigma_{k=1}^{27}p_k=1$. From $E_{LHV}(\vec{\phi}_i,\vec{\theta}_j)$, we build the matrix $\hat{E}_{LHV}$.

Quantum predictions in the form of the matrix $\hat{Q}^F$ may be recovered by local hidden variables if and only if $\hat{Q}^F=\Sigma_{n=1}^{27}p_n\hat{H}_n$. Now, we want to find the minimal possible F for which it is still possible to recover matrix $\hat{Q}^F$ using the probability distribution p_n and factorizable matrices $\hat{H}_n$. For convenience, we define a parameter $V=1-F$. Then the minimal F refers to the maximal V.

Theorem. The maximal V equals $V_{thr}=6\sqrt{3}-9/2$.

Proof. First, we observe that matrix $\hat{Q}^V$ may be written as

$$\hat{Q}^V=V\left[\frac{2\sqrt{3}+1}{6}-i\,\frac{2-\sqrt{3}}{6}\right]\hat{I}+V\left[-\frac{2\sqrt{3}-1}{6}+i\,\frac{2+\sqrt{3}}{6}\right]\vec{n}\cdot\vec{\sigma}, \tag{4}$$

where $\vec{n}=(n_x,n_y,n_z)=(-1/2,\sqrt{3}/2,0)$, $\hat{I}$ is unit matrix, and $\vec{\sigma}$ are Pauli matrices. One observes that $\hat{Q}^V$ commutes with the matrix $\hat{\mathcal{U}}=\vec{n}\cdot\vec{\sigma}$, which has only two nonzero entries $\mathcal{U}_{12}=\alpha^2,\mathcal{U}_{21}=\alpha$ and is unitary and Hermitian. Furthermore, $\hat{\mathcal{U}}$ preserves the structure of matrices $\hat{H}_n$ in the sense that for every $n=1,2,\ldots,27$, $\hat{\mathcal{U}}\hat{H}_n\hat{\mathcal{U}}=\hat{H}_m$ for some $m=1,2,\ldots,27$. This is an injective mapping. One may also find that some matrices $\hat{H}_n$ are invariants with respect to transformation $\hat{\mathcal{U}}$. For further considerations, it is necessary to list all the pairs $(n,m)\in\{(1,8),(2,10),(3,24),(4,17),(5,19),(7,26),(9,15),(11,13),(12,27),(14,22),(16,20),(23,25)\}$. By considering the ith pair (n,m), we may define matrices $\hat{G}_i=\hat{H}_n+\hat{H}_m$. Thus, $\hat{G}_1=\hat{H}_1+\hat{H}_8$, $\hat{G}_2=\hat{H}_2+\hat{H}_{10}$, and so forth. The remaining invariant matrices are $\hat{H}_6,\hat{H}_{18},\hat{H}_{21}$. For convenience, we label them as $\hat{G}_{13}$ to $\hat{G}_{15}$, viz. $\hat{G}_{13}=\hat{H}_6,\hat{G}_{14}=\hat{H}_{18},\hat{G}_{15}=\hat{H}_{21}$.

Suppose that we have the optimal solution (the solution for which $V=V_{thr}$), i.e., we have the probability distribution p_n so that $\hat{Q}^{V_{thr}}=\Sigma_{n=1}^{27}p_n\hat{H}_n$. Acting on both sides of this equation with matrix $\hat{\mathcal{U}}$, we get another optimal solution with the same V_{thr} (matrix $\hat{\mathcal{U}}$ commutes with $\hat{Q}^{V_{thr}}$) but with the probability distribution p_k', which may be obtained from the previous one by swapping probabilities belonging to the same pair, for instance, $p_{10}'=p_2$, and so on. Therefore, due to the above property, we may assume *without any loss of generality* that in the optimal solution the probabilities referring to the same pair are equal. Therefore, we have reduced the number of relevant probabilities from 27 to 15. One may observe that *every* matrix $\hat{G}_k$ may be expressed by matrices $\hat{G}_1,\hat{G}_{10},\hat{G}_{13}$ by multiplying them by α, α^2, and -1. For instance, $\hat{G}_6=\alpha\hat{G}_{10},\hat{G}_{11}=-\hat{G}_{13}$, etc. Three matrices are the same, $\hat{G}_5=\hat{G}_3,\hat{G}_7=\hat{G}_4,\hat{G}_{11}=\hat{G}_9$, which further reduces the number of relevant probabilities from 15 to 12.

Recalling the above properties, we may write the optimal solution in the form $\hat{Q}^{V_{thr}}=\Sigma_{k\neq 5,7,11}w_k\hat{G}_k$, with the normalization condition of probabilities $2(w_1+w_2+w_3+w_4+w_6+w_8+w_9+w_{10}+w_{12})+w_{13}+w_{14}+w_{15}=1$. Since all $\hat{G}_k$ may be expressed by $\hat{G}_1$, $\hat{G}_{10}$, and $\hat{G}_{13}$, we then have $\hat{Q}^{V_{thr}}=(w_1+\alpha w_8+\alpha^2w_{12})\hat{G}_1+(w_{10}+\alpha w_6+\alpha^2w_2)\hat{G}_{10}+[(w_{13}-w_9)+\alpha(w_{14}-w_3)+\alpha^2(w_{15}-w_4)]\hat{G}_{13}$. Notice that $\hat{G}_1+\hat{G}_{10}-\hat{G}_{13}=0$, thus,

$$\begin{aligned}\hat{Q}^{V_{thr}}=&[(w_1+w_{13}-w_9)+\alpha(w_8+w_{14}-w_3)\\&+\alpha^2(w_{12}+w_{15}-w_4)]\hat{G}_1\\&+[(w_{10}+w_{13}-w_9)+\alpha(w_6+w_{14}-w_3)\\&+\alpha^2(w_2+w_{15}-w_4)]\hat{G}_{10}.\end{aligned} \tag{5}$$

Matrix $\hat{G}_1$ (with entries $G_1^{11}=2$, $G_1^{12}=-\alpha^2$, $G_1^{21}=-\alpha$, $G_1^{22}=2$) is a sum of matrices $\hat{H}_1$ and $\hat{H}_8$, whereas matrix $\hat{G}_{10}$ (with entries $G_{10}^{11}=-1$, $G_{10}^{12}=2\alpha^2$, $G_{10}^{21}=2\alpha$, $G_{10}^{22}=-1$) is a sum of matrices $\hat{H}_{14}$ and $\hat{H}_{22}$. These four matrices are linearly independent so they form a basis in four-dimensional space of 2×2 complex matrices. The expansion of $\hat{Q}^{V_{thr}}$ in this basis reads

$$\hat{Q}^{V_{thr}}=\lambda_1\hat{G}_1+\lambda_{10}\hat{G}_{10}, \tag{6}$$

where $\lambda_1=V_{thr}[(1/6+1/3\sqrt{3})+i(-1/9+1/2\sqrt{3})]$ and $\lambda_{10}=V_{thr}[(1/6-1/3\sqrt{3})+i(1/9+1/2\sqrt{3})]$. Because both λ_1 and λ_{10} lie on the complex plane between complex numbers 1 and α, they may be uniquely expressed by 1 and α with positive coefficients, i.e., $\lambda_1=V_{thr}[1/27(9+2\sqrt{3})+\alpha(1/27)(9-2\sqrt{3})]$ and $\lambda_{10}=V_{thr}[1/27(9-2\sqrt{3})+\alpha(1/27)(9+2\sqrt{3})]$.

We may rewrite the formula (5) using the identity $1+\alpha+\alpha^2=0$ as

$$\lambda_1=(w_1+w_4+w_{13}-w_9-w_{12}-w_{15})$$
$$+\alpha(w_4+w_8+w_{14}-w_3-w_{12}-w_{15}),$$
$$\lambda_{10}=(w_4+w_{10}+w_{13}-w_2-w_9-w_{15})$$
$$+\alpha(w_4+w_6+w_{14}-w_2-w_3-w_{15}). \tag{7}$$

After comparing Eqs. (6) and (7), we have

$$w_1+w_4+w_{13}-w_9-w_{12}-w_{15}=\frac{V_{thr}}{27}(9+2\sqrt{3}),$$

$$w_4+w_8+w_{14}-w_3-w_{12}-w_{15}=\frac{V_{thr}}{27}(9-2\sqrt{3}),$$

$$w_4+w_{10}+w_{13}-w_2-w_9-w_{15}=\frac{V_{thr}}{27}(9-2\sqrt{3}),$$

$$w_4+w_6+w_{14}-w_2-w_3-w_{15}=\frac{V_{thr}}{27}(9+2\sqrt{3}). \tag{8}$$

Because we deal with the optimal solution for which V is maximal $V=V_{thr}$, all the probabilities with a negative sign in Eq. (8) must be zero (note that none of the probabilities that come into Eq. (8) with negative sign appears in any equation with a positive sign). We get

$$w_1+w_4+w_{13}=\frac{V_{thr}}{27}(9+2\sqrt{3}),$$

$$w_4+w_8+w_{14}=\frac{V_{thr}}{27}(9-2\sqrt{3}),$$

$$w_4+w_{10}+w_{13}=\frac{V_{thr}}{27}(9-2\sqrt{3}),$$

$$w_4+w_6+w_{14}=\frac{V_{thr}}{27}(9+2\sqrt{3}). \tag{9}$$

Now the whole probability distribution consists of $w_1,w_4,w_6,w_8,w_{10},w_{13},w_{14}$. By subtracting the fourth equation from the second one and the third one from the first one we arrive at

$$w_6-w_8=\frac{4\sqrt{3}}{27}V_{thr}, \qquad w_1-w_{10}=\frac{4\sqrt{3}}{27}V_{thr}.$$

Again, because V_{thr} is maximal, it must be $w_8=w_{10}=0$. Thus, $w_1=w_6=(4\sqrt{3}/27)V_{thr}$. The second and the third equations in Eq. (9) then become

$$w_4+w_{14}=\frac{V_{thr}}{27}(9-2\sqrt{3}), \quad w_4+w_{13}=\frac{V_{thr}}{27}(9-2\sqrt{3}),$$

which clearly implies $w_{13}=w_{14}=q$ and $q+w_4=V_{thr}/27(9-2\sqrt{3})$. The normalization condition now reads $2(w_1+w_4+w_6)+w_{13}+w_{14}=1$. A simple algebra gives $q+w_4=1/2-(8\sqrt{3}/27)V_{thr}$. Therefore,

$$\frac{1}{2}-\frac{8\sqrt{3}}{27}V_{thr}=\frac{V_{thr}}{27}(9-2\sqrt{3}), \tag{10}$$

which gives $V_{thr}=6\sqrt{3}-9/2$. This ends the proof.

IV. FINAL REMARKS

We have shown analytically that for the Bell experiment with the four trichotomic observables (2) (two at each side of the experiment) defined by the sets of phase shifts $\vec{\phi}_1=(0,\pi/3,-\pi/3)$, $\vec{\phi}_2=(0,0,0)$, $\vec{\theta}_1=(0,\pi/6,-\pi/6)$, $\vec{\theta}_2=(0,-\pi/6,\pi/6)$, the minimal noise admixture F_{thr} above which local and realistic description exists is $F_{thr}=1-V_{thr}=11-6\sqrt{3}/2$. For two maximally entangled qubits, this number is $(2-\sqrt{2}/2)<F_{thr}$. Therefore, two entangled qutrits are more robust against the local and realistic description than two entangled qubits.

Although our proof cannot be easily extended to the set of arbitrary observables defined in Eq. (2) as it relies on the symmetry properties of matrix $\hat{Q}^F$, it may be considered as the first step towards the Bell theorem for two entangled qutrits.

ACKNOWLEDGMENTS

M.Z. thanks Anton Zeilinger and Alipasha Vaziri for discussions. M.Z. and D.K. are supported by KBN Grant No. 5 P03B 088 20. This paper is also supported in part under NUS research Grant No. R-144-000-054-112.

[1] J. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[2] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969); E.P. Wigner, Am. J. Phys. **38**, 8 (1970); J.F. Clauser and M.A. Horne, Phys. Rev. D **10**, 526 (1974); N.D. Mermin, *ibid.* **22**, 2 (1980); S.L. Braunstein and C.M. Caves, Ann. Phys. (N.Y.) **202**, 22 (1990). N.D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990); M. Ardehali, Phys. Rev. D **44**, 10 (1991); N. Gisin and A. Peres, Phys. Lett. A **162**, 1 (1992); A.V. Belinskii and D.N. Klyshko, Usp. Fiz. Nauk. **36**, 653 (1993) [Sov. Phys. Usp. **36**, 653 (1993)]; M. Żukowski, Phys. Lett. A **177**, 4 (1993); N. Gisin, H. Bechmann-Pasquinucci, *ibid.* **246**, 1 (1998); N. Gisin, *ibid.* **260**, 1 (1999); D. Kaszlikowski and M. Żukowski, Phys. Rev. A **61**, 2 (2000); I. Pitovsky and K. Svozil, e-print quant-ph//0011060.
[3] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[4] H. Weinfurter and M. Żukowski, Phys. Rev. A (to be published); R.F. Werner and M.M. Wolf, e-print quant-ph//0102024; M. Żukowski and C. Brukner, e-print quant-ph//0102039.
[5] N.D. Mermin and G. Schwarz, Found. Phys. **12**, 101 (1982).
[6] I. Pitovsky, Math. Program. **50**, 395 (1991).
[7] A. Peres, Found. Phys. **29**, 589 (1999).
[8] M. Żukowski, D. Kaszlikowski, A. Baturo, and J-A. Larsson, e-print quant-ph//9910058.

[9] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, Phys. Rev. Lett. **85**, 4418 (2000).
[10] T. Durt, D. Kaszlikowski, and M. Żukowski, e-print quant-ph//0101084.
[11] N. Gisin and A. Peres, Phys. Lett. A **162**, 15 (1992).
[12] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[13] M. Żukowski, A. Zeilinger, and M.A. Horne, Phys. Rev. A **55**, 2564 (1997).
[14] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).