Detector inefficiencies in the Einstein-Podolsky-Rosen experiment

Anupam Garg
Department of Physics, University of Illinois, Urbana, Illinois 61801
*and Institute for Theoretical Physics, University of California, Santa Barbara, California 93106**

N. D. Mermin
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853
(Received 3 March 1986)

The question of how to deal with inefficient detectors in actual experiments of the Einstein-Podolsky-Rosen type is studied. We derive the necessary *and sufficient* condition for compatibility with local realism of data collected in experiments with two settings of each detector, without making auxiliary assumptions about undetected events. For the conventional experiment with particles in the singlet state (or its photon analogue), the data predicted by the quantum theory do not violate this condition unless the quantum efficiency of the detectors exceeds 83%.

I. INTRODUCTION

The experiments of Aspect and co-workers[1] have drawn considerable attention in the broader scientific literature,[2] because they bear on whether or not a common-sense description of the world, first proposed by Einstein, Podolsky, and Rosen,[3] is valid. This description, sometimes termed "local realism," assigns an independent and objective reality to physical properties of parts of a many-component system when these parts are well separated from one another. Local realism is in patent conflict with the conceptual apparatus of quantum mechanics. Bell's inequality[4] and its generalizations[5–7] show that it is also in conflict with the quantitative predictions of quantum mechanics.

The experiments by Aspect *et al.* and by others[8,9] are all variants of Bohm's version[10] of the Einstein-Podolsky-Rosen experiment, and they entail the measurement of the spin projections (or polarizations) of two spatially well-separated spin-$\frac{1}{2}$ particles (or photons) in a highly correlated state such as the singlet state, along independently chosen directions. Loosely speaking, local realism implies an upper limit on the spin correlations which is exceeded by quantum-mechanical predictions. In all the experiments that have been performed so far, not all the particles are detected, and so the correlations are measured using only those events in which both members of a pair are observed, and it is these correlations which exceed the upper limit. In defense of local realism, however, one could argue that the excessive correlations might be possessed only by those particles actually detected in pairs. If the detected particles were not representative of the whole ensemble, then the experiments could not be regarded as having conclusively refuted local realism.

This loophole for local realism has, in fact, long been noted[5–7,11] and various auxiliary assumptions have been proposed, whose conjunction with local realism can be tested in a real experiment. Lo and Shimony[12] have studied how efficient the detectors must be in order not to make such assumptions and still permit a discriminating experiment to be carried out, assuming that the pairs that are observed obey quantum mechanics. The necessary condition for local realism that they derive can be shown to require a minimum detector quantum efficiency of 86% (more precisely 85.97%) to be violated in the most favorable ideal case,[13] and a particular experiment they propose requires an efficiency of 90% in order to reveal a violation.

Our point of view in this paper is very similar to that of Lo and Shimony. We shall improve upon their analysis and obtain the strongest possible inequality implied by local realism for the standard experiment where there are two choices for the direction along which each particle's spin can be measured. (We shall call such an experiment a "2×2 experiment.") In practical terms, alas, this does not help very much: the limit on the detectors' efficiency is lowered to 83% (more precisely 82.84%) in the ideal case.

The condition that we obtain is, in fact, contained in a somewhat disguised form in Ref. 5, although it appears there only as a necessary condition for local realism, and the authors do not use it to discuss the question we are asking in this paper, namely, how efficient must the detectors be in order to test local realism without invoking subsidiary assumptions. Our demonstration of its sufficiency for a local realistic model of data obtained from 2×2 experiments implies that such experiments done with detectors less than 83% efficient will always be explicable in terms of some hidden-variable model exploiting detector inefficiencies. It is possible that $n\times n$ experiments with n larger than 2 can refute local realism with lower detector efficiencies, but the generalization of the analysis we describe below even to the 3×3 case has not yet been done, and promises to be an exceedingly cumbersome exercise.[14] Until 3×3 or higher-order experiments can be shown to refute local realism with lower efficiencies, the critical efficiency will have to remain at 83%.

To formulate the question more precisely, consider an experiment with ideal detectors. Two spin-$\frac{1}{2}$ particles in the singlet state fly apart toward Stern-Gerlach apparatuses, and through repeated runs of the experiment, we measure the probability $p_{ik}(m,m')$ for the first particle to have a spin projection m along a direction $\hat{\mathbf{a}}_i$, $i=1,2$, and for the second particle to have a spin projection m' along a direction $\hat{\mathbf{a}}_k$, $k=3,4$. It is convenient to rescale the m's to take on values ± 1 rather than $\pm\frac{1}{2}$. The data are said to be compatible with local realism if they can be represented in the form

$$p_{ik}(m,m')=\int d\lambda\,\rho(\lambda)p_i(m\mid\lambda)\bar{p}_k(m'\mid\lambda)\ . \tag{1.1}$$

Here, λ is a set of hidden variables with distribution $\rho(\lambda)$, and $p_i(m\mid\lambda)$ and $\bar{p}_k(m'\mid\lambda)$ are conditional probabilities for m and m', respectively, given a particular choice for the hidden variable λ. The representation (1.1) implies the inequality[5,7]

$$|E_{13}\pm E_{23}|+|E_{14}\mp E_{24}|\le 2\ , \tag{1.2}$$

where E_{ik} is the correlation function

$$E_{ik}=\sum_{m,m'} mm'p_{ik}(m,m')\ . \tag{1.3}$$

Fine[15] subsequently showed that these inequalities are in fact *sufficient* for a representation (1.1) of the data in a 2×2 experiment.

An example of a subsidiary assumption that is often made in order to analyze realistic experiments is the no-enhancement assumption of Clauser and Horne.[5] This assumption is formulated in the context of photon experiments with polarizers for detectors in which only one value of the polarization is detected. (In the spin language, this corresponds to shutting off one of the beams in which particles can emerge from the Stern-Gerlach magnets.) The assumption states that for *each* value of λ, the probability of detecting a photon when the polarizer is in place is not more than the probability of detecting it when the polarizer is absent. Although quantum theory asserts (and experiments will almost surely confirm) that this is true of the *ensemble* of photons as a whole, there is no reason to impose it pointwise with respect to λ in an acceptable hidden-variable representation of the data. In principle there is no reason why the probability of triggering a counter might not be correlated with the value of the hidden variable carried by the photon and the orientation of the polarizer it has just passed through.

It should be noted that the inequality we derive is applicable only to an experiment in which particles are detected in both spin (or polarization) channels. Such a two-channel experiment has in fact been done with photons.[1(b)]

To exclude hidden-variable representations that exploit detector inefficiencies without untestable subsidiary conditions, we require a generalization of the necessary and sufficient conditions (1.2) to the case in which nondetection of a particle is treated as an outcome on the same footing as finding the values $m=1$ or $m=-1$. We assign to such nondetection the value $m=0$, so that for any pair of axes $\hat{\mathbf{a}}_i$ and $\hat{\mathbf{a}}_k$ for which the experiment is performed, the data consist of the numbers of runs $n_{ik}(m,m')$ in which the outcomes (m,m') are obtained. (Note that there are eight possible outcomes, corresponding to each m being $+1$, -1, or 0, except $m=m'=0$.) We can now ask if it is possible to supplement these data by numbers $n_{ik}(0,0)$ in such a way that the resulting probabilities $p_{ik}(m,m')$ (which now have nine possible values for their argument) can be represented in the form (1.1). In this paper, we treat only the 2×2 case, deriving the necessary and sufficient condition for the existence of a representation (1.1) when there are four such probability distributions (corresponding to two choices for each of the magnets' orientation), subject only to certain symmetry requirements discussed below. The problem thus posed is formally identical to a spin-1 problem where all the particles are detected. A slightly simpler version of it, containing somewhat different symmetry restrictions, has already been solved by Mermin and Schwarz,[16] and we summarize in Sec. II the requisite extension of their analysis.

The symmetries that we shall impose on the extended distributions are

$$\begin{aligned}
&\text{(i)}\ \ p_{ik}(m,m')=p_{ik}(-m,-m')\ ,\\
&\text{(ii)}\ \ p_{ik}(0,1)=p_{ik}(1,0)\equiv p(1,0)\ \ \text{independent of } i,k\ ,\\
&\text{(iii)}\ \ \sum_m p_{ik}(0,m)=\sum_m p_{ik}(m,0)\ \ \text{independent of } i,k\ .
\end{aligned} \tag{1.4}$$

Note that if one assumes (or verifies by some other means) that the source in an actual experiment produces pairs at a steady, fixed rate, all these relations can be subjected to direct experimental test. Since they also follow from quantum mechanics when the pairs are in a singlet state, we expect that they will be satisfied in reality, within experimental error. Note also that assumptions (ii) and (iii) imply that the (unmeasurable) rate or probability with which both members of a pair escape detection [denoted $p(0,0)$] is independent of the magnet orientations.

The inequalities that follow from the analysis of Sec. II can be written in a form superficially similar to the inequality (1.2) above:

$$|e_{13}\pm e_{23}|+|e_{14}\mp e_{24}|\le 2\ , \tag{1.5}$$

where e_{ik} is defined as

$$e_{ik}=E'_{ik}/[1-p(0,0)]\ , \tag{1.6a}$$

with

$$E'_{ik}=\sum_{m,m'} mm'p_{ik}(m,m')\ . \tag{1.6b}$$

Note that the correlation function E'_{ik}—distinct from the ideal case function E_{ik}, defined in Eq. (1.3)—is defined in terms of the extended distribution including probabilities for nondetection, which cannot be determined absolutely since $p(0,0)$, the probability of neither particle being detected, is unknown. This unknown probability drops out of the condition (1.5) entirely, however, since e_{ik} is normalized by the total number of events in which at least one particle is detected, and so can be expressed solely in terms of the observed counting rates for single and double events as

$$e_{ik}=\sum_{m,m'} mm'n_{ik}(m,m')\Big/\sum_{m,m'} n_{ik}(m,m') , \tag{1.7}$$

where the unmeasurable case $m=m'=0$ is omitted from the sums.

How hard will it be to violate condition (1.5) in an actual experiment if the results of such experiments are consistent with current theoretical expectations? If particle detection depends solely on the quantum efficiency η of the detectors, which operate independently of each other, then it is easy to show that e_{ik} and the ideal correlation function E_{ik} pertaining to the underlying distribution are related by

$$e_{ik}=\frac{\eta}{2-\eta}E_{ik} , \tag{1.8}$$

so that the condition (1.5) can be written as

$$|E_{13}\pm E_{23}|+|E_{14}\mp E_{24}|\le 4\eta^{-1}-2 . \tag{1.9}$$

If, in addition, the underlying distributions are those given by quantum mechanics for pairs in a singlet state, then

$$E_{ik}=\hat{\mathbf{a}}_i\cdot\hat{\mathbf{a}}_k . \tag{1.10}$$

The left-hand side of (1.9) then has a maximum value of $2^{3/2}$, which exceeds the right-hand side if $\eta\ge 2(\sqrt{2}-1)\simeq 0.83$.

The close similarity between conditions (1.5) and (1.2) suggests that the latter ought to follow from a straightforward generalization of the arguments leading to the former or that it ought to be possible to map the case with inefficient detectors back on to the ideal case before carrying out the analysis. We know of no simpler way to establish that (1.5) is a sufficient condition for a representation (1.1) (which is the basis for the somewhat demanding implications for subsequent experimental tests), other than by going through the rather elaborate analysis described in Sec. II. The necessity is considerably easier to show. A proof (suggested by Appendix B of Ref. 5) is as follows.

Let us define the function $f(m)$ to be 1 if m equals 1, and to be -1 otherwise. Denoting $f(m_i)$ by f_i, we have the obvious inequality

$$(f_1+f_2)f_3+(f_1-f_2)f_4\le 2 . \tag{1.11}$$

If we now multiply this inequality by the conditional probabilities $p_1(m_1|\lambda)$, $p_2(m_2|\lambda)$, $\bar{p}_3(m_3|\lambda)$, and $\bar{p}_4(m_4|\lambda)$, sum over all the m's, and average over the distribution $\rho(\lambda)$, we get

$$E'_{13}+E'_{23}+E'_{14}-E'_{24}+2p(0,0)\le 2 , \tag{1.12}$$

which via Eq. (1.6) is easily seen to be one of the eight inequalities encapsulated in (1.5). Three others are similarly obtained by appropriate changes of signs in Eq. (1.11), and the remaining four are obtained by averaging inequalities such as

$$(f_1+f_2)g_3+(f_1-f_2)g_4\le 2 , \tag{1.13}$$

where $g_i\equiv f(-m_i)$.

The above proof that if the data are compatible with local realism then the inequality (1.5) must hold, does not require the data to satisfy any of the symmetries (1.4). These are needed only to establish the rather more difficult converse: that if the inequality does hold then the data are compatible with local realism.

II. DERIVATION OF THE NECESSARY AND SUFFICIENT CONDITION

The inequality (1.5) can be derived by a straightforward generalization of the analysis of Sec. 5 of Mermin and Schwarz.[16] We refer the reader to their paper for a discussion of how to formulate the solution and we shall use their notation throughout, referring to their equations by prefacing the equation numbers by the letters MS.

If local realism holds, then just as the measured distributions (or "two-axis" functions) are written in Eq. (1.1) as averages over λ of a product of two conditional probabilities, we can construct three-axis functions $p_{12,3}$ and $p_{12,4}$ by averaging products of *three* conditional probabilities. These functions will be non-negative, and, respectively, return $p_{i,3}$ and $p_{i,4}$, $i=1,2$, when summed over the appropriate variable. In addition, summing $p_{12,3}$ over variable number 3 will give the same function as summing $p_{12,4}$ over variable number 4. Conversely, if two three-axis functions with all these properties can be found, then one can easily construct[14,15] a four-axis function $p_{12,34}$, which is non-negative and returns all four pair distributions as marginals. This in turn can be shown[15,17] to be equivalent to the existence of a representation (1.1) for the data in a 2×2 experiment. The existence of two such three-axis functions is therefore necessary *and sufficient* for the existence of a representation (1.1) for the data in the 2×2 case.

Very briefly, the method of Mermin and Schwarz derives necessary and sufficient conditions for the existence of a pair of three-axis functions as follows. The most general functions with these properties can be expressed as linear combinations of the specified correlation coefficients such as the E_{ik}, and a finite number of unknown quantities which drop out of all observable quantities. The requirement that the three-axis functions be non-negative leads to a set of linear inequalities in these unknowns. One picks any one of the unknowns and rewrites the inequalities as a set of upper and lower bounds for it. By demanding that every upper bound exceed every lower bound one obtains a new set of inequalities in one fewer unknown. Repeating this process, one finally arrives at inequalities in the known coefficients alone, which are the desired necessary and sufficient conditions for the existence of three-axis functions, and therefore, for the compatibility of the data at hand with local realism. (They are clearly necessary, and they are also sufficient, since if they hold, the method of their derivation shows how to construct the three-axis functions.)

The most general observed distributions compatible with the symmetries (1.4) can be written as[18]

$$p_{ik}(m,m')=\tfrac{1}{9}[1+\tfrac{3}{2}k(3m^2+3m'^2-4)+\tfrac{3}{2}\bar{c}_{ik}mm' + \tfrac{1}{4}(3c_{ik}{}^2-1)(3m^2-2)(3m'^2-2)] . \tag{2.1}$$

To extend the analysis of Ref. 16, we need only to introduce one more b coefficient, namely, $b_{200}=b_{020}=b_{002}$, since the probability of detection of one particle (irrespective of what happens to the other) will not in general be twice the probability of nondetection. Further, the assumption that $p_{ij}(0,0)$ does not depend on the choice of axes means that $c_{12}^2=c_{13}^2$ but it is convenient not to exploit this fact immediately. With the definition

$$k=\sqrt{2}b_{002}/3 , \tag{2.2}$$

Eqs. (MS5.4)—(MS5.7) are thus replaced, respectively, by[19]

$$p=\tfrac{1}{24}(h+m_1m_2d_{12}+m_1m_3d_{13}+m_2m_3d_{23}) \quad (\text{no } m_i \text{ are zero}) , \tag{2.3}$$

$$p=\tfrac{1}{12}[1+c_{12}^2-h+4k+(2\overline{c}_{12}-d_{12})m_1m_2] \quad (\text{only } m_3 \text{ is zero}), \text{ etc.} , \tag{2.4}$$

$$p=\tfrac{1}{6}(h-5k-c_{13}^2-c_{23}^2) \quad (\text{only } m_3 \text{ not zero}), \text{ etc.} , \tag{2.5}$$

$$p=\tfrac{1}{3}(c_{12}^2+c_{13}^2+c_{23}^2+3k-h) \quad (\text{all } m_i \text{ are zero}) . \tag{2.6}$$

Note that these differ from the corresponding equations in Mermin and Schwarz only by the additional terms in k. [See Eq. (MS5.9) for the definitions of h, d_{12}, etc.]

The conditions for the three-axis distribution to be nonnegative are then given by (MS5.10)—(MS5.13) with the following replacements:

$$\begin{aligned} &h \to h-4k \text{ in (MS5.11)} ,\\ &h \to h-5k \text{ in (MS5.12)} ,\\ &h \to h-3k \text{ in (MS5.13)} , \end{aligned} \tag{2.7}$$

Eq. (MS5.10) is unchanged.

We now proceed to eliminate the unknowns h, d_{12}, d_{13}, and d_{23} in the same order as in Appendix C of 16 (Ref. 20). This leads to the following generalization of (MS5.14):

$$\min\left\{\begin{matrix} 3+2(\overline{c}_{12}+\overline{c}_{13}+\overline{c}_{23})+2k \\ 3+2(\overline{c}_{12}-\overline{c}_{13}-\overline{c}_{23})+2k \\ 3+2(-\overline{c}_{12}+\overline{c}_{13}-\overline{c}_{23})+2k \\ 3+2(-\overline{c}_{12}-\overline{c}_{13}+\overline{c}_{23})+2k \\ 1-k \end{matrix}\right\} \ge \max\left\{\begin{matrix} c_{12}^2+c_{13}^2-c_{23}^2 \\ c_{12}^2-c_{13}^2+c_{23}^2 \\ -c_{12}^2+c_{13}^2+c_{23}^2 \end{matrix}\right\} . \tag{2.8}$$

One must also retain the conditions that the unobservable two-axis distribution be positive: namely,

$$1\pm 2\overline{c}_{12}+c_{12}^2+4k\ge 0 , \tag{2.9a}$$

$$1-k\ge c_{12}^2\ge 2k . \tag{2.9b}$$

Finally, we get one more inequality which has no analogue in Ref. 16 since it is an automatic consequence of the positivity of the observed distributions and the condition (2.9b) when $k=0$:

$$c_{12}^2+c_{13}^2+c_{23}^2+3k\ge 0 . \tag{2.10}$$

The next step consists of eliminating the coefficients $\overline{c}_{12}$ and c_{12}^2. They cannot be related to observed quantities, but we do know that they must not depend on the choice of the axis $\hat{\mathbf{a}}_3$. We also set $c_{13}^2=c_{23}^2=c^2$ at this stage. The inequalities (2.8)—(2.10) can then be written more compactly as

$$1-k\ge c_{12}^2\ge \max\left\{\begin{matrix} 2k \\ -1+k+2c^2 \\ -3k-2c^2 \end{matrix}\right\} , \tag{2.11}$$

$$C_k\equiv 3+2k-2\,|\,\overline{c}_{1k}-\overline{c}_{2k}\,|\ge c_{12}^2+2\overline{c}_{12} , \tag{2.12}$$

$$c_{12}^2+2\overline{c}_{12}\ge -3-2k+2c^2+|\,\overline{c}_{1k}+\overline{c}_{2k}\,|\equiv D_k , \tag{2.13}$$

$$E_k\equiv 3+2k-2\,|\,\overline{c}_{1k}+\overline{c}_{2k}\,|\ge c_{12}^2-2\overline{c}_{12} , \tag{2.14}$$

$$c_{12}^2-2\overline{c}_{12}\ge -3-2k+2c^2+|\,\overline{c}_{1k}-\overline{c}_{2k}\,|\equiv F_k , \tag{2.15}$$

$$c_{12}^2\pm 2\overline{c}_{12}\ge -(1+4k) , \tag{2.16}$$

where the indices k,l can take on values 3,4 independently.

If we now eliminate $\overline{c}_{12}$ from the last five inequalities we get four inequalities that do not contain c_{12}^2,

$$C_k\ge D_l,\quad E_k\ge F_l,\quad C_l\ge -(1+4k),\quad E_k\ge -(1+4k) , \tag{2.17}$$

and five more that do:

$$2c_{12}^2\ge \{D_k+F_l, D_k-(1+4k), F_k-(1+4k), -2(1+4k)\} , \tag{2.18}$$

$$C_k+E_l\ge 2c_{12}^2 . \tag{2.19}$$

Similarly eliminating c_{12}^2 we get 14 more inequalities all of which can be shown to be consequences either of the positivity of the observed distributions or of (2.17), as can the last two inequalities in (2.17) themselves. The remaining inequalities, i.e., the first two in (2.17), can be written concisely as

$$|\,\overline{c}_{1k}\pm\overline{c}_{2k}\,|+|\,\overline{c}_{1l}\mp\overline{c}_{2l}\,|\le 3+2k-c^2 . \tag{2.20}$$

From Eq. (2.1) we can relate $\overline{c}_{ij}$ to E'_{ij}, and k and c^2 to $p(0,0)$ and $p(0,1)$ and thus write this inequality as

$$|\,E'_{1k}\pm E'_{2k}\,|+|\,E'_{1l}\mp E'_{2l}\,|\le 2[1-p(0,0)] , \tag{2.21}$$

which is nothing but (1.5).

ACKNOWLEDGMENTS

This research was supported by the John D. and Catherine T. MacArthur Foundation, and by the National Science Foundation under Grants Nos. DMR83-14625 and PHY82-17853, supplemented by funds from the National Aeronautics and Space Administration, at the University of California at Santa Barbara.

---

*Present address.

[1](a) A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981); (b) **49**, 91 (1982); (c) A. Aspect, J. Dalibard, and G. Roger, *ibid.* **49**, 1804 (1982).

[2]See, for example, A. L. Robinson, Science **217**, 436 (1982); **219**, 40 (1983), or N. D. Mermin, Physics Today **38** (No. 4), 38 (1985). See also, B. d'Espagnat, *In Search of Reality* (Springer, New York, 1983), the review of this book by T. W. Marshall, Nature **308**, 669 (1984), and the subsequent correspondence from B. d'Espagnat and T. W. Marshall, *ibid.* **313**, 483 (1985).

[3]A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[4]J. S. Bell, Ann. Phys. (N.Y.) **1**, 195 (1965).

[5]J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974). See especially Appendix B.

[6]J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[7]J. S. Bell, in *Foundations of Quantum Mechanics,* edited by B. d'Espagnat (Academic, New York, 1971).

[8]J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1882 (1978) contains a discussion of all the experiments performed prior to 1978. This review also contains a discussion of many other considerations important in an actual experiment, including subsidiary assumptions for dealing with undetected events.

[9]W. Perrie, A. J. Duncan, H. J. Beyer, and H. Kleinpoppen, Phys. Rev. Lett. **54**, 1790 (1985).

[10]D. Bohm, *Quantum Theory* (Prentice Hall, Englewood Cliffs, N.J., 1951), pp. 614–622.

[11]The earliest paper we are aware of that points this out is by Phillip M. Pearle, Phys. Rev. D **2**, 1418 (1970). Other discussions of this loophole are given in Ref. 5 above, and by A. Fine, Synthese **50**, 279 (1982), and T. W. Marshall, E. Santos, and F. Selleri, Phys. Lett. **98A**, 5 (1983).

[12]T. K. Lo and A. Shimony, Phys. Rev. A **23**, 3003 (1981); see also the Comment on this work by E. Santos, *ibid.* **30**, 2128 (1984), and the Response by A. Shimony, *ibid.* **30**, 2130 (1984).

[13]This figure is not contained anywhere in Ref. 12, but can be easily deduced from Eq. (12) there.

[14]The analysis for the 3×3 experiment with perfect detectors for the spin-$\frac{1}{2}$ case is given by Anupam Garg, Phys. Rev. D **28**, 785 (1983). A partial analysis for the 4×4 and some higher-order experiments has been performed by J. H. B. Kemperman (unpublished).

[15]A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[16]N. D. Mermin and Gina M. Schwarz, Found. Phys. **12**, 101 (1982).

[17]A simple demonstration of this equivalence is given by Anupam Garg (Ref. 15), under the heading "Step (a)" of Sec. II. The proof is given in the context of the $3\times N$ case where the variables each take on only two values, but exactly the same argument works for any number of values in the $M\times N$ case and hence, in particular, to the present case (2×2 with three-valued variables). The same argument is also given (for two-valued variables in the $N\times N$ case) in N. D. Mermin, Philos. Sci. **50**, 359 (1983), also as "Step (a)" of Sec. 2, along with a general discussion of the relevance of the numbers of variables, M and N, to the question of sufficient conditions for the existence of representations such as (1.1).

[18]This follows from the observation that 1, m, and $3m^2-2$ are a complete (and orthogonal) set of functions on the domain $\{-1,0,1\}$. The nine products of these with 1, m', and $3m'^2-2$ are thus a complete set on the domain of values assumed by the pair (m,m'). The form (2.1) follows from writing such a expansion for $p_{ik}(m,m')$ and imposing on the coefficients the constraints that follow from the symmetries (1.4).

[19]The quantities p in Eqs. (2.3)–(2.6) are the three- (and not two-) axis functions. Further, in contrast to Ref. 16, the c_{ij}^2 are no longer squares of real quantities, and represent general numbers which may be negative. We have chosen to retain this somewhat peculiar notation for continuity with Ref. 16.

[20]In carrying out the elimination of unknown variables like h, d_{12}, etc., we repeatedly discard inequalities that follow from the non-negativity of *both* the observed and the unobserved two-axis distributions. One must be careful, however, not to discard the latter conditions themselves. This point was overlooked in Sec. 5 of Ref. 16, but it does not affect any of the results quoted there. In particular, Eqs. (MS5.23)–(MS5.28) are unchanged and do not need to be supplemented.