Relevance of Bell's theorem as a signature of nonlocality: Case of classical angular momentum distributions

A. Matzkin

Laboratoire de Spectrométrie Physique (CNRS Unité 5588), Université Joseph-Fourier Grenoble-1, Boîte Postale 87, 38402 Saint-Martin d'Hères, France

(Received 2 October 2007; revised manuscript received 19 March 2008; published 20 June 2008)

For a system composed of two particles, Bell's theorem asserts that averages of physical quantities determined from local variables must conform to a family of inequalities. In this work we show that a classical model containing a local probabilistic interaction in the measurement process can lead to a violation of the Bell inequalities. We first introduce two-particle phase-space distributions in classical mechanics constructed to be the analogs of quantum-mechanical angular momentum eigenstates. These distributions are then employed in four schemes characterized by different types of detectors measuring the angular momenta. When the model includes an interaction between the detector and the measured particle leading to ensemble dependencies, the relevant Bell inequalities are violated if the total angular momentum is required to be conserved. The violation is explained by identifying assumptions made in the derivation of Bell's theorem that are not fulfilled by the model, in particular noncommutativity of single-particle measurements. We discuss to what extent a violation of these assumptions is a faithful marker of nonlocality.

DOI: 10.1103/PhysRevA.77.062110 PACS number(s): 03.65.Ud, 03.65.Ta, 45.20.dc

I. INTRODUCTION

Bell's theorem was originally introduced [1,2] to examine quantitatively the consequences of postulating hidden variable distributions on the incompleteness of quantum mechanics put forward by Einstein, Podolsky, and Rosen [3] (EPR). In particular, the hidden variables were supposed to locally and causally complete quantum mechanics by making sense of the "reality" of physical quantities described by noncommuting operators relative to two spatially separated particles in an entangled state. Bell showed that a correlation function obtained from averages over the hidden variables of these physical quantities must satisfy certain inequalities (the Bell inequalities), and that these inequalities are violated by quantum-mechanical averages. Given that experiments have confirmed with increasing precision the correctness of the quantum formalism, it is generally stated that the violation of the Bell inequalities contradicts locality. The strong version of such statements asserts that quantum mechanics itself is nonlocal [4]. This vocable is quite popular (in particular among nonspecialists as well as in quantum information papers) but there is a general agreement among most specialists that this strong assertion is unsubstantiated [5–7]. Instead, the received view is the weak version following which Bell's theorem asserts the incompatibility of local hidden variables with quantum mechanics. Nevertheless, it can objected, in principle [8,9] or through abstract models [10,11], whether the assumptions made in order to derive Bell's theorem are necessary in order to enforce locality, or whether they only rule out a certain manner of ascribing local variables to the measurement outcomes.

In this work we will show that statistical distributions in classical mechanics can violate Bell-type inequalities. Moreover, the statistical distributions we will employ are not exotic objects but the *classical analogs* of the quantum-mechanical coupled angular momenta eigenstates, so that our model is essentially the classical version of the paradigmatic 2-particles singlet state. The violation of the inequalities can of course be achieved only provided the model falls outside the assumptions necessary in order to prove Bell's theorem. This role will be played by a probabilistic interaction that is assumed to take place between the measured particle and the detector, combined with the requirement that the total angular momentum be conserved. Although this interaction is local, it nevertheless spoils the derivation of Bell's theorem, because it introduces an ensemble dependency of the outcomes: the resulting averages then involve correlations given by conditional probabilities between ensembles rather than between the individual phase-space positions. As a consequence, the different expectation values employed in Bell's inequalities cannot be derived jointly, as required in the derivation of the theorem.

The paper is organized as follows. We will start by introducing the classical phase-space distributions (Sec. II), first for a single particle, then for two particles with total zero angular momentum. We will explain why these distributions are the classical analogs of the quantum angular momentum eigenstates. In Sec. III we will investigate three different examples of Bell-type models. Each of the examples will be characterized by the same phase-space distribution but by differing detection schemes. In the first case, the projections of the angular momentum of each of the particles along arbitrary axes are directly measured by the detectors, leading to a straightforward application of Bell's theorem (which will be briefly derived). In the second example the detectors yield discrete outcomes, depending on the values of the angular momenta; this example, which also abides by Bell's theorem, will allow us to introduce conditional probabilities to account for the correlated angular momenta. The third example will illustrate the same situation with stochastic variables (the angular momenta specify probabilities of obtaining an outcome). In Sec. IV, we will introduce an example falling outside the class of Bell-type models. This example will also involve discrete measurement outcomes, but the presence of

 ©2008 The American Physical Society

an interaction leading to ensemble dependencies will be introduced. We will see that ensemble dependencies lead to noncommutative measurements for a single particle, and to the violation of the Bell inequalities for initially correlated two-particle systems. In Sec. V we will discuss these results, insisting on the role played by the existence of joint distributions and on the relationship between locality and conservation laws. A short summary and our conclusion are given in Sec. VI.

II. CLASSICAL DISTRIBUTIONS ANALOGS OF ANGULAR MOMENTA EIGENSTATES

A. One-particle angular momentum distributions

A quantum-mechanical angular momentum eigenstate $|j_0 m\rangle$ is characterized by a well-defined value $\sqrt{j_0(j_0+1)}$ of the modulus of the angular momentum $\mathbf{J}$ and of its projection J_z (of value m) along a given axis z. In configuration space the spherical harmonic $|\langle\theta,\phi|j_0 m\rangle|^2$ gives the probability distribution corresponding to a fixed value of J and J_z as θ and ϕ (the polar and azimuthal angles) span the unit sphere. The classical statistical distributions can be considered either in phase-space, defined by $\Omega=\{\theta,\phi,p_\theta,p_\phi\}$ where p_θ and p_ϕ are the conjugate canonical momenta, or in configuration space. Let us assume the modulus J of the angular momentum is fixed. Let $\rho_z(\Omega)$ be the distribution in phase-space given by

$$\rho_{z_0}(\theta,\phi,p_\theta,p_\phi) = N\delta(J_z(\Omega) - J_{z_0})\,\delta(J^2(\Omega) - J_0^2). \quad (1)$$

ρ_{z_0} defines a distribution in which every particle has an angular momentum with the same magnitude, namely J_0, and the same projection on the z axis J_{z_0}, without any additional constraint. Hence ρ_{z_0} can be considered as a classical analog of the quantum-mechanical density matrix $|j_0 m\rangle\langle j_0 m|$. Equation (1) can be integrated over the conjugate momenta to yield the *configuration space* distribution

$$\rho(\theta,\phi) = N[\sin(\theta)\sqrt{J_0^2 - J_{z_0}^2/\sin^2(\theta)}]^{-1}, \quad (2)$$

where we have used the defining relations $J_z(\Omega)=p_\phi$ and $J^2(\Omega)=p_\theta^2+p_\phi^2/\sin^2\theta$ to perform the integration. Further integrating over θ and ϕ and requiring the phase-space integration of ρ to be unity allows us to set the normalization constant $N=J_0/2\pi^2$.

$\rho(\theta,\phi)$ gives the statistical distribution of the particles in configuration space. Its standard graphical representation (parametrization on the unit sphere) is shown in Fig. 1(a) along with the quantum-mechanical orbital momentum eigenstate (a spherical harmonic taken for the same values of j and m) in Fig. 1(b). The similarity of both figures is a statement of the quantum-classical correspondence in the semiclassical regime, since $\sqrt{\rho(\theta,\phi)}$ is approximately the amplitude of the configuration space quantum-mechanical eigenstate for high quantum numbers. Note that rather than working with the particle distributions in configuration space, it will also be convenient to visualize the distribution of the angular momentum in *physical* space corresponding to a given particle distribution [see Fig. 1(c)]; θ and ϕ will then denote the position of $\mathbf{J}$ on the angular momentum sphere.

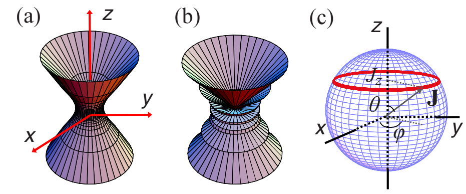

FIG. 1. (Color online). Normalized angular distribution for a single particle in configuration space. (a) *Classical* distribution $\rho(\theta,\phi)$ of Eq. (2). (b) Corresponding *quantum* distribution [spherical harmonic $|Y_{JM}(\theta,\phi)|^2$ with $J/\eta=40$, $\eta=\hbar$, and $M/J=5/8$ as in (a)]. (c) Distribution of the angular momentum on the sphere for a distribution of the type $\rho(\theta,\phi)$, invariant around the z axis with a fixed value of J_z.

Let us take a second axis a making an angle θ_a relative to the z axis (in this paper we will take all the axes to lie in the zy plane). We can define a distribution by fixing the projection J_a of the angular momentum on a to be constant, $\rho_{a_0} = \delta(J_a - J_{a_0})\,\delta(J-J_0^2)$. In configuration space, this distribution may be shown to be obtained by rotating the distribution of Eq. (2) by the angle θ_a toward the a axis. We will be interested below in determining the average projection J_a on the a axis for a distribution of the type (2) corresponding to a well defined value of J_z. Using $J_a=J_z\cos\theta_a+J_y\sin\theta_a$, $J_z=p_\phi$, and

$$\langle J_y\rangle_{J_{z_0}} = \int J\sin\theta\sin\phi\sin\theta_a\,\delta(p_\phi - J_{z_0})d\Omega = 0 \quad (3)$$

by rotational invariance, we obtain

$$\langle J_a\rangle_{J_{z_0}} = \int p_\phi\cos\theta_a\,\delta(p_\phi - J_{z_0})d\Omega = J_{z_0}\cos\theta_a. \quad (4)$$

Note that a given $\mathbf{J}$ can belong jointly to several distributions ρ_{a_0} and ρ_{b_0} (a and b being different directions). But if we require that any distribution must correspond to a well-defined value of the angular momentum projection along a given axis, then distributions such as ρ_{a_0} and ρ_{b_0} become mutually exclusive. The ringlike distribution of the angular momentum on the angular momentum sphere represented in Fig. 1(c) [corresponding to the configuration space distribution shown in Fig. 1(a)] can be generalized to cover the entire hemisphere centered on the z axis [see Fig. 3(a)]. Then properties such as J_{1a} and J_{1b} being of the same sign on such hemispheres become mutually exclusive properties.

B. Two-particle angular momentum distributions

The situation we will consider below, by analogy with the well-known EPR-Bohm pairs in quantum mechanics, is that of the fragmentation of an initial particle with a total angular momentum $\mathbf{J}_T=0$ into two particles carrying angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$. Conservation of the total angular momentum imposes $J_1=J_2\equiv J$ and

$$\mathbf{J}_1 + \mathbf{J}_2 = 0. \quad (5)$$

Equation (5) implies a *correlation*, imposed initially at the

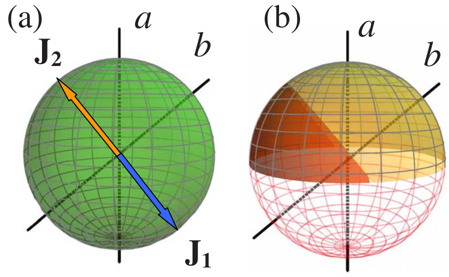

FIG. 2. (Color online). (a) Uniform distribution of $\mathbf{J}_1$ and $\mathbf{J}_2$ on the unit sphere; the angular momenta are correlated via the conservation law (5) and must thus point in opposite directions. In the first example (Sec. III B), the detectors measure J_{1a} and the correlated J_{2b} as the angular momenta span the sphere. (b) Example 2 (Sec. III C): Distribution of $\mathbf{J}_2$, when $D_{1a}=-1/2$ was obtained. A measurement of D_{2b} will yield $\pm 1/2$ depending on the position of $\mathbf{J}_2$: if $\mathbf{J}_2$ lies within the light shaded region [intersection of the two positive hemispheres centered on a and on b, denoted $\mathcal{D}(-1/2,1/2)$ in the text], $D_{2b}=1/2$ will be obtained, $-1/2$ when $\mathbf{J}_2$ belongs to the dark-shaded region [$\mathcal{D}(-1/2,-1/2)$].

source, between the angular momenta of the two particles and of their projections along any axis a: the knowledge of the value of J_{1a} allows us to infer the value of J_{2a}, $J_{2a}=-J_{1a}$. Without further constraints (or additional knowledge), the classical distribution in the two-particle phase-space is given by

$$\rho(\Omega_1,\Omega_2) = N\delta(\mathbf{J}_1+\mathbf{J}_2)\,\delta(J_1^2-J^2), \tag{6}$$

where N is again a normalization constant. The corresponding distributions of the angular momenta in physical space—easier to visualize than ρ—are uniform on the sphere, with $\mathbf{J}_1$ and $\mathbf{J}_2$ pointing in opposite directions [see Fig. 2(a)], reflecting the isotropic character of the fragmentation as well as the correlation (5).

III. BELL-TYPE MODELS

A. Setting

The Bell inequalities are obtained by computing average values of measurement outcomes performed independently on each of the two particles. Three examples are studied below, all involving the initial fragmentation of a particle with zero angular momentum (Sec. II B). In the first example, we assume that the measurements give directly the value of the projection of the angular momentum of each particle along an arbitrarily chosen axis. In the second example we introduce detectors having a threshold, resulting in discrete measurement outcomes depending solely on the position of the particles' angular momenta. The third example is a repetition of the second but with stochastic variables. Bell's theorem, which is derived in Sec. III B, is verified in all these cases. To alight the notation, we will choose units such that $J=1$.

B. Bell's theorem

1. Example 1: Direct measurement of the classical angular momenta

Two particles with initial total angular momentum $\mathbf{J}_T=0$ flow apart. Let a and b be two axes in the zy plane. The projection of particle 1's angular momentum along the a axis, J_{1a} and that of particle 2 along b, J_{2b} are measured. The average of the joint measurement outcomes on the two particles is directly given by the values of J_{1a} and J_{2b} and the probability distribution given by Eq. (6). All these quantities depend on the phase-space position of the particles, i.e., on the position of the angular momenta on the sphere [see Fig. 2(a)]. The average is computed from

$$\langle J_{1a}J_{2b}\rangle = \int J_{1a}(\Omega_1)J_{2b}(\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2. \tag{7}$$

Given the rotational symmetry, z is chosen along a, hence $J_{1a}=p_{\phi_1}$ and

$$J_{2b} = p_{\phi_2}\cos(\theta_b-\theta_a) + \{J\sin\theta_2\sin\phi_2\sin(\theta_b-\theta_a)\}. \tag{8}$$

One first integrates over ϕ_2 (the term between $\{\cdots\}$ vanishes), then over p_{ϕ_2} [yielding $p_{\phi_2}=-p_{\phi_1}$ because of the correlation $\delta(J_{1a}+J_{2a})$]. The last nontrivial integration is over p_{ϕ_1},

$$\langle J_{1a}J_{2b}\rangle = \int_{-1}^{1} dp_{\phi_1} - p_{\phi_1}^2\cos(\theta_b-\theta_a)\left[2\pi N\int d\tilde{\Omega}\right], \tag{9}$$

where $d\tilde{\Omega}$ represents the variables remaining after the integration of the delta functions. Since ρ is normalized, we have

$$\int_{-1}^{1} dp_{\phi_1} 2\pi N\int d\tilde{\Omega} = 1. \tag{10}$$

Integrating Eq. (10) over p_{ϕ_1} allows us to obtain the value between the $[\cdots]$ in Eq. (9) thereby avoiding the explicit calculation of the normalization constant. The result for the expectation is

$$E(a,b) \equiv \langle J_{1a}J_{2b}\rangle = -\frac{1}{3}\cos(\theta_b-\theta_a). \tag{11}$$

2. Derivation of the Bell inequality

The correlation function $C(a,b,a',b')$ involved in Bell's inequality is given by

$$C(a,b,a',b') = [|E(a,b) - E(a,b')| + |E(a',b) + E(a',b')|]/V_{\max}^2, \tag{12}$$

where a' and b' are arbitrary axes in the xy plane and $V_{\max}$ is the maximal absolute value that can be obtained in a measurement outcome. Let us denote by $V_{1a}(\Omega_1)$, $V_{2b}(\Omega_2)$, etc., the detected values along the relevant axes, with the two-particle average being

$$E(a,b) = \int V_{1a}(\Omega_1)V_{2b}(\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2. \quad (13)$$

The Bell inequality

$$C(a,b,a',b') \le 2 \quad (14)$$

puts a bound on the value of the correlation function. It is obtained [12] by forming the difference

$$E(a,b) - E(a,b') = \int V_{1a}(\Omega_1)[V_{2b}(\Omega_2) - V_{2b'}(\Omega_2)] \times \rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2, \quad (15)$$

where V_{1a} has been factored. Likewise,

$$E(a',b) + E(a',b') = \int V_{1a'}[V_{2b} + V_{2b'}]\rho d\Omega_1 d\Omega_2. \quad (16)$$

We now use $|V_{2\beta}| \le V_{\max}$ $(\beta = b,b')$ to derive

$$|V_{2b} - V_{2b'}| + |V_{2b} + V_{2b'}| \le 2V_{\max}. \quad (17)$$

Take the absolute values and use $|V_{1\alpha}| \le V_{\max}$ $(\alpha = a,a')$ in each of Eqs. (15) and (16) to obtain two inequalities. Adding these inequalities and using Eq. (17) leads to the Bell inequality

$$|E(a,b) - E(a,b')| + |E(a',b) + E(a',b')| \le 2V_{\max}^2. \quad (18)$$

In the present example, $V_{\max}=1$, and $C(a,b,a',b')$ is bounded by $2\sqrt{2}/3$, so that the Bell inequality (14) is verified.

As a corollary, note that the factorization made in Eqs. (15) and (16) is equivalent [13] to the existence of joint distributions of the form

$$\mathcal{F}_{aba'b'} = \int V_{1a}(\Omega_1)V_{2b}(\Omega_2)V_{1a'}(\Omega_1)V_{2b'}(\Omega_2) \times \rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2. \quad (19)$$

Indeed, Bell's inequality can be proved [14] by adding and subtracting $\mathcal{F}_{aba'b'}$ from Eq. (15) and then factorizing $V_{1a}V_{2b}$ and $V_{1a}V_{2b'}$, respectively. The term $\mathcal{F}_{aba'b'}$ is the average obtained when four measurements are made—two outcomes are obtained for each particle (particle 1's V property is measured along the axes a and a' whereas particle 2 is measured along the axes b and b'). The factorization, or equivalently the existence of $\mathcal{F}_{aba'b'}$, is an important assumption in the derivation of the inequalities.

3. Derivation in the stochastic case and joint distributions

In the stochastic case, a given phase-space position (Ω_1,Ω_2) does not determine a unique valued outcome $(V_{1a}(\Omega_1),V_{2b}(\Omega_2))$ as above (corresponding to what is usually termed "deterministic case") but determines instead well-defined probabilities $p(V_{1a},V_{2b},\Omega_1,\Omega_2)$ of obtaining (V_{1a},V_{2b}). The counterpart of the factorization made in Eq. (15) lies in the factorization of the probabilities,

$$p(V_{1a},V_{2b},\Omega_1,\Omega_2) = p(V_{1a},\Omega_1)p(V_{2b},\Omega_2), \quad (20)$$

where $p(V_{1a},\Omega_1)$ is the single-particle elementary probability such that

$$P(V_{1a}) = \int p(V_{1a},\Omega_1)\rho(\Omega_1)d\Omega_1. \quad (21)$$

The expectation value (13) is then replaced by

$$E(a,b) = \int \bar{V}_{1a}(\Omega_1)\bar{V}_{2b}(\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2 \quad (22)$$

with

$$\bar{V}_{1a}(\Omega_1) = \sum V_{1a}p(V_{1a},\Omega_1), \quad (23)$$

$$\bar{V}_{2b}(\Omega_2) = \sum V_{2b}p(V_{2b},\Omega_2). \quad (24)$$

The derivation leading to Eq. (18) proceeds as above by replacing the value of the outcomes by their respective averages $\bar{V}_{1a}$ and $\bar{V}_{2b}$. The factorization (20) allows us to obtain a joint probability for an arbitrary number of events from the elementary probabilities $p(V,\Omega)$; the counterpart to Eq. (19) is

$$P_{aba'b'} = \int p(V_{1a},\Omega_1)p(V_{2b},\Omega_2)p(V_{1a'},\Omega_1) \times p(V_{2b'},\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2. \quad (25)$$

Note that the existence of a joint probability $P_{aba'b'}$ [that appears here as a consequence of the factorization (20)] leads immediately to the inequality (18) irrespective of *any other assumption* concerning the dependence of the outcomes or probabilities on supplementary variables (here the phase-space positions, the "hidden variables" in quantum mechanics). Indeed, using expressions of the type

$$E(a,b) = \sum_{V_{1a},V_{2b}} V_{1a}V_{2b} \sum_{V_{1a'},V_{2b'}} P_{aba'b'} \quad (26)$$

for the average values, we have

$$|E(a,b) - E(a,b')| \le \sum P_{aba'b'}|V_{1a}(V_{2b} - V_{2b'})| \quad (27)$$

and an analog inequality for $|E(a',b)+E(a',b')|$. Adding both inequalities yields

$$|E(a,b) - E(a,b')| + |E(a',b) + E(a',b')| \le \sum P_{aba'b'}(|V_{1a}(V_{2b} - V_{2b'})| + |V_{1a'}(V_{2b} + V_{2b'})|) \le 2V_{\max}^2, \quad (28)$$

since the expression between $(\cdots)$ is bounded by $2V_{\max}^2$ and the joint probability sums to 1.

C. Discrete outcomes

In this second example, we take over the setup of the first example except for the measurement outcomes: we now assume that a given detector placed on an axis can only give two values, depending on the sign of the angular momentum's projection. The outcomes are given by

$$D_{1a}(\Omega_1) = \begin{cases} \frac{1}{2} & \text{if } J_{1a} > 0 \\ -\frac{1}{2} & \text{if } J_{1a} < 0 \end{cases} \tag{29a}$$

$$D_{2b}(\Omega_2) = \begin{cases} \frac{1}{2} & \text{if } J_{2b} > 0 \\ -\frac{1}{2} & \text{if } J_{2b} < 0 \end{cases} \tag{29b}$$

and depend only on the positions $\mathbf{J}_1$ and $\mathbf{J}_2$ of the angular momentum (hence on the phase-space position of the measured particles). The average value

$$\langle D_{1a}D_{2b}\rangle = \int D_{1a}(\Omega_1)D_{2b}(\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2 \tag{30}$$

takes the form

$$\langle D_{1a}D_{2b}\rangle = \sum_{k,k'=-1/2}^{1/2} kk' \int_{\mathcal{D}(k,k')} \rho d\Omega_1 d\Omega_2, \tag{31}$$

where $\mathcal{D}(k,k')$ represents the domain of integration on which the joint conditions $\text{sgn}(J_{2a}) = -\text{sgn}(J_{1a}) = -\text{sgn}(k)$ *and* $\text{sgn}(J_{2b}) = \text{sgn}(k')$ hold [see Fig. 2(b)]. The integral gives the probability

$$\begin{aligned} P_{kk'} &\equiv P(D_{1a}=k \cap D_{2b}=k') \\ &= P(D_{1a}=k)P(D_{2b}=k' | D_{1a}=k), \end{aligned} \tag{32}$$

where $P(D_{2b}=k' | D_{1a}=k)$ is the probability of obtaining $D_{2b}=k'$ conditioned on the knowledge that $D_{1a}=k$. The conditional probability appears because of the initial correlation (5)—the positions of the angular momenta are not independent. The conditional probability can more easily be determined on the angular momentum sphere by computing the area where $\text{sgn}(J_{2b}) = \text{sgn}(k')$ relative to the area of the hemisphere where $\text{sgn}(J_{2a}) = -\text{sgn}(k)$ (of area 2π). This area is given by the intersection of the two relevant hemispheres [see Fig. 2(b)], i.e., a spherical lune whose area can be put under the form $2\pi k(k-k') + 4kk'(\theta_b - \theta_a)$. Since ρ is uniform on the sphere, we have $P(D_{1a}=k) = 1/2$ from where

$$P_{kk'} = k(k-k') + \frac{2kk'}{\pi}|\theta_b - \theta_a|, \tag{33}$$

and the average $\langle D_{1a}D_{2b}\rangle$ becomes

$$E(a,b) = -\frac{1}{4} + \frac{|\theta_b - \theta_a|}{2\pi}. \tag{34}$$

The maximal detected value here is $V_{\max} = 1/2$. The correlation function is computed from Eq. (12) and it may be verified that $C(a,b,a',b')$ is bounded by 2: Bell's inequality (14) is again verified.

D. Discrete outcomes: a stochastic model

We now elaborate on the preceding example to give a model in line with the stochastic version of Bell-type variables. A given position of the angular momentum of a particle in phase-space does not specify the outcome S, as in Eq. (29), but the probabilities $p(S_{1a}=k,\Omega_1)$ of obtaining the outcome k. For definiteness we will replace Eqs. (29) by

$$p\left(S_{1a} = \frac{1}{2}, \Omega_1\right) = \begin{cases} \frac{3}{4} & \text{if } J_{1a} > 0 \\ \frac{1}{4} & \text{if } J_{1a} < 0 \end{cases}, \tag{35a}$$

$$p\left(S_{2b} = \frac{1}{2}, \Omega_2\right) = \begin{cases} \frac{3}{4} & \text{if } J_{2b} > 0 \\ \frac{1}{4} & \text{if } J_{2b} < 0 \end{cases}, \tag{35b}$$

$$p\left(S_{1a} = -\frac{1}{2}, \Omega_1\right) = \begin{cases} \frac{1}{4} & \text{if } J_{1a} > 0 \\ \frac{3}{4} & \text{if } J_{1a} < 0 \end{cases} \tag{36a}$$

$$p\left(S_{2b} = -\frac{1}{2}, \Omega_2\right) = \begin{cases} \frac{1}{4} & \text{if } J_{2b} > 0 \\ \frac{3}{4} & \text{if } J_{2b} < 0 \end{cases}. \tag{36b}$$

The expectation value involves first averaging, for each phase-space position, over the two possible outcomes, before averaging over the distribution ρ of the angular momenta:

$$\langle S_{1a}S_{2b}\rangle = \int \bar{S}_{1a}(\Omega_1)\bar{S}_{2b}(\Omega_2)\rho(\Omega_1,\Omega_2)d\Omega_1 d\Omega_2 \tag{37}$$

with [cf. Eqs. (22)–(24)]

$$\bar{S}_{1a}(\Omega_1) = \sum_k k p(S_{1a}=k,\Omega_1), \tag{38}$$

$$\bar{S}_{2b}(\Omega_2) = \sum_{k'} k' p(S_{2b}=k',\Omega_2). \tag{39}$$

Taking into account the correlation at the source [Eqs. (5) and (6)], we proceed as in the preceding example, except that now each probability $P_{kk'}$ contains several contributions with a weight given by $p(S_{1a}=k,\Omega_1)p(S_{2b}=k',\Omega_2)$ that depends, through Eqs. (35) and (36), on the domains $\mathcal{D}(\pm 1/2, \pm 1/2)$ over which $\text{sgn}(J_{1a}) = \mp 1$ and $\text{sgn}(J_{2b}) = \pm 1$. For example for $k,k' = \frac{1}{2}$, we have

$$P_{1/21/2} = \frac{9}{16}\int_{\mathcal{D}(1/2,1/2)} \rho d\Omega_1 d\Omega_2 + \frac{3}{16}\int_{\mathcal{D}(1/2,-1/2)} \rho d\Omega_1 d\Omega_2 + \frac{1}{16}\int_{\mathcal{D}(-1/2,1/2)} \rho d\Omega_1 d\Omega_2 + \frac{3}{16}\int_{\mathcal{D}(-1/2,-1/2)} \rho d\Omega_1 d\Omega_2; \quad (40)$$

now each integral represents a probability $P(\mathrm{sgn}(J_{1a}) = \mp 1 \cap \mathrm{sgn}(J_{2b}) = \pm 1)$. Comparing with Eqs. (31) and (32), we see that in the stochastic case, the probabilities $P_{kk'}$ depend as in the preceding example on the areas on the angular momentum sphere occupied by the individual positions of each angular momentum compatible with the outcomes (although in the stochastic case there are many more such areas, each contributing with a given weight). Overall, Eq. (37) yields

$$\langle S_{1a} S_{2b} \rangle = \sum_{k,k'=-1/2}^{1/2} kk' P_{kk'} = \frac{1}{8}\left(\frac{\theta_b - \theta_a}{\pi} - 1 \right).$$

$C(a,a',b,b')$ is readily computed and is again, in line with Bell's theorem, bounded by 2.

IV. A DETECTION MODEL VIOLATING THE INEQUALITIES

The fourth example has similarities and differences with the models studied in Secs. III C and III D. A given detector on an axis measures the angular momentum's projection but only delivers the outcomes $\pm 1/2$. However, the outcomes depend on a probabilistic random interaction between the detected particle and the detector. This interaction has a specific property (it vanishes on average) that results in the introduction of an ensemble dependency. We will see that this feature combined with the conservation of the angular momentum between ensembles prevents the factorization that was seen above to be necessary in order to derive Bell's theorem.

A. Particle-detector interaction for a single particle

1. Basic properties

Let $\rho_1(\Omega_1)$ be the phase-space distribution for the single particle 1 and $R_{1a} = \pm 1/2$ denote the outcome obtained by placing a detector on the a axis. Let $P(R_{1a} = k, \rho_1)$ be the probability of obtaining the reading k on the detector if the statistical distribution of particle 1 (or equivalently, the distribution of $\mathbf{J}_1$) is known to be ρ_1. We will impose the following constraint on the interaction: the average $\langle J_{1a} \rangle_{\rho_1}$ over phase-space of the measured value is the one obtained by averaging over the measurement outcomes. This constraint takes the form

$$\langle R_{1a} \rangle_{\rho_1} = \sum_{k=-1/2}^{1/2} kP(R_{1a} = k, \rho_1) = \langle J_{1a} \rangle_{\rho_1}, \quad (41)$$

meaning that whereas individual outcomes depend on the interaction, on average the net effect of this interaction is zero. The models leading to Eq. (41) are not unique—any model verifying Eq. (41) and obeying $\Sigma_k P(R_{1a} = k, \rho_1) = 1$ will do. Depending on the specific model, Eq. (41) will not be verified for an arbitrarily chosen distribution ρ_1; only a class of distributions can be consistent within a given model. In the present model, we will assume as in the previous examples that ρ_1 can only be a uniform distribution occupying one (or both) of the two hemispheres of the angular momentum sphere.

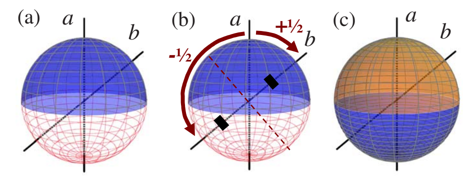

FIG. 3. (Color online). (a) The ensemble ρ_{1a+} for the single-particle model described in Sec. IV A. Any $\mathbf{J}_1$ in this ensemble has a positive projection J_{1a}; measuring R_{1a} gives the outcome $+1/2$ with certainty, without changing the ensemble, since ρ_{1a+} is symmetric relative to the a axis and $\langle J_{1a} \rangle_{\rho_{1a+}} = 1/2$. (b) In the same situation R_{1b} is measured. Now the symmetry axis of the ensemble ρ_{1a+} does not coincide with the b axis. Hence measuring R_{1b} can yield either $+1/2$ or $-1/2$ with probabilities proportional to $\langle H(\pm J_{1b}) J_{1b} \rangle_{\rho_{1a+}}$. The ensemble is modified during the measurement, undergoing a rotation toward the positive or negative b axis as indicated by the arrows. (c) The two-particle distribution after R_{1a} was measured and the outcome is known to have been $R_{1a} = -1/2$, in which case particle 1 is described by the ensemble ρ_{1a-}. Conservation of the angular momentum then requires that particle 2 be described by ρ_{2a+}, so that if R_{2a} were measured, the outcome $+1/2$ would be obtained with certainty [Eq. (58)]. If instead R_{2b} is measured, we have a single-particle problem for particle 2 identical to the one portrayed in Fig. 2(b).

Let us examine for such distributions the consequences of Eq. (41). Assume that ρ_1 corresponds to a uniform distribution of $\mathbf{J}_1$ on the positive hemisphere centered on the a axis, to be denoted ρ_{1a+} [Fig. 3(a)]. If a measurement is made along the b axis, a direct computation of $\langle J_{1b} \rangle_{\rho_{1a+}}$ gives

$$\langle R_{1b} \rangle_{\rho_{1a+}} = \sum_k kP(R_{1b} = k, \rho_{1a+}) = \frac{1}{2}\cos(\theta_b - \theta_a). \quad (42)$$

If one measures R_{1a} the average (42) becomes $+1/2$, i.e., the only positive detected outcome. Therefore since the probabilities are positive, we must have

$$P(R_{1a} = 1/2, \rho_{1a+}) = 1, \quad (43)$$

$$P(R_{1a} = -1/2, \rho_{1a+}) = 0. \quad (44)$$

Conversely if the distribution is ρ_{1a-} (uniform on the lower hemisphere) we obtain the opposite probabilities,

$$P(R_{1a} = 1/2, \rho_{1a-}) = 0, \quad (45)$$

$$P(R_{1a} = -1/2, \rho_{1a-}) = 1, \quad (46)$$

and

$$\langle R_{1b}\rangle_{\rho_{1a-}} = \sum_k kP(J_{1b}=k,\rho_{1a-}) = -\frac{1}{2}\cos(\theta_b-\theta_a). \quad (47)$$

Note that Eq. (42) along with the normalization of the probabilities uniquely determines the value of the probabilities,

$$P\left(R_{1b}=\pm\frac{1}{2},\rho_{1a+}\right) = \frac{\cos(\theta_b-\theta_a)\pm 1}{2}, \quad (48)$$

as well as the equality between the relative expectation value corresponding to positive (respectively, negative) outcomes R_{1b} and the average of the angular momentum projection over the regions where J_{1b} is positive (respectively, negative), i.e.,

$$\pm\frac{1}{2}P\left(R_{1b}=\pm\frac{1}{2},\rho_{1a+}\right) = \langle H(\pm J_{1b})J_{1b}\rangle_{\rho_{1a+}}, \quad (49)$$

H denoting the unit-step function.

The main property of this particle–detector interaction based model is that the detected result does not depend on a phase-space point (or on a given individual position of the particle's angular momentum on the sphere), be it through a deterministic value ascription or through well-defined probabilities. Indeed, if this were the case, then Eqs. (43)–(46) would imply that

$$R_{1a}(\Omega_1)=1/2 \Leftrightarrow J_{1a}>0 \quad \text{and} \quad R_{1a}(\Omega_1)=-1/2 \Leftrightarrow J_{1a}<0, \quad (50)$$

as in the example studied in Sec. III D. But then assume that R_{1b} is measured and the ensemble is known to be ρ_{1a+} (uniform distribution on the positive hemisphere centered on the a axis). On the angular momentum sphere ρ_{1a+} can be seen as being composed of the intersections with ρ_{1b+} and ρ_{1b-}, $\rho_{1a+}=(\rho_{1a+}\cap\rho_{1b+})\cup(\rho_{1a+}\cap\rho_{1b-})$. The respective integration domains are $\mathcal{D}(-\frac{1}{2},\frac{1}{2})$ and $\mathcal{D}(-\frac{1}{2},-\frac{1}{2})$ [we use the notation introduced in Sec. III C; see Fig. 2(b)]. Hence

$$\begin{aligned}\langle R_{1b}\rangle_{\rho_{1a+}} &= \int R_{1b}(\Omega_1)\rho_{1a+}(\Omega_1)d\Omega_1\\ &= \frac{1}{2}\int_{\mathcal{D}(-1/2,1/2)}\rho_{1b+}(\Omega_1)d\Omega_1\\ &\quad -\frac{1}{2}\int_{\mathcal{D}(-1/2,-1/2)}\rho_{1b-}(\Omega_1)d\Omega_1, \end{aligned} \quad (51)$$

yielding $[1-2(\theta_b-\theta_a)/\pi]/2$ in contradiction with the constraint (41) defining the model, $\langle J_{1b}\rangle_{\rho_{1a+}}=\cos(\theta_b-\theta_a)/2$.

We see therefore that the value ascription given by Eq. (50) *does not fit* with the main property of the model [23] The reason is that Eq. (41) introduces an ensemble dependency on the model: the probabilities do not depend on the phase-space position but on the ensemble, as if the particle's angular momentum effectively occupied an entire hemisphere (physically, this may happen, for example, if the particle follows a stochastic motion with its angular momentum constrained to remain in the ensemble, the time scale of the measurement being significantly larger than the time scale of the stochastic motion). Equation (50) should thus be replaced by

$$R_{1a}(\Omega_1)=\pm 1/2 \Leftrightarrow J_{1a}\gtrless 0 \ \text{ for } \textit{every } J_{1a}\in\rho_1. \quad (52)$$

2. Further considerations

Although this has no effect on the computations, it will be convenient, in order to provide a physical interpretation, to detail the consequences arising from the model. Equation (52) associates an outcome R_{1a} along an axis a with J_{1a} being of the same sign for every member of the hemispheric ensemble $\rho_{1a\pm}$ [see Fig. 3(a)]. Since $R_{1a}=\pm 1/2=\langle J_{1a}\rangle_{\rho_{1a\pm}}$, we can envisage that the random interaction occurring during a measurement effectively changes the distribution of the angular momentum: for instance, if initially the distribution is on a given hemisphere, say ρ_{1a+}, Eq. (52) is realized and $R_{1a}=1/2$ is obtained with certainty, reflecting $\langle J_{1a}\rangle_{\rho_{1a+}}$. If R_{1b} is measured, the final distribution is ρ_{1b+} (respectively, ρ_{1b-}) if the outcome $k=1/2$ (respectively, $-1/2$) is obtained [see Fig. 3(b)]. The outcome thus appears as the average value of the angular momentum projection in the post-measurement distribution and Eq. (42) can be written as

$$\langle R_{1b}\rangle_{\rho_{1a+}} = \sum_{k=\pm 1}\langle J_{1b}\rangle_{\rho_{1b(k)}}P(R_{1b}=k,\rho_{1a+}) = \langle J_{1b}\rangle_{\rho_{1a+}}. \quad (53)$$

Note that this implies that consecutive measurements involving projections along different axes *do not commute*: the condition (52) cannot be realized jointly along two different directions, like the classical analogs of the angular momenta eigenstates presented in Sec. II [24]. If the initial distribution is ρ_{1a+} measuring R_{1b} then $R_{1a'}$ entails that R_{1b} is measured over ρ_{1a+} but $R_{1a'}$ over $\rho_{1b\pm}$ depending on the outcome R_{1b}. In the reverse order, $R_{1a'}$ is measured first, the average being given by $\langle J_{1a'}\rangle_{\rho_{1a+}}$ and R_{1b} then involves the values of J_{1b} over one of the distributions $\rho_{1a'\pm}$. Equation (53) also allows us to compute the change in the angular momentum projection due the measurement,

$$\begin{aligned}\Delta\langle J_{1b}\rangle &\equiv \langle J_{1b}\rangle_{\rho_{1b(2k)}} - \langle J_{1b}\rangle_{\rho_{1a+}}\\ &= -2kP(R_{1b}=k,\rho_{1a+}), \quad k=\pm\frac{1}{2}. \end{aligned} \quad (54)$$

Consider now the uniform distribution on the entire sphere $\rho_{1\Sigma}$. It can first be envisaged as the angular momentum occupying the upper or lower hemispheres along a definite direction (say a) so that

$$\rho_{1\Sigma} = (\rho_{1a+}+\rho_{1a-})/2. \quad (55)$$

Since distributions in classical mechanics obey the principle of linear superposition, $\rho_{1\Sigma}$ can also be taken as a sum of the expressions given by Eq. (55) over different directions a. Alternatively the angle a in Eq. (55) can be taken to vary in time (then the measurement does not involve a change in the distribution but rather a selection of the angular momenta such that $J_{1a}>0$ or $J_{1a}<0$), or $\mathbf{J}$ can be distributed on the entire spherical surface (then the measurement induces a change in the distribution $\rho_{1\Sigma}\to\rho_{1a\pm}$). Only in these latter cases is the distribution spherically symmetric; all these possibilities lead to the same probabilities and average values, yielding $P(R_{1a}=\pm 1/2,\rho_{1_\Sigma})=1/2$ for *any* axis a as well as a vanishing average (41) as required.

B. Two-particle expectation

1. Distribution and conservation of the angular momentum

Before computing the two-particle averages and correlation functions, we make explicit the initial distribution and the conservation of the angular momentum for the model. We have seen that the defining property Eq. (41) implied that value ascription depended on distributions (taken to be ensembles on given hemispheres) and not on individual phase-space positions. The two-particle distribution given above by Eq. (6) is (i) spherically symmetric and (ii) anticorrelates the individual positions of the angular-momenta $\mathbf{J}_2=-\mathbf{J}_1$, so that we have $J_{2a}=-J_{1a}$ for projections along arbitrary axes a on the sphere. We require the extension of these two properties so that they hold over the initial distribution, to be denoted by ρ_Σ. We must thus have for any of the two particles i and axis direction a

$$\langle R_{ia}\rangle_{\rho_\Sigma}=\langle J_{ia}\rangle_{\rho_\Sigma}=0, \tag{56}$$

so that both outcomes $R_{ia}=\pm 1/2$ can be obtained with equal probability. The correlation between the outcomes for the two particles is obtained by applying Eq. (49) to $\rho_{\Sigma,}$ giving

$$\langle H(J_{2a})J_{2a}\rangle_{\rho_\Sigma}=\langle H(-J_{1a})J_{1a}\rangle_{\rho_\Sigma}. \tag{57}$$

By Eq. (53) we have $R_{1a}=\langle J_{1a}\rangle_{\rho_{1\pm a}}$ so that by way of Eq. (57) the anticorrelation $J_{2a}=-J_{1a}$ implies that the outcomes and the distributions for the particles along the same axis are anticorrelated,

$$\langle J_{2a}\rangle_{\rho_{2a\mp}}\equiv R_{2a}=-R_{1a}\equiv-\langle J_{1a}\rangle_{\rho_{1a\pm}}. \tag{58}$$

Equations (56)–(58) hold for any arbitrary axis a. The anticorrelation for the measurement outcomes, based on the conservation of the angular momentum over the ensembles, implies anticorrelations *between these ensembles*. Measuring R_{1a} links the outcome to one of the two ensembles $\rho_{1a\pm}$ depending on whether $R_{1a}=\pm 1/2$. In turn, this also fixes $\rho_2=\rho_{2a\mp}$. Note that contrary to the correlation between individual phase-space positions (for which one has $J_{2a}=-J_{1a}$ and $J_{2b}=-J_{1b}$ jointly for any axes a and b), Eq. (58) cannot hold jointly along several directions. This is a consequence of Eq. (52) not holding simultaneously along several axes.

There are different possibilities for choosing explicit realizations of ρ_Σ: all these possibilities lead to the same results and all hinge on the conservation of the *total* angular momentum along an arbitrary axis demanded by Eq. (58). For example, ρ_Σ can be taken as proportional to $\rho_{1b+}\rho_{2b-}+\rho_{1b-}\rho_{2b+}$. Equation (58) is then ensured provided the change in the angular momentum (54) after the first measurement is taken into account in the angular momentum balance for the other particle. Alternatively as in Eq. (55), b can be taken as varying in time, giving

$$\rho_\Sigma=\frac{1}{2}(\rho_{1b(t)+}\rho_{2b(t)-}+\rho_{1b(t)-}\rho_{2b(t)+}). \tag{59}$$

As for the single-particle distribution $\rho_{1\Sigma}$, measuring R_{1a} then selects the individual positions of J_{1a} such that $J_{1a}\gtrless 0$, correlated with the individual positions $J_{2a}\lessgtr 0$. Another possibility for ρ_Σ would be to take the distribution (6) and consider R_{1a} as inducing a change in the distribution $\rho_\Sigma\to\rho_{1a\pm}$.

2. Computation of the correlation

Since the measurement outcomes do not depend on the individual phase-space positions, the average $E(a,b)\equiv\langle R_{1a}R_{2b}\rangle_{\rho_\Sigma}$ cannot be obtained as in the preceding example from the phase-space averages (30) and (31), but from the probabilities of detecting a given outcome as a function of the distribution. $E(a,b)$ is computed from the general formula, also employed in Sec. III C,

$$\langle R_{1a}R_{2b}\rangle_{\rho_\Sigma}=\sum_{k,k'=-1/2}^{1/2}kk'P_{kk'}, \tag{60}$$

where as in Eq. (32) $P_{kk'}$ is given by

$$\begin{aligned}P_{kk'}&=P(R_{1a}=k\cap R_{2b}=k',\rho_\Sigma)\\&=P(R_{1a}=k)P(R_{2b}=k'|R_{1a}=k)\end{aligned} \tag{61}$$

and the two particle expectation takes the form

$$\langle R_{1a}R_{2b}\rangle_{\rho_\Sigma}=\sum_{k=-1/2}^{1/2}kP(R_{1a}=k)[\sum_{k'=-1/2}^{1/2}k'P(R_{2b}=k'|R_{1a}=k)]. \tag{62}$$

For any particle i and direction a, we have

$$P(R_{ia}=\pm 1/2,\rho_\Sigma)=1/2. \tag{63}$$

The conditional probability $P(R_{2b}=k'|R_{1a}=k)$ is, as in the example involving discrete outcomes studied above in Sec. III C, the probability of obtaining $R_{2b}=k'$ if it is known that $R_{1a}=k$. But we have just seen that obtaining an outcome $R_{1a}=k$ is linked to the respective densities $\rho_{1a[\mathrm{sgn}(k)]}$ and $\rho_{2a[\mathrm{sgn}(-k)]}$. The conditional probability is therefore given by

$$P(R_{2b}=k'|R_{1a}=k)=P(R_{2b}=k',\rho_{2a[\mathrm{sgn}(-k)]}), \tag{64}$$

which is a single-particle probability of the type given by Eq. (48). Note that in order to compute the expectation value, we do not need to know the values of these individual probabilities, as the knowledge of the conditional expectation—the expression between brackets in Eq. (62)—is sufficient. The two-particle conditional expectation is given by the single-particle average $\langle J_{2b}\rangle_{\rho_{2a[\mathrm{sign}(-k)]}}$ whose expression was determined above [Eqs. (41), (42), and (47)]. We can rewrite the average in the form

$$\sum_{k'=-1/2}^{1/2}k'P(R_{2b}=k'|R_{1a}=k)=-k\cos(\theta_b-\theta_a). \tag{65}$$

We can now compute the expectation $E(a,b)\equiv\langle R_{1a}R_{2b}\rangle$ from Eqs. (62) and (65). The result is easily seen to be

$$E(a,b) = -\frac{1}{4}\cos(\theta_b - \theta_a). \quad (66)$$

In the present derivation, we have assumed that the knowledge of particle 1's outcome was obtained first, hence the appearance of the conditional probability regarding the outcomes of particle 2. But obviously by Bayes' theorem the result is the same if we assume instead that R_{2b} is known first, and the conditional probability then concerns the computation of the outcomes of particle 1 [15].

The correlation function $C(a,b,a',b')$ is again given by Eq. (12) with $V_{\max}=1/2$. $C(a,b,a',b')$ violates the Bell inequality (14) for a wide range of values, the maximal violation being obtained for $C(0,\frac{\pi}{4},\frac{\pi}{2},\frac{3\pi}{4})=2\sqrt{2}$. This correlation function, with $E(a,b)$ given by Eq. (66), is familiar from quantum mechanics—it is precisely the correlation obtained for the two particles with spin 1/2 in the singlet state. It was shown in this case that Eq. (66) can be seen as a consequence of a particular correlation between vectors whose projection is conserved on average [15,16].

V. DISCUSSION

A. Ensemble dependence

We have seen in our fourth example (Sec. IV) that correlation functions obtained from two-particle distributions in classical mechanics can lead to a violation of the Bell inequalities, without nonlocality being explicitly involved (it may play an implicit role, see Sec. V C below). The main difference between this model and the other examples we have given consists in the ensemble dependencies: probabilities, average values, and conservation laws are relative to a collective property (a given distribution) and do not depend, as in the other cases, on the individual phase-space positions. Indeed, the constraint (41) cuts the link between a definite phase-space position of a particle and a given measurement outcome (be it in a probabilistic or deterministic way).

In this respect, it is noteworthy to compare the interpretation of the conditional probabilities appearing in examples 2 (Sec. III C) and 4 (Sec. IV). In both cases $P(V_{2b}=k'|V_{1a}=k)$ is grounded on the correlation (5) and represents the probability of obtaining $V_{2b}=k'$ given the knowledge that $V_{1a}=k$. In both cases the distribution of $\mathbf{J}_2$ is modified once the outcome $V_{1a}=k$ is known [25] (it changes from a uniform distribution on the sphere to a uniform distribution on the positive or negative hemisphere centered on a, depending on k). However, in example 2 the probabilities depend on the individual phase-space positions of the particles: although it may be unknown in practice, $\mathbf{J}_1$ has in principle a definite position that determines $V_{1a}=k$, and to this position corresponds the definite position $\mathbf{J}_2=-\mathbf{J}_1$ that will determine the outcome V_{2b}; so the conditional probability is computed by finding the individual positions of $\mathbf{J}_2$ such that $V_{2b}=k'$ compatible with the positions of $\mathbf{J}_1$ imposed by $V_{1a}=k$ (namely $J_{1a}>0$). In example 4 an outcome $V_{1a}=k$ cannot be linked in principle to an individual position of $\mathbf{J}_1$ and thus we can only infer from the outcome the *ensemble* to which $\mathbf{J}_1$ must belong; then from the conservation law we know the distribution for $\mathbf{J}_2$, which allows us to compute $P(V_{2b}=k'|V_{1a}=k)$ from the probability $P(R_{2b}=k',\rho_{2a[\mathrm{sgn}(-k)]})$. Hence we can only correlate observable outcomes with ensembles, not with individual positions of the angular momenta. Assuming that a given phase-space position determines probabilities, as in the stochastic model of Sec. III D, only brings in several combinations of possible outcomes allowed by the definite positions of $\mathbf{J}_1$ and $\mathbf{J}_2=-\mathbf{J}_1$ on the angular momentum sphere, but still allows us to correlate these individual positions with measurement outcomes.

B. Joint distributions and noncommutative measurements

We had remarked in Sec. III B that the existence of a joint probability $P_{aba'b'}$ is sufficient to ensure that a Bell-type inequality holds, irrespective of whether the assumption that measurement outcomes and probabilities depend on the individual phase–space positions is made. But if that specific assumption is made, then one is led to the existence of a joint probability by imposing the factorization (20). Along these lines, there are two ways of seeing why Bell's theorem does not apply to our fourth example.

First, the ensemble dependence can formally be thought of as arising from elementary phase-space probability functions *specific* to a given ensemble, i.e.,

$$P(R_{1a}=k,\rho_1) = \int p(\Omega_1;\rho_1)\rho_1(\Omega_1)d\Omega_1 \quad (67)$$

[compare with Eq. (21)]. By employing expressions such as Eq. (67) in the expectation value as given by Eqs. (60) and (61), it can be seen directly that the ensemble dependence of the elementary probabilities spoils the factorization (20)—for example, instead of $p(\Omega_2)$, one has outcome dependent expressions such as $p(\Omega_2;\rho_2(R_{1a}))$ [15].

The second manner starts with the remark made above concerning the noncommutation of the R measurements introduced in our model; in classical mechanics, measurements usually commute, but this is not the case if they arise from collective phenomena (encapsulated in the ensemble dependency). By requiring that the angular momentum be conserved between ensembles (just as it is when individual positions are considered), the consequences of the noncommutation are carried over from one particle to the other. As seen in Sec. IV A 2 in the single-particle case, the probabilities and outcomes when R_{1b} is measured after a first measurement is made will be different depending on whether R_{1a} or $R_{1a'}$ was measured first. Because Eq. (58) links the outcomes with the ensembles, this is also the case in the two-particle problem when R_{2b} (or $R_{2b'}$) is determined after R_{1a} *or* $R_{1a'}$ were measured. Put differently, although Eq. (58) holds for the a, a', b, and b' axes, it cannot hold jointly for all the axes because the single-particle ensembles $\rho_{1a\pm}$ and $\rho_{1a'\pm}$ are mutually exclusive, as well as $\rho_{2b\pm}$ and $\rho_{2b'\pm}$ (see Secs. II and IV A 2). Hence a joint probability $P_{aba'b'}$ cannot be defined and the model is not constrained by the inequality (28). This is consistent with the equivalence [13,17] shown between the verification of Bell's inequality and the commutation of the four observables entering Eq. (19). The ensemble dependence introduced in our model appears as a tool in order to enforce, in a classical context, the noncommutation of the measurements along different axes made on the same particle.

C. Conservation laws and locality

Factorization, enforcing the existence of joint distributions, and as such a necessary assumption in the derivation of Bell's theorem, is usually argued to be intimately linked to locality. According to Bell [18], factorization is a *consequence* of local causality, given that spacelike separated events can only have common causes in their backward light cone: therefore the probability of obtaining a certain result in an event regarding one of the particles cannot depend on what has been measured on the other. It is known, however, that factorization can be seen [19,20] as a consequence of two separate conditions, outcome independence (the conditional probability of one event does not depend on the outcome obtained in the other event) and parameter independence (dependence on the measurement direction of the other event). Only the violation of outcome independence can result in a genuine violation of local causality (it would permit superluminal signaling), whereas the violation of parameter independence allows a "peaceful coexistence" [20] between local causality and other types of correlations preventing the factorization.

The present model—like typical quantum-mechanical entangled systems—respects parameter independence [Eq. (63)] but violates outcome independence [Eq. (64)] (the dependencies here must be understood relative to the ensembles and not relative to the individual positions of the angular momenta). This outcome dependence of the conditional probabilities is due to the conservation of the angular momentum, as encapsulated by $\mathbf{J}_2=-\mathbf{J}_1$ (anticorrelation between individual positions), Eq. (57) (correlation between ensembles occupying opposite hemispheres centered on the same arbitrary axis) and Eq. (58) (anticorrelation between the outcomes made on the same arbitrary axis). Parameter independence on the other hand guarantees that the predictions relative to R_{1a} do not depend on what measurement or whether a measurement is carried out on particle 2, and vice versa [Eq. (63)]. It is clear, nevertheless, that the angular momentum conservation affects the distributions of both particles. For example, if ρ_Σ is given by Eq. (59), made up from rotating distributions, measuring R_{1a} not only freezes the rotation of particle 1's distribution, but also that of particle 2 (precisely because the angular momenta are correlated and need to be conserved). If instead ρ_Σ is taken as a uniform distribution on the sphere for the individually anticorrelated angular momenta, measuring R_{1a} changes the distribution $\rho_{1\Sigma}\to\rho_{1a\pm}$ but also $\rho_{2\Sigma}\to\rho_{2a\mp}$. Hence it can be argued that the conservation of the angular momentum as implemented in our model actually results from an implicit implementation of nonlocality. There are several answers to this question, depending on the status one gives to conservation laws and ensemble distributions, or on how nonlocality or causality are defined. The three following positions can be singled out:

(i) The changes in the distributions are real physical effects, but the conservation of the angular momentum results from a symmetry that is intrinsically linked to space-time. Indeed the correlation (58) arises by generalizing the angular momentum conservation for individual positions to ensembles accounting for noncommutative measurements. There is no need to invoke a specific mechanism for a conservation law—conservation laws and symmetry principles are just postulated. But if desired, a field can be can be ascribed the role of transporting the angular momentum; in this respect, it may be useful to make the analogy with Feynman's paradox in which mechanical angular momentum is transmitted between two charged particles through the electromagnetic field [21].

(ii) The changes in the distributions are real physical effects due to a nonlocal form of causation. The requirement given by Eq. (58) is sufficient to imply nonlocality. Action at a distance effects are quite common in nonrelativistic classical mechanics, although the modern view is to ascribe such effects (like gravity or several phenomena in electrostatics) to the action of fields. Here the nonlocal effect would consist in accounting for angular momentum conservation. This does not necessarily contradict the preceding position since it can be argued that symmetries can give rise to nonlocality, a position leading to a holistic vision of symmetries as holding beyond a space-time framework.

(iii) The changes in the distributions are not physical effects: one must distinguish the observed frequencies (which are measured) from the calculus of probabilities (whose role is to make logical inferences given a certain state of information [22]). Conditional probabilities do not therefore express causation and the factorization of the probabilities does not follow from the requirement of local causality. In Bell's term, the variables entering the probabilities are not beables, an argument that may be supported by the fact that individual angular momenta positions do not ascribe values and that the status of ensembles as beables is questionable; moreover, these correlations cannot be employed to communicate (no signaling guaranteed by parameter independence). The ensembles and their correlations are theoretical constructs encapsulating the state of knowledge we have of the situation, including the constraints (like conservation laws).

VI. CONCLUSION

To summarize, we have first constructed classical distributions analogs of the quantum-mechanical angular momentum eigenstates; these classical distributions are characterized by being mutually exclusive, leading, with appropriate assumptions to noncommutative measurements. We have then derived Bell's theorem in the deterministic and stochastic cases; both cases are characterized by the fact that an individual position of the angular momentum ascribes a value (with certainty or with a given probability) to a measurement of the projection along any axis. As a result, a joint probability distribution for an arbitrary number of events can be defined. Three different examples of Bell-type models were studied. A non-Bell-type model was introduced in Sec. IV: in this model, individual positions of the angular momenta are irrelevant to determine the measurement outcomes, that only depend on ensembles. As a result, single-particle measurements do not commute, and a distribution for joint measurements along different axes cannot be defined. If it is further assumed that the total angular momentum must be conserved, the Bell inequalities are violated.

The present results do not disprove Bell's theorem—as we have seen, in these circumstances the assumptions made in the derivation of the theorem are not fulfilled. We have argued that the violation of the inequalities in our classical model is due to the conservation of the total angular momentum in the context of noncommutative measurements; nonlocality does not need to be invoked (although it may). From this perspective, the violation of the Bell inequalities would not necessarily constitute a marker of nonlocality. It still remains to be investigated what type of collective or individual phenomena are compatible with the type of model introduced in this work.

[1] J. S. Bell, Physics **1**, 195 (1964). Reprinted in [2].
[2] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, England, 2004).
[3] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[4] H. P. Stapp, Am. J. Phys. **65**, 300 (1997).
[5] I. C. Percival, Phys. Lett. A **279**, 105 (2001).
[6] N. D. Mermin, Found. Phys. **29**, 571 (1999).
[7] W. Unruh, Phys. Rev. A **59**, 126 (1999).
[8] A. Fine, in *From Physics to Philosophy*, edited by J. Butterfield and C. Pagonis (Cambridge University Press, Cambridge, England, 1999), p. 1.
[9] L. Accardi and M. Regoli, in *QP-XIII, Foundations of Probability and Physics*, edited by A. Khrennikov (World Scientific, Singapore, 2002), p. 1.
[10] Y. F. Orlov, Phys. Rev. A **65**, 042106 (2002).
[11] J. Christian, e-print arXiv:quant-ph/0703179.
[12] J. S. Bell, J. Phys. (Paris), Colloq. **42**, 41 (1981). Reprinted in [2].
[13] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[14] J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).
[15] A. Matzkin, J. Phys. A **41**, 085303 (2008).
[16] C. S. Unnikrishnan, Europhys. Lett. **69**, 489 (2005).
[17] J. D. Malley, Phys. Rev. A **69**, 022118 (2004).
[18] J. S. Bell, in *Between Science and Technology*, edited by A. Sarlemijn and P. Kroes (Elsevier, New York, 1990), reprinted in [2], Chap. 24.
[19] J. P. Jarrett, Nous **18**, 569 (1984).
[20] A. Shimony, *Search for a Naturalistic World View* (Cambridge University Press, Cambridge, England, 1993), Vol. II, Part B.
[21] J. M. Aguirregabiria, A. Hernandez, and M. Rivas, Am. J. Phys. **58**, 635 (1990).
[22] E. T. Jaynes, in *Maximum-Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer, Dordrecht, 1989), p. 1.
[23] Since value ascriptions given by $V(\Omega)$ or $p(V,\Omega)$ are characteristic of Bell models, it could be said that even for a single particle, Bell-type models are inconsistent with the present model. On the other hand, it could be argued that the ensemble ρ_1 should be taken as the "hidden variable," given that this is the variable ascribing values to the outcomes and probabilities, even though ρ_1 may not qualify following Bell's terminology as being a beable (see Sec. V).
[24] We have chosen distributions on hemispheres rather than the ring distributions of Sec. II for continuity with the examples investigated in Sec. III; the model studied here would also hold if ringlike distributions were employed.
[25] It seems it is necessary to stress that the change in the probability distribution of particle 2 when the outcome of particle 1 is known is not a physical phenomenon involving action at a distance, but the result of the information brought by the knowledge of the first outcome, given the conservation law. This point, unrelated to Bell's theorem, is an elementary inference common in the calculus of probabilities.