State-space dimensionality in short-memory hidden-variable theories

Alberto Montina
Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada
(Received 7 October 2010; published 10 March 2011)

Recently we have presented a hidden-variable model of measurements for a qubit where the hidden-variable state-space dimension is one-half the quantum-state manifold dimension. The absence of a short memory (Markov) dynamics is the price paid for this dimensional reduction. The conflict between having the Markov property and achieving the dimensional reduction was proved by Montina [A. Montina, Phys. Rev. A **77**, 022104 (2008)] using an additional hypothesis of trajectory relaxation. Here we analyze in more detail this hypothesis introducing the concept of invertible process and report a proof that makes clearer the role played by the topology of the hidden-variable space. This is accomplished by requiring suitable properties of regularity of the conditional probability governing the dynamics. In the case of minimal dimension the set of continuous hidden variables is identified with an object living an N-dimensional Hilbert space whose dynamics is described by the Schrödinger equation. A method for generating the economical non-Markovian model for the qubit is also presented.

DOI: 10.1103/PhysRevA.83.032107 PACS number(s): 03.65.Ta, 03.67.−a

I. INTRODUCTION

One of the most peculiar features of quantum mechanics is the exponential growth of resources required to define the quantum state $|\psi\rangle$ of a composite system. It makes the direct simulation of even a handful of particles impossible in practice. This growth is due to the fact that $|\psi\rangle$ contains the full statistical information about the probabilities of any possible event, such as the joint probability $\rho(s_1,\ldots,s_N)$ of obtaining the outcomes $\{s_1,\ldots,s_N\}$ by measuring the z-axis components of N 1/2-spins. The information in $\rho(s_1,\ldots,s_N)$ grows exponentially for a given accuracy and considerably exceeds the classical information required to specify the actual measurement outcome. Since the quantum state is not a physically accessible observable, but can be statistically reconstructed only by performing many measurements on different replicas [1], it is natural to wonder if this resource excess is strictly necessary to describe the actual state of a single realization. The quantum probabilities could be reproduced by a hidden-variable theory where a single system carries less information than the quantum state. In such a theory, the quantum state $|\psi\rangle$ is mapped to a probability distribution on a space X of hidden-variable states, that is,

$$|\psi\rangle \to \rho(X|\psi). \tag{1}$$

It is clear that the sampling space X can have, in principle, a smaller dimension than the quantum-state manifold. For example, the space of functions on a one-dimensional domain is infinite dimensional and any finite-dimensional Hilbert space can be embedded within it.

In accordance with recent terminology, we will refer to the actual state X of a quantum system and the corresponding space as *ontic state* and *ontological space* [2], respectively, and name the dimensional reduction of the ontological space *ontological shrinking*. It is interesting to note that, in any known short memory (Markov) hidden-variable theory, the dimension of the state space is never smaller than the quantum-state manifold dimension. As an example, in the de-Broglie–Bohm model the wave function has the role of a field of physical quantities and is supplied by additional variables describing the particle positions.

The ontological shrinking has a connection with the concept of classical "weak simulation" [3,4] in quantum information theory. In a classical "strong simulation" of a quantum computer, the goal is to evaluate the measurement probabilities with high accuracy. This requirement is stronger than necessary since in a real quantum computer a single run does not give the measurement probabilities and the output is a precise event. The probabilities concern the behavior of many experimental realizations. The goal of a classical "weak simulation" is not to compute the probability weights, but the outcomes in accordance with the weights. There are examples of quantum circuits that cannot be efficiently simulated in the strong way, but whose weak simulation is nevertheless tractable [4]. In a hidden-variable theory with reduced sampling space, the evaluation of the actual dynamics of a single realization would require less resources than the computation of the quantum-state dynamics. Thus, the ontological shrinking could offer, in a natural way, an efficient method of "weak simulation" of quantum computers.

In recent years the possibility of a statistical representation of quantum states on a reduced sampling space was discussed by various authors [5–10]. The problem of the smallest dimension of the ontological space was posed in Ref. [6]. It was subsequently proved that the ontological dimension cannot be smaller than the quantum-state manifold dimension in the case of a Markov hidden-variable theory with an additional hypothesis of trajectory relaxation [7]. We will refer to this result as the *no-shrinking theorem*. Recently we reported an example of hidden-variable model of measurements for a qubit whose state space is one-dimensional (i.e., smaller than the two-dimensional Bloch sphere [10]). As a consequence of the dimensional reduction, the dynamics is not a Markov process. This counterexample makes evident that the short memory hypothesis is strictly necessary for the proof of the theorem in Ref. [7]. In this article, we review the one-dimensional model providing a method for generating it and analyze in more detail the hypothesis of trajectory relaxation. We define the concept of *invertible process* and show that it is always possible to find a subregion of a compact ontological space where all the processes are invertible. Discarding insignificant transient states and considering only invertible processes, we present a different version of the *no-shrinking* theorem that makes clearer the role played by the topology of the hidden-variable space. This is accomplished by requiring reasonable properties of regularity of the conditional probability governing the dynamics and explicitly using them in the theorem proof. It is useful to remind that the dimension is a topological property and is not defined by the space cardinality.

In Sec. II we introduce the general framework of a hidden-variable theory. In Sec. III the economical ontological model in Ref. [10] is reviewed. The properties of regularity and the concept of invertible process are introduced in Sec. IV, where we also prove the no-shrinking theorem and discuss its consequences in terms of resource cost. In the same section we show that in any Markov hidden-variable theory with minimal space dimension the set of continuous hidden variables can be identified with an object, living in an N-dimensional Hilbert space, whose dynamics is described by the Schrödinger equation. In the Appendix we report a systematic construction method to generate the model in Sec. III, starting from a particular form of the probability distribution.

II. GENERAL FRAMEWORK

In a general hidden-variable theory the quantum state is translated into a classical language by replacing it with a probability distribution on a sampling space X of ontic states. We assume that the ontological space is an M-dimensional manifold described by an M-tuple $\vec{x}$ of continuous variables and a possible discrete index n. The mapping (1) is not the most general since the probability distribution could depend on the preparation context. For example, a 1/2-spin can be prepared in the *up* state by merely selecting the beam outgoing from a Stern-Gerlach apparatus or can be prepared in the state $|\uparrow\rangle + |\downarrow\rangle$ and then suitably rotated. To account for this possible dependence, we add suitable parameters η that identify the preparation context [11],

$$|\psi\rangle \to \rho(\vec{x},n|\psi,\eta). \tag{2}$$

The set of ontological variables $X = (\vec{x},n)$ contains the whole information about a single realization, thus the probability of any event is conditioned only by it. In particular, for the measurement of the trace-one projector $|\phi\rangle\langle\phi|$, it is assumed that there exists a conditional probability P for the event $|\phi\rangle$ given the ontic state X. In general, P also could depend on the context of the measurement, thus we introduce additional parameters τ in the conditional probability,

$$|\phi\rangle \to P(\phi|\vec{x},n,\tau). \tag{3}$$

The probability of the event $|\phi\rangle$ given $|\psi\rangle$ has to be equal to the quantum mechanical probability

$$\sum_n \int d^M x\, P(\phi|\vec{x},n,\tau)\rho(\vec{x},n|\psi,\eta) = |\langle\phi|\psi\rangle|^2. \tag{4}$$

Finally, it is assumed that the dynamics at the ontological level is Markovian [12]. The ontic state evolves from X_1 at time t_1 to X_2 at time t_2 according to a conditional probability $K(X_2,t_2|X_1,t_1)$. K satisfies the Chapman-Kolmogorov equation [12] and is a δ distribution for $t_2 = t_1$

$$K(\vec{x}_2,n_2,t_1|\vec{x}_1,n_1,t_1) = \delta_{n_1,n_2}\delta(\vec{x}_2 - \vec{x}_1). \tag{5}$$

For a time-homogeneous process the transition probability depends only on the time difference $t_2 - t_1$

$$K(X_2,t_2|X_1,t_1) = K(X_2,t_2 - t_1|X_1,0) \equiv K(X_2|X_1,t_2 - t_1). \tag{6}$$

To link the quantum language with the classical one, we can label K with the corresponding unitary operator $\hat{U}$. As for the state preparation and measurement, in general, the conditional probability can depend on additional parameters χ, the transformation context,

$$\hat{U} \to K_{\hat{U},\chi}(X_2|X_1). \tag{7}$$

Indeed, an operator $\hat{U}$ can be physically implemented in different ways. For example, the spin rotation $e^{-it\hat{\sigma}_x}$ can be performed directly rotating along the x axis or implementing the three-step rotation $e^{-i\frac{\pi}{4}\hat{\sigma}_y}e^{-it\hat{\sigma}_z}e^{i\frac{\pi}{4}\hat{\sigma}_y}$ along the z and y axes. These two schemes are physically different and not necessarily described by the same conditional probability. It is useful to note that $K_{\hat{U},\chi}(X|\bar{X})$ for $\hat{U} = \mathbb{1}$ is not necessarily the δ distribution $\delta(X - \bar{X})$ for all the contexts. For example, the identity evolution can correspond to physically performing the three-step rotation $e^{-i\frac{\pi}{4}\hat{\sigma}_y}e^{-it\hat{\sigma}_z}e^{i\frac{\pi}{4}\hat{\sigma}_y}$, followed by the inverse transformation $e^{it\hat{\sigma}_x}$. The overall operation is not equal to doing nothing and does not necessarily correspond to a δ-peaked conditional probability.

If the quantum state $|\psi\rangle$ evolves to $\hat{U}|\psi\rangle \equiv |\bar{\psi}\rangle$, the associated probability $\rho(X|\psi,\eta)$ evolves to

$$\rho(X|\bar{\psi},\bar{\eta}) \equiv \int dY K_{\hat{U},\chi}(X|Y)\rho(Y|\psi,\eta). \tag{8}$$

Some regularity properties of $K_{\hat{U},\chi}$ will be discussed in Sec. IV. Any short memory hidden-variable theory has this general structure.

III. ECONOMICAL ONTOLOGICAL MODEL OF MEASUREMENTS

In this section we show that if the Markov condition is not assumed, then the ontological space can be smaller than the quantum-state manifold. This goal is achieved by explicitly providing a one-dimensional hidden-variable model for a qubit [10]. Its systematic construction is discussed in the Appendix.

For the moment, we introduce a model working only for a subset of preparation states. The extension to the whole quantum-state manifold will be discussed later on. The ontological space is given by a continuous real variable x and a discrete index n that takes the two values 0 and 1. It is convenient to represent the quantum state $|\psi\rangle$ and the event $|\phi\rangle$ by means of the Bloch vectors $\vec{v} \equiv \langle\psi|\vec{\sigma}|\psi\rangle$ and $\vec{w} \equiv \langle\phi|\vec{\sigma}|\phi\rangle$, where $\vec{\sigma} = (\hat{\sigma}_x,\hat{\sigma}_y,\hat{\sigma}_z)$, $\hat{\sigma}_i$ being the Pauli matrices.

The probability distribution associated with the state $\vec{v}$ is

$$\rho(x,n|\vec{v}) = \sin\theta\,\delta_{n,0}\delta(x - \varphi) + (1 - \sin\theta)\delta_{n,1}\delta(x - \theta), \tag{9}$$

where θ and φ are, respectively, the zenith and azimuth angles in the spherical coordinate system

$$v_x = \sin\theta\cos\varphi, \quad v_y = \sin\theta\sin\varphi, \quad v_z = \cos\theta. \tag{10}$$

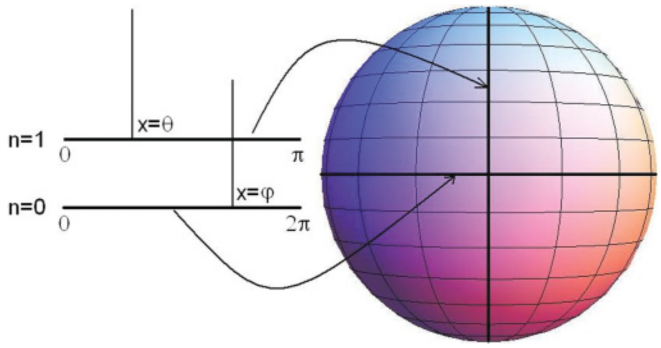

FIG. 1. (Color online) One-dimensional ontological space $\{x,n\}$ at left. According to Eq. (9), each point of the Bloch sphere is associated with a probability distribution on $\{x,n\}$ having two δ peaks.

Thus, when the quantum state $\vec{v}$ is prepared, the index n takes the value 0 or 1 with probability $\sin\theta$ or $1-\sin\theta$ and the continuous variable is equal to the zenith or azimuth angle according to the value of n (see Fig. 1).

The conditional probability $P(\vec{w}|x,n)$ for an event $\vec{w}$ with $w_z>0$ is defined as follows

$$P(\vec{w}|x,0)=1+\frac{w_x\cos x+w_y\sin x-\sqrt{1-w_z^2}}{2}, \quad (11)$$

$$P(\vec{w}|x,1)=1+\frac{\sqrt{1-w_z^2}\sin x+w_z\cos x-1}{2(1-\sin x)}. \quad (12)$$

The events $\vec{w}$ with $w_z<0$ correspond simply to the nonoccurrence of the events $-\vec{w}$ with $w_z>0$ [i.e, $P(-\vec{w}|x,n)=1-P(\vec{w}|x,n)$].

It is easy to prove that these probability functions fulfill the condition (4), that is,

$$P(\vec{w}|\varphi,0)\sin\theta+P(\vec{w}|\theta,1)(1-\sin\theta)=\tfrac{1}{2}(1+\vec{w}\vec{v}). \quad (13)$$

As shown in Ref. [10], the conditional probabilities $P(\vec{w}|x,0)$ and $P(\vec{w}|x,1)$ are always smaller than or equal to 1, but $P(\vec{w}|x,1)$ is positive only if $\theta<\theta_0\equiv\arccos(\frac{3}{5})\simeq 53.13^\circ$.

Thus, the model in this form works only for a set of prepared states whose Bloch vector lies inside a cone with aperture $2\theta_0$, the z axis being the symmetry axis. Since the positive region has nonzero measure in the quantum-state manifold, it is possible to extend the model to the whole Bloch sphere by covering the manifold with a sufficiently large number of patch regions with different coordinate systems and enriching the ontic state with a finite quantity of information. This was accomplished in Ref. [10] by adding a discrete index m taking 12 possible values, labeling 12 different regions of the Bloch sphere. See the referred paper for further details.

This is a concrete example of ontological shrinking, where the ontic state space is smaller than the quantum-state manifold, which is, in this case, the two-dimensional Bloch sphere. It is a remarkable fact that a single realization of the ontic state $\{x,n\}$ contains less information than the quantum state. The whole information on $|\psi\rangle$ is contained in the probability distribution ρ.

The absence of a short memory description of the dynamics is the price paid for the dimensional reduction. Indeed it is impossible to associate a positive conditional probability with each unitary evolution, that is, it is impossible to satisfy the identity

$$\rho(x,n|\hat{U}\psi)=\sum_{\bar{n}}\int d\bar{x}\,K_{\hat{U}}(x,n|\bar{x},\bar{n})\rho(\bar{x},\bar{n}|\psi), \quad (14)$$

with the probability distribution in Eq. (9), apart from the unitary evolution $e^{it\hat{\sigma}_3}$. Indeed, in the other cases the dynamical equation of $\rho(x,n|\psi)$ is, in general, nonlinear. Let us consider the unitary evolution with the Pauli matrix $\hat{\sigma}_y$ as generator. The dynamical equations of the Bloch vector are

$$\frac{\partial v_x}{\partial t}=v_z,\quad \frac{\partial v_y}{\partial t}=0,\quad \frac{\partial v_z}{\partial t}=-v_x, \quad (15)$$

which correspond in spherical coordinates to

$$\frac{\partial\varphi}{\partial t}=-\cot\theta\sin\varphi,\quad \frac{\partial\theta}{\partial t}=\cos\varphi. \quad (16)$$

Let $K(x,n|\bar{x},\bar{n},t)$ be the transition probability associated with the evolution in Eq. (15), we have from Eq. (14) that

$$\frac{\partial\rho(x,n,t)}{\partial t}=\sum_{\bar{n}}\int d\bar{x}\frac{\partial K}{\partial t}(x,n|\bar{x},\bar{n},t)\rho(\bar{x},\bar{n}), \quad (17)$$

which becomes by means of Eq. (9)

$$\begin{aligned}&\sin\theta\delta_{n,0}\frac{\partial\delta(x-\varphi)}{\partial\varphi}\frac{\partial\varphi}{\partial t}+(1-\sin\theta)\delta_{n,1}\frac{\partial\delta(x-\theta)}{\partial\theta}\frac{\partial\theta}{\partial t}\\&\quad+\cos\theta[\delta_{n,0}\delta(x-\varphi)-\delta_{n,1}\delta(x-\theta)]\frac{\partial\theta}{\partial t}\\&=\frac{\partial K}{\partial t}(x,n|\varphi,0,t)\sin\theta+\frac{\partial K}{\partial t}(x,n|\theta,1,t)(1-\sin\theta). \quad (18)\end{aligned}$$

In particular, for $n=0$ and using Eq. (16)

$$\begin{aligned}&-\cos\theta\frac{\partial\delta(x-\varphi)}{\partial\varphi}\sin\varphi+\cos\theta\delta(x-\varphi)\cos\varphi\\&=\frac{\partial K}{\partial t}(x,0|\varphi,0,t)\sin\theta+\frac{\partial K}{\partial t}(x,0|\theta,1,t)(1-\sin\theta). \quad (19)\end{aligned}$$

Dividing both sides by $\sin\theta$ and differentiating with respect to θ, we obtain that

$$\begin{aligned}&\frac{\partial}{\partial\theta}\left[-\cot\theta\frac{\partial\delta(x-\varphi)}{\partial\varphi}\sin\varphi+\cot\theta\delta(x-\varphi)\cos\varphi\right]\\&\quad=\frac{\partial}{\partial\theta}\left[\frac{\partial K}{\partial t}(x,0|\theta,1,t)\frac{1-\sin\theta}{\sin\theta}\right], \quad (20)\end{aligned}$$

This equation is not satisfied by any function K since the left-hand side is a function of both θ and φ, whereas the right-hand side depends only on θ. Thus, the dynamical equation of the probability distribution (9) is non-Markovian.

IV. ONTOLOGICAL SHRINKING AND MARKOV PROCESSES

In this section we will prove that the ontological shrinking is in contradiction with a Markov dynamics. This purpose is achieved by means of a very reasonable hypothesis concerning the support of the conditional probability $K(X_2|X_1,t)$. It will be introduced in the following section. In Sec. IV B we will introduce the concept of invertible process and will show that it is always possible to find a subregion of a compact ontological space where all the processes are invertible. Then we prove the no-shrinking theorem in Sec. IV C. In Sec. IV D, it is shown that in the case of minimal ontological dimension it is possible to identify the set of continuous ontological variables with a vector of the Hilbert space whose dynamics is given by the Schrödinger equation. In the last section we discuss the consequences of the no-shrinking theorem in terms of resource cost.

A. Conditional probabilities and associated unitary operators

For the sake of simplicity, from now on we will omit without loss of generality the discrete index n in the definition of the ontic state and assume that the ontological space is a differentiable manifold whose points are identified by the vector $\vec{x}$. When the local Euclidean structure is not required, we will use the more generic symbol X to indicate the ontic state.

In order that a transition probability describing a Markov process physically makes sense, it has to satisfy some conditions of regularity. For our purpose, it is sufficient to require a very weak condition. As a reasonable hypothesis, we assume that there exists a subset of the support of the conditional probability $K(X_2|X_1,t)$ that changes smoothly with respect to the condition X_1 and the evolution time t. Let us refine this statement in a more precise way.

Property 1. Given a time-homogeneous process $X_1 \to X_2$ with nonzero probability $K(\vec{x}_2|\vec{x}_1,t)$, there exist an $M \times M$ matrix $\hat{\lambda}$ and a vector $\vec{\alpha}$ such that $K(\vec{x}_2 + \hat{\lambda}\delta\vec{x}_1 + \vec{\alpha}\delta t|\vec{x}_1 + \delta\vec{x}_1,t+\delta t) \neq 0$ for any infinitesimal variation $\delta\vec{x}_1$ and δt. The matrix $\hat{\lambda}$ and the vector $\vec{\alpha}$ are functions of t and the process $X_1 \to X_2$.

It is important to note that this property is fulfilled for the very large class of Markov processes that involve drift, diffusion, and jumps (they are discussed, for example, in Ref. [12]). Furthermore, the matrix $\hat{\lambda}$ and the vector $\vec{\alpha}$ are not necessarily unique. For example, in the case of a stochastic process any choice of $\hat{\lambda}$ and $\vec{\alpha}$ is suitable since the conditional probability is a multidimensional smooth function whose support is the whole manifold. For a pure deterministic process both $\hat{\lambda}$ and $\vec{\alpha}$ are unique. The latter gives the drift velocity of the ontic state. In the case of pure jumps with finite transition probability, one choice is $\vec{\alpha} = 0$ and is unique if the jump distance cannot be arbitrarily small. One can imagine more complicated cases that involve diffusion in some direction, drift in the other ones, and jumps, however, also in these situations the stated property is fulfilled for some $\hat{\lambda}$ and $\vec{\alpha}$.

Property 1 strictly depends on the topology of the ontological space since it involves its local Euclidean structure. Using this structure in the proof of the no-shrinking theorem is fundamental because the space dimension is not a property defined merely by the cardinality of the space.

The quantum unitary evolution is a continuous process and corresponds to a trajectory in a Lie group manifold with the Hamiltonian as generator. For a time-homogeneous process, the unitary operator of the evolution has the form $e^{-it\hat{H}}$, where $\hat{H}$ is the transformation generator and t is the evolution time. In general, it is not possible to directly implement every generator of the su(N) algebra by means of a purely time-homogeneous process. In practice, only a small set of evolutions (building blocks) can be directly generated by a physical process. The other evolutions are obtained by suitably concatenating the building blocks. Given any generator $\hat{G}$, it will be possible, in principle, to experimentally implement a unitary evolution $\hat{U}_1$, a physically attainable time-homogeneous process $e^{-it\hat{H}}$, and another unitary evolution $\hat{U}_2$ such that

$$\hat{U}_2 e^{-it\hat{H}} \hat{U}_1 = e^{-it\hat{G}}. \tag{21}$$

The operators $\hat{U}_i$ can be constructed by suitably concatenating physically attainable processes. This allows us to associate with any unitary operator $e^{-it\hat{G}}$ a conditional probability $K(X_2|X_1,t)$ satisfying Property 1. Note that t in K does not correspond, in general, to the evolution time of a purely time-homogeneous process. It is the evolution time of only a part of the overall process involving also the transformations $\hat{U}_i$. This implies that for $t=0$ the conditional probability $K(X_2|X_1,t)$ is not necessarily a δ distribution. Furthermore, we can add a shift time t_0 to t and absorb the extra-term $e^{-it_0\hat{H}}$ in Eq. (21) into the operators $\hat{U}_1$ and $\hat{U}_2$. Thus, it is not necessary to require that t is a positive quantity, as in the case of stochastic Markov processes where the propagation kernels are elements of a semi-group. For our purpose, it is sufficient that $K(X_2|X_1,t)$ is defined in a neighborhood of $t=0$.

Let $\{\hat{G}_i\}$ with $i=1,\ldots,D$ be a set of $D \equiv N^2-1$ generators of the Lie algebra. The unitary operators $e^{-it\hat{G}_i}$ are associated with the conditional probabilities $K_i(X_2|X_1,t)$.

Any unitary evolution $\hat{U}$ in a region around the identity can be constructed in the following way

$$\hat{U}(t_1,\ldots,t_D) = e^{-it_1\hat{G}_1},\ldots,e^{-it_D\hat{G}_D} = \prod_{i=1}^{D} e^{-it_i\hat{G}_i}, \tag{22}$$

where the variables t_i parametrize the SU(N) manifold.

Let the linear operator

$$\rho(X) \to \int dY\, K_i(X|Y,t)\rho(Y),$$

be denoted by $K_i(t)$. The overall evolution $\hat{U}(\vec{t})$ is associated with the conditional probability

$$K(X_2|X_1,t_1,\ldots,t_D) = \left[\prod_{i=1}^{D} K_i(t_i)\right](X_2|X_1), \tag{23}$$

the product order is such that the sum index grows from left to right.

It is easy to prove a general property of the conditional probability $K(X_2|X_1,\vec{t})$.

Property 2. Suppose that $K(\vec{x}_2|\vec{x}_1,\vec{t})$ is different from zero for some process $X_1 \to X_2$, then there exists an $M \times D$ matrix $\hat{\eta}$ such that, for any small variation of the time parameters

$$\delta\vec{t} = \begin{pmatrix} \delta t_1 \\ \delta t_2 \\ \ldots \\ \delta t_D \end{pmatrix},$$

the conditional probability $K(\vec{x}_2 + \hat{\eta}\delta\vec{t}|\vec{x}_1,\vec{t}+\delta\vec{t})$ is different from zero. In the following, $\hat{\eta}$ will be called the *shift matrix* of $K(X_2|X_1,\vec{t})$. As with the matrix $\hat{\lambda}$ and the vector $\vec{\alpha}$, it is a function of t and the process $X_1 \to X_2$.

Proof. First, we consider the concatenation K_{12} of two conditional probabilities, that is,

$$K_{12}(X_2|X_1,t_1,t_2) \equiv \int dZ\, K_1(X_2|Z,t_1)K_2(Z|X_1,t_2). \tag{24}$$

Let $X_1 \to X_2$ be a process with nonzero transition probability $K_{12}(X_2|X_1,t_1,t_2)$. There exists a state Z such that $K_1(X_2|Z,t_1)$ and $K_2(Z|X_1,t_2)$ are different from zero. Property 1 implies that there exist two vectors $\vec{\alpha}_1$ and $\vec{\alpha}_2$ and a matrix $\hat{\lambda}_1$ such that

$$K_1(\vec{y} + \hat{\lambda}_1\vec{\alpha}_2\delta t_2 + \vec{\alpha}_1\delta t_1|\vec{z} + \vec{\alpha}_2\delta t_2, t_1 + \delta t_1) \neq 0,$$
$$K_2(\vec{z} + \vec{\alpha}_2\delta t_2|\vec{x}, t_2 + \delta t_2) \neq 0.$$

Thus, the conditional probability $K_{12}(\vec{x}_2 + \hat{\eta}\delta\vec{t}|\vec{x}_1, t_1 + \delta t_1, t_2 + \delta t_2)$ is different from zero, where the columns of the $M \times 2$ matrix $\hat{\eta}$ are $\vec{\alpha}_1$ and $\hat{\lambda}_1\vec{\alpha}_2$. The property can be proved by induction for any concatenation of processes K_i.

There is an important direct consequence of this property.

Lemma 1. A process $X_1 \to X_2$ with nonzero probability $K(X_2|X_1,\vec{0})$ is associated with a D_s-dimensional manifold of unitary transformations $\hat{U}(\delta\vec{t})$, where $D_s \geqslant D - M$, M being the ontological space dimension. The manifold is identified by the M equations

$$\sum_{j=1}^{D} \eta_{ij}\delta t_j = 0, \quad \text{with} \quad i = 1, \ldots, M, \tag{25}$$

where $\hat{\eta}$ is the shift matrix of $K(X_2|X_1,\vec{0})$. In particular, $D_s = D - N_I$, $N_I(\leqslant M)$ being the number of independent equations in the constraints (25).

Proof. Because of Property 2, there exists an $M \times D$ matrix $\hat{\eta}$ such that $K(\vec{x}_2 + \hat{\eta}\delta\vec{t}|\vec{x}_1,\delta\vec{t})$ is different from zero [i.e, $\vec{x}_1 \to \vec{x}_2 + \hat{\eta}\delta\vec{t}$ is a process associated with the unitary evolutions $\hat{U}(\delta\vec{t})$]. In particular, the submanifold of unitary evolutions with $\hat{\eta}\delta\vec{t} = 0$ is associated with the process $X_1 \to X_2$. Its dimension D_s is equal to $D - N_I$, N_I being the number of independent equations in $\hat{\eta}\delta\vec{t} = 0$. Since $N_I \leqslant M$, we have $D_s \geqslant D - M$.

The actual value of D_s depends on the number N_I of independent constraints in the vectorial equation $\hat{\eta}\delta\vec{t} = 0$. For example, D_s is equal to $N^2 - 1$ for $\hat{\eta} = 0$, which is the case if $K_i(\vec{y}|\vec{x},t)$ are M-dimensional diffusive processes.

B. Set $\mathcal{S}(X)$ and its symmetry property

In an ontological model the quantum state is associated with a probability distribution $\rho(X|\psi,\eta)$ according to the mapping (2). It is useful to introduce the following definition of the set $\mathcal{S}(X)$.

Definition 1. A vector $|\psi\rangle$ of the Hilbert space is in $\mathcal{S}(X)$ if and only if there exists a context η such that the probability $\rho(X|\psi,\eta)$ is different from zero for the state X.

In other words, the set $\mathcal{S}(X)$ contains every quantum state that is compatible with the occurrence of the ontic state X. As discussed in Ref. [7], the set $\mathcal{S}$ cannot lose vectors along its evolution. More precisely, if $X \overset{\hat{U}}{\to} Y$ is a nonzero probability process associated with the unitary evolution $\hat{U}$, then $\hat{U}\mathcal{S}(X) \subseteq \mathcal{S}(Y)$. This is a direct consequence of the definition. Indeed, if $|\psi\rangle$ is in $\mathcal{S}(X)$, then there exists a probability distribution associated with $|\psi\rangle$ such that X is in its support. Since $X \to Y$ is a nonzero probability process, then Y is in the support of a probability distribution associated with the evolved quantum state $\hat{U}|\psi\rangle$, that is, $|\psi\rangle \in \mathcal{S}(X) \Rightarrow \hat{U}|\psi\rangle \in \mathcal{S}(Y)$. The opposite implication $|\psi\rangle \in \mathcal{S}(X) \Leftarrow \hat{U}|\psi\rangle \in \mathcal{S}(Y)$ is not trivially satisfied. Thus, the set $\mathcal{S}$ cannot lose vectors, but, in principle, it could grow acquiring vectors.

The opposite implication can be deduced by assuming that each process is invertible.

Definition 2. A nonzero probability process $X_1 \overset{\hat{U}_1}{\to} X_2$ is said to be *invertible* if there exists a unitary operator $\hat{U}_2$ such that $X_2 \overset{\hat{U}_2}{\to} X_1$ is a nonzero probability process.

Property 3. Every process is invertible.

Note that the operator $\hat{U}_2$ associated with the inverse process is not required to be necessarily the inverse of $\hat{U}_1$. Property 3 is very reasonable and is satisfied by any known hidden-variable theory. Indeed, a state connected to other states by means of a noninvertible process would be only transient and could be safely eliminated by the theory. The fact that the transient states are insignificant is made clearer if it is assumed that the ontological space is compact.

Proposition. If the ontological space is compact, then the processes in any series $X_1 \overset{\hat{U}_1}{\to} X_2 \overset{\hat{U}_2}{\to} X_3 \overset{\hat{U}_3}{\to} \cdots$ become closer and closer to being invertible.

Proof. a metric ontological space is compact if it is closed and bounded. Let $B_R(X)$ be M-dimensional balls with radius R and center X. Suppose that for any R there exists a series $X_1 \overset{\hat{U}_1}{\to} X_2 \overset{\hat{U}_2}{\to} X_3 \overset{\hat{U}_3}{\to} \cdots$ with an infinite number of processes that take a state X away from its ball $B_R(X)$. Then there exists a subseries $Y_1 \to Y_2 \to Y_3 \to \cdots$ where $Y_m \notin B_R(Y_n)$ for every $m > n$, but this is impossible because the space is bounded. Along the series, the processes connecting an element to the following ones become, in fact, closer and closer to being invertible.

Discarding insignificant transient states and taking for granted Property 3, we can prove the second lemma.

Lemma 2. Assuming Property 3, if $X_1 \overset{\hat{U}_1}{\to} X_2$ is a nonzero probability process, then $\hat{U}_1\mathcal{S}(X_1) = \mathcal{S}(X_2)$.

Proof. The process $X_1 \overset{\hat{U}_1}{\to} X_2$ is allowed, thus

$$\hat{U}_1\mathcal{S}(X_1) \subseteq \mathcal{S}(X_2). \tag{26}$$

Since there exists an operator $\hat{U}_2$ such that $X_2 \to^{\hat{U}_2} X_1$ is a nonzero probability process, then we have also that

$$\hat{U}_2\mathcal{S}(X_2) \subseteq \mathcal{S}(X_1). \tag{27}$$

These two relations imply that

$$\hat{U}\mathcal{S}(X_2) \subseteq \mathcal{S}(X_2), \tag{28}$$

where $\hat{U} = \hat{U}_1\hat{U}_2$. By iteration we obtain that

$$\hat{U}^n\mathcal{S}(X_2) \subseteq \mathcal{S}(X_2), \tag{29}$$

for any integer n. For a finite-dimensional Hilbert space, it is always possible to find an integer n such that $\hat{U}^n$ is very close to the inverse operator $\hat{U}^{-1}$. Thus, $\hat{U}^{-1}\mathcal{S}(X_2) \subseteq \mathcal{S}(X_2)$, that is,

$$\hat{U}\mathcal{S}(X_2) \supseteq \mathcal{S}(X_2). \tag{30}$$

From inclusions (28) and (30) we have that

$$\hat{U}\mathcal{S}(X_2) = \mathcal{S}(X_2). \tag{31}$$

Applying the operator $\hat{U}_1$ to both sides of inclusion (27) and using Eq. (31), the inclusion

$$\mathcal{S}(X_2) \subseteq \hat{U}_1 \mathcal{S}(X_1), \tag{32}$$

is obtained. This relation and inclusion (26) imply that

$$\hat{U}_1 \mathcal{S}(X_1) = \mathcal{S}(X_2), \tag{33}$$

and the lemma is proved. ■

Note that, for symmetry reasons, a similar equation holds also for the inverse process, that is,

$$\hat{U}_2 \mathcal{S}(X_2) = \mathcal{S}(X_1). \tag{34}$$

In Ref. [7] we proved the following property for $\mathcal{S}(X)$.

Lemma 3. The set $\mathcal{S}(X)$ cannot contain every vector of the Hilbert space. Equivalently, the set $\mathcal{S}(X)$ is not invariant with respect to the group SU(N).

Proof by contradiction. Suppose that $\mathcal{S}(X)$ contains every vector of the Hilbert space, then it contains, in particular, also two orthogonal vectors. This means that there exist two overlapping distributions associated with two orthogonal quantum states. But this is impossible because two orthogonal states can be perfectly discriminated by a measurement [5,11]. Indeed the probability of obtaining $|\psi\rangle$ given $|\psi\rangle$ is

$$\int dX\, P(\psi|X,\tau)\rho(X|\psi,\eta) = |\langle\psi|\psi\rangle|^2 = 1. \tag{35}$$

This implies that $P(\psi|X,\tau)$ is equal to 1 in the support of $\rho(X|\psi,\eta)$. However, the probability of $|\psi\rangle$ given an orthogonal state $|\psi_\perp\rangle$ is

$$\int dX\, P(\psi|X,\tau)\rho(X|\psi_\perp,\bar{\eta}) = |\langle\psi|\psi_\perp\rangle|^2 = 0. \tag{36}$$

This implies that $\rho(X|\psi_\perp,\bar{\eta})$ cannot be different from zero if X is in the support of $\rho(X|\psi,\eta)$, where $P(\psi|X,\tau) = 1$. ■

Finally, we enunciate the last lemma.

Lemma 4. Let G be a Lie subgroup of the group SU(N) acting on an N-dimensional Hilbert space $\mathcal{H}$ and $\mathcal{S}$ be a set of vectors in $\mathcal{H}$. If the manifold dimension of G is larger than $(N-1)^2$ and $\mathcal{S}$ is invariant with respect to G, then $\mathcal{S}$ contains every vector of the Hilbert space, that is, $\mathcal{S}$ is invariant with respect to SU(N).

Proof. Any compact Lie group and their linear representations on $\mathbb{C}^N$ are well known. One can check that, for $N \neq 4$, the proper Lie subgroup of SU(N) with largest manifold dimension is SU($N-1$) × U(1). Its dimension is equal to $(N-1)^2$. Thus, if the dimension of G is larger than $(N-1)^2$, then G is, in fact, SU(N) and $\mathcal{S}$ contains every vector. In the special case $N = 4$, the symplectic group $Sp(2)$ is the subgroup of SU(4) with largest dimension. A set of generators in the representation space $\mathbb{C}^4$ is

$$\hat{\sigma}_i^{(1)},\ \hat{\sigma}_i^{(1)}\hat{\sigma}_1^{(2)},\ \hat{\sigma}_i^{(1)}\hat{\sigma}_2^{(2)},\ \hat{\sigma}_3^{(2)}, \tag{37}$$

with $i = 1,2,3$, $\hat{\sigma}_i^{(k)}$ being two sets of Pauli matrices acting on the tensor space $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$. It is easy to check that these generators form the basis of a Lie algebra with dimension equal to 10, which is larger than $(N-1)^2 = 9$. However, also in this case if the set $\mathcal{S}$ is invariant with respect to Sp(2), then it contains every vector of the Hilbert space. Indeed, it is possible to show that the only orbit of Sp(2) is the whole Hilbert space, the orbit of a subgroup being the set of states connected by means of some subgroup element. It is sufficient to prove that any vector is connected to $|\uparrow\rangle|\uparrow\rangle$. A generic vector $|\psi\rangle$ has the form

$$|\psi\rangle = \cos\theta|\uparrow\rangle|\phi_1\rangle + \sin\theta|\downarrow\rangle|\phi_2\rangle, \tag{38}$$

where $|\phi_1\rangle$ and $|\phi_2\rangle$ are two-dimensional vectors. This vector is connected through the generators $\hat{\sigma}_i^{(1)}$ to a vector with the form

$$|\tilde{\psi}\rangle = \cos\theta|\uparrow\rangle|\uparrow\rangle + \sin\theta|\downarrow\rangle|\tilde{\phi}_2\rangle. \tag{39}$$

Through the unitary operator $e^{\frac{\mathbb{1}-\hat{\sigma}_3^{(1)}}{2}(\theta_1\hat{\sigma}_1^{(2)}+\theta_2\hat{\sigma}_2^{(2)})}$ and a suitable choice of θ_i, it is possible to connect $|\tilde{\psi}\rangle$ to

$$|\bar{\psi}\rangle = \cos\theta|\uparrow\rangle|\uparrow\rangle + e^{i\varphi}\sin\theta|\downarrow\rangle|\downarrow\rangle. \tag{40}$$

This last vector is connected to $|\uparrow\rangle|\uparrow\rangle$ through the generators $\hat{\sigma}_3^{(1)}$ and $(\hat{\sigma}_1^{(1)}\hat{\sigma}_2^{(2)} + \hat{\sigma}_2^{(1)}\hat{\sigma}_1^{(2)})/2 = -i\hat{\sigma}_+^{(1)}\hat{\sigma}_+^{(2)} + i\hat{\sigma}_-^{(1)}\hat{\sigma}_-^{(2)}$, $\hat{\sigma}_+$ ($\hat{\sigma}_-$) being the raising (lowering) operators. ■

This theorem implies that the minimal number of parameters required to define the orientation of a set $\mathcal{S}$ cannot be smaller than $2N-2$, apart from the trivial case of a completely symmetric set $\mathcal{S}$, which does not have an orientation. This property was intuitively introduced in Ref. [7], where we noted that a set with the highest symmetry SU($N-1$) × U(1) requires $2N-2$ variables to specify the orientation of its symmetry axis.

C. No-shrinking theorem

At this point we have sufficient tools to prove the no-shrinking theorem. We will show that if the ontological space dimension would be smaller than the quantum-state manifold dimension, then the set $\mathcal{S}(X)$ would be invariant with respect to the group SU(N), in contradiction with Lemma 3.

Theorem. If the dynamics in a Markov ontological theory satisfies Property 1 and all the processes are invertible, then the ontological space dimension M is not smaller than the quantum-state manifold dimension, that is, $M \geqslant 2N-2$, N being the Hilbert space dimension.

Proof by contradiction. Suppose that $M < 2N-2$. Let us consider the unitary evolution $\hat{U}(\vec{t})$ defined in Eq. (22) and the associated conditional probability K given by Eq. (23). We have from Lemma 1 that a process $X \to Y$ with nonzero probability $K(Y|X,\vec{0})$ is associated with a D_s-dimensional manifold of unitary evolutions, say $\mathcal{U}$, with

$$D_s \geqslant N^2 - 1 - M > N^2 - 1 - 2N + 2 = (N-1)^2. \tag{41}$$

The manifold contains the identity. As a consequence, the set $\mathcal{S}(X)$ evolves, for any unitary evolution in $\mathcal{U}$, to the same $\mathcal{S}(Y)$. By means of Lemma 2 we have that

$$\hat{U}\mathcal{S}(X) = \mathcal{S}(Y), \quad \forall \hat{U} \in \mathcal{U}. \tag{42}$$

Since $\mathbb{1} \in \mathcal{U}$,

$$\mathcal{S}(X) = \mathcal{S}(Y). \tag{43}$$

Thus,

$$\hat{U}\mathcal{S}(X) = \mathcal{S}(X), \quad \forall \hat{U} \in \mathcal{U}, \tag{44}$$

that is, $\mathcal{S}(X)$ is symmetric with respect to a group whose generators are algebraically generated by D_s generators. The manifold dimension of this group is equal to or larger than $D_s > (N-1)^2$. Because of Lemma 4, $\mathcal{S}(X)$ is invariant with respect to SU(N), but this is in contradiction with Lemma 3. We conclude that the ontological space dimension M cannot be smaller than the quantum-state manifold dimension $2N-2$. ■

D. Minimal dimension and Schrödinger equation

It is possible to show that, in the case of minimal ontological dimension, the set of continuous variables can be identified with a vector living in the Hilbert space whose dynamics is described by the Schrödinger equation. It does not necessarily coincide with the quantum state, but moves rigidly with respect to it.

Suppose that the ontological dimension is minimal, that is, $M = 2(N-1)$. From Lemmas 1 through 4 we have that the set $\mathcal{S}(X)$ has to be symmetric with respect to a subgroup of SU(N) with manifold dimension equal to $N^2 - 1 - M = (N-1)^2$. The subgroup is SU($N-1$) × U(1). Thus, the set has a symmetry axis $|\phi\rangle$, which is a vector in the Hilbert space and identifies the orientation of $\mathcal{S}$. Since each ontic state X is associated with a set $\mathcal{S}(X)$, we have the mapping

$$X \to |\phi\rangle. \tag{45}$$

Because of Lemma 2, in a process associated with the unitary evolution $\hat{U}(t)$, the set $\mathcal{S}$ rotates rigidly to $\hat{U}(t)\mathcal{S}$. In particular, its symmetry axis satisfies the Schrödinger equation

$$i\frac{\partial|\phi\rangle}{\partial t} = \hat{H}(t)|\phi\rangle, \tag{46}$$

where $\hat{H}(t)$ is the Hamiltonian that generates $\hat{U}(t)$.

The mapping (45) is surjective, but is not necessarily injective. However, since the number of continuous ontological variables is just sufficient to label $|\phi\rangle$ up to a global phase, an additional discrete index is sufficient to make the mapping bijective

$$X \leftrightarrow (|\phi\rangle, n). \tag{47}$$

Thus, the ontic state in a minimal Markov theory is identified with a vector in the Hilbert space and a possible additional discrete index. The dynamics of the vector is given by the Schrödinger equation. $|\phi\rangle$ is equal to the quantum state if the set $\mathcal{S}(X)$ contains only one state that coincides with the symmetry axis. This is the case in a wave-pilot theory such as the de-Broglie–Bohm mechanics.

It is interesting to note that requirements (46) and (47) are fulfilled by the Kochen-Specker model [13], where the ontic state is identified by a Bloch vector having the same dynamical equation of the quantum state.

E. Resource cost with round off error

We have proved that in any hidden-variable theory with short memory dynamics the dimension of the state space cannot be smaller than the quantum-state manifold dimension $2N-2$. There is a relation between dimension and resource cost required to identify an ontic state. By resource cost we mean the quantity of information required to identify an ontological state. Obviously the information carried by an ontic state is infinite since the ontological space is continuous. Thus, the resource cost as a function of the dimension makes sense only in the presence of a fixed roundoff error of the continuous variables.

One can find a scaling law between dimension and resource cost in the following way. Suppose that the ontic state is identified by M continuous variables x_i, which are defined in a finite interval. With a suitable rescaling, we can assume that they run in the interval between 0 and 1. The probability $P(\mathcal{E}|\vec{x})$ of an event $\mathcal{E}$ is conditioned by the ontic state $\vec{x}$. It is assumed that $P(\mathcal{E}|\vec{x})$ is a smooth function of $\vec{x}$, but the discussion could be extended to the case of a finite number of discontinuities. Let g_i be the mean magnitude of $|\partial_{x_i} P(\mathcal{E}|\vec{x})|$.

Let each continuous variable be discretized by n_i points. This introduces a roundoff error ΔE in P that scales as $\sqrt{\sum_{i=1}^M (g_i/n_i)^2}$ for sufficiently large values of n_i. The number of bits required to identify the ontic state on the lattice is proportional to the information

$$\mathcal{I} = \sum_{i=1}^{M} \log n_i. \tag{48}$$

For a fixed $\mathcal{I}$, using the Lagrange multiplier method we find that the optimal choice of n_i that gives the smallest error is

$$n_i = \frac{g_i e^{\frac{\mathcal{I}}{M}}}{\bar{g}}, \tag{49}$$

where $\bar{g} \equiv (\prod_i g_i)^{1/M}$ is the log-average of g_i. Thus, we have the scaling law

$$\Delta E \sim \bar{g} M^{\frac{1}{2}} e^{-\frac{\mathcal{I}}{M}}. \tag{50}$$

The error exponentially decreases with $\mathcal{I}$ at a rate inversely proportional to M. For a fixed error, the information scales as

$$\mathcal{I} \sim M \log \frac{M^{\frac{1}{2}}\bar{g}}{\Delta E}. \tag{51}$$

The no-shrinking theorem states that M is not smaller than the quantum-state manifold dimension. Thus, in a composite system, $\mathcal{I}$ grows at least exponentially in the number of parts (for example, the number of qubits in a quantum computer).

It is possible to estimate a lower bound for $\bar{g}$. Since the probability of an event goes from 0 to 1 and x_i rambles about the interval [0:1], it is reasonable to assume that g_i cannot be smaller than a value around 1, that is,

$$\bar{g} \gtrsim 1.$$

Thus the linear growth of $\mathcal{I}$ with respect to M cannot be mitigated by an exponential decrease of $\bar{g}$.

V. CONCLUSION

We have shown that the ontological space dimension in any hidden-variable theory with a short memory dynamics cannot be smaller than the quantum-state manifold dimension. Thus, like the quantum state, the ontic state necessarily carries for a given accuracy an amount of information that grows exponentially in the number of subsystems. In comparison with Ref. [7], we have provided a better justification of the relaxation hypothesis, showing that in a compact ontological space it is possible to find a trajectory along which the system unavoidably converges toward a region where all the processes are invertible. Furthermore we have presented a hidden-variable model of measurement for a qubit whose ontological space is one-dimensional, which is one-half the dimension of the Bloch sphere. The corresponding dynamics is not Markovian, in accordance with the no-shrinking theorem.

This model provides a counterexample making it evident that the hypotheses of short memory is strictly necessary to prove the no-shrinking theorem. By dropping it, we have shown that a single realization can carry less information than the quantum state. More drastically, we could drop the causality hypothesis. Indeed the noncausality is implied also by the Bell theorem and the Lorentz invariance of the hidden-variable theory. Thus, there are two signs that point to the same direction, that is, the rejection of causality at the ontological level. The possibility of an ontological shrinking for a general N-dimensional Hilbert space in a theory without a Markov dynamics is an open question whose answer could provide a deeper understanding of the computational complexity in quantum mechanics.

ACKNOWLEDGMENTS

I wish to thank J. Wallman and R. W. Spekkens for the careful reading of the manuscript and useful suggestions. Research at the Perimeter Institute for Theoretical Physics is supported, in part, by the Government of Canada through NSERC and by the Province of Ontario through MRI.

APPENDIX: DERIVING THE ECONOMICAL MODEL FOR A QUBIT

In this Appendix we will present a systematic method to generate the model reported in Sec. II. We will consider the class of probability distributions with the form

$$\rho(x,n|\vec{v}) = r(n|\vec{v})\delta[x - f_n(\vec{v})], \tag{A1}$$

where the pair $(x \in \mathbb{R}, n \in \{0,1\})$ is the ontic state and $\vec{v} \equiv \langle\psi|\vec{\sigma}|\psi\rangle$ is the Bloch vector corresponding to the quantum state $|\psi\rangle$. The quantities $f_n(\vec{v})$ are real functions. The normalization of the distribution gives

$$r(0|\vec{v}) + r(1|\vec{v}) = 1. \tag{A2}$$

When a system is prepared in $\vec{v}$, at the ontological level there is a probability $r(0|\vec{v})$ $[r(1|\vec{v})]$ that the discrete index takes the value $n = 0$ $(n = 1)$. Correspondingly, the continuous variable takes with certainty the value $f_0(\vec{v})$ $[f_1(\vec{v})]$. This class is the simplest one that fulfils some requirements. As noted in Sec. IV B, two orthogonal states cannot have overlapping probability distributions, implying that the support of $\rho(x,n|\vec{v})$ cannot be the whole ontological space. This rules out the class of smooth analytical distributions and leads us to consider the probability distributions with zero-measure support as the simplest case. In particular, the distributions with a two-point support are parametrized with three real parameters [the positions $f_n(\vec{v})$ of the points and the relative probability weight $r(0|\vec{v}) - r(1|\vec{v})$], which are sufficient to cover the two-dimensional Bloch sphere.

For our purposes, it is convenient to define the variables $x_n \equiv f_n(\vec{v})$ and to use them to parametrize the quantum state. Thus, the probability distribution becomes

$$\rho(x,n|x_0,x_1) = r(n|x_0,x_1)\delta(x - x_n). \tag{A3}$$

Let $|\phi\rangle$ be the event of some projective measurement. As for the preparation state, we introduce the Bloch vector $\vec{w} \equiv \langle\phi|\vec{\sigma}|\phi\rangle$ to label the event $|\phi\rangle$. $P_n(\vec{w}|x)$ is the conditional probability of the event $\vec{w}$ given the ontic state (x,n). Equation (4) becomes

$$\sum_{n=0}^{1}\int dx\, P_n(\vec{w}|x)\rho(x,n|\vec{x}) = \frac{1+\vec{w}\vec{v}(\vec{x})}{2} \equiv S, \tag{A4}$$

where $\vec{v}(\vec{x})$ gives the Bloch vector $\vec{v}$ as a function of the parameters $(x_0,x_1) \equiv \vec{x}$. The quantity $1/2(1+\vec{w}\vec{v})$ is the Born probability $|\langle\phi|\psi\rangle|^2$ in terms of the Bloch vectors.

Using Eq. (A3), Eq. (A4) becomes after integration

$$P_0(\vec{w}|x_0)r(0|x_0,x_1) + P_1(\vec{w}|x_1)r(1|x_0,x_1) = S. \tag{A5}$$

Note that the conditional probability $P_n(\vec{w}|x_n)$ in this equation depends only on the variable x_n. Thus, to find it, we could consider the four pairs (x_0,x_1), (x_0,y_1), (y_0,x_1), and (y_0,y_1). Correspondingly, we have four equations with the four unknown functions $P_0(\vec{w}|x_0)$, $P_0(\vec{w}|y_0)$, $P_1(\vec{w}|x_1)$, and $P_1(\vec{w}|y_1)$. These equations are the row elements of the vector equation

$$\hat{R}\vec{P} = \vec{S}, \tag{A6}$$

where

$$\hat{R} \equiv \begin{pmatrix} r(0|x_0,x_1) & 0 & r(1|x_0,x_1) & 0 \\ r(0|x_0,y_1) & 0 & 0 & r(1|x_0,y_1) \\ 0 & r(0|y_0,x_1) & r(1|y_0,x_1) & 0 \\ 0 & r(0|y_0,y_1) & 0 & r(1|y_0,y_1) \end{pmatrix}, \tag{A7}$$

$$\vec{P} \equiv \begin{pmatrix} P_0(\vec{w}|x_0) \\ P_0(\vec{w}|y_0) \\ P_1(\vec{w}|x_1) \\ P_1(\vec{w}|y_1) \end{pmatrix}, \tag{A8}$$

and

$$\vec{S} \equiv \frac{1}{2}\begin{pmatrix} 1+\vec{w}\cdot\vec{v}(x_0,x_1) \\ 1+\vec{w}\cdot\vec{v}(x_0,y_1) \\ 1+\vec{w}\cdot\vec{v}(y_0,x_1) \\ 1+\vec{w}\cdot\vec{v}(y_0,y_1) \end{pmatrix}. \tag{A9}$$

If $\hat{R}$ would be invertible, $P_n(\vec{w}|x_n)$ would be a linear function of $\vec{w}$, but this is impossible. It can be proved by contradiction. Suppose that the conditional probabilities are linear functions of $\vec{w}$, that is, $P_n(\vec{w}|x_n) = C_n(x_n) + \vec{d}_n(x_n)\vec{w}$. They have to be nonnegative and smaller than or equal to 1. It is clear that $P_n(\vec{w}|x_n)$ are equal to 1 at most for only one vector $\vec{w}$, but this is in contradiction with the fact that the conditional probabilities have to be equal to 1 in the one-dimensional manifold of vectors $\vec{w}$ such that $f_n(\vec{w}) = x_n$. Thus, the equation

$$\det \hat{R}(x_0,x_1,y_0,y_1) = 0, \tag{A10}$$

has to be satisfied for any x_n and y_n. Its solution is found by differentiating in y_0 and y_1

$$\frac{\partial^2}{\partial y_0 \partial y_1} \det \hat{R}(x_0,x_1,y_0,y_1)|_{y_0=x_0,y_1=x_1} = 0, \tag{A11}$$

that is,

$$r(0|x_0,x_1)^2 r(1|x_0,x_1)^2 \partial_{x_0}\partial_{x_1} \log \frac{r(1|x_0,x_1)}{r(0|x_0,x_1)} = 0. \tag{A12}$$

The functions $r(n|\vec{x})$ cannot be identically equal to zero, thus the only acceptable solution that satisfies also the constraint (A2) is

$$r(0|x_0,x_1) = \frac{k_1(x_1)}{k_0(x_0)+k_1(x_1)}, \tag{A13}$$

$$r(1|x_0,x_1) = \frac{k_0(x_0)}{k_0(x_0)+k_1(x_1)}, \tag{A14}$$

where $k_i(x_i)$ are positive functions of the variable x_i. It is easy to check that this is also a solution of Eq. (A10). Thus, the fact that the conditional probabilities $P_n(\vec{w}|x_n)$ cannot be linear in $\vec{w}$ allows us to find a constraint for the probabilities $r(n|\vec{x})$.

Since the determinant of $\hat{R}$ is equal to zero, there is a constraint also for the $\vec{S}$ because of Eq. (A6). Let the row vector $\vec{u}^T$ be the left eigenvector of $\hat{R}$ with eigenvalue 0, then we have from Eq. (A6) that

$$\vec{u}^T \vec{S} = 0. \tag{A15}$$

By means of Eqs. (A13) and (A14), after a bit of calculation we find that the left eigenvector of $\hat{R}$ with zero eigenvalue is

$$\vec{u} = \begin{pmatrix} k_0(x_0)^{-1} + k_1(x_1)^{-1} \\ -k_0(x_0)^{-1} - k_1(y_1)^{-1} \\ -k_0(y_0)^{-1} - k_1(x_1)^{-1} \\ k_0(y_0)^{-1} + k_1(y_1)^{-1} \end{pmatrix}. \tag{A16}$$

From Eqs. (A9), (A15), and (A16) we obtain by differentiation the condition

$$\begin{gathered} \partial_{y_0,y_1}(\vec{u}^T\vec{S})|_{y_n=x_n} = 0 \\ \Longrightarrow \partial_{y_0,y_1}\left[\frac{1}{k_0(x_0)} + \frac{1}{k_1(x_1)}\right]\vec{v}(x_0,x_1)\bigg|_{y_n=x_n} = 0, \end{gathered} \tag{A17}$$

that is satisfied if

$$\vec{v}(x_0,x_1) = \left[\frac{1}{k_0(x_0)} + \frac{1}{k_1(x_1)}\right]^{-1} [\vec{g}_0(x_0) + \vec{g}_1(x_1)], \tag{A18}$$

where $\vec{g}_0(x_0)$ and $\vec{g}_1(x_1)$ are generic vectorial functions. This is the inverse of the equation $\vec{x} = \vec{f}(\vec{v})$. The functions $g_n(x_n)$ and $k_n(x_n)$ are constrained by the equation $\vec{v}^2 = 1$, that is,

$$[\vec{g}_0(x_0) + \vec{g}_1(x_1)]^2 = \left[\frac{1}{k_0(x_0)} + \frac{1}{k_1(x_1)}\right]^2. \tag{A19}$$

We will return to it later on.

From Eqs. (A5), (A13), (A14), and (A18) and the identity $S = [1 + \vec{w}\vec{v}(\vec{x})]/2$ we find that

$$\sum_{n=0}^{1}\left\{\left[P_n(\vec{w}|x_n) - \frac{1}{2}\right]k_n^{-1}(x_n) - \frac{1}{2}\vec{w}\vec{g}_n\right\} = 0. \tag{A20}$$

Note that each term of the summation depends only on one of the variables x_0 and x_1. Thus, the conditional probabilities have the form

$$P_0(\vec{w}|x_0) = k_0(x_0)\left[\tfrac{1}{2}\vec{w}\vec{g}_0(x_0) + H(\vec{w})\right] + \tfrac{1}{2}, \tag{A21}$$

$$P_1(\vec{w}|x_1) = k_1(x_1)\left[\tfrac{1}{2}\vec{w}\vec{g}_1(x_1) - H(\vec{w})\right] + \tfrac{1}{2}, \tag{A22}$$

where $H(\vec{w})$ is an additional function independent of $\vec{x}$. Since the probability $P_0(\vec{w}|x_0)$ has to be equal to 1 for $\vec{w} = \vec{v}$, we find that

$$H(\vec{v}) = \frac{1}{2}\left[\frac{1}{k_0(x_0)} - \vec{v}\vec{g}_0(x_0)\right], \tag{A23}$$

where $x_0 = f_0(\vec{v})$. Similarly, from the condition $P_0(\vec{v}|x_0) = 1$ we obtain the equation

$$H(\vec{v}) = -\frac{1}{2}\left[\frac{1}{k_1(x_1)} - \vec{v}\vec{g}_1(x_1)\right]. \tag{A24}$$

This last equation can be derived by Eqs. (A18) and (A23) and the constraint $\vec{v}^2 = 1$. It is interesting to note that the model we are constructing works only for a subset of preparation states $\vec{v}$, as shown in Sec. III, thus the function $H(\vec{w})$ is not necessarily given by Eqs. (A23) and (A24) if $\vec{w}$ is outside that subset.

At this point we have almost everything, the last step is to find the functions $k_n(x_n)$ and $\vec{g}_n(x_n)$ that solve Eq. (A19). Differentiating this equation with respect to x_0 and x_1, we have that

$$\frac{\partial \vec{g}_0(x_0)}{\partial x_0}\frac{\partial \vec{g}_1(x_1)}{\partial x_1} - \left(\frac{\partial}{\partial x_0}\frac{1}{k_0(x_0)}\right)\left(\frac{\partial}{\partial x_1}\frac{1}{k_1(x_1)}\right) = 0, \tag{A25}$$

that is, the Minkowski inner product between the two four-vectors

$$\alpha(x_0) \equiv \begin{pmatrix} \vec{g}_0(x_0) + \vec{\gamma}_0 \\ \frac{1}{k_0(x_0)} + \chi_0 \end{pmatrix}; \quad \beta(x_1) \equiv \begin{pmatrix} \vec{g}_1(x_1) + \vec{\gamma}_1 \\ \frac{1}{k_1(x_1)} + \chi_1 \end{pmatrix}, \tag{A26}$$

is equal to 0, $\vec{\gamma}_n$ and χ_n being constant vectors and scalars, respectively. Using the Einstein notation, on the index contraction, the constraint is

$$\alpha_\mu(x_0)\beta^\mu(x_1) = 0. \tag{A27}$$

It is important to note that α and β depend only on one of the variables x_0 and x_1. Using Eq. (A27), Eq. (A19) becomes $(\vec{g}_0 - \vec{\gamma}_1)^2 - (\frac{1}{k_0} - \chi_1)^2 - (\vec{\gamma}_0 + \vec{\gamma}_1)^2 + (\chi_0 + \chi_1)^2 = -(\vec{g}_1 - \vec{\gamma}_0)^2 + (\frac{1}{k_1} - \chi_0)^2$. The left-hand and right-hand sides depend only on x_0 and x_1, respectively, thus they have to be equal to a constant r_0

$$(\vec{g}_0-\vec{\gamma}_1)^2-\left(\frac{1}{k_0}-\chi_1\right)^2=(\vec{\gamma}_0+\vec{\gamma}_1)^2-(\chi_0+\chi_1)^2+r_0,$$
$$(\vec{g}_1-\vec{\gamma}_0)^2-\left(\frac{1}{k_1}-\chi_0\right)^2=-r_0. \tag{A28}$$

These equations and Eq. (A27) are a convenient resettlement of Eq. (A19).

It is interesting to observe that the conditional probabilities $P_n(\vec{w}|x_n)$ and Eqs. (A18) and (A19) are invariant with respect to the transformation

$$\vec{g}_0\to\vec{g}_0+\vec{t},\quad \vec{g}_1\to\vec{g}_1-\vec{t}, \tag{A29}$$

for a generic vector $\vec{t}$. This means that the constant vectors $\vec{\gamma}_0$ and $\vec{\gamma}_1$ are redundant and, for example, one could set $\vec{\gamma}_1=0$.

To solve Eq. (A27), we have to consider two possibilities: (a) one of the two four-vectors α and β spans a one-dimensional vectorial subspace and the other one lives in the orthogonal subspace; (b) the two vectors span orthogonal two-dimensional spaces. In the former case no solution exists such that $\vec{v}(x_0,x_1)$ is locally invertible (only a one-dimensional subspace of the Bloch sphere is represented). Thus, we consider the latter case. The most general solution up to rotations, transformation (A29), and variable change $x_n\to F_n(x_n)$ is

$$\vec{g}_0(x_0)=\begin{pmatrix}\cos x_0\\ \sin x_0\sin\theta_0\\ 0\end{pmatrix},\quad \frac{1}{k_0(x_0)}=\cos\theta_0\cos x_0+s,$$
$$\vec{g}_1(x_1)=\begin{pmatrix}\cos\theta_0\csc x_1\\ 0\\ \cot x_1\sin\theta_0\end{pmatrix},\quad \frac{1}{k_1(x_1)}=\csc x_1-s, \tag{A30}$$

where θ_0 and s are two free parameters. Only the former is present in the mapping $\vec{x}\to\vec{v}(\vec{x})$, given by Eq. (A18). We have that

$$\vec{v}(x_0,x_1)=\frac{1}{1+\cos\theta_0 u_x}\begin{pmatrix}\cos\theta_0+u_x\\ \sin\theta_0 u_y\\ \sin\theta_0 u_z\end{pmatrix}, \tag{A31}$$

where

$$lu_x=\sin x_1\cos x_0,\quad u_y=\sin x_1\sin x_0,\quad u_z=\cos x_1. \tag{A32}$$

The mapping is bijective by removing the poles at $x_1=0$ and $x_1=\pi$

$$0\leqslant x_0\leqslant 2\pi,\quad 0<x_1<\pi. \tag{A33}$$

The inverse of the vectorial function $\vec{v}(\vec{x})$ in Eq. (A31) gives the functions $f_n(\vec{v})=x_n$, whose trigonometric functions are

$$\sin x_0=v_y\sin\theta_0/\Delta,\quad \cos x_0=(v_x-\cos\theta_0)/\Delta,$$
$$\sin x_1=\frac{\Delta}{1-v_x\cos\theta_0},\quad \cos x_1=\frac{v_z\sin\theta_0}{1-v_x\cos\theta_0}, \tag{A34}$$

with $\Delta\equiv\sqrt{(w_x-\cos\theta_0)^2+w_y^2\sin^2\theta_0}$.

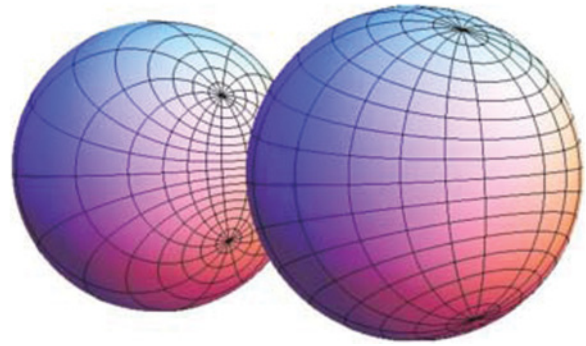

FIG. 2. (Color online) Two orthogonal coordinate systems for $\theta_0=0.5$ rad (left) and $\theta_0=1$ rad (right).

Equation (A31) provides a set of orthogonal coordinate systems of the sphere. The spherical coordinate system is obtained with $\theta_0=\pi/2$. Each system is mapped to another one by means of the Möbius transformation [14]. In Fig. 2 we report two coordinate systems for $\theta_0=0.5, 1$. Both of them have two poles, but with different angular distance that is equal to $2\theta_0$.

From Eqs. (A13), (A14), and (A30) we have that

$$r(0|\vec{x})=\frac{\sin x_1(s+\cos\theta_0\cos x_0)}{1+\cos\theta_0\cos x_0\sin x_1}, \tag{A35}$$
$$r(1|\vec{x})=\frac{1-s\sin x_1}{1+\cos\theta_0\cos x_0\sin x_1}. \tag{A36}$$

They are positive if

$$|\cos\theta_0|\leqslant s\leqslant 1. \tag{A37}$$

$H(\vec{w})$ is obtained by Eq. (A23)

$$H(\vec{w})=\frac{s-\Delta}{2}. \tag{A38}$$

Finally, the conditional probabilities for the events are given by Eqs. (A21) and (A22)

$$P_0(\vec{w}|x)=1+\frac{(w_x-\cos\theta_0)\cos x+w_y\sin x\sin\theta_0-\Delta}{2(s+\cos\theta_0\cos x)}, \tag{A39}$$

$$P_1(\vec{w}|x)=1+\frac{(w_x\cos\theta_0-1)+w_z\cos x\sin\theta_0+\Delta\sin x}{2(1-s\sin x)}. \tag{A40}$$

The model in Sec. III is obtained for $\theta_0=\pi/2$ and $s=1$.

[1] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[2] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[3] R. Jozsa and A. Miyake, Proc. R. Soc. A **464**, 3089 (2008).
[4] M. Van den Nest, Quant. Inf. Comp. **10**, 0258 (2010); e-print arXiv:0911.1624.
[5] L. Hardy, Stud. Hist. Phil. Sci. B **35**, 267 (2004).
[6] A. Montina, Phys. Rev. Lett. **97**, 180401 (2006); J. Phys.: Conf. Ser. **67**, 012050 (2007).

[7] A. Montina, Phys. Rev. A **77**, 022104 (2008).
[8] B. Dakić, M. Suvakov, T. Paterek, and C. Brukner, Phys. Rev. Lett. **101**, 190402 (2008).
[9] E. F. Galvão, Phys. Rev. A **80**, 022106 (2009).
[10] A. Montina, Phys. Lett. A **375**, 1385 (2011).
[11] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[12] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 2004).
[13] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[14] S. G. Krantz, *Geometric Function Theory: Explorations in Complex Analysis* (Birkhäuser, Boston, 2006).