Local model of a qudit: Single particle in optical circuits

Pawel Blasiak*
Institute of Nuclear Physics, Polish Academy of Sciences, PL-31342 Kraków, Poland

(Received 12 March 2018; published 13 July 2018)

It has been said about quantum interference, that "*in reality, it contains the only mystery*". Together with nonlocality, it is often considered the characteristic feature of quantum theory challenging our classical understanding of the world. In this work, we are concerned with the restricted setting of a single particle propagating in multipath interferometric circuits—that is, the physical realization of a qudit—which is host to many typically quantum mechanical effects including collapse of the wave function and contextuality. In this paper, we show that this framework can be simulated with classical resources without violating the locality principle. We present a *local ontological model* whose predictions are indistinguishable from the quantum case. In the model, 'nonlocality' appears merely as an epistemic effect arising from a level of description by agents whose knowledge is incomplete. It is notably different from the multiparticle scenarios where entanglement leads to nonlocal correlations on an ontological level. This result exposes the conceptual difference between single- and multiparticle phenomena, pointing to the latter as a deeper quantum mystery.

DOI: 10.1103/PhysRevA.98.012118

I. INTRODUCTION

In the Feynman Lectures on Physics, quantum interference is described as "*a phenomenon which is impossible, absolutely impossible, to explain in any classical way, and which has in it the heart of quantum mechanics*" [1]. Broadly speaking, it concerns behavior of a particle in the interferometric circuits, and the problem consists in reconciling the wave and particle character of the phenomenon. Another difficulty of quantum formalism is a common-sense explanation for the collapse of the wave function upon measurement. In some mysterious way, behavior of a quantum particle depends on knowledge of what is happening in distant parts of the experimental setup. Notably, nonlocality of the collapse already manifests itself in single-particle scenarios, as Einstein first pointed out during the Fifth Solvay Conference [2] when he metaphorically called such an influence "*spooky action at a distance*" [3,4]. A full-fledged argument against local realism in quantum theory is due to the profound insight of Bell [5,6], who noted that it requires two particles to show nonlocal correlations between measurements in distant arms of an interferometric setup. Remarkably, all further refinements of the argument exploit properties of entangled states in multiparticle scenarios—see, e.g., Refs. [7–9]. This leaves open the question of a possible local explanation of quantum-interferometric phenomena in the single-particle case—cf. [10–13].

Quantum mechanics of single-particle phenomena is a rich source of paradoxes and surprising effects which challenge our classical understanding of the world. Apart from quantum interference [14], they include, e.g., interaction-free measurements [15–17], quantum Zeno effect [16–19], Wheeler's delayed-choice experiment [20,21], violation of Leggett-Garg inequalities [22,23], pre- and postselection paradoxes [24,25], and contextuality [26–31]. These phenomena are often considered strictly quantum mechanical effects and some of them, like contextuality or Leggett-Garg inequalities, are sometimes treated as signatures of the quantum regime. However, as suggestive as it might appear, it is not at all clear to what extent these features are unique to the quantum realm. On the one hand, there are various models indicating analogies on the grounds of classical probabilistic theories—see, e.g., Refs. [32–43]. On the other hand, none of these results fully reconstruct quantum predictions for general single-particle scenarios. Altogether, this makes the question about the distinctive quantum features an interesting problem. Specifically, it is not clear whether nonlocality in a single-particle framework is on par with a multiparticle case, i.e., does not admit existence of a local hidden variable model [10–13]. A decisive answer would require either a rigorous no-go proof, such as Bell's theorem is for two particles, or a counterexample encompassing all relevant aspects of quantum interferometric setups.

In this paper we are concerned with a single particle propagating in general multipath interferometric circuits—that is, a physical realization of finite-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^N$ (qudit) [44]—and explicitly construct a local ontological model which faithfully imitates all quantum-mechanical predictions. This suggests being cautious of statements about the absolute impossibility of the classical explanation of single-particle interferometric phenomena. Indeed, the model shows that a local explanation is conceivable and hence the real quantum mystery should be sought in multiparticle behavior [5–9].

Our construction is made within the ontological model framework [27,28,45] which encompasses a broad range of hidden-variable scenarios. It makes a crucial distinction between the ontic and epistemic level of description, allowing for systematic treatment of situations in which access to information is constrained. The basic idea of the model consists in defining ontic variables which propagate locally through

*pawel.blasiak@ifj.edu.pl

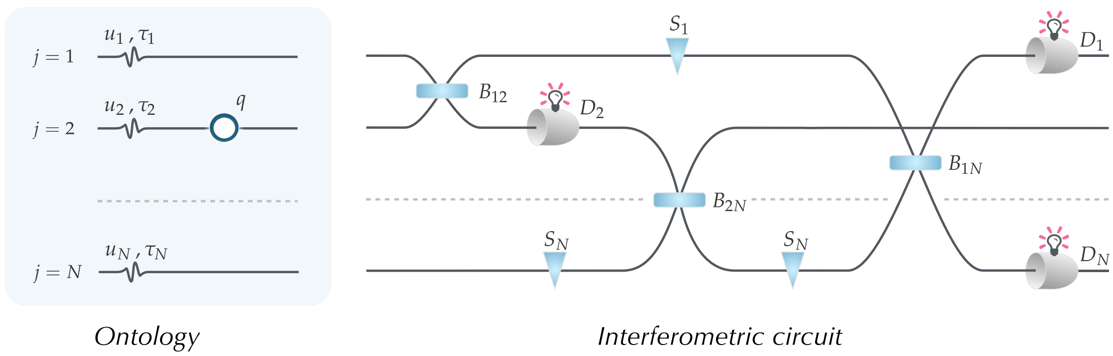

FIG. 1. Ontology of the model and interferometric circuits. On the right, the circuits describe propagation of a particle through a network of (spatially separated) paths and gates which represent a sequence of transformations. A basic interferometric toolkit consists of *free evolution* (empty path), *phase shifters* S_j, *beam splitters* B_{st} (on which two paths cross), and *detectors* D_j, which inform (CLICK or NO CLICK) about the presence or absence of a particle in a given path. This selection of gates is general enough to provide a physical realization of any unitary and projective measurement described by quantum formalism in $\mathcal{H} = \mathbb{C}^N$ [44]. On the left, ontology of the model consists of a *single particle* and *local fields* propagating in each path of the circuit. Each time the particle has the well-defined *position* $q = 1, \ldots, N$ and the fields are characterized by *amplitude* u_j and *strength* τ_j with $j = 1, \ldots, N$ labeling the paths. In this paper, we show that this ontology, completed with appropriately defined local stochastic gates, fully reconstructs quantum mechanical predictions for a single particle in the interferometric circuits.

interferometric circuits built of gates, and the latter determine the evolution of the system by affecting only those variables which go into a given gate—see Sec. III. In Sec. IV we take the perspective of an observer who is unaware or indifferent to the underlying ontology of the model and analyze the structure of information which is available to the agents investigating the model by conducting experiments in every conceivable circuit. We show that the operational account of the model is indistinguishable from the quantum-mechanical description of a qudit ($\mathcal{H} = \mathbb{C}^N$). This demonstrates that constraints on knowledge play an important role in describing a system under study. In the model, agents whose resources are constrained and receive incomplete information report a variety of quantumlike effects on an epistemic level, which faithfully imitate all quantum-mechanical predictions for a single particle in the interferometric circuits. For the sake of clarity, rigorous treatment within the ontological model framework is given in Appendix A and the proofs are delegated to Appendix B.

II. QUANTUM INTERFEROMETRY IN A NUTSHELL

We begin with a brief account of quantum interferometric circuits. This is meant to introduce the notation and provide a basis for comparison of the model constructed in this paper with the standard quantum-mechanical description.

In the following, we consider a standard interferometric framework for a single particle propagating through a network of spatially separated paths. Evolution of the system is implemented by gates attached to the paths which represent nontrivial transformations (with empty paths corresponding to free evolution). See Fig. 1 (on the right) for illustration. It is sufficient to consider only a few kinds of gates which form the basis for construction of complex interferometric circuits [44]. These gates include phase shifters S_j and detectors D_j which are attached to individual paths and beam splitters B_{st} on which two paths cross, with j and s,t indicating the respective paths. A special role of detectors is to provide an outcome CLICK or NO CLICK, which attests to the presence or absence of a particle in a given path.

The quantum description of a single particle in an interferometric circuit, which consists of N paths associates the position of the particle with the vectors of computational basis $|1\rangle, \ldots, |N\rangle$, where $|j\rangle$ represents the fact of the particle being in the jth path. Generally, the state of the system is a superposition with the complex coefficients ψ_j defining a vector (ray) in $\mathcal{H} = \mathbb{C}^N$, i.e.,

$$|\psi\rangle = \sum_{j=1}^{N} \psi_j |j\rangle = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} = \vec{\psi}, \tag{1}$$

with normalization $\||\psi\rangle\|^2 = \sum_j |\psi_j|^2 = 1$ and vectors differing by an overall phase being equivalent. Evolution implemented by gates corresponds to a sequence of unitary and projective transformations described as follows. *Free evolution* in the jth path acts trivially and the *phase shifter* S_j introduces the phase $e^{i\omega}$ in the relevant path, i.e., we have

$$\psi_j \xrightarrow{free} \psi_j \quad \text{and} \quad \psi_j \xrightarrow{S_j} e^{i\omega}\psi_j. \tag{2}$$

The *beam splitter* B_{st} located at the crossing of paths s and t implements a unitary in the subspace spanned by kets $|s\rangle$ and $|t\rangle$ given by

$$\begin{pmatrix} \psi_s \\ \psi_t \end{pmatrix} \xrightarrow{B_{st}} \begin{pmatrix} \psi_s' \\ \psi_t' \end{pmatrix} = \begin{pmatrix} i\sqrt{R} & \sqrt{T} \\ \sqrt{T} & i\sqrt{R} \end{pmatrix} \begin{pmatrix} \psi_s \\ \psi_t \end{pmatrix}, \tag{3}$$

where R,T are reflectivity and transitivity coefficients (with the usual convention that the particle gains phase i upon reflection). Finally, according to the measurement postulate (von Neumann–Lüders rule), *detector* D_j is described by the PVM $\{\mathbb{P}_j, \mathbb{1} - \mathbb{P}_j\}$, where $\mathbb{P}_j \equiv |j\rangle\langle j|$, i.e., depending on the outcome it effects the projection

$$|\psi\rangle \xrightarrow{D_j} \begin{cases} |j\rangle & \text{CLICK}, \\ \frac{(\mathbb{1}-\mathbb{P}_j)|\psi\rangle}{\|(\mathbb{1}-\mathbb{P}_j)|\psi\rangle\|} & \text{NO CLICK}, \end{cases} \tag{4}$$

with the probability that detector D_j CLICKS given by Born's rule $\text{Prob}\,(D_j|\,\psi) = |\langle j|\psi\rangle|^2 = |\psi_j|^2$. Note that the projection postulate Eq. (4) affects the whole space $\mathcal{H} = \mathbb{C}^N$ in spite of the fact that the detector D_j is localized only in the jth path. An explanation of this behavior leads to the notorious problem concerning the ontological status of quantum states and the issue of nonlocality of the collapse of the wave function.

These rules provide a mathematical description of a single particle in the interferometric circuit. It was shown in Ref. [44] that any unitary and projective measurement in $\mathcal{H} = \mathbb{C}^N$ can be experimentally realized in a circuit composed of N paths as a sequence of interferometric gates as defined above. Thus it provides a convenient physical framework for explorations in quantum foundations.

III. CONSTRUCTION OF THE MODEL

Our main goal in this paper is the explicit construction of a classical analog with the same structural components (comprised of paths and gates arranged into circuits) which mimics quantum behavior of a particle in the interferometric circuits as described above. The crux of the matter is to provide a model with a well-defined underlying ontology which does not violate the locality principle and yet on an operational level its predictions are indistinguishable from the quantum case.

A. Ontology of the model

Let us consider circuits composed of N paths labeled with index $j = 1,\ldots,N$. When defining the model, we assume that in the circuit a *single particle* propagates which has the well-defined *position* $q = 1,\ldots,N$. Additionally, we postulate that along each path propagates a local field characterized by two degrees of freedom: (complex) *amplitude* u_j such that $|u_j| \leq 1$ and (real) *strength* τ_j such that $0 \leq \tau_j \leq 1$. This means that at each time the system of N paths is fully specified by a point $\lambda = (q,\vec{u},\vec{\tau})$ in the *ontic state space*

$$\begin{aligned}\Lambda = &\left\{q\,:\,q = 1,\ldots,N\right\}\\ &\times\left\{\vec{u}\in\mathbb{C}^N\,:\,|u_j|\leq 1\right\}\\ &\times\left\{\vec{\tau}\in\mathbb{R}^N\,:\,0\leq\tau_j\leq 1\right\},\end{aligned}\tag{5}$$

where u_j and τ_j describe the field in the jth path. See Fig. 1 (on the left) for illustration.

In the following, we explore the stochastic evolution which requires a probabilistic description and hence consider the set of all possible probability distributions over the ontic states, i.e.,

$$\mathcal{P}(\Lambda) = \left\{\boldsymbol{p}\,:\,\Lambda\longrightarrow[0,1]\,:\,\int_\Lambda \boldsymbol{p}(\lambda)\,d\lambda = 1\right\},\tag{6}$$

which will be called an *epistemic state space*. Accordingly, a general stochastic *transformation (or gate)* is defined as a mapping $T\,:\,\Lambda\longrightarrow\mathcal{P}(\Lambda)$, where $T(\lambda)$ specifies the distribution of final states given the system was in state $\lambda\in\Lambda$. In the model we will be concerned with a limited choice of transformations (gates) which are described below.

B. Local interferometric gates

For such defined ontology we need to specify the stochastic counterparts of the interferometric gates. Note that, to obey the locality principle, action of the gates should be restricted to the paths they are attached to, i.e., to modify degrees of freedom only in the respective paths and the effected transformation not being dependent on the situation (configuration of gates, outcomes, or fields) in the other paths.

We start with the description of paths without gates which correspond to *free evolution*. It will be assumed that the amplitude of the field in such a path remains unchanged and its strength decreases subject to so-called '*natural ageing*'. We make the following definition of *free evolution* in the jth path:

$$u_j \xrightarrow{\textit{free}} u_j \quad\text{and}\quad \tau_j \xrightarrow{\textit{free}} {}^{\tau_j}/_2.\tag{7}$$

Phase shifter S_j is a deterministic gate which acts in the jth path by rotating the phase of the field by $e^{i\omega}$ and the strength '*ageing naturally*', i.e., we have

$$u_j \xrightarrow{S_j} e^{i\omega}u_j \quad\text{and}\quad \tau_j \xrightarrow{S_j} {}^{\tau_j}/_2.\tag{8}$$

Detector D_j checks for the presence of the particle in the jth path (i.e., detector CLICKS only if $q = j$). Furthermore, we postulate that the detection modifies the amplitude and strength of the field in the jth path depending on the result (CLICK or NO CLICK) in the following way:

$$u_j \xrightarrow{D_j} \begin{cases}1 & \text{if}\quad q = j,\\ u_j & \text{if}\quad q\neq j,\end{cases}\tag{9}$$

and

$$\tau_j \xrightarrow{D_j} \begin{cases}1 & \text{if}\quad q = j,\\ 0 & \text{if}\quad q\neq j.\end{cases}\tag{10}$$

In the above definitions it is implicitly assumed that the particle cannot jump between paths. In other words, if the particle happens to be in path $q = j$, then it stays there $q\longrightarrow q$, and otherwise for $q\neq j$ it remains outside $q\longrightarrow q\neq j$. The particle may change its location only at the crossing points, i.e., where the beam splitters are placed, as explained below.

Beam splitter B_{st} is a gate which brings paths s and t together and implements the following transformation. Amplitude and strength of the fields are modified according to the recipe

$$\begin{pmatrix}u_s\\ u_t\end{pmatrix} \xrightarrow{B_{st}} \begin{pmatrix}i\sqrt{R} & \sqrt{T}\\ \sqrt{T} & i\sqrt{R}\end{pmatrix}\begin{pmatrix}\delta_{\tau_s\tau^{(st)}} & 0\\ 0 & \delta_{\tau_t\tau^{(st)}}\end{pmatrix}\begin{pmatrix}u_s\\ u_t\end{pmatrix},\tag{11}$$

and

$$\tau_s,\tau_t \xrightarrow{B_{st}} {}^{\tau^{(st)}}/_2,\quad\text{where }\tau^{(st)} = \max\{\tau_s\,,\tau_t\}.\tag{12}$$

In plain words, the role of Kronecker δ's in the diagonal matrix in Eq. (11) is to suppress the field with weaker strength so that it does not contribute to the transformed amplitudes at the output. Note that the strengths of the outgoing fields are subsequently levelled up to ${}^{\tau^{(st)}}/_2$—see Eq. (12). Additionally, if the particle happens to be in one of the crossing paths, i.e. $q = s$ or $q = t$, then it may change its position following the probabilistic rule

$$q \xrightarrow{B_{st}} \begin{cases}q' = s & \text{with probability}\quad \frac{|u'_s|^2}{|u'_s|^2+|u'_t|^2},\\ q' = t & \text{with probability}\quad \frac{|u'_t|^2}{|u'_s|^2+|u'_t|^2},\end{cases}\tag{13}$$

and otherwise, for $q\neq s$ and $q\neq t$, it remains outside.

All gates defined above are local (with the interaction between the paths on the beam splitter being allowed, since it is placed at the crossing point). We also note that transformations effected by free evolution, phase shifters S_j, and detectors D_j are *deterministic*, while the beam splitters B_{st} are nontrivial *stochastic* gates.

Notice that the structure of circuits constructed in the model is analogous to those in the quantum interferometric framework—see Fig. 1. The difference rests in the underlying ontology, which in the presented model is given explicitly with the locality being built in from the outset. In the following Sec. IV it is shown that the statistical predictions for any experimental circuit in the model are the same as for its quantum-mechanical counterpart.

IV. RECONSTRUCTION OF QUANTUM PREDICTIONS

A. Operational desideratum

Imagine an agent without any prior knowledge of the model trying to understand how it works only by analyzing results of experiments which they can perform. Clearly, their conception of the model may diverge from the '*true*' ontology described above. This is because their choice of gates in constructing experimental circuits is constrained, and hence their knowledge is indirect and to a certain extent incomplete.[1] In the following we are interested in determining a minimal account of the model as seen by the agent, therefore avoiding any unfounded interpretational commitments. For this reason, we adopt an *operational approach* and restrict our attention solely to the description of experimental predictions in the circuits built according to the rules of the model.

To construct such a minimal account, we need to identify what information is in fact available to the agent by investigating the model in every possible way. The following questions provide guidance in this process.

(i) Which distributions in $\mathcal{P}(\Lambda)$ can be prepared by the agent with limited resources at hand?

Generally, it may be the case that the agent explores only a restricted range of distributions in $\mathcal{P}(\Lambda)$, meaning that some distributions are beyond their reach. It is therefore natural to ask the following.

(ii) How do these distributions transform under action of the gates in the model?

What remains is to remove any of the ontological structure that is redundant. Here is the key to the operational account.

Operational indifference principle. *Distributions that are not distinguishable by means available to the agent, that is which give the same probabilistic predictions for any conceivable experiment (circuit), are equivalent from an operational point of view.*

This allows discarding any ontological details which are irrelevant (or inaccessible) to the agent by treating all indistinguishable distributions as a single entity. At this point one should be able to identify the underlying mathematical framework and answer the following question.

(iii) What is the minimal operational account which correctly describes the model's predictions?

In short, we seek for a bare-bones' description without preference to any particular interpretation, with the sole purpose of providing a tool for prediction of experimental results. Such an account should specify the set of possible operational states which correspond to inequivalent preparation procedures and provide transformation rules describing the evolution in conceivable experimental circuits (including measurement outcomes). In the following points, we show how to construct such an operational account of the model which makes no reference to the underlying ontology.

B. Main theorem

Closer analysis of the model reveals the significance of special classes of distributions $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$ labeled with complex vectors (rays) $\vec{z} \in \mathbb{C}^N$, that is

$$\vec{z} = \sum_{j=1}^{N} z_j\, \boldsymbol{e}_j = \begin{pmatrix} z_1 \\ \vdots \\ z_N \end{pmatrix}, \tag{14}$$

with normalization $\|\vec{z}\| = \sum_j |z_j|^2 = 1$ and equivalence up to the overall phase [see Definition 1 in Appendix A 3 b for details]. These classes can be shown to form a disjoint family of subsets in $\mathcal{P}(\Lambda)$, i.e., we have

$$[\,\vec{z}\,] \cap [\,\vec{z}\,'] \neq \varnothing \quad \Leftrightarrow \quad \vec{z} = \vec{z}\,' \text{ (up to phase)}. \tag{15}$$

For explicit construction and proofs, see Appendix A.

Interest in these very special classes of distributions $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$ is due to their behavior under action of the gates defined in the model as explained below.[2] See Fig. 2 for illustration and Appendix A for the proof.

Theorem 1. Transformations implemented by free evolution, phase shifters S_j, detectors D_j, and beam splitters B_{st} act congruently on the family of classes $\{[\,\vec{z}\,] \subset \mathcal{P}(\Lambda) : \vec{z} \in \mathbb{C}^N, \|\vec{z}\| = 1\}$ defined in Eq. (A16). This means that classes transform as a whole, i.e., all distributions in a given class map into distributions in some other class

$$[\,\vec{z}\,] \ni \boldsymbol{p} \longrightarrow \boldsymbol{p}' \in [\,\vec{z}\,'], \tag{16}$$

where the mapping $\vec{z} \longrightarrow \vec{z}\,'$ is determined by the gates implemented in the circuit according to the following rules.

– *Free evolution* acts trivially and *phase shifter S_j* introduces the phase in the relevant component of vector $\vec{z}$,

$$z_j \xrightarrow{\mathit{free}} z_j \quad \text{and} \quad z_j \xrightarrow{S_j} e^{i\omega} z_j. \tag{17}$$

– *Detector D_j* placed in the jth path CLICKS with probability

$$\text{Prob}\,(D_j\,|\,\vec{z}) = |z_j|^2, \tag{18}$$

[1] Following philosophical grounds, it is strongly evocative of Plato's *Allegory of the Cave*, where prisoners chained in a cave experience the outside world only by observing shadows cast on the wall and come up with a distorted picture of the '*true*' reality outside the cave.

[2] We note that Theorem 1 is also valid for parallel configurations of gates, in which case the evolution is described by the joint transformation of vector $\vec{z}$ given by Eqs. (17)–(20); see Theorem 2 in Appendix A 3 d.

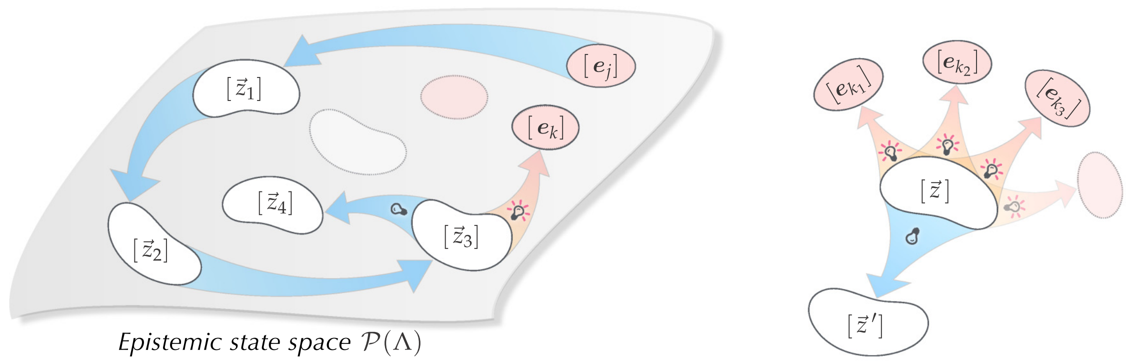

FIG. 2. Evolution of classes $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$. On the left is an illustration of the whole space of probability distributions over Λ, denoted by $\mathcal{P}(\Lambda)$, with disjoint subsets representing classes of interest $[\,\vec{z}\,]$. These classes transform congruently (as a whole) under action of the gates in the model, as explained in Theorem 1. In the picture, the initial class $[\,\boldsymbol{e}_j\,]$ undergoes a sequence of transformations $[\,\boldsymbol{e}_j\,] \to [\,\vec{z}_1] \to [\,\vec{z}_2] \to [\,\vec{z}_3] \prec^{[\,\boldsymbol{e}_k\,]\,\ldots}_{[\,\vec{z}_4]\,\ldots}$. On the right, evolution of a single class $[\,\vec{z}\,]$ depends on the composition of gates at a given time step in the experiment: conditioned on a CLICK in detector D_k we have $[\,\vec{z}\,] \longrightarrow [\,\boldsymbol{e}_k\,]$; otherwise, in a case when all detectors remain silent NO CLICK (or there are no detectors at all) we get $[\,\vec{z}\,] \longrightarrow [\,\vec{z}\,'\,]$ with $\vec{z}\,'$ determined by the implemented gates.

and depending on the outcome implements the projection of vector $\vec{z}$,

$$\vec{z} \xrightarrow{D_j} \begin{cases} \boldsymbol{e}_j & \text{CLICK}, \\ \frac{(\mathbb{1}-\mathbb{P}_j)\,\vec{z}}{\|(\mathbb{1}-\mathbb{P}_j)\,\vec{z}\,\|} & \text{NO CLICK}. \end{cases} \tag{19}$$

– *Beam splitter* B_{st} at the crossing of two paths s and t implements the following unitary on the corresponding components of vector $\vec{z}$,

$$\begin{pmatrix} z_s \\ z_t \end{pmatrix} \xrightarrow{B_{st}} \begin{pmatrix} i\sqrt{R} & \sqrt{T} \\ \sqrt{T} & i\sqrt{R} \end{pmatrix} \begin{pmatrix} z_s \\ z_t \end{pmatrix}. \tag{20}$$

Additionally, we have the following property of possible initial preparations of the system [see Eq. (A19) in Appendix A 3 c].

Observation 1 (Initial preparations). The agent start with distributions contained in one of the initial classes $[\,\boldsymbol{e}_1\,],\ldots,[\,\boldsymbol{e}_N\,]$, where $\boldsymbol{e}_j = (0,\ldots,1,\ldots,0)^T$ has a single 1 in the jth position, indicating the position of the particle (CLICK) ascertained by the initial preparation.

Combining these facts together provides answers to questions *(i)* and *(ii)* from the operational desideratum discussed above. Since the family of classes is closed under available transformations, we infer that an agent with a limited choice of gates at their command remains confined by their explorations to a restricted subset of distributions in $\mathcal{P}(\Lambda)$ given by the union of all classes, i.e.,

$$\bigcup \left\{ [\,\vec{z}\,] \,:\, \vec{z} \in \mathbb{C}^N ,\, \|\vec{z}\| = 1 \right\} \subsetneq \mathcal{P}(\Lambda). \tag{21}$$

See Fig. 2 (on the left). Note that this set has natural coarse graining (partitioning) into classes $[\,\vec{z}\,]$, which have the property that action of the gates in the model is concisely described as transformation of the labeling vectors $\vec{z} \longrightarrow \vec{z}\,'$. A crucial observation is that, on the level of classes, these transformation rules are exactly the same as for the quantum interferometric gates [compare Eqs. (2)–(4) with Eqs. (17)–(20) in Theorem 1].[3]

Such a coarse-grained description is just enough for our purposes. This is because distributions in the same class $[\,\vec{z}\,]$ give identical measurement predictions, i.e., the probability of a CLICK in detector D_j is equal to $|z_j|^2$. Moreover, since classes transform as a whole there is no way for the agent to differentiate between two distributions in the same class by arranging any complicated circuit from the gates available in the model. This permits using the *operational indifference principle* to observe that all information relevant for predicting behavior of a system is held by the class itself, that is knowledge of a particular distribution in $[\,\vec{z}\,]$ is redundant. This means that label $\vec{z}$ plays the role of an *operational state* which encodes complete information available to the agent, thereby answering question *(iii)* from the operational desideratum discussed above. See Fig. 2 for an illustration. Notice that we obtain a full analogy with the quantum description of the interferometric circuits given in Sec. II, i.e., we have the same structure (geometry) of operational states which are complex vectors (rays) in $\mathcal{H} = \mathbb{C}^N$ with identical transformation and measurement rules given in Eqs. (2)–(4) and Eqs. (17)–(20), respectively. All things considered, both descriptions are indistinguishable and hence for all practical purposes we can make the identification

$$|\psi\rangle \overset{equiv}{\rightsquigarrow} [\,\vec{z}\,] \quad (\text{or} \quad \vec{\psi} \overset{equiv}{\rightsquigarrow} \vec{z}\,). \tag{22}$$

In conclusion, we have the following result.

Corollary 1. An operational account of the model boils down to specification of the state given by the complex vector (ray) $\vec{z} \in \mathbb{C}^N$ with the transformation rules and the statistics of outcomes (CLICKS) being the same as for the quantum gates.

This means that for all practical purposes, from the perspective of an agent unaware or indifferent to the underlying ontology, the behavior of quantum interferometric circuits and their counterparts in the presented model are indistinguishable.

[3]Note that this is a generic feature of any ontological model reconstructing quantum theory as discussed in Appendix A 2 b.

V. DISCUSSION

To summarize, we have constructed a local ontological model which faithfully imitates quantum predictions for a single particle in the interferometric circuits. The distinction between the two levels of description is crucial for the analysis of the model. On the one hand, we have an *ontological description* from an omniscient observer having access to all details of the model, i.e., seeing the structure of the ontic state space and being familiar with definitions of the gates. On the other hand, we have an *epistemic description* concerned only with the information which is actually available. The latter adopts an *operational perspective* of an agent unaware of the underlying ontology who investigates the model only with the tools at hand, i.e., building interferometric circuits and analyzing experimental results (statistics of CLICKS). We have shown that the operational predictions of the constructed model are indistinguishable from the quantum-mechanical behavior. This illustrates the fact that properly chosen constraints for gaining knowledge can modify the picture on the epistemic level. In our case, from the local ontology with a classical probabilistic description in $\mathcal{P}(\Lambda)$ we see emergent geometry of the projective space $\mathcal{H} = \mathbb{C}^N$ and the quantum-mechanical account of a qudit [46].

This result is an explicit counterexample showing the impossibility of proving nonlocality for a single particle in the interferometric setups. For the sake of clarity, we address the question of (non)locality in a different context than the Bell-type scenario. The latter is concerned with correlations between measurements on a pair of quantum particles, whereas here we are concerned with a single quantum particle interacting with a classical apparatus (free evolution, phase shifters, beam splitters, and detectors) as described by quantum theory in $\mathcal{H} = \mathbb{C}^N$ without a tensor product structure [44].

At first sight, our conclusion seems to contradict proofs claiming nonlocality of a single particle [10–13]. However, we note that these arguments exploit an additional quantum resource, namely coherent states whose properties rely on superposition of multiparticle states. This requires the presence of other particles in the system making the claim of having a single-particle character of the considered phenomena open to question [12]. A similar objection applies to recent demonstration of the collapse of the wave function using homodyne detection [47]. In view of the presented model, these proofs seem to illustrate the nontrivial aspect of an '*almost*' classical resource provided by local oscillators (coherent states), as compared with '*clean*' single-particle scenarios considered in this paper.

The single-particle framework is a rich source of paradoxes and weird phenomena which are often considered as typically quantum effects without a classical explanation [14–31]. The latter assertion should be treated with caution, since any argument for nonclassicality of an effect always depends on additional assumptions whose plausibility should be properly assessed. For example, interaction-free measurements assume a null effect of negative measurement results [15–17], Leggett-Garg inequalities require noninvasive measurements [22,23], pre- and postselection paradoxes rest upon contextual effects [24,25,48–50], etc. Our model illustrates the nontrivial aspect of these assumptions and shows that mere local state disturbance by detectors ushers in the possibility of a classical-like explanation of single-particle phenomena. A strong point of the model is that the presented ontology is made ready for any kind of circuit with an arbitrary number of paths. As such, it provides exhaustive reconstruction of single-particle phenomena in a unified framework as opposed to separate models devised for simulation of particular effects—cf. [32–42].

Let us remark that in the classification of Harrighan-Spekkens [45] our construction is ψ-ontic (or more precisely ψ-supplemented), that is, distributions corresponding to different quantum states have nonoverlapping supports. Another important feature of the model is that it is not outcome-deterministic. To see this, we observe that measurement of a general observable is implemented by first evolving the system and then registering the outcome by an array of detectors, exactly in the same way as in the real quantum interferometric experiments [44]. In the model, it is the first stage where beam splitters introduce randomness into the measurement outcomes (detectors respond deterministically to the ontic variables which are stochastically distributed due to the first preparatory stage of the measurement procedure). We note that in this way the model goes beyond the standard Kochen-Specker notion of contextuality [9,26–31], which assumes outcome-determinism for every observable (our construction is outcome deterministic only for measurements in the computational basis). It is therefore consistent with the results on contextuality in the ontological model framework [27,28,51], showing that with a lack of outcome-determinism only preparation and transformation contextuality is upheld. Clearly, our model allows for different representations of preparation procedures associated with the same quantum state (preparation contextuality), where any distribution $\boldsymbol{p} \in [\,\vec{z}\,]$ is a valid representation of the same state $|\psi\rangle$ (with the identification $\psi \leftrightsquigarrow \vec{z}$). This variety is necessary to accommodate contextual effects which abound in the quantum regime [9,26–31].

To give a broader perspective, we hasten to note that there is only a handful of ontological models which reconstruct a qudit. One of them is the ψ-ontic model by Beltrametti and Bugajski [52] which essentially is a restatement of the standard Copenhagen interpretation (where nonlocality of the collapse of the wave function is built in from the outset). There is also an interesting proposal by Lewis-Jennings-Barrett-Rudolph [53] built within the framework of ψ-epistemic models. It is explicitly nonlocal and, in addition, violates the so-called preparation independence principle—the latter seems to be a generic feature of any successful ψ-epistemic approach—see Refs. [54,55]. The de Broglie–Bohm interpretation of quantum mechanics [56,57] which postulates local guidance of particles by a quantum potential should also be mentioned. For a single particle quantum potential (directly related to the wave function) lives in a 3D space, and its dependence on the configuration of the apparatus is a source of nonlocal effects. Additionally, the de Broglie–Bohm model has many weird features, such as a complicated spatial description, "surrealistic" trajectories [58] or excessive contextual effects [59], which persist even in the simple interferometric setups whose relevant degrees of freedom reduce to a qudit. In summary, all these models have built-in nonlocal effects in the description and therefore do not make a case against the nonlocality of single-particle interferometry discussed in this paper.

To conclude, let us quote Jaynes [60] on the current understanding of quantum-mechanical formalism: "*But our present QM formalism is not purely epistemological; it is a peculiar mixture describing in part realities of Nature, in part incomplete human information about Nature—all scrambled up by Heisenberg and Bohr into an omelette that nobody has seen how to unscramble. Yet we think that the unscrambling is a prerequisite for any further advance in basic physical theory. For, if we cannot separate the subjective and objective aspects of the formalism, we cannot know what we are talking about; it is just that simple*." Following this line of thought, our model is an illustration of the idea that careful distinction between the epistemic aspect of the description and the underlying ontological account provides a way towards understanding weird quantum phenomena as an effect of incomplete knowledge—which is tenable, at least for the single-particle framework as the model demonstrates. This gives support to the belief that *unscrambling the quantum omelette* should in principle be possible, albeit it is not yet evident how to construct such a theory. It seems that nonlocal effects should play a role in the full reconstruction—as Bell's theorem suggests—however, it is not clear to what extent and in what form (see [61] for some hints). The presented model points to multiparticle behavior as the real source of the quantum mystery in comparison to single-particle phenomena, which are less problematic in this respect. In particular, we have shown that a single-particle framework in itself is not enough to establish quantum nonlocality, since in this case "*spooky action at a distance*" could be understood as an epistemic effect due to a lack of knowledge, with the underlying ontology being local from the construction.

APPENDIX A: METHODOLOGY AND ANALYSIS OF THE MODEL

Here we take an operational perspective of the interferometric experiments and recast our model in the so-called ontological model framework [27,28,45]. Then we construct certain classes of distributions in $\mathcal{P}(\Lambda)$ and analyze, in Theorem 2, how they evolve under transformations implemented by a single step in the interferometric circuit. This result allows one to characterize distributions available to agents exploring the model and reveals the geometric pattern of transformations and measurements which is instrumental for the discussion in Sec. IV.

1. Interferometric experiment: Operational account

It is convenient to think of an experiment in terms of preparation, transformation, and measurement procedures which in the interferometric setups build up in a series of sequential steps. Accordingly, each experiment starts with some well-defined *initial preparation* $\mathcal{P}_j^{in}$ which ascertains the presence of a particle in a given path $j = 1,\dots,N$. Then the circuit implements a sequence of transformations, with each step specified by the parallel configuration of gates acting in *different* paths at the same time as described in Sec. II. See Fig. 3 for illustration. For the purpose at hand we will characterize each step by grouping paths with gates of the same kind and define the respective subsets as

$$\begin{aligned}
&\mathcal{F} - \text{paths without gates (empty paths)},\\
&\mathcal{D} - \text{paths with detectors},\\
&\mathcal{S} - \text{paths with phase shifters},\\
&\mathcal{B} - \text{pairs of paths crossing on beam splitters.}
\end{aligned} \tag{A1}$$

In the following, we refer to each step in the circuit by specifying parallel configuration of gates by the partition $\mathcal{G} \equiv \{\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\}$. In a case when $\mathcal{D} = \varnothing$, we denote the corresponding *transformation* by $\mathcal{T}^{\mathcal{G}}$. In a case when there are one or more detectors in the paths (i.e., $\mathcal{D} \neq \varnothing$), we speak of a *measurement* at the given time step and denote it by $\mathcal{M}^{\mathcal{G}} \equiv \{\mathcal{M}_k^{\mathcal{G}}\}$, where $k \in \mathcal{D} \cup \{\varnothing\}$ labels possible outcomes with $k \in \mathcal{D}$ associated to a CLICK in the respective detector D_k and $k = \varnothing$ standing for the negative measurement result (i.e., none of the detectors CLICK).

Clearly, the definition of preparations $\mathcal{P}$ extends to any sequence of transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$ which start from a given initial preparation $\mathcal{P}_j^{in}$. Similarly, general transformation $\mathcal{T}$ is understood as any sequence of transformations $\mathcal{T}^{\mathcal{G}}$, and a general measurement $\mathcal{M}$ obtains by first transforming the system by $\mathcal{T}$ and then measuring $\mathcal{M}^{\mathcal{G}}$. It was shown in Ref. [44] that a single-particle interferometric framework is a practical way of physical realization of a qudit, i.e., every quantum state $|\psi\rangle$, unitary U, and PVM $\{\Pi_k\}$ in $\mathcal{H} = \mathbb{C}^N$ can be realized as an appropriate preparation $\mathcal{P}$, transformation $\mathcal{T}$, and measurement $\mathcal{M}$ procedure in the N-path interferometric circuit.

Let us be more specific about each step in the interferometric circuit. See Fig. 3 for illustration. Quantum theory describes initial preparation $\mathcal{P}_j^{in}$ by vector $|j\rangle$ from the computational basis $|1\rangle,\dots,|N\rangle$, and general preparations $\mathcal{P}$ are associated with vectors (rays) $|\psi\rangle$ in $\mathcal{H} = \mathbb{C}^N$—cf. Eq. (1). According to the definitions of interferometric gates in Eqs. (2)–(4) transformation $\mathcal{T}^{\mathcal{G}}$ is described as follows:

$$|\psi\rangle \xrightarrow{\mathcal{T}^{\mathcal{G}}} |\psi'\rangle = \prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}|\psi\rangle. \tag{A2}$$

For a measurement $\mathcal{M}^{\mathcal{G}} \equiv \{\mathcal{M}_k^{\mathcal{G}}\}$, we have

$$|\psi\rangle \xrightarrow{\mathcal{M}_k^{\mathcal{G}}} |\psi'\rangle = |k\rangle, \tag{A3}$$

if detector D_k CLICKS, which happens with probability equal to $\text{Prob}\,(D_k|\,\psi) = |\psi_k|^2$. Otherwise, in case of negative measurement result denoted by $\varnothing$ (i.e., NO CLICK in all detectors D_k for $k \in \mathcal{D}$), we get[4]

$$|\psi\rangle \xrightarrow{\mathcal{M}_\varnothing^{\mathcal{G}}} |\psi'\rangle \sim \prod_{j\in\mathcal{D}}(\mathbb{1} - \mathbb{P}_j) \prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}|\psi\rangle, \tag{A4}$$

which happens with probability equal to $\text{Prob}\,(\varnothing|\,\psi) = 1 - \sum_{k\in\mathcal{D}} |\psi_k|^2$. Above we have used matrix representation of the

[4]To simplify the notation we use the proportionality symbol "$\sim$" which expresses the need of subsequent renormalization $|\psi'\rangle \rightsquigarrow {}^{|\psi'\rangle}/_{\||\psi'\rangle\|}$ due to nonunitary projections in the representation of measurements—cf. Eq. (4).

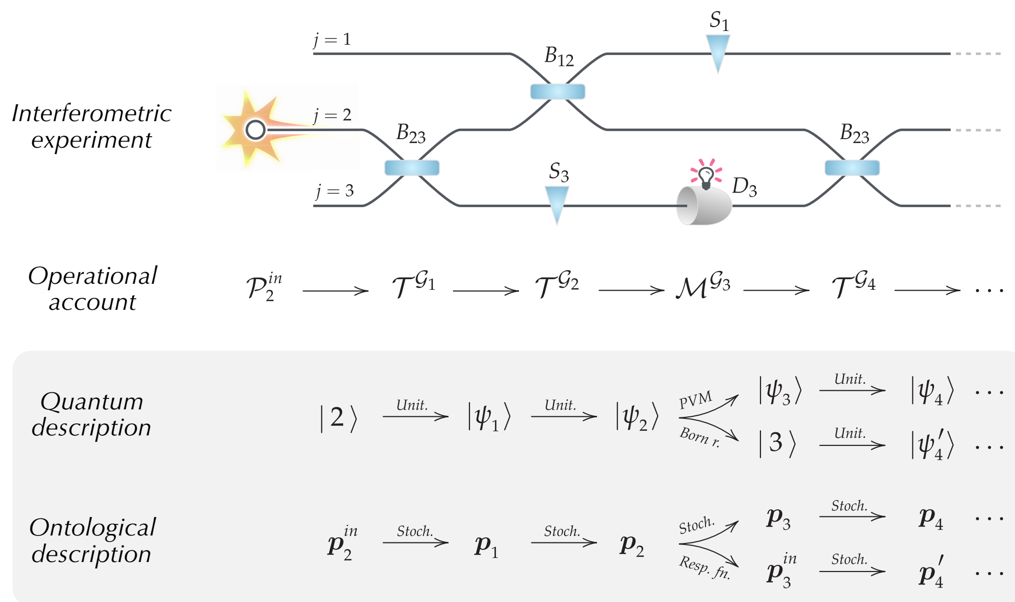

FIG. 3. Operational account of interferometric experiment. A source ascertains the presence of a particle in a given path $j = 2$. After initial preparation $\mathcal{P}_j^{in}$ the system undergoes a sequence of transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$ which depend on the composition of gates $\mathcal{G}$ at a given time in the circuit. In the box at the bottom, a schematic illustration of quantum vs ontological model description is given. *Quantum theory* describes preparation procedures by state vectors $|\psi\rangle \in \mathbb{C}^N$ (with initial preparations associated with vectors in the computational basis), transformations $\mathcal{T}^{\mathcal{G}}$ are represented by unitaries U and measurements $\mathcal{M}^{\mathcal{G}} \equiv \{\mathcal{M}_k^{\mathcal{G}}\}$ by PVM's $\{\Pi_k\}$, with Born's rule describing the statistics of outcomes—see Eqs. (A2)–(A5). The *ontological model* describes the experiment by evolution of a distribution $\boldsymbol{p} \in \mathcal{P}(\Lambda)$ over the underlying ontic state space Λ which starts from some well-defined initial distribution $\boldsymbol{p}_j^{in}$ with transformations $\mathcal{T}^{\mathcal{G}}$ represented by stochastic mappings $\Gamma_{\mathcal{T}^{\mathcal{G}}}$ and measurements $\mathcal{M}^{\mathcal{G}} \equiv \{\mathcal{M}_k^{\mathcal{G}}\}$ represented by collections of stochastic mappings $\{\Gamma_{\mathcal{M}^{\mathcal{G}},k}\}$ with outcomes specified by the response functions $\xi_k^{\mathcal{M}^{\mathcal{G}}}$—see Eqs. (A7)–(A10).

respective gates in $\mathcal{H} = \mathbb{C}^N$, which follows from Eqs. (2)–(4), i.e.,

$$\mathbb{P}_j = \begin{pmatrix} 0 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 0 \end{pmatrix}, \quad \mathbb{S}_k = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & e^{i\omega} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix},$$

$$\mathbb{B}_{st} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & i\sqrt{R} & \dots & \sqrt{T} & & \\ & & \vdots & \ddots & \vdots & & \\ & & \sqrt{T} & \dots & i\sqrt{R} & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}, \quad \text{(A5)}$$

where $\mathbb{P}_j \equiv |j\rangle\langle j|$ is a projector on the jth component, $\mathbb{S}_k$ stands for phase shifter S_k which introduces phase $e^{i\omega}$ in the kth component leaving the remaining ones unchanged, and $\mathbb{B}_{st}$ represents beam splitter B_{st} mixing components $\{s,t\}$ without affecting the rest with some real coefficients R and T satisfying $R + T = 1$. Clearly, matrices $\mathbb{S}_j$ and $\mathbb{B}_{st}$ are unitary.

Any successful model of an interferometric experiment should have the same operational structure of preparations, transformations, and measurements encoded in the sequential design of circuits and recover quantum predictions as described in Eqs. (A2)–(A4).

2. Ontological model framework: General remarks

In this work we follow the ontological model framework which recasts hidden variable models in the operational setting [27,28,45]. It assumes that the system is always in a well-defined state picked from the underlying ontic state space, i.e., $\lambda \in \Lambda$, but due to constraints depending on the specifics of the model's preparations, transformations and measurement procedures are described in probabilistic terms. Accordingly, in the ontological model each preparation procedure $\mathcal{P}$ is specified by a distribution $\boldsymbol{p}_{\mathcal{P}} \in \mathcal{P}(\Lambda)$, where $\mathcal{P}(\Lambda)$ denotes the space of probability distributions over Λ. Transformation $\mathcal{T}$ is represented stochastically by a mapping $\Gamma_{\mathcal{T}} : \Lambda \longrightarrow \mathcal{P}(\Lambda)$ with $\Gamma_{\mathcal{T}}(\lambda)(\lambda')$ describing probability of transition from state λ to state λ', i.e., a system prepared in $\boldsymbol{p}_{\mathcal{P}}$ after the transformation is described by $\boldsymbol{p}_{\mathcal{P}}(\lambda) \longrightarrow \boldsymbol{p}_{\mathcal{P}'}(\lambda') = \int_\lambda \Gamma_{\mathcal{T}}(\lambda)(\lambda')\, \boldsymbol{p}(\lambda)\, d\lambda$. Measurement $\mathcal{M} \equiv \{\mathcal{M}_k\}$ is modeled via response functions $\xi_k^{\mathcal{M}}(\lambda)$ of the apparatus to the underlying ontic state, i.e., for a system described by distribution $\boldsymbol{p}_{\mathcal{P}}$ probability of outcome k is given by $\text{Prob}\,(k|\mathcal{P},\mathcal{M}) = \int_\Lambda \xi_k^{\mathcal{M}}(\lambda)\, \boldsymbol{p}_{\mathcal{P}}(\lambda)\, d\lambda$ and after the measurement the state transforms via the associated stochastic mapping $\Gamma_{\mathcal{M},k} : \Lambda \longrightarrow \mathcal{P}(\Lambda)$ which depends on the outcome k.[5]

[5]Clearly, a probabilistic framework imposes natural constraints on the components of the models. We require that for preparations $\boldsymbol{p}_{\mathcal{P}}(\lambda) \geq 0$ and $\int_\Lambda \boldsymbol{p}_{\mathcal{P}}(\lambda)\, d\lambda = 1$, for transformations $\Gamma_{\mathcal{T}}(\lambda)(\lambda') \geq 0$ and $\int_\Lambda \Gamma_{\mathcal{T}}(\lambda)(\lambda')\, d\lambda = 1\ \forall\, \lambda'$, and for measurements $\xi_k^{\mathcal{M}}(\lambda) \geq 0$ and $\sum_k \xi_k^{\mathcal{M}}(\lambda) = 1\ \forall\, \lambda$.

Operational interpretation of a physical theory treats preparations, transformations, and measurements as primitive elements. It is merely concerned with specification of probabilities $\mathrm{Prob}\,(k|\mathcal{P},\mathcal{T},\mathcal{M})$ of different outcomes k resulting from a given measurement procedure $\mathcal{M} \equiv \{\mathcal{M}_k\}$ for a system prepared according to procedure $\mathcal{P}$ and transformation procedure $\mathcal{T}$ without further inquiry about the underlying reality. For example, quantum theory formulated in operational terms associates preparations $\mathcal{P}$ with kets $|\psi\rangle$ in the Hilbert space $\mathcal{H}$, while transformations $\mathcal{T}$ are represented by unitaries $|\psi\rangle \longrightarrow U|\psi\rangle$ and measurements $\mathcal{M} \equiv \{\mathcal{M}_k\}$ are described by PVMs $\{\Pi_k\}$. The statistics of outcomes is given by Born's rule $\mathrm{Prob}(k|\psi) = |\langle\psi|\Pi_k|\psi\rangle|^2$ with the von Neumann–Lüders rule $|\psi\rangle \xrightarrow{k} \Pi_k|\psi\rangle/\|\Pi_k|\psi\rangle\|$ specifying postmeasurement states—cf. Fig. 3 (at the bottom).

The ontological model framework seeks to explain predictions of operational theory by postulating the existence of ontic states $\lambda \in \Lambda$ (hidden variables) which mediate between the consecutive stages of the experiment according to the prescriptions given above. For example, in the ontological model of quantum theory the statistics of outcomes which are obtained from the formula $\mathrm{Prob}(k|\mathcal{P},\mathcal{T},\mathcal{M}) = \int_\Lambda \xi_k^{\mathcal{M}}(\lambda')\, \Gamma_{\mathcal{T}}(\lambda)(\lambda')\, \boldsymbol{p}_{\mathcal{P}}(\lambda)\, d\lambda\, d\lambda'$ should reproduce the quantum statistics given by $\mathrm{Prob}(k|\psi,U,\{\Pi_k\}) = |\langle\psi|U^\dagger\Pi_k U|\psi\rangle|^2$ with state $|\psi\rangle$, unitary U, and PVM $\{\Pi_k\}$ representing the respective preparation $\mathcal{P}$, transformation $\mathcal{T}$, and measurement $\mathcal{M}$ procedures—cf. Fig. 3 (at the bottom).

a. Contextuality issue

In the operational formulation of a theory there is a natural notion of equivalence which groups preparations, transformations, and measurements into equivalence classes of indistinguishable procedures, i.e., two procedures are equivalent if their exchange does not change statistical predictions in any conceivable experiment [27]. This makes the specifics of equivalent procedures redundant and it is enough to keep the label of the whole class to derive predictions of the theory. Note that this is exactly how the quantum theory is laid out: there are many ways to prepare a system in state $|\psi\rangle$, realize unitary transformation U, or implement a measurement PVM $\{\Pi_k\}$, and quantum theory only uses the class labels $|\psi\rangle$, U, and $\{\Pi_k\}$ to provide the statistics of outcomes in any experiment.

In the ontological model we can have two situations. If equivalent procedures are represented in the same way, i.e., for preparations by the same distribution $\boldsymbol{p}_{\mathcal{P}} \in \mathcal{P}(\Lambda)$, for transformations by the same stochastic mapping $\Gamma_{\mathcal{T}} : \Lambda \longrightarrow \mathcal{P}(\Lambda)$, and for measurements by the same response function $\xi_k^{\mathcal{M}}(\lambda)$, then the model is called *noncontextual*. Otherwise, when the representation within the equivalence class varies, then the model is *contextual*. One can make a further distinction between preparation, transformation, and measurement (non)contextuality if the situation concerns the respective type of procedure. It has been shown that quantum theory is preparation and transformation contextual [27,28,51] and measurement contextual [26,29–31] (with the latter under additional assumption of outcome-determinism).

b. Generic structure of distributions in $\mathcal{P}(\Lambda)$

For the purpose at hand let us look closer into the contextual structure of preparations and group together distributions $\boldsymbol{p}_{\mathcal{P}}$ associated with equivalent preparation procedures into separate classes in $\mathcal{P}(\Lambda)$. Since we are interested in modeling the quantum theory, in which case equivalent preparations are represented by states $|\psi\rangle$, we will denote the corresponding class of distributions with the same label ψ in square brackets and define

$$[\psi] \equiv \big\{\boldsymbol{p}_{\mathcal{P}} :\ \mathcal{P} \text{ is represented by } |\psi\rangle\big\} \subset \mathcal{P}(\Lambda). \tag{A6}$$

We can make three simple observations which are valid in every ontological model that are derived from the very notion of equivalent preparations. First, these classes are *disjoint* in $\mathcal{P}(\Lambda)$, i.e., $[\psi] \cap [\psi'] \neq \varnothing \Leftrightarrow |\psi\rangle \neq |\psi'\rangle$. Second, these classes transform *congruently* $[\psi] \longrightarrow [\psi']$ under any transformation $\mathcal{T}$ or measurement $\mathcal{M}$. This is derived from the fact that distributions in the same equivalence class correspond to indistinguishable preparations and therefore after any experimental procedure have to be indistinguishable as well, i.e., belong to the same class again $[\psi] \ni \boldsymbol{p}_1, \boldsymbol{p}_2 \longrightarrow \boldsymbol{p}_1', \boldsymbol{p}_2' \in [\psi']$. Third, since the model reproduces quantum predictions, the mapping $\psi \longrightarrow \psi'$ must *follow the rules* of quantum theory with the unitaries U and PVMs $\{\Pi_k\}$ associated to the respective transformations $\mathcal{T}$ and measurements $\mathcal{M}$ with the statistics of outcomes for every distribution $\boldsymbol{p}_{\mathcal{P}} \in [\psi]$ given by Born's rule.

This is a generic situation for any ontological model of quantum theory, i.e., on the level of classes $[\psi] \subset \mathcal{P}(\Lambda)$ the structure is always the same, no matter what the underlying ontology is. This comes from the different perspectives taken by the operational and ontological interpretations of the theory. Namely, from the operational point of view only the relations between whole equivalence classes $[\psi]$ are important (which provides enough information to predict experimental results), while the ontological account is concerned with the fine-grained description in terms of distributions in $\mathcal{P}(\Lambda)$ (which explain behavior of the system by the properties of the underlying ontic state $\lambda \in \Lambda$).

We shall take these observations as a clue in the analysis of the model described in the paper. Namely, by having a well-defined ontology of the model we will characterize distributions accessible to agents investigating the model, distinguish classes of indistinguishable distributions, and analyze their transformation properties. It should be enough to come up with the minimal operational account of the model which will be compared with the quantum description of interferometric experiments; see Sec. IV.

3. Analysis of the model

We begin with a brief recount of the model presented in the paper in the ontological model framework. We then distinguish certain classes of distributions in $\mathcal{P}(\Lambda)$ which will be used to characterize the full set of distributions available to agents exploring the model (i.e., distributions describing general preparation procedures $\mathcal{P}$ which can be prepared in conceivable experiments), and particularly identify distributions associated with initial preparation procedures $\mathcal{P}_j^{in}$. Finally, we state our main result, Theorem 2, concerning transformation properties of these classes following the evolution implemented at each step of the interferometric experiment.

a. Ontological model of interferometric experiments

Let us be more precise concerning the ontological framework for the model described in the paper. Note that our construction is made to exactly mimic the design of interferometric experiments. That is, each experiment starts from some initial preparation procedure $\mathcal{P}_j^{in}$ which asserts the presence of a particle (CLICK) in a given path followed by a sequence of transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$; see Appendix A 1 and Fig. 3 for illustration. These components determine the operational structure of the model. The ontology postulated in the model is defined in Eq. (5) and interpreted as a single particle (with position q) and fields (characterized by amplitudes u_j and strengths τ_j) propagating locally through the circuit. Accordingly, the state of the system at a given time is fully characterized by the ontic state $\lambda = (q,\vec{u},\vec{\tau}) \in \Lambda$. Definitions of interferometric gates in Eqs. (7)–(13) are crucial for our construction. All these gates act locally in the respective paths and combine into transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$ implemented in the consecutive steps of the interferometric experiment.

To be complete, we will write out explicitly in the ontological framework the form of transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$ implemented by a parallel configuration of gates $\mathcal{G} = \{\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\}$ corresponding to a single step in the circuit. For transformation $\mathcal{T}^{\mathcal{G}}$ (i.e., no detectors $\mathcal{D} = \varnothing$) we get deterministic behavior

$$\Gamma_{\mathcal{T}^{\mathcal{G}}}(\lambda) = \delta_{(q,\vec{u}',\vec{\tau}')}, \tag{A7}$$

if $q \notin \bigcup \mathcal{B}$ (i.e., when the particle does not hit any of the beam splitters), and otherwise due to the stochastic behavior in Eq. (13), we have

$$\Gamma_{\mathcal{T}^{\mathcal{G}}}(\lambda) = \frac{|u'_s|^2}{|u'_s|^2 + |u'_t|^2}\, \delta_{(s,\vec{u}',\vec{\tau}')} + \frac{|u'_t|^2}{|u'_s|^2 + |u'_t|^2}\, \delta_{(t,\vec{u}',\vec{\tau}')}, \tag{A8}$$

if $q \in \{s,t\} \in \bigcup \mathcal{B}$ (i.e., when the particle hits the beam splitter B_{st}). In both cases components of vectors $\vec{u}'$ and $\vec{\tau}'$ are the same—they are determined by the configuration of gates $\mathcal{G}$ and given by the definitions in Eqs. (7)–(12). For the above we have used standard notation for the Dirac δ-functions $\delta_{\lambda'}(\lambda) \equiv \delta(\lambda - \lambda') \equiv \delta(q - q')\,\delta(\vec{u} - \vec{u}')\,\delta(\vec{\tau} - \vec{\tau}')$. Now, for measurement $\mathcal{M}^{\mathcal{G}} \equiv \{\mathcal{M}_k^{\mathcal{G}}\}$ (i.e., when detectors are present $\mathcal{D} \neq \varnothing$) response functions associated with the respective outcomes $k \in \mathcal{D} \cup \{\varnothing\}$ are as follows:

$$\xi_k^{\mathcal{M}^{\mathcal{G}}}(\lambda) = \begin{cases} 1 & \text{if } q = k, \\ 0 & \text{if } q \neq k, \end{cases} \tag{A9}$$

for $k \in \mathcal{D}$ (associated with CLICK or NO CLICK in detector D_k), and for the negative measurement result $k = \varnothing$ (none of the detectors CLICK), we have

$$\xi_\varnothing^{\mathcal{M}^{\mathcal{G}}}(\lambda) = \begin{cases} 1 & \text{if } q \notin \mathcal{D}, \\ 0 & \text{if } q \in \mathcal{D}. \end{cases} \tag{A10}$$

Notice that the response functions at a given time step are deterministic[6] and no coincidences in different detectors occur. Finally, evolution of the system after the measurement $\Gamma_{\mathcal{M}^{\mathcal{G}},k}(\lambda)$ is described by the same Eqs. (A7) and (A8), where the dependence on the outcome k is implicit in the description of the detectors in Eqs. (9) and (10).

Having defined all these components, it is possible using the standard rules of probability calculus to trace the evolution of any distribution in $\mathcal{P}(\Lambda)$ via any complicated circuit one can think of. In the following, we will be interested in some specific initial distributions associated with the initial preparation procedures $\mathcal{P}_j^{in}$ which have to be explicitly characterized in the model. It is convenient to postpone their description until we will have defined in the next section some special classes of distributions in $\mathcal{P}(\Lambda)$; see Appendix A 3 c.

In order to say that the ontological model described in the paper reconstructs behavior of a single quantum particle in the interferometric circuit we need to show that the statistics of outcomes in every conceivable experiment, calculated with the rules of Eqs. (A7)–(A10), are exactly the same as those calculated by the rules of quantum theory, Eqs. (A2)–(A4), for the corresponding gates and initial preparations $\mathcal{P}_j^{in}$. In other words, this means that on an operational level, i.e., disregarding the underlying ontology, predictions of the model are indistinguishable from the predictions of quantum theory. Our discussion in the paper is based on the operational indifference principle desideratum described in Sec. IV A. It consists in the analysis of a full set of distributions in $\mathcal{P}(\Lambda)$ accessible to agents exploring the model (i.e., distributions describing general preparation procedures $\mathcal{P}$ which can be prepared in conceivable experiments). In the following part we characterize this set which appears to be composed from classes of distributions with certain properties, and investigate its structure under transformations $\mathcal{T}^{\mathcal{G}}$ and measurements $\mathcal{M}^{\mathcal{G}}$ given in Eqs. (A7)–(A10).

b. Classes of interest in $\mathcal{P}(\Lambda)$

What is crucial for the analysis of the model is some distinguished classes of distributions $[\,\vec{z}\,]$ in $\mathcal{P}(\Lambda)$ labeled with (normalized) complex vectors $\vec{z} \in \mathbb{C}^N$. Our construction proceeds in three steps. In ***Step 1*** certain subsets in the ontic state space Λ are defined, which are then used in ***Step 2*** as the supports of auxiliary distributions in $\mathcal{P}(\Lambda)$. Finally, in ***Step 3*** probabilistic mixtures of the latter will define the distributions in classes of interest $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$. See Fig. 4 for illustration.

Step 1. Let us consider special subsets of the ontic state space $\Lambda_{\vec{z}}^i \subset \Lambda$ labeled by integers $i \in \{1,\ldots,N\}$ and complex vectors $\vec{z} \in \mathbb{C}^N$ defined as follows:

$$(q,\vec{u},\vec{\tau}) \in \Lambda_{\vec{z}}^i \overset{def}{\iff} \begin{cases} (a) & q = i, \\ (b) & \tau_i = \tau > 0, \\ (c) & \Delta_\tau\,\vec{u} \sim \vec{z}, \end{cases} \tag{A11}$$

where $\Delta_\tau\,\vec{u}$ is a vector obtained from $\vec{u}$ by retaining field amplitudes corresponding to the highest field strength

$$\tau := \max\{\tau_1,\ldots,\tau_N\},$$

[6] However, for the general measurement procedure $\mathcal{M}$ the situation is different. If the system is subject to some transformation $\mathcal{T}$ before the measurement $\mathcal{M}^{\mathcal{G}}$, then in general, due to the stochastic character of transformations, the response function for such a compound measurement is probabilistic as well.

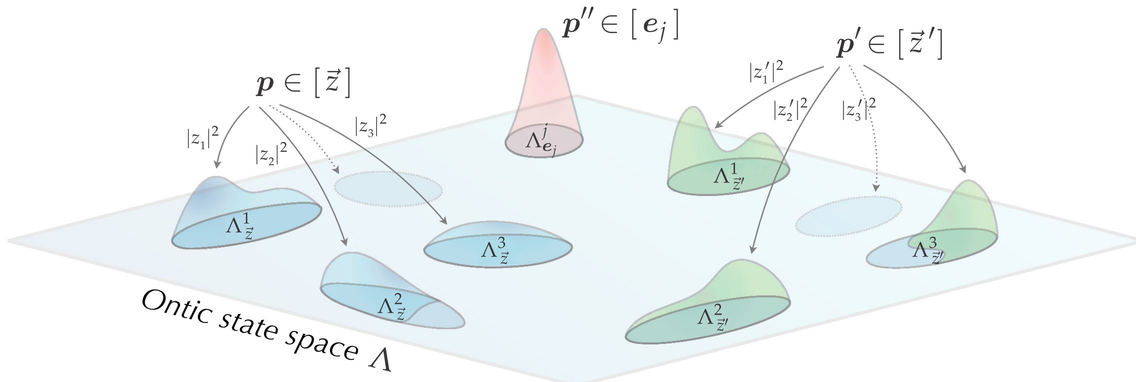

FIG. 4. Construction of classes of interest $[\vec{z}\,] \subset \mathcal{P}(\Lambda)$. Distributions $\boldsymbol{p} \in [\,\vec{z}\,]$ are defined to have support in $\bigcup_{i=1}^N \Lambda^i_{\vec{z}}$ with cumulative probability over the respective subsets $\Lambda^i_{\vec{z}}$ equal to $|z_i|^2$, where $\vec{z} \in \mathbb{C}^N$ is a normalized vector defined up to the overall phase—see Definition 1 and construction in ***Steps 1–3***. Since all $\Lambda^i_{\vec{z}}$ are disjoint, distributions in different classes have nonoverlapping supports, e.g., compare $\boldsymbol{p} \in [\,\vec{z}\,]$ and $\boldsymbol{p}' \in [\,\vec{z}\,']$. Among all these classes there are a few special ones $[\,\boldsymbol{e}_j\,]$ for $j = 1, \ldots, N$, which describe initial preparations $\mathcal{P}^{in}_j$ with the particle in a given path.

and the remaining ones equal to zero. Technically, it is implemented by the diagonal matrix

$$\Delta_\tau := \mathrm{diag}\,(\delta_{\tau_1\tau}, \ldots, \delta_{\tau_N\tau}), \tag{A12}$$

which picks out those entries of $\vec{u}$ which correspond to the highest strength τ. In our notation, the symbol "$\sim$" stands for proportionality, i.e., $\vec{z} \sim \vec{z}\,'$ iff $\vec{z} = \alpha\,\vec{z}\,'$ for some $\alpha \in \mathbb{C}$, $\alpha \neq 0$. In plain words, these conditions express the following requirements:

(a) the particle is present in the ith path,

(b) the field in the ith path has maximal strength (nonvanishing, possibly equal to strengths in other paths), and

(c) the vector of field amplitudes with highest strengths $\Delta_\tau\,\vec{u}$ is proportional to $\vec{z}$.

Clearly, for different labels i and $\vec{z}$ (up to proportionality) these subsets are disjoint, i.e., we have

$$\Lambda^i_{\vec{z}} \cap \Lambda^j_{\vec{z}\,'} \neq \varnothing \quad \Leftrightarrow \quad i = j \quad \text{and} \quad \vec{z} \sim \vec{z}\,', \tag{A13}$$

and $\Lambda^i_{\vec{z}} = \Lambda^i_{\vec{z}\,'}$ for $\vec{z} \sim \vec{z}\,'$.

Step 2. Then, we introduce auxiliary classes of probability distributions with support in $\Lambda^i_{\vec{z}}$ and denote

$$[\,\vec{z}\,]_i := \left\{\boldsymbol{p} \in \mathcal{P}(\Lambda) : \,\mathrm{supp}\,\boldsymbol{p} \subset \Lambda^i_{\vec{z}}\right\} \subset \mathcal{P}(\Lambda). \tag{A14}$$

By virtue of Eq. (A13) these classes form a disjoint family of subsets in $\mathcal{P}(\Lambda)$, i.e., we have

$$[\,\vec{z}\,]_i \cap [\,\vec{z}\,']_j \neq \varnothing \quad \Leftrightarrow \quad i = j \quad \text{and} \quad \vec{z} \sim \vec{z}\,', \tag{A15}$$

and $[\,\vec{z}\,]_i = [\,\vec{z}\,']_i$ for $\vec{z} \sim \vec{z}\,'$.

Step 3. Now, we are ready to define classes of distributions which play a central role in our analysis of the model.

Definition 1. With each normalized vector $\vec{z} \in \mathbb{C}^N$, such that $\|\vec{z}\| := \sum_{i=1}^N |z_i|^2 = 1$, we associate the following class of probability distributions:

$$[\,\vec{z}\,] := \left\{ \sum_{i=1}^N |z_i|^2\, \boldsymbol{p}_i \,:\, \boldsymbol{p}_i \in [\,\vec{z}\,]_i \right\} \subset \mathcal{P}(\Lambda). \tag{A16}$$

See Fig. 4 for illustration. This means that distributions in $[\,\vec{z}\,]$ have support in $\bigcup_{i=1}^N \Lambda^i_{\vec{z}}$ with cumulative probability over the respective subsets $\Lambda^i_{\vec{z}}$ equal to $|z_i|^2$ (otherwise the shape of distributions are arbitrary). Another way to characterize classes of interest is to write $[\,\vec{z}\,] = \sum_{i=1}^N |z_i|^2\,[\,\vec{z}\,]_i$, which means that its elements are convex combinations of distributions in $[\,\vec{z}\,]_i$'s with weights $|z_i|^2$. As a consequence of Eq. (A15) we observe that such defined classes are disjoint subsets in $\mathcal{P}(\Lambda)$, i.e., we have

$$[\,\vec{z}\,] \cap [\,\vec{z}\,'] \neq \varnothing \quad \Leftrightarrow \quad \vec{z} \sim \vec{z}\,', \tag{A17}$$

and $[\,\vec{z}\,] = [\,\vec{z}\,']$ for $\vec{z} \sim \vec{z}\,'$.

c. Initial preparations

Any prediction of experimental behavior rests upon knowledge of initial preparation of the system. In general, it is an intrinsic characteristic of the source which provides an ensemble of systems with a given distribution of the ontic states. However, if no information about the source is available, then the agent who is given some unknown (possibly random) source must prepare initial ensembles of the systems on their own. Here is a generic scheme of how to proceed in such a case.

Since we are interested in single-particle scenarios, first the presence of a single particle (CLICK) in the system should be verified. This property can be confirmed by sieving an unknown ensemble through the array of detectors $D_1, D_2, \ldots, D_N$ placed in each path and retaining only those cases when a single detection occurred. In this way the agent carries out an effective initial preparation $\mathcal{P}^{in}_j$ which attests to the presence of a single particle (CLICK) in a given path $j = 1, \ldots, N$. Note that, on the ontological level, selection of events with a single CLICK in detector D_j results in an ensemble distributed over the ontic states $(q, \vec{u}, \vec{\tau}) \in \Lambda$ subject to the following conditions:

$$\begin{aligned} q = j, \quad u_j &= 1, \quad \tau_j = 1, \\ u_k &= ?, \quad \tau_k = 0, \quad \text{for } k \neq j, \end{aligned} \tag{A18}$$

where u_k's depend on the unknown source; see Eqs. (9) and (10). A quick look at definitions in Eqs. (A11), (A14), and (A16) reveals that such distributions have support in $\Lambda^j_{\boldsymbol{e}_j}$, and hence are included in class

$$[\,\boldsymbol{e}_j\,] \subset \mathcal{P}(\Lambda) \quad (\text{if } D_j \text{ CLICKS}), \tag{A19}$$

where $\boldsymbol{e}_j = (0,\ldots,1,\ldots,0)^T$ has a single 1 in the jth position. In conclusion, the agent starts in one of the classes $[\,\boldsymbol{e}_1\,],\ldots,[\,\boldsymbol{e}_N\,]$, which correspond to the initial preparation of the system with a single particle (CLICK) in a given path. It is the content of Observation 1 is Sec. IV B.[7]

d. Operational account of the model: Geometry of classes

It appears that the structure of classes $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$ given in Definition 1, Eq. (A16), is closed under transformations (circuits) considered in the model. Here is the key result describing behavior of these classes under action of the gates for any transformation $\mathcal{T}^{\mathcal{G}}$ and measurement $\mathcal{M}^{\mathcal{G}}$ as defined by Eqs. (A7)–(A10). See Fig. 2 for illustration.

Theorem 2. Parallel configuration of gates $\mathcal{G}$ acts congruently on the family of classes $\{[\,\vec{z}\,] \subset \mathcal{P}(\Lambda) : \vec{z} \in \mathbb{C}^N,\ \|\vec{z}\| = 1\}$ defined in Eq. (A16). This means that each class transforms as a whole with all distributions in a given class mapping into distributions in some other class

$$[\,\vec{z}\,] \ni \boldsymbol{p} \longrightarrow \boldsymbol{p}' \in [\,\vec{z}\,'\,], \tag{A20}$$

where the mapping $\vec{z} \longrightarrow \vec{z}\,'$ depends on the configuration of gates $\mathcal{G} \equiv \{\,\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\,\}$ and measurement outcomes (CLICKS). It is specified by the following rules.

– For the system described by a distribution in class $[\,\vec{z}\,]$ detector D_j CLICKS with probability

$$\text{Prob}\,(D_j|\,\vec{z}\,) = |z_j|^2\,, \tag{A21}$$

and conditioning (postselecting) on a CLICK in D_j leaves the system in a state described by a distribution $\boldsymbol{p}' \in [\,\boldsymbol{e}_j\,]$, i.e.,

$$\vec{z} \xrightarrow{D_j} \vec{z}\,' = \boldsymbol{e}_j \quad (\text{CLICK}). \tag{A22}$$

With each run of the experiment, either one of the detectors CLICK or all detectors remain silent (negative measurement result), with the latter happening with probability $\text{Prob}\,(\varnothing|\,\vec{z}\,) = 1 - \sum_{j\in\mathcal{D}} |z_j|^2$.

– In the case of a negative measurement result (NO CLICK in all detectors D_j for $j \in \mathcal{D}$) or no measurement at all (no detectors $\mathcal{D} = \varnothing$) transformation implemented by the gates is given by

$$\vec{z} \longrightarrow \vec{z}\,' \sim \prod_{j\in\mathcal{D}} (\mathbb{1} - \mathbb{P}_j) \prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}\,\vec{z}, \tag{A23}$$

with the order of matrices in the product being irrelevant.[8]

It is straightforward to convince oneself that Theorem 1 of Sec. IV B follows from Theorem 2. Both are in fact equivalent, with the latter being a more rigorous version for parallel configurations of gates $\mathcal{G}$ formulated in the operational framework for interferometric experiments. The proof of Theorem 2 is given in Appendix B.

APPENDIX B: PROOFS

In the following we prove our main result Theorem 2 stated in Appendix A about transformation properties of classes $[\,\vec{z}\,] \subset \mathcal{P}(\Lambda)$ implemented by a single step in the circuit. Note that classes $[\,\vec{z}\,]$ are defined in Eq. (A16) as convex combinations of distributions from auxiliary classes $[\,\vec{z}\,]_i$ with supports in subsets $\Lambda^i_{\vec{z}}$ [see Eqs. (A11) and (A14)]. It is thus natural to look first at the evolution of distributions in auxiliary classes $[\,\vec{z}\,]_i$ and in particular their supports $\Lambda^i_{\vec{z}}$. We give such a description in Lemma 1. Then, by linearity of transformations, we shall extend this result to arbitrary distributions in $[\,\vec{z}\,]$ concluding with the proof of Theorem 2.

Remark. Recall that gates defined in the model act locally in the respective path of the circuit with the interaction between the paths allowed only on the beam splitters. All the gates are deterministic apart from the beam splitters which introduce a stochastic component into the evolution. Each step in the circuit is defined by the parallel configurations of gates $\mathcal{G} \equiv \{\,\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\,\}$ whose individual (local) action is given in Eqs. (7)–(13). Joint action of the gates in the respective paths combines into the evolution of the whole ontic state space Λ as discussed in Appendix A 3 a; see Eqs. (A7)–(A10).

Remark. In the proofs, we consider N to be fixed and assume all complex vectors $\vec{z} = \sum_j z_j\,\boldsymbol{e}_j \in \mathbb{C}^N$ to be normalized, i.e., $\|\vec{z}\| = \sum_j |z_j|^2 = 1$.

1. Helpful Lemma

We begin with a technical lemma describing transformation of distributions in auxiliary classes $[\,\vec{z}\,]_i \subset \mathcal{P}(\Lambda)$ defined in Eq. (A14) implemented by a single step in the circuit. See Fig. 5 for illustration.

Lemma 1. Suppose that the system is described by a distribution $\boldsymbol{p} \in [\,\vec{z}\,]_i$. Then, the parallel configuration of gates specified by $\mathcal{G} \equiv \{\,\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\,\}$ implements a transformation with the following properties.

– If there is a detector placed in the ith path (i.e., when $i \in \mathcal{D}$), then D_i CLICKS with certainty and afterwards the system is described by the distribution $\boldsymbol{p}' \in [\,\boldsymbol{e}_i\,]_i$, i.e., we get

$$[\,\vec{z}\,]_i \ni \boldsymbol{p} \xrightarrow{D_i} \boldsymbol{p}' \in [\,\boldsymbol{e}_i\,]_i \quad (\text{CLICK}), \tag{B1}$$

and all other detectors D_j with $j \neq i$ remain silent.

– If there is no detector in the ith path (i.e., when $i \notin \mathcal{D}$), then none of the detectors CLICK and afterwards the system is described by a distribution $\boldsymbol{p}'$ characterized by

[7] In the above we have assumed no prior knowledge of the source and hence the need of initial filtering of the unknown ensemble. We note that it could have been bypassed if the agent was granted access to a single-particle source with all paths blocked except one (as is usually assumed in the quantum scenarios). This can be easily realized within the model by postulating that the source injects particles (with nonvanishing amplitudes and strengths) into a given path and the blocks remove particles, resetting the strength of the field to zero. It can be observed that it boils down to preparation of distributions in one of the classes in Eq. (A19) again.

[8] Due to nonunitary projections, the length of $\vec{z}\,'$ in Eq. (A23) may be less than 1. Hence the proportionality symbol "$\sim$" expresses the need for subsequent renormalization $\vec{z}\,' \rightsquigarrow \vec{z}\,'/\|\vec{z}\,'\|$ [cf. analogous issue in the description of quantum measurement in Eq. (4)].

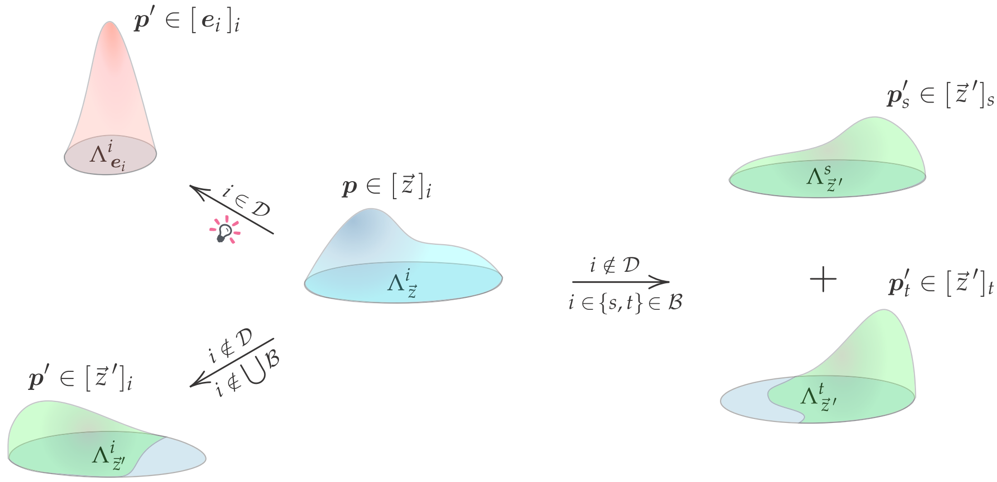

FIG. 5. Evolution of distributions in auxiliary classes $[\vec{z}]_i \subset \mathcal{P}(\Lambda)$. Distributions $\boldsymbol{p} \in [\vec{z}]_i$ are defined to have support in $\Lambda^i_{\vec{z}}$ and their evolution $\boldsymbol{p} \longrightarrow \boldsymbol{p}'$ depends on the composition of gates $\mathcal{G} \equiv \{\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\}$. If there is a detector in the ith path (i.e., for $i \in \mathcal{D}$), then it CLICKS and afterwards $\boldsymbol{p}' \in [\boldsymbol{e}_i]_i$ (at the top left). In the remaining cases (i.e., for $i \notin \mathcal{D}$) all the detectors remain silent and the evolution is as follows. If the ith path does not go into a beam splitter (i.e., for $i \notin \bigcup\mathcal{B}$), then we get $\boldsymbol{p}' \in [\vec{z}']_i$. Otherwise, if the ith path goes into the beam splitter B_{st} (i.e., for $i \in \{s,t\} \in \mathcal{B}$), then distribution $\boldsymbol{p}$ evolves into a mixture of distributions $\boldsymbol{p}'_s \in [\vec{z}']_s$ and $\boldsymbol{p}'_t \in [\vec{z}']_t$, which is weighted with probabilities proportional to $|z'_s|^2$ and $|z'_t|^2$.

vector[8]

$$\vec{z}' \sim \prod_{j\in\mathcal{D}}(\mathbb{1}-\mathbb{P}_j)\prod_{k\in\mathcal{S}}\mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}}\mathbb{B}_{st}\ \vec{z}, \tag{B2}$$

with the following two cases.

- In the case when the ith path does not go into a beam splitter (i.e., when $i \notin \bigcup\mathcal{B}$), then we get

$$[\vec{z}]_i \ni \boldsymbol{p} \longrightarrow \boldsymbol{p}' \in [\vec{z}']_i. \tag{B3}$$

- Otherwise, if the ith path goes into the beam splitter B_{st} (i.e., when $i \in \{s,t\} \in \mathcal{B}$), then

$$[\vec{z}]_i \ni \boldsymbol{p} \longrightarrow \boldsymbol{p}' = \frac{|z'_s|^2}{|z'_s|^2+|z'_t|^2}\,\boldsymbol{p}'_s + \frac{|z'_t|^2}{|z'_s|^2+|z'_t|^2}\,\boldsymbol{p}'_t, \tag{B4}$$

with distributions $\boldsymbol{p}'_s \in [\vec{z}']_s$ and $\boldsymbol{p}'_t \in [\vec{z}']_t$. It holds for $|z'_s|^2+|z'_t|^2 \neq 0$.

Proof of Lemma 1. Throughout the proof we take $\boldsymbol{p} \in [\vec{z}]_i$. From the definition of Eq. (A14), the ontic state $(q,\vec{u},\vec{\tau}) \in \Lambda$ of the system described by distribution $\boldsymbol{p}$ is certainly in the subset $\Lambda^i_{\vec{z}} \subset \Lambda$, meaning that it satisfies three conditions of Eq. (A11),

$$(a)\quad q=i,\quad (b)\quad \tau_i=\tau>0,\quad (c)\quad \Delta_\tau\,\vec{u}\sim\vec{z}, \tag{B5}$$

where

$$\tau := \max\{\tau_1,\dots,\tau_N\}, \tag{B6}$$

$$\Delta_\tau := \mathrm{diag}\,(\delta_{\tau_1\tau},\dots,\delta_{\tau_N\tau}). \tag{B7}$$

In the following we seek the form of distribution $\boldsymbol{p}'$ obtained from $\boldsymbol{p}$ as a result of transformation implemented by the parallel configuration of gates $\mathcal{G} \equiv \{\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\}$. Our strategy is to take an ontic state in support of $\boldsymbol{p}$, i.e., satisfying conditions in Eq. (B5), and check its properties after the transformation. This provides knowledge about the support of distribution $\boldsymbol{p}'$ which compared with conditions in Eqs. (A11) and (A14) will prove the result.

First part: Eq. (B1) (Case $i \in \mathcal{D}$). We begin by noting that $\mathrm{Prob}(q=i)=1$ and $\mathrm{Prob}(q\neq i)=0$. This means that detector D_i placed in the ith path CLICKS with certainty and detectors in other paths D_j with $j\neq i$ remain silent (since by definition detectors react only to the particle present in the respective path). Moreover, after detection the particle remains in the same path

$$q \longrightarrow q' = q = i. \tag{B8}$$

Second, along with a CLICK, detector D_i modifies the amplitude and strength of the field in the ith path as described by Eqs. (9) and (10), i.e., we get

$$u_i \longrightarrow u'_i = 1,\quad \tau_i \longrightarrow \tau'_i = 1. \tag{B9}$$

Third, a quick look at Eqs. (7)–(12) reveals that the strength of the fields in other paths decreases, which entails that $\tau_m \longrightarrow \tau'_m < 1$ for $m \neq i$. Together with the previous equation it gives

$$\tau' := \max\{\tau'_1,\dots,\tau'_n\} = 1 = \tau'_i, \tag{B10}$$

and hence after the transformation we obtain $\Delta_{\tau'} := \mathrm{diag}\,(\delta_{\tau'_1\tau'},\dots,\delta_{\tau'_N\tau'}) = \mathrm{diag}\,(0,\dots,1,\dots 0)$ with a single 1 in the ith place. This gives the identity

$$\Delta_{\tau'}\,\vec{u}' = \boldsymbol{e}_i. \tag{B11}$$

Putting all this together, we infer that for any configuration of gates with D_i in the ith path (i.e., for $i \in \mathcal{D}$) after the transformation the system is left in the ontic state $(q',\vec{u}',\vec{\tau}')$ satisfying the conditions

$$(a)\quad q'=i,\quad (b)\quad \tau'_i=\tau'>0,\quad (c)\quad \Delta_{\tau'}\vec{u}'\sim\boldsymbol{e}_i. \tag{B12}$$

In consequence, any distribution $\boldsymbol{p} \in [\,\vec{z}\,]_i$ is transformed into a distribution $\boldsymbol{p}' \in [\,\boldsymbol{e}_i\,]_i$ [see Eqs. (A11) and (A14)]. This proves the first part of Lemma 1, Eq. (B1).

Second part: Eqs. (B2)–(B4) (Case $i \notin \mathcal{D}$). We look at the second part of Lemma 1 when there is no detector in the ith path (i.e., $i \notin \mathcal{D}$). Clearly, in this situation all detectors remain silent (NO CLICK), since the particle is in the ith path (i.e., $q = i \notin \mathcal{D}$).

Let us begin by writing explicitly how the strength of the field changes in each path for a given configuration of gates $\mathcal{G} \equiv \{\,\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\,\}$. From Eqs. (7)–(12) we have

$$\tau_l \longrightarrow \tau_l' = {}^{\tau_l}\!/2 \qquad \text{for} \quad l \in \mathcal{F}, \tag{B13}$$

$$\tau_k \longrightarrow \tau_k' = {}^{\tau_k}\!/2 \qquad \text{for} \quad k \in \mathcal{S}, \tag{B14}$$

$$\tau_j \longrightarrow \tau_j' = 0 \qquad \text{for} \quad j \in \mathcal{D}, \tag{B15}$$

$$\tau_r \longrightarrow \tau_r' = {}^{\tau^{(st)}}\!/2 \qquad \text{for} \quad r \in \{s,t\} \in \mathcal{B}, \tag{B16}$$

where $\tau^{(st)} := \max\{\tau_s,\tau_t\}$. Along with the defining condition $\tau_i = \tau$ in Eq. (B5) and Eq. (B6), it entails that $\tau_m' \leqslant {}^{\tau}\!/2$ for all paths $m = 1,\ldots,N$. Furthermore, it follows that $\tau_i' = {}^{\tau_i}\!/2$ (since $i \notin \mathcal{D}$ and in case $i \in \{r,s\} \in \mathcal{B}$ we have $\tau^{(st)} = \tau_i$). Therefore, we get

$$\tau' := \max\{\tau_1',\ldots,\tau_N'\} = \tau_i' = {}^{\tau}\!/2 > 0. \tag{B17}$$

This fact, together with Eqs. (B13)–(B16), will be used in the following analysis to compare $\Delta_{\tau'} := \mathrm{diag}\,(\delta_{\tau_1'\tau'},\ldots,\delta_{\tau_N'\tau'})$ with Δ_τ of Eq. (B7).

Next, we investigate transformation properties of the vector of field amplitudes $\vec{u}$. Since action of each gate is limited to the path(s) it is attached to, the effect of each separate gate can be written in the following way [see Eqs. (7) and (12)]:

$$\vec{u} \xrightarrow{free} \vec{u}' = \vec{u}, \tag{B18}$$

$$\vec{u} \xrightarrow{S_k} \vec{u}' = \mathbb{S}_k\,\vec{u}, \tag{B19}$$

$$\vec{u} \xrightarrow{D_j} \vec{u}' = \vec{u}, \tag{B20}$$

$$\vec{u} \xrightarrow{B_{st}} \vec{u}' = \mathbb{B}_{st}\,\Delta_\tau^{(st)}\,\vec{u}, \tag{B21}$$

where

$$\Delta_\tau^{(st)} := \mathrm{diag}\,(1,\ldots,\delta_{\tau_s\tau^{(st)}},\ldots,\delta_{\tau_t\tau^{(st)}},\ldots,1), \tag{B22}$$

with all 1's on the diagonal except entries s and t which depend on $\tau^{(st)} := \max\{\tau_s,\tau_t\}$. Recall that we consider the case $q = i \notin \mathcal{D}$, and hence for all $j \in \mathcal{D}$ we have $q \neq j$, which explains the trivial action of the detectors in Eq. (B20). Taking all this together, joint transformation implemented by a parallel configuration of gates $\mathcal{F}$, $\mathcal{D}$, $\mathcal{S}$, and $\mathcal{B}$ boils down to the product

$$\vec{u} \longrightarrow \vec{u}' = \prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}\,\Delta_\tau^{(st)}\,\vec{u}. \tag{B23}$$

Now, we will justify the following matrix identity:

$$\underbrace{\Delta_{\tau'}\Bigg(\prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}\,\Delta_\tau^{(st)}\Bigg)}_{\mathbb{L}} = \underbrace{\Bigg(\prod_{j\in\mathcal{D}}(\mathbb{1}-\mathbb{P}_j)\prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}\Bigg)\Delta_\tau}_{\mathbb{R}}, \tag{B24}$$

where $\Delta_{\tau'} := \mathrm{diag}\,(\delta_{\tau_1'\tau'},\ldots,\delta_{\tau_N'\tau'})$. For proof, we observe that on both sides all matrices in the products are diagonal except for matrices $\mathbb{B}_{st}$ with 2×2 blocks acting in entries $\{s,t\}$ (without overlap for different $\mathbb{B}_{st}$). Therefore, we have the same block-diagonal structure of the matrix both on the left ($\mathbb{L}$) and on the right ($\mathbb{R}$) hand side, which consists of 1×1 blocks in entries $l \in \mathcal{F}$, $j \in \mathcal{D}$, $k \in \mathcal{S}$ and 2×2 blocks in entries $\{s,t\} \in \mathcal{B}$. Thus it is enough to verify each block separately in the identity. For 1×1 blocks, in the respective entries we have

$$\mathbb{L}_{ll} = \delta_{\tau_l'\tau'}\cdot 1 \overset{(B13)}{\underset{(B17)}{=}} 1\cdot\delta_{\tau_l\tau} = \mathbb{R}_{ll}, \tag{B25}$$

$$\mathbb{L}_{kk} = \delta_{\tau_k'\tau'}\cdot(\mathbb{S}_k)_{kk} \overset{(B14)}{\underset{(B17)}{=}} (\mathbb{S}_k)_{kk}\cdot\delta_{\tau_k\tau} = \mathbb{R}_{kk}, \tag{B26}$$

$$\mathbb{L}_{jj} = \delta_{\tau_j'\tau'}\cdot 1 \overset{(B15)}{\underset{(B17)}{=}} 0 \overset{(A5)}{=} 0\cdot\delta_{\tau_j\tau} = \mathbb{R}_{jj}. \tag{B27}$$

For 2×2 blocks, in the subspace restricted to the respective entries $\{r,s\} \in \mathcal{B}$, we have

$$\mathbb{L}_{\{s,t\}} = \begin{pmatrix}\delta_{\tau_s'\tau'} & 0\\ 0 & \delta_{\tau_t'\tau'}\end{pmatrix}\begin{pmatrix}i\sqrt{R} & \sqrt{T}\\ \sqrt{T} & i\sqrt{R}\end{pmatrix}\begin{pmatrix}\delta_{\tau_s\tau^{(st)}} & 0\\ 0 & \delta_{\tau_t\tau^{(st)}}\end{pmatrix} \tag{B28}$$

and

$$\mathbb{R}_{\{s,t\}} = \begin{pmatrix}i\sqrt{R} & \sqrt{T}\\ \sqrt{T} & i\sqrt{R}\end{pmatrix}\begin{pmatrix}\delta_{\tau_s\tau} & 0\\ 0 & \delta_{\tau_t\tau}\end{pmatrix}. \tag{B29}$$

In order to show $\mathbb{L}_{\{s,t\}} = \mathbb{R}_{\{s,t\}}$ we use Eqs. (B16) and (B17) to check the following three cases.

Case $\tau_s\,,\tau_t < \tau$. We have $\tau_s' = \tau_t' = {}^{\tau^{(st)}}\!/2 < {}^{\tau}\!/2 = \tau'$, and hence

$$\mathbb{L}_{\{s,t\}} = 0 = \mathbb{R}_{\{s,t\}}. \tag{B30}$$

Case $\tau_s = \tau_t = \tau$. We have $\tau_s' = \tau_t' = {}^{\tau^{(st)}}\!/2 = {}^{\tau}\!/2 = \tau'$, which gives

$$\mathbb{L}_{\{s,t\}} = \begin{pmatrix}i\sqrt{R} & \sqrt{T}\\ \sqrt{T} & i\sqrt{R}\end{pmatrix} = \mathbb{R}_{\{s,t\}}. \tag{B31}$$

Case $\tau_s < \tau_t = \tau$. We have $\tau_s' = \tau_t' = {}^{\tau^{(st)}}\!/2 = {}^{\tau}\!/2 = \tau'$, and consequently

$$\mathbb{L}_{\{s,t\}} = \begin{pmatrix}i\sqrt{R} & \sqrt{T}\\ \sqrt{T} & i\sqrt{R}\end{pmatrix}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = \mathbb{R}_{\{s,t\}}. \tag{B32}$$

(For the case $\tau_t < \tau_s = \tau$ similar reasoning holds.)

Having verified the identity in Eq. (B24) we can write

$$\Delta_{\tau'}\vec{u}' \overset{(B23)}{\underset{(B24)}{=}} \left(\prod_{j\in\mathcal{D}}(\mathbb{1}-\mathbb{P}_j)\prod_{k\in\mathcal{S}}\mathbb{S}_k\prod_{\{s,t\}\in\mathcal{B}}\mathbb{B}_{st}\right)\Delta_\tau\vec{u}$$

$$\overset{(B5)}{\sim} \prod_{j\in\mathcal{D}}(\mathbb{1}-\mathbb{P}_j)\prod_{k\in\mathcal{S}}\mathbb{S}_k\prod_{\{s,t\}\in\mathcal{B}}\mathbb{B}_{st}\,\vec{z}\,, \tag{B33}$$

which proves that after transformation the following condition holds:

$$(c)\quad \Delta_{\tau'}\vec{u}' \sim \vec{z}\,', \tag{B34}$$

with vector $\vec{z}\,'$ given in Eq. (B2).

Finally, let us check the position of the particle after the transformation if we know that at the beginning it is in path $q=i$. Clearly, change of the path is possible only on a beam splitter and hence we have two cases.

Case 1. If the ith path does not go into a beam splitter (i.e., for $i\notin\bigcup\mathcal{B}$), then the particle remains in the path $q\longrightarrow q'=q=i$. Hence, together with Eq. (B17), we get

$$(a)\quad q'=i,\quad (b)\quad \tau_i'=\tau'>0. \tag{B35}$$

A quick comparison of Eqs. (B34) and (B35) with definitions in Eqs. (A11) and (A14) reveals that, in that case, distribution $\boldsymbol{p}\in[\,\vec{z}\,]_i$ becomes transformed into a distribution with support in $\Lambda^i_{\vec{z}\,'}$, meaning that $\boldsymbol{p}'\in[\,\vec{z}\,']_i$. This proves Eq. (B3) with vector $\vec{z}\,'$ given in Eq. (B2).

Case 2. Otherwise, if the ith path crosses with another path on the beam splitter $\mathbb{B}_{st}$, i.e., for $i\in\{s,t\}\in\mathcal{B}$, then the particle may change its position either to path s or t with the respective probabilities specified by Eq. (13), i.e., we get

$$q\longrightarrow\begin{cases} q'=s & \text{with probability } \frac{|u_s'|^2}{|u_s'|^2+|u_t'|^2},\\ q'=t & \text{with probability } \frac{|u_t'|^2}{|u_s'|^2+|u_t'|^2}.\end{cases} \tag{B36}$$

Let us note that in this case Eqs. (B16) and (B17) entail that

$$(b)\quad \tau_s'=\tau_t'=\tau'>0. \tag{B37}$$

Specifically, this means that $\Delta_{\tau'}$ restricted to entries $\{s,t\}$ equals the identity. Therefore, from Eq. (B34) we have

$$\begin{pmatrix}u_s'\\u_t'\end{pmatrix}\sim\begin{pmatrix}z_s'\\z_t'\end{pmatrix}, \tag{B38}$$

which allows one to replace u's with z's in Eqs. (B36) whenever $|z_s'|^2+|z_t'|^2\neq 0$. We thus obtain

$$(a)\quad\begin{cases} q'=s & \text{with probability } \frac{|z_s'|^2}{|z_s'|^2+|z_t'|^2},\\ q'=t & \text{with probability } \frac{|z_t'|^2}{|z_s'|^2+|z_t'|^2}.\end{cases} \tag{B39}$$

This result, along with Eqs. (B34) and (B37), should be compared with definitions in Eqs. (A11) and (A14). We can conclude that, in the case of the beam splitter placed in the ith path, the system initially described by distribution $\boldsymbol{p}\in[\,\vec{z}\,]_i$ after the transformation has support in $\Lambda^s_{\vec{z}\,'}\cup\Lambda^t_{\vec{z}\,'}$, with cumulative probability over the respective sets given by Eq. (B39). In other words, the system is described by a probabilistic mixture

$$\boldsymbol{p}'=\frac{|z_s'|^2}{|z_s'|^2+|z_t'|^2}\,\boldsymbol{p}_s'+\frac{|z_t'|^2}{|z_s'|^2+|z_t'|^2}\,\boldsymbol{p}_t', \tag{B40}$$

with $\boldsymbol{p}_s'\in[\,\vec{z}\,']_s$ and $\boldsymbol{p}_t'\in[\,\vec{z}\,']_t$, which proves Eq. (B4) with vector $\vec{z}\,'$ given in Eq. (B2). ■

2. Proof of Theorem 2

Now, we are ready to prove our main result.

Proof of Theorem 2. Let us consider a situation where the system is described by a distribution $\boldsymbol{p}\in[\,\vec{z}\,]$. Recall that Definition 1, Eq. (A16) specifies distributions in $[\,\vec{z}\,]$ as convex combinations of distributions in auxiliary classes $\boldsymbol{p}_i\in[\,\vec{z}\,]_i$, i.e., we have

$$\boldsymbol{p}=\sum_{i=1}^{N}p_i\,\boldsymbol{p}_i, \tag{B41}$$

with

$$p_i=|z_i|^2,\quad \text{supp}\,\boldsymbol{p}_i\subset\Lambda^i_{\vec{z}}. \tag{B42}$$

Clearly, since $\vec{z}$ is normalized we have $\sum_i p_i=\|\vec{z}\|^2=1$.

In the following we are interested in the shape of distribution $\boldsymbol{p}'$ which is obtained from $\boldsymbol{p}$ because of transformation via the parallel configuration of gates $\mathcal{G}\equiv\{\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\}$. For proof of Theorem 2 we will find convex decomposition of $\boldsymbol{p}'=\sum_i p_i'\,\boldsymbol{p}_i'$ into a mixture of distributions $\boldsymbol{p}_i'\in[\,\vec{z}\,']_i$ and then compare it with Definition 1, Eq. (A16).

We begin with two simple observations about distribution $\boldsymbol{p}$ in Eq. (B41). Since distributions $\boldsymbol{p}_i$ have disjoint supports, see Eq. (A13), and $q=i$ only for the ontic state $(q,\vec{u},\vec{\tau})\in\Lambda$ in support of $\boldsymbol{p}_i$, then we have

$$\text{Prob}\,(q=i)=p_i, \tag{B43}$$

$$\text{Prob}\,(q,\vec{u},\vec{\tau}\,|\,q=i)=\boldsymbol{p}_i(q,\vec{u},\vec{\tau}). \tag{B44}$$

Furthermore, the sum in Eq. (B41) can be split into four groups by collecting together the terms associated with the same kind of gate attached to the relevant path, i.e.,

$$\boldsymbol{p}=\sum_{l\in\mathcal{F}}p_l\,\boldsymbol{p}_l+\sum_{j\in\mathcal{D}}p_j\,\boldsymbol{p}_j+\sum_{k\in\mathcal{S}}p_k\,\boldsymbol{p}_k+\sum_{r\in\bigcup\mathcal{B}}p_r\,\boldsymbol{p}_r, \tag{B45}$$

since $\mathcal{F}$, $\mathcal{D}$, $\mathcal{S}$, $\bigcup\mathcal{B}$ partition the set of all paths labeled with $\{1,\ldots,N\}$.

Throughout the proof we will use the following auxiliary notation:

$$\vec{w}\,':=\prod_{j\in\mathcal{D}}(\mathbb{1}-\mathbb{P}_j)\prod_{k\in\mathcal{S}}\mathbb{S}_k\prod_{\{s,t\}\in\mathcal{B}}\mathbb{B}_{st}\;\vec{z}\,, \tag{B46}$$

which relates to vector $\vec{z}\,'$ in Eqs. (A23) and (B2) as follows:

$$\vec{z}\,'={}^{\vec{w}'}\!/\!_{\|\vec{w}'\|}. \tag{B47}$$

Recall that in both Theorem 2 and Lemma 1 we use the convention in which vector $\vec{z}\,'$ is normalized.[8]

First part: Eqs. (A21) and (A22). Because detectors react only to particles in a given path we get that detector D_j CLICKS if and only if $q=j$. From Eqs. (B42) and (B43) it happens with probability

$$\text{Prob}\,(D_j\,|\,\vec{z})=\text{Prob}\,(q=j)=|z_j|^2. \tag{B48}$$

Note that because the system is always in a well-defined ontic state (which means that the position of the particle is definite), simultaneous detection in different detectors at the same time (i.e., more than one CLICK) is impossible. Moreover, a negative outcome in all detectors NO CLICK occurs only if $q \notin \mathcal{D}$, which happens with probability equal to $\text{Prob}\,(\varnothing|\,\vec{z}\,) = 1 - \sum_{j\in\mathcal{D}} \text{Prob}\,(q = j) = 1 - \sum_{j\in\mathcal{D}} |z_j|^2$.

To conclude, we observe that a CLICK in detector D_j provides additional knowledge that $q = j$. This entails update of the initial probability distribution $\boldsymbol{p}$ to the conditional distribution $\boldsymbol{p}_j \in [\,\vec{z}\,]_j$ given in Eq. (B44). Then from Lemma 1, Eq. (B1) we infer that after detection (i.e., postselection on a CLICK in detector D_j) the system is described by a distribution $\boldsymbol{p}' \in [\,\boldsymbol{e}_j\,]_j$. Since in this case from Definition 1, Eq. (A16) we have $[\,\boldsymbol{e}_j\,]_j = [\,\boldsymbol{e}_j\,]$, then

$$[\,\vec{z}\,] \ni \boldsymbol{p} \xrightarrow{D_j} \boldsymbol{p}' \in [\,\boldsymbol{e}_j\,] \quad (\text{CLICK}). \tag{B49}$$

This proves the first part of Theorem 2, Eqs. (A21) and (A22).

Second part: Eq. (A23). Now we consider a case when none of the detectors CLICKS (i.e., $q \notin \mathcal{D}$) or there are no detectors at all (i.e., $\mathcal{D} = \varnothing$). Observe that since distributions $\boldsymbol{p}_i \in [\,\vec{z}\,]_i$ in Eq. (B45) are supported in $\Lambda^i_{\vec{z}}$ (for which $q = i$), then additional knowledge of $q \notin \mathcal{D}$ entails an update of the initial probability distribution $\boldsymbol{p}$ to the following form:

$$\boldsymbol{p} \longrightarrow \tilde{\boldsymbol{p}} = \sum_{l\in\mathcal{F}} \tilde{p}_l\, \boldsymbol{p}_l + \sum_{k\in\mathcal{S}} \tilde{p}_k\, \boldsymbol{p}_k + \sum_{r\in\bigcup\mathcal{B}} \tilde{p}_r\, \boldsymbol{p}_r, \tag{B50}$$

where we have used renormalized coefficients,

$$\tilde{p}_i = \frac{p_i}{\sum_{j\notin\mathcal{D}} p_j} \overset{(B42)}{=} \frac{|z_i|^2}{\sum_{j\notin\mathcal{D}} |z_j|^2}, \tag{B51}$$

for $i \in \mathcal{F} \cup \mathcal{S} \cup \bigcup\mathcal{B}$. It will be convenient to rewrite the last sum more explicitly as

$$\tilde{\boldsymbol{p}} = \sum_{l\in\mathcal{F}} \tilde{p}_l\, \boldsymbol{p}_l + \sum_{k\in\mathcal{S}} \tilde{p}_k\, \boldsymbol{p}_k + \sum_{\{s,t\}\in\mathcal{B}} (\tilde{p}_s\, \boldsymbol{p}_s + \tilde{p}_t\, \boldsymbol{p}_t). \tag{B52}$$

In the following we are interested in the transformation of $\tilde{\boldsymbol{p}}$ under action of the parallel configuration of gates $\mathcal{G} \equiv \{\,\mathcal{F},\mathcal{D},\mathcal{S},\mathcal{B}\,\}$. Since supports of distributions $\boldsymbol{p}_i$ are disjoint, we can simplify our task by individually analyzing each term in Eq. (B52) and then collecting together all the results.

From Lemma 1, Eq. (B3) for $l \in \mathcal{F}$ and $k \in \mathcal{S}$, we get

$$\begin{aligned} \tilde{p}_l\, \boldsymbol{p}_l &\longrightarrow \tilde{p}_l\, \boldsymbol{p}'_l, \\ \tilde{p}_k\, \boldsymbol{p}_k &\longrightarrow \tilde{p}_k\, \boldsymbol{p}'_k, \end{aligned} \tag{B53}$$

where $\boldsymbol{p}'_l \in [\,\vec{z}\,']_l$ and $\boldsymbol{p}'_k \in [\,\vec{z}\,']_k$ with label $\vec{z}\,'$ given in Eq. (B2). As for the coefficients $\tilde{p}_i$ defined in Eq. (B51), we observe that

$$\begin{aligned} \sum_{j\notin\mathcal{D}} |z_j|^2 &= \Big\| \prod_{j\in\mathcal{D}} (1 - \mathbb{P}_j)\, \vec{z}\, \Big\|^2 \\ &= \Big\| \prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st} \prod_{j\in\mathcal{D}} (\mathbb{1} - \mathbb{P}_j)\, \vec{z}\, \Big\|^2 \\ &\overset{(B46)}{=} \|\vec{w}\,'\|^2, \end{aligned} \tag{B54}$$

where the first equality comes from the definition of projectors $\mathbb{P}_j$, the second is due to preservation of the norm under unitary $\prod_{k\in\mathcal{S}} \mathbb{S}_k \prod_{\{s,t\}\in\mathcal{B}} \mathbb{B}_{st}$, and the third draws on commutativity of all factors in the product (matrices are either diagonal or block-diagonal with nonoverlapping blocks). Now, because of the fact that the product in Eq. (B46) consists of factors acting separately in the respective components, we have

$$\begin{aligned} w'_l &= z_l \quad \text{for } l \in \mathcal{F}, \\ |w'_k| &= |z_k| \quad \text{for } k \in \mathcal{S}. \end{aligned} \tag{B55}$$

Substitution of Eqs. (B54) and (B55) into Eq. (B51), gives

$$\begin{aligned} \tilde{p}_l &= \frac{|w'_l|^2}{\|\vec{w}\,'\|^2} \overset{(B47)}{=} |z'_l|^2 =: p'_l, \\ \tilde{p}_k &= \frac{|w'_k|^2}{\|\vec{w}\,'\|^2} \overset{(B47)}{=} |z'_k|^2 =: p'_k. \end{aligned} \tag{B56}$$

Therefore, for the first two sums in Eq. (B52), we get

$$\sum_{l\in\mathcal{F}} \tilde{p}_l\, \boldsymbol{p}_l + \sum_{k\in\mathcal{S}} \tilde{p}_k\, \boldsymbol{p}_k \longrightarrow \sum_{l\in\mathcal{F}} p'_l\, \boldsymbol{p}'_l + \sum_{k\in\mathcal{S}} p'_k\, \boldsymbol{p}'_k. \tag{B57}$$

Now, we proceed to the analysis of terms in the last sum in Eq. (B52). From Lemma 1, Eq. (B4), we obtain (note that we are interested in the case when $\tilde{p}_s + \tilde{p}_t \neq 0$, which entails $|z'_s|^2 + |z'_t|^2 \neq 0$)

$$\begin{aligned} \tilde{p}_s\, \boldsymbol{p}_s + \tilde{p}_t\, \boldsymbol{p}_t \longrightarrow\; & \tilde{p}_s \left(\frac{|z'_s|^2}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{q}'_s + \frac{|z'_t|^2}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{q}'_t \right) \\ & + \tilde{p}_t \left(\frac{|z'_s|^2}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{r}'_s + \frac{|z'_t|^2}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{r}'_t \right), \end{aligned} \tag{B58}$$

with $\boldsymbol{q}'_s,\boldsymbol{r}'_s \in [\,\vec{z}\,']_s$ and $\boldsymbol{q}'_t,\boldsymbol{r}'_t \in [\,\vec{z}\,']_t$ and the label $\vec{z}\,'$ given by formula (B2). By regrouping terms on the right side, we get

$$\begin{aligned} \tilde{p}_s\, \boldsymbol{p}_s + \tilde{p}_t\, \boldsymbol{p}_t \longrightarrow\; & \frac{|z'_s|^2}{|z'_s|^2 + |z'_t|^2} (\tilde{p}_s \boldsymbol{q}'_s + \tilde{p}_t \boldsymbol{r}'_s) \\ & + \frac{|z'_t|^2}{|z'_s|^2 + |z'_t|^2} (\tilde{p}_s \boldsymbol{q}'_t + \tilde{p}_t \boldsymbol{r}'_t), \end{aligned} \tag{B59}$$

and observe that

$$\boldsymbol{p}'_s := \frac{\tilde{p}_s \boldsymbol{q}'_s + \tilde{p}_t \boldsymbol{r}'_s}{\tilde{p}_s + \tilde{p}_t}, \quad \boldsymbol{p}'_t := \frac{\tilde{p}_s \boldsymbol{q}'_t + \tilde{p}_t \boldsymbol{r}'_t}{\tilde{p}_s + \tilde{p}_t} \tag{B60}$$

are properly normalized distributions with the property that $\boldsymbol{p}'_s \in [\,\vec{z}\,']_s$ and $\boldsymbol{p}'_t \in [\,\vec{z}\,']_t$. Therefore, Eq. (B58) can be rewritten in the form

$$\tilde{p}_s\, \boldsymbol{p}_s + \tilde{p}_t\, \boldsymbol{p}_t \longrightarrow \frac{|z'_s|^2(\tilde{p}_s + \tilde{p}_t)}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{p}'_s + \frac{|z'_t|^2(\tilde{p}_s + \tilde{p}_t)}{|z'_s|^2 + |z'_t|^2}\, \boldsymbol{p}'_t. \tag{B61}$$

A closer look at the first coefficient reveals that

$$\begin{aligned} \frac{|z'_s|^2(\tilde{p}_s + \tilde{p}_t)}{|z'_s|^2 + |z'_t|^2} &\overset{(B51)}{=} \frac{|z'_s|^2}{\sum_{j\notin\mathcal{D}} |z_j|^2} \frac{|z_s|^2 + |z_t|^2}{|z'_s|^2 + |z'_t|^2} \\ &\overset{(B47)(B54)}{=} \frac{|w'_s|^2}{\|\vec{w}\,'\|^2} \frac{|z_s|^2 + |z_t|^2}{|w'_s|^2 + |w'_t|^2} \\ &= \frac{|w'_s|^2}{\|\vec{w}\,'\|^2} \overset{(B47)}{=} |z'_s|^2 =: p'_s, \end{aligned} \tag{B62}$$

where in the penultimate equality the last fraction cancels out. The latter is due to the fact that the only nontrivial action on components $\{s,t\}$ in Eq. (B46) comes from the matrix $\mathbb{B}_{st}$, which gives

$$\begin{pmatrix} w'_s \\ w'_t \end{pmatrix} = \begin{pmatrix} i\sqrt{R} & \sqrt{T} \\ \sqrt{T} & i\sqrt{R} \end{pmatrix} \begin{pmatrix} z_s \\ z_t \end{pmatrix}, \tag{B63}$$

and since it is a unitary transform it preserves the norm $|w'_s|^2 + |w'_t|^2 = |z_s|^2 + |z_t|^2$. Clearly, the same reasoning applies to the second term in Eq. (B61), which equals to

$$\frac{|z'_t|^2(\tilde{p}_s + \tilde{p}_t)}{|z'_s|^2 + |z'_t|^2} = |z'_t|^2 =: p'_t. \tag{B64}$$

Hence, for the last sum in Eq. (B52), we get

$$\sum_{\{s,t\}\in\mathcal{B}} (\tilde{p}_s\, \boldsymbol{p}_s + \tilde{p}_t\, \boldsymbol{p}_t) \longrightarrow \sum_{\{s,t\}\in\mathcal{B}} (p'_s\, \boldsymbol{p}'_s + p'_t\, \boldsymbol{p}'_t). \tag{B65}$$

Having analyzed all terms in Eq. (B52) we use results from Eqs. (B56), (B57), (B62), (B64), and (B65), which together with Eq. (B50) provide the following decomposition:

$$\boldsymbol{p} \longrightarrow \boldsymbol{p}' = \sum_{i=1}^{N} p'_i\, \boldsymbol{p}'_i, \tag{B66}$$

where $\boldsymbol{p}'_i \in [\vec{z}\,']_i$ and $p'_i = |z'_i|^2$ with vector $\vec{z}\,'$ given in Eq. (B2). (Clearly, for $i \in \mathcal{D}$ we have $p'_i = |z'_i|^2 = 0$.) By comparing with Definition 1, Eq. (A16), we can determine that

$$[\vec{z}\,] \ni \boldsymbol{p} \longrightarrow \boldsymbol{p}' \in [\vec{z}\,'], \tag{B67}$$

which concludes the proof of the second part of Theorem 2, Eq. (A23). ■

[1] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Redwood City, 1965), Vol. III .

[2] G. Bacciagaluppi and A. Valentini, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference* (Cambridge University Press, Cambridge, UK, 2009).

[3] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[4] A. Einstein, Letter to Max Born, 3rd March 1947, in *The Born-Einstein Letters* (Macmillan, London, 1971).

[5] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[6] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, UK, 1987).

[7] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Bell's theorem without inequalities, Am. J. Phys. **58**, 1131 (1990).

[8] L. Hardy, Quantum Mechanics, Local Realistic Theories, and Lorentz-Invariant Realistic Theories, Phys. Rev. Lett. **68**, 2981 (1992).

[9] N. D. Mermin, Hidden variables and the two theorems of John Bell, Rev. Mod. Phys. **65**, 803 (1993).

[10] S. M. Tan, D. F. Walls, and M. J. Collett, Nonlocality of a Single Photon, Phys. Rev. Lett. **66**, 252 (1991).

[11] L. Hardy, Nonlocality of a Single Photon Revisited, Phys. Rev. Lett. **73**, 2279 (1994).

[12] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Nonlocality of a Single Photon? Phys. Rev. Lett. **75**, 2064 (1995).

[13] J. Dunningham and V. Vedral, Nonlocality of a Single Particle, Phys. Rev. Lett. **99**, 180404 (2007).

[14] V. Scarani and A. Suarez, Introducing quantum mechanics: One-particle interferences, Am. J. Phys. **66**, 718 (1998).

[15] A. Elitzur and L. Vaidman, Quantum mechanical interaction-free measurements, Found. Phys. **23**, 987 (1993).

[16] P. Kwiat, H. Weinfurter, T. Herzog, A. Zeilinger, and M. A. Kasevich, Interaction-Free Measurement, Phys. Rev. Lett. **74**, 4763 (1995).

[17] A. J. DeWeerd, Interaction-free measurement, Am. J. Phys. **70**, 272 (2002).

[18] B. Misra and E. C. G. Sudarshan, The Zeno's paradox in quantum theory, J. Math. Phys. **18**, 756 (1977).

[19] A. Peres, Zeno paradox in quantum theory, Am. J. Phys. **48**, 931 (1980).

[20] J. A. Wheeler, The 'Past' and the 'Delayed-Choice Double-Slit Experiment', in *Mathematical Foundations of Quantum Theory*, edited by A. R. Marlow (Academic Press, New York, 1978).

[21] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, Experimental realization of Wheeler's delayed-choice gedanken experiment, Science **315**, 966 (2007).

[22] A. J. Leggett and A. Garg, Quantum Mechanics Versus Macroscopic Realism: Is the Flux There When Nobody Looks? Phys. Rev. Lett. **54**, 857 (1985).

[23] C. Emary, N. Lambert, and F. Nori, Leggett-Garg inequalities, Rep. Prog. Phys. **77**, 016001 (2014).

[24] Y. Aharonov, D. Z. Albert, and L. Vaidman, How the Result of a Measurement of a Component of the Spin of a Spin-1/2 Particle Can Turn Out to be 100, Phys. Rev. Lett. **60**, 1351 (1988).

[25] R. E. George, L. M. Robledo, O. J. E. Maroney, M. S. Blok, H. Bernien, M. L. Markham, D. J. Twitchen, J. J. L. Morton, G. A. D. D. Briggs, and R. Hanson, Opening up three quantum boxes causes classically undetectable wavefunction collapse, Proc. Natl. Acad. Sci. USA **110**, 3777 (2013).

[26] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

[27] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[28] N. Harrigan and T. Rudolph, Ontological models and the interpretation of contextuality, arXiv:0709.4266 [quant-ph].

[29] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Simple Test for Hidden Variables in Spin-1 Systems, Phys. Rev. Lett. **101**, 020403 (2008).

[30] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wiesniak, and A. Zeilinger, Experimental non-classicality of an indivisible quantum system, Nature (London) **474**, 490 (2011).

[31] J. Thompson, P. Kurzyński, S.-Y. Lee, A. Soeda, and D. Kaszlikowski, Recent advances in contextuality tests, Open Syst. Inf. Dyn. **23**, 1650009 (2016).

[32] L. Hardy, Disentangling nonlocality and teleportation, arXiv:quant-ph/9906123.
[33] A. Daffertshofer, A. R. Plastino, and A. Plastino, Classical No-Cloning Theorem, Phys. Rev. Lett. **88**, 210601 (2002).
[34] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110 (2007).
[35] M. Kleinmann, O. Gühne, J. R. Portillo, J.-A. Larsson, and A. Cabello, Memory cost of quantum contextuality, New J. Phys. **13**, 113011 (2011).
[36] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction, Phys. Rev. A **86**, 012103 (2012).
[37] J. J. Wallman and S. D. Bartlett, Non-negative subtheories and quasiprobability representations of qubits, Phys. Rev. A **85**, 062121 (2012).
[38] P. Blasiak, Quantum cube: A toy model of a qubit, Phys. Lett. A **377**, 847 (2013).
[39] C. Ferrie and J. Combes, How the Result of a Single Coin Toss Can Turn Out to be 100 Heads, Phys. Rev. Lett. **113**, 120404 (2014).
[40] P. Blasiak, Classical systems can be contextual too: Analogue of the Mermin-Peres square, Ann. Phys. (N.Y.) **353**, 326 (2015).
[41] P. Blasiak, Local model of a qubit in the interferometric setup, New J. Phys. **17**, 113043 (2015).
[42] A. Karanjai, E. G. Cavalcanti, S. D. Bartlett, and T. Rudolph, Weak values in a classical theory with an epistemic restriction, New J. Phys. **17**, 073015 (2015).
[43] P. Blasiak, Ontological models with epistemic constraints: Local reconstruction of a dual-rail qubit, Ann. Phys. (N.Y.) **385**, 667 (2017).
[44] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental Realization of Any Discrete Unitary Operator, Phys. Rev. Lett. **73**, 58 (1994).
[45] N. Harrigan and R. W. Spekkens, Einstein, incompleteness, and the epistemic view of quantum states, Found. Phys. **40**, 125 (2010).
[46] I. Bengtsson and K. Życzkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, Cambridge, UK, 2006).
[47] M. Fuwa, S. Takeda, M. Zwierz, H. M. Wiseman, and A. Furusawa, Experimental proof of nonlocal wavefunction collapse for a single particle using homodyne measurements, Nat. Commun. **6**, 6665 (2015).
[48] M. S. Leifer and R. W. Spekkens, Pre- and Post-Selection Paradoxes and Contextuality in Quantum Mechanics, Phys. Rev. Lett. **95**, 200405 (2005).
[49] O. J. E. Maroney, Measurements, disturbances and the quantum three box paradox, Stud. Hist. Philos. Mod. Phys. **58**, 41 (2017).
[50] M. F. Pusey, Anomalous Weak Values are Proofs of Contextuality, Phys. Rev. Lett. **113**, 200401 (2014).
[51] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, Nat. Commun. **7**, 11780 (2016).
[52] E. G. Beltrametti and S. Bugajski, A classical extension of quantum mechanics, J. Phys. A **28**, 3329 (1995).
[53] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph, Distinct Quantum States Can Be Compatible with a Single State of Reality, Phys. Rev. Lett. **109**, 150404 (2012).
[54] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, Nat. Phys. **8**, 476 (2012).
[55] M. S. Leifer, Is the quantum state real? An extended review of ψ-ontology theorems, Quanta **3**, 67 (2014).
[56] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables. Parts I and II, Phys. Rev. **85**, 166 (1952).
[57] *Bohmian Mechanics and Quantum Theory: An Appraisal*, edited by J. T. Cushing, A. Fine, and S. Goldstein, Boston Studies in the Philosophy of Science Vol. 184 (Kluwer Academic Publishers, Dordrecht, 1996).
[58] B.-G. Englert, M. O. Scully, G. Süssmann, and H. Walther, Surrealistic Bohm trajectories, Z. Naturforsch. A **47**, 1175 (1992).
[59] L. Hardy, Contextuality in Bohmian mechanics, in *Bohmian Mechanics and Quantum Theory: An Appraisal*, edited by J. T. Cushing, A. Fine, and S. Goldstein, Boston Studies in the Philosophy of Science Vol. 184 (Springer, New York, 1996), pp. 67–76.
[60] E. T. Jaynes, Probability in Quantum Theory, in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Redwood City, CA, 1990), p. 381.
[61] B. F. Toner and D. Bacon, Communication Cost of Simulating Bell Correlations, Phys. Rev. Lett. **91**, 187904 (2003).